Indium Antimonide Plasmonic Nanostructures for Tunable Terahertz Sources

by

Trevor Blaikie

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics (Nanotechnology)

Waterloo, Ontario, Canada, 2022

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this work, procedures were successfully created and deployed for the development, characterization, and study of indium antimonide nanostructures as well as their terahertz plasmonic response. Using molecular beam epitaxy, indium antimonide was successfully grown on Gallium Arsenide (100) substrates of various surface misorientations. The resulting indium antimonide showed a reflection high energy electron diffraction pattern characteristic of a single-crystal epitaxial layer. These layers were then characterized through, Nomarski interference contrast microscopy, atomic force microscopy, high-resolution x-ray diffraction, and electron channeling contrast imaging. Cleaved samples from these growths were also used in developing a nanofabrication procedure to produce structures where the largest dimension was $4\mu m$. Ultraviolet photolithography and inductively coupled plasma etching were used to shape the indium antimonide material. Electromagnetic simulations were also carried out to demonstrate the tunable response of a localized surface plasmon resonance. The localized surface plasmon resonance frequency is demonstrated to depend on the temperature of the indium antimonide. This project will serve as a stepping stone for the pathway into the development of tunable indium antimonide terahertz plasmonic devices for use in conjunction with terahertz sources.

Acknowledgements

Don't be fooled by the fact that I am the sole author of this document, the completion of this research was a group effort and it could not have been done without my supervisor, professors, colleagues, friends and family.

First, I would like to express my deepest gratitude to *Professor* Zbig. Nearly two years ago when Professor Zbig welcomed me into his research group I was uncertain if graduate school would be the right place for me. Now, I find myself thriving under Professor Zbig's supervision as I extend my graduate studies into a PhD. That is thanks, in part, to the welcoming research environment that Professor Zbig cultivates within his research group as well as Professor Zbig's insatiable appetite for research, which has induced a similar response in myself.

In addition to Professor Zbig, I would also like to thank the others who directly contributed to this research. Many thanks go to my fellow group members Hadiya, Peyton, Ahmed, and Alan who all contributed in different ways to this research. Without Alan's help with the initial preparation of the wafers, none of the subsequent studies would have been completed. Hadiya contributed to the development of the methods used to map reciprocal space. Hadiya was also a great moral support, being at similar points in our programs Hadiya and I always had relatable experiences to share. Peyton's help in the collection and analysis of AFM, rocking curve, and ECCI data was invaluable. And finally, I would like to acknowledge Ahmed for his support in nanofabrication. Ahmed shared his superior experience with me in the cleanroom which greatly accelerated and improved overall the nanofabrication process.

Outside of the MBE research group, support was also received from the knowledgeable staff of the Brockhouse x-ray beamline at the CLS. I am grateful for the wonderful experiences I had at the CLS and I appreciate all of the "behind the scenes" work that goes into preparing for beamline users' experiments. Thank you Beatriz, Narayan and Al, for providing 'round the clock support and troubleshooting.

Plenty of others contributed to the completion of this work indirectly by providing me with support in many ways. I thank my friends who have always been around to provide a stress-relieving escape from my work. I tend to become all-consumed with large projects like this and without the occasional distraction and reminders to take breaks I might just work myself to death.

For supporting me in all of my decisions and being there whenever I needed her, for my whole life, I will forever be indebted to and forever appreciative of my mother.

Dedication

I dedicate this work to those who have come and gone.

Table of Contents

List of Figures **ix**

List of Tables **xv**

List of Abbreviations **xvi**

1 Introduction **1**

1.1 Motivations & Applications 1

1.2 Goals 2

2 Modelling & Simulation **4**

2.1 Theory 4

2.1.1 THz Radiation 4

2.1.2 Drude-Lorentz Model 5

2.1.3 Surface Plasmons 7

2.1.4 Indium Antimonide 12

2.2 Methods 15

2.2.1 COMSOL Multiphysics 15

2.2.2 The Simulated Model 15

2.3 Results 18

3 Molecular Beam Epitaxy **22**
3.1 Theory . . . 22
3.1.1 MBE Systems . . . 22
3.1.2 Reflection High Energy Electron Diffraction . . . 25
3.1.3 Other In-situ Characterization . . . 27
3.2 Methods . . . 29
3.2.1 Growing InSb . . . 29
3.3 Results . . . 32

4 Atomic Force Microscopy **35**
4.1 Theory . . . 35
4.2 Methods . . . 38
4.2.1 Surface Roughness . . . 38
4.2.2 Effects of Varying Offcuts . . . 40
4.3 Results . . . 41

5 X-Ray Diffraction **46**
5.1 Theory . . . 46
5.1.1 Crystal Structure . . . 47
5.1.2 Miller Indices . . . 47
5.1.3 Reciprocal Lattice . . . 49
5.1.4 Diffraction . . . 49
5.2 Methods . . . 51
5.2.1 High-resolution X-ray Diffraction . . . 51
5.2.2 Experimental Setups . . . 53
5.2.3 Reciprocal Space Maps . . . 54
5.2.4 Rocking Curves . . . 55
5.3 Results . . . 58
5.3.1 3D Reciprocal Space Maps . . . 59
5.3.2 Rocking Curves . . . 62

6 Nanofabrication **70**
6.1 Theory . . . 70
6.1.1 Sample Cleaning . . . 71
6.1.2 Spin Coating . . . 71
6.1.3 UV Photolithography . . . 73
6.1.4 Inductively Coupled Plasma Etching . . . 74
6.2 Methods . . . 76
6.2.1 Cleanroom . . . 76
6.2.2 Procedure . . . 76
6.3 Results . . . 80
6.3.1 Photoresist Patterns . . . 80
6.3.2 Etching . . . 81

7 Conclusion **85**

References **88**

APPENDICES **94**

A TE Polarization SPP Excitation **95**

B Reciprocal Space Maps **97**
B.1 Mapping to reciprocal space . . . 97
B.2 3D Plots . . . 100
B.3 2D Polar Plots . . . 101

C Electron Chanelling Contrast Imaging **103**

D Rocking Curve Fitting **106**

List of Figures

2.1 Diagram of the surface charge oscillations (electron oscillations) coupled with electric field oscillations in a SPP propagating along a dielectric-metal interface. Figure sourced from [1]. 8

2.2 A plot of the dispersion relations of the possible SPPs that could be excited at the interfaces of prism, metal, and air in the prism coupling technique. The straight lines are the light lines in the labelled material. Figure sourced from [2, pg. 43]. 10

2.3 The schematic of a simple prism coupling set up to excite SPPs. Figure adapted from [3]. 11

2.4 A simplified visual representation of how an oscillating electric field will interact with a metal nanoparticle to excite an LSP. Figure sourced from [4]. 12

2.5 A stereoscopic projection of the 3D model used in the simulation. There are three domains containing different materials. The mesh used in the simulation is displayed on the surface. 17

2.6 Plots of the imaginary (a) and real (b) parts of the refractive index for GaAs. Anywhere outside of the range of 0.2 THz to 2.0 THz was linearly extrapolated. The data for these plots was sourced from [5]. 18

2.7 Two plots of the total transmittance through the simulated nanostructure. (a) The total transmittance when the incident radiation is TE Polarized at different temperatures of 300 K and 315 K. (b) The total transmittance when the incident radiation is TM Polarized at different temperatures of 300 K and 315 K. 19

2.8 Four 2D plots of the amplitude of the electric field across a cross-section of the nanostructure when the incident light frequency is 1.56 THz with different polarizations and InSb temperatures. The colour indicates the amplitude according to the colour bar. The maximum and minimum values for the electric field amplitude is displayed above and below the colour bar, respectively. (a) TE Polarized at 300 K. (b) TE Polarized at 315 K. (c) TM Polarized at 300 K. (d) TM Polarized at 315 K. 20

3.1 A schematic of a typical MBE chamber equipped for RHEED analysis. Figure sourced from [6, pg. 113]. Used with permission of Taylor & Francis Group LLC - Books, from Heteroepitaxy of Semiconductors : Theory, Growth, and Characterization, Ayers, John E., Kujofsa, Tedi, Rago, Paul, Raphael, Johanna E., volume, 2nd edition, © 2017; permission conveyed through Copyright Clearance Center, Inc. Permission was also granted for the original figure from Molecular beam epitaxy: from research to manufacturing, Henini, M. © 1997, with permission from Elsevier. 24

3.2 RHEED images (a) during the initial stages of the growth, (b) Shows an intermediate stage where rings form indicting a polycrystalline surface, and (c) after the MBE growth of InSb is completed. 27

3.3 (a) A photograph of a Bayard-Alpert Ion Gauge. (b) A schematic of the Bayard-Alpert Ion Gauge. Figures sourced from [7, pg. 29]. Used with permission of Springer, from Surface science : an introduction, Oura, K., © 2003; permission conveyed through Copyright Clearance Center, Inc. 29

3.4 A schematic representation of the layers grown by MBE. 31

3.5 Three images of representative crystal surfaces. (a) A crystal surface cleaved with no misorientation to the (100) planes. (b) When the surface is cleaved with a misorientation from the (100) planes along a single crystallographic direction. (c) When the surface is cleaved with a misorientation from the (100) planes along two crystallographic directions. Figure sourced from [7, pg. 233]. Used with permission of Springer, from Surface science : an introduction, Oura, K., © 2003; permission conveyed through Copyright Clearance Center, Inc. 32

3.6 Three photographs from Nomarski interference contrast microscopy of the surfaces of all three wafers, G0845, G0846, and G0859. Each photo is taken with a 40 X lens. (a) Surface of G0845 (0° misorientation). (b) Surface of G0846 (2° misorientation). (c) Surface of G0859 (0.55° misorientation) . . 33

3.7 Three photographs showing the reflectance of the surfaces of all three wafers, G0845, G0846, and G0859. Each photo is taken with the same exposure time and along the same crystallographic orientation. (a) Surface of G0845 (0° misorientation). (b) Surface of G0846 (2° misorientation). (c) Surface of G0859 (0.55° misorientation). The primary flat is on the opposite side in (c) because it is the only EJ cut wafer of the three. 34

4.1 A graph of the force response curve from the combination of the long range attractive van der Waals forces and the short range repulsive Coulomb interaction forces between two atoms. Figure sourced from [7, pg. 166]. Used with permission of Springer, from Surface science : an introduction, Oura, K., © 2003; permission conveyed through Copyright Clearance Center, Inc. 36

4.2 39

4.3 (a) A composite of images from a Nomarski microscope showing the bowing of the G0845 substrate surface at the primary flat. The vertical lines that seem to bisect the entire image are the result of combining multiple images that had slightly different exposures and are not a feature of the surface. (b) An illustration emphasizing the rounded edge of a one side polished substrate wafer. 42

4.4 10 μm by 10 μm AFM scans of (a) G0845, (b) G0846, and (c) G0859 . . . 42

4.5 (a) A plot of the RMS roughness as the effective offcut changes. (b) An arrangement of AFM scans proceeding closer and closer to the substrate edge. 44

4.6 Three tighter view, 1 μm across, AFM scans of the surface of G0845. (a) Image number 1 showing a hillock at a 0° offcut with RMS roughness 0.727 nm (b) Image number 15 with an effective offcut of 0.86° and RMS roughness 0.258 nm. (c) Image number 21 with an effective offcut of 1.24° and RMS roughness 0.351 nm. 45

5.1 (a) The unit cell of cubic Bravais lattice types. Figure adapted from [8]. © 2008 Napy1kenobi, some rights reserved. Attribution-ShareAlike 3.0 Unported. (b) Zincblende crystal structure. Indium is yellow and antimony is grey. Figure sourced from [9].This work has been released into the public domain by its author, Solid State. 47

5.2 Three graphical representations of crystal structure. (a) The unit cell of a cubic crystal structure. The arrows indicate different crystallographic directions. (b) and (c) the (001) plane and the (111) plane, respectively, and their respective normal vectors. 48

5.3 A schematic of the diffractometer and the angles that relate to the system in a four circle geometry. 52

5.4 A cropped selection of the image captured by the x-ray detector for the InSb 004 Bragg peak in G0859. An anomalous second reflection appears separated from the first. 58

5.5 Two 3D scatter plots of the RSMs of the InSb 004 Bragg peaks in (a) sample G0845 and (b) sample G0846. The data is sliced by the xy plane to show the inner structure. 59

5.6 Two polar plots of the 3D RSMs of the InSb 004 Bragg peaks in (a) sample G0845 and (b) sample G0846. ψ is measured as the radius and ϕ is the angle measured counter-clockwise from the positive x axis. 61

5.7 Two plots of the FWHMs measured versus the azimuthal angle ϕ for (a) double axis measurements and (b) triple axis measurements. Each plot shows the measurements for both samples G0845 and G0846. 62

5.8 SEM ECCI images for (a) sample G0845 and (b) sample G0846. Small dots of increased contrast caused by TDDs are circled in red. The corresponding crystallographic directions are indicated by the blue arrows on the right. . 64

5.9 A schematic representation of the atomic structure at a twin boundary. Figure sourced from [10]. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer, Nature Reviews Materials. Stabilizing nanostructures in metals using grain and twin boundary architectures, Lu, K. © 2016. 65

5.10 A graphical representation of the active slip systems that intersect the (001) plane in a zincblende crystal structure. The triangular faces of the pyramid represent the {111} glide planes and the edges of these faces that intersect the (001) plane are along the ⟨110⟩ Burgers vectors. 67

6.1 A plot of the spin curves for the three nLof 2000 series photoresists. Final film thickness values correspond to a sample spun for 60 seconds at the indicated rotational velocity. 72

6.2 Some examples of the UV exposure patterns that are used to shape photoresists. Red areas indicate where exposure will occur. (a) A pattern of an array of lines with different line widths and gap widths. (b) A pattern with line width of $4\mu m$ and gap width of $2\mu m$ which is the pattern used in the final nanofabrication process. 78

6.3 Two SEM images of a developed photoresist pattern that had a gap width of $1.3\mu m$. Images (a) and (b) are in adjacent locations on the pattern. (b) shows that the photoresist was not properly developed in some areas. . . . 80

6.4 Two SEM images of different developed photoresist patterns that correspond to a pattern file that has a gap width of $0.5\mu m$. Both images show that many of the gaps were not developed fully. Where gap widths were measureable the widths were larger than $1\mu m$. 81

6.5 An SEM image of a developed photoresist pattern that almost perfectly matched its corresponding pattern file. The pattern file has a gap width of $2.0\mu m$ and a line width of $4.0\mu m$. All of the gaps were clearly well developed with sharp edges. The line width was only $0.1\mu m$ different from the pattern file which is an acceptable amount of deviation. 82

6.6 Three SEM images of the etching results. Images (a) and (c) correspond to the 44 minute etched sample. (a) shows a close up view of the gap between the etched lines and (c) shows a large area where the etched pattern was most homogeneous. Image (b) corresponds to the 10 minute etched sample and shows a close view of the gap between two lines. 83

6.7 Three images are shown that were taken after etching was complete. Images (a) and (c) correspond to the 44 minute etched sample. Image (a) is taken with an optical microscope immediately after etching was complete. The dark tangled lines are parts of the photoresist pattern that separated from the sample during etching. Image (c) is a SEM image that shows a large area of densely packed "pillars" that formed around the entire edge of the pattern. Image (b) corresponds to the 10 minute etched sample and is a SEM image of the photoresist separating from the sample. 84

B.1 3D Schematic showing the relevant vectors in calculating a RSM. (a) Side angle view. (b) Top Down View. 98

B.2 A schematic showing the shape of the region of reciprocal space that is sampled in an ω-2θ scan with an array detector. 100

B.3 Spherical Coordinates 101

C.1 SEM ECCI images for (a) sample G0845 and (b) sample G0846. The corresponding crystallographic directions are indicated by the blue arrows on the right. 105

D.1 Rocking Curves (a) double axis rocking curve. (b) triple axis rocking curve. (c) both triple and double axis rocking curves for scale. 107

List of Tables

2.1 Typical values for parameters used in modelling InSb at 300K. 13

2.2 Parameters of the simulated geometry. 16

5.1 A table showing the measured FWHMs and TDDs in the triple axis configuration for both samples G0845 and G0846. The values are averaged between the measurements for azimuthal angles that are separated by 180° in ϕ. . . 63

5.2 A table showing the measured FWHMs and TDDs in the double axis configuration for both samples G0845 and G0846. The values are averaged between the measurements for azimuthal angles that are separated by 180° in ϕ. 63

List of Abbreviations

AFM atomic force microscope 35, 37, 39

AFM atomic force microscopy 35–38, 40, 41, 43, 46, 65, 85

AlAs aluminum arsenide 30, 31, 87

AlInSb aluminum indium antimonide 44

AMI acetone-methanol-isopropyl alcohol 71, 77

BET band edge thermometry 27, 28, 30

CLS Canadian Light Source 53, 54, 59, 86

CVD chemical vapour deposition 22, 23

ECCI electron channeling contrast imaging 64, 66, 86, 103, 104

EM electromagnetic 4, 5, 7, 8, 15, 17, 25, 35, 49, 53

FC frequency comb 2, 87

FWHM full width at half maximum 56, 57, 62, 66, 68, 69, 86, 107

GaAs Gallium Arsenide 16, 17, 19, 29–32, 44, 56, 58, 60, 68, 85, 87

HRXRD high-resolution x-ray diffraction 51, 54, 60, 86, 100

HT high temperature 31

ICP inductively coupled plasma 70, 74, 76, 77, 79, 86

InSb indium antimonide 1–4, 7, 12–22, 29–31, 35, 41, 43, 44, 47, 55–58, 60, 61, 68, 70, 74–77, 79, 81, 82, 84–87

IPA isopropyl alcohol 71

LSP localized surface plasmon 10, 11

LSPR localized surface plasmon resonance 11, 20

LT low temperature 31

MBE molecular beam epitaxy 22, 23, 25, 27–31, 35, 41, 51, 70, 71, 76, 85, 87

PAC photoactive compound 73

PEB post-exposure bake 74, 78

PVD physical vapour deposition 22, 23

QCL quantum cascade laser 1–3, 5, 87

QNFCF Quantum-Nano Fabrication and Characterization Facility 79

RGA residual gas analysis 28, 30

RHEED reflection high energy electron diffraction 24–27, 30, 31, 46, 85

RMS root mean square 38, 39, 41–43, 85

RPM rotations per minute 77

RSM reciprocal space map 54, 55, 58–60, 66, 87, 97, 100, 101

RSS resolved shear stress 68

SEM scanning electron microscope 78, 80, 103

SPP surface plasmon polarition 7–12, 96

TD threading dislocation 65, 69

TDD threading dislocation density 57, 62, 64–66, 68, 69, 86

TE transverse electric 17, 19, 20, 95, 96

THz terahertz 1–5, 7, 12–15, 17–22, 31, 70, 85, 87

TLK terrace ledge kink 38

TM transverse magnetic 8, 9, 17, 19, 20

UV ultraviolet 70, 71, 73, 74, 77, 86

VGF vertical gradient freeze 22

XRD x-ray diffraction 25, 43, 46, 50, 51, 53, 55, 62, 64–66, 86, 87, 99, 103

Chapter 1

Introduction

Terahertz (THz) radiation, often reffered to simply as THz, is a type of light that has, historically, been a challenge for scientists and engineers to generate, detect or control. Current technologies have thus been unable to utilize THz radiation to its full potential. This thesis is a foundational study into the development of novel materials for the purpose of improving generation, detection, and overall control of THz radiation. These new materials are nanostructures based on the III-V semiconductor indium antimonide (InSb). The intention is to exploit the plasmonic nature of InSb when it interacts with THz radiation. This will eventually lead to the development of devices that have a dynamically tunable response to THz frequency light.

1.1 Motivations & Applications

Unlike other bands of electromagnetic radiation, generating THz radiation has been challenging. Therefore, many important scientific and technological applications of THz radiation remain heavily underdeveloped. However, recent developments such as the record-breaking demonstration of a portable and powerful THz quantum cascade laser (QCL) have opened up many previously inaccessible pathways in THz research and development of THz applications [11].

Now that higher power THz radiation has become more accessible in laboratory environments, there has been a surge in research involving THz radiation. Researchers are now exploiting the properties of THz radiation for applications in many areas including, but not limited to: chemical sensing, astronomy, communications, security, and environmental

monitoring [11]. THz radiation also offers unique imaging capabilities because it can penetrate many materials, including plastics, ceramics, textiles, paints, and wood. This can be used to non-destructively reveal structural defects in objects ranging in size from small microchips to large spacecraft [12].

Another important and also life-saving application of THz radiation is for the detection and diagnosis of skin cancer. Water molecules readily absorb THz radiation so, it can be used to study hydration levels in skin allowing for differentiation between healthy and malignant tissues [13]. This reveals malignant abnormalities in a non-invasive procedure. Also, in spectroscopy, many important biological molecules such as proteins and DNA have rotational or vibrational resonance modes that correspond to THz frequencies [14] . Additionally, THz radiation finds many applications in the field of astronomy because of the information contained about life sustaining molecules such as oxygen and water.

Nevertheless, THz QCLs, in their current state, are not ready for many of their potential applications. This research will lay the groundwork for improving the functionality of THz QCLs by combining them with InSb nanostructures to dynamically shape, steer and concentrate THz radiation for greatly improved performance.

For example, creating ultrafast pulsing THz QCLs by way of mode-locking has been deployed for THz frequency comb (FC) generation [15]. By augmenting THz QCLs with InSb nanostructures we can improve the performance of THz FCs. This is possible thanks to the unique THz plasmonic properties of InSb, which can be tuned through temperature control, carrier injection via applied voltage, or photogeneration [16]. Such post-processing tunability is expected to also benefit other THz photonic structures. THz FCs can be readily deployed in ultrafast sciences, data storage, high-speed communication, and ultrahigh-resolution spectroscopy. They are also important for the calibration of many kinds of THz equipment and in quantum metrology [15, 17].

1.2 Goals

This project is a foray into unknown territory. As a pioneering work, the goal of this project is not to extend a previous research project or to completely design these "dynamically tunable" devices mentioned in the chapter opening. The goal is to lay the foundation for future work in this area. The goals of this project are to prepare procedures for developing, characterizing, and understanding the intricacies of these InSb nanostructures and also their interactions with THz radiation. This includes discovering which tools are needed to study these subjects, learning how to best utilize these tools and overcoming

the unexpected issues that are revealed along the way. These tools do not only include the various equipments required to physically produce and study these nanostructures but this also includes the software and theoretical machinery used to understand this subject. The hope is that the experience gained now will be utilized in a future project directly following this one. With this groundwork, research will continue developing these InSb nanostructures for use with other technologies such as THz QCLs.

Chapter 2

Modelling & Simulation

This chapter will cover the theory used to model the THz response of InSb and the software that uses this model to simulate an InSb nanostructure. The purpose of this part of the thesis is to understand the model that is used to explain the behaviour of plasmonic materials, such as InSb. Then, use the model to predict the outcome of an experiment with an actual InSb nanostructure.

2.1 Theory

2.1.1 THz Radiation

Although the range is somewhat arbitrarily defined and may differ between sources, the THz band is typically defined as all electromagnetic (EM) radiation with a frequency between 0.1 THz and 10 THz [18, pg.303]. The corresponding wavelengths of EM radiation in this regime would be between 3.0 mm and 0.03 mm. This definition places THz radiation between the microwave (< 0.1 THz) and infrared (> 10 THz) regimes.

In the past, THz radiation was difficult to generate in any way other than through the black-body radiation of sources like mercury lamps [18,19]. These sources output EM radiation in a wide range of frequencies that does extend into the THz region. However, in the THz region, the output power of these devices is on the order of nanowatts.

In more recent times, the most convenient methods of generating THz radiation are using photoconductive antennas or optical rectification [11, 19]. These sources are still limited in power to the range of tens of nanowatts. High power sources of THz radiation do

exist, such as synchrotron radiation. However, to be able make use of synchrotron radiation, large facilities with infrastructure dedicated to the generation of synchrotron radiation must be constructed. For example, synchrotrons require large ring structures, greater than 50 m in diameter, to accelerate electrons to nearly the speed of light. Synchrotrons can be made to produce light with nearly any desirable frequency by precisely altering the path of these accelerating electrons [18]. Other high power sources of THz radiation, such as free electron lasers are also not compact and inconvenient.

The lack of THz sources that are both high-power and convenient made progress in the study of material interactions with THz radiation very slow and this subject fell out of focus. This has created what is commonly referred to as the 'THz gap' because technologies that would operate in the THz band have been underdeveloped [11].

Now however, development of THz sources that are both high power and compact is gaining momentum. Advancements have been made in creating higher power sources of THz radiation. One such device was the first THz QCL [20]. THz QCLs can generate THz radiation with output powers greater than one watt [11]. It is only very recently that new THz QCLs have been able to extend their operational temperature range above cryogenic temperatures that would require bulky cryogenic cooling systems [11]. As development continues, smaller and more convenient sources of THz radiation will allow THz radiation to become more accessible outside of laboratories. This will open up many previously inaccessible pathways in THz research and applications.

2.1.2 Drude-Lorentz Model

To understand how semiconductor nanostructures interact with incident THz radiation, we need to know the permittivity and permeability of the material. The Drude-Lorentz model is a classical model that applies kinetic theory to electrons to explain how metals interact with EM waves.

First, consider a general medium which is either a metal or a semiconductor. This medium has both a conduction band and valence band of electrons. The Drude-Lorentz model considers only two types of electrons, unbound conduction band electrons that move freely throughout the material and bound valence band electrons which are localized to an individual atom's nucleus [21, pg.15-27]. Electrons can become unbound by going through an interband transition if they are excited by an incident photon of energy greater than the band gap energy. In this model, these interband transitions will create an electron-hole pair in the medium.

The first half of the model, the Drude model, considers the motion of free electrons in the conduction band when excited by an incident oscillating electric field along with a damping force caused by collisions [21]. The solution of the equation of motion gives the instantaneous position of electrons as they depend on the frequency of the incident light. This position can then be used to find the polarization density of the material which is then used to find the permittivity in Equation 2.1 [21].

$$\epsilon(\omega) = 1 - \frac{\omega_P^2}{\omega^2 + i\omega\gamma} = 1 - \frac{\omega_P^2}{\omega^2 + \gamma^2} + i\frac{\gamma\omega_P^2}{\omega(\omega^2 + \gamma^2)} \tag{2.1}$$

ω_P is the plasma frequency and $\gamma = \nu/l$ is the damping coefficient. ν is the Fermi velocity and l is the mean free path [21]. The plasma frequency is an important characteristic of the material as it determines what behaviour the material has at the incident light frequency ω.

At the frequencies of visible light, metals are more dispersive than dielectrics [21]. When ω is below ω_P, the Drude model predicts that the response behaviour of the material will be similar to most metals at visible frequencies. When ω is above ω_P, then the material acts like a typical dielectric [21].

Equation 2.1 only includes the contribution of the free conduction band electrons to the permittivity, but the bound electrons in the valence band also have a contribution through interband transitions. This is where the Lorentz model comes in.

In the Lorentz model, the bound electrons are treated as damped harmonic oscillators attached to a solid positive nucleus and are driven by the incoming electric field. These oscillators have resonant modes each with a resonant frequency ω_i where the imaginary part of the permittivity peaks, indicating stronger absorption of energy [21]. The contribution of these electron oscillators to the permittivity is given by Equation 2.2 [21].

$$\varepsilon(\omega) = 1 + \sum_{i=1}^{K} \frac{f_i\omega_i^2}{\omega_i^2 - \omega^2 - i\gamma_i\omega} \tag{2.2}$$

Here, K is the number of oscillators (electrons), f_i is a weighting coefficient and γ_i is the damping coefficient for the ith oscillator. The Drude-Lorentz model combines both the effects from the free electrons with the effects from the bound electrons to give the permittivity as Equation 2.3 [21].

$$\varepsilon(\omega) = 1 - \frac{\omega_P^2}{\omega^2 + i\omega\gamma} + \sum_{i=1}^{K} \frac{f_i\omega_i^2}{\omega_i^2 - \omega^2 - i\gamma_i\omega} \tag{2.3}$$

At frequencies very far from ω_i, the Lorentz component of the permittivity reaches an asymptotic value which is unique for each type of material. The modified Drude-Lorentz model replaces this interband contribution with its asymptotic value to get Equation 2.4 [21]. ε_∞ is the high frequency permittivity that is unique for each type of material.

$$\varepsilon(\omega) = \varepsilon_\infty - \frac{\omega_P^2}{\omega^2 + \gamma^2} + i\frac{\gamma\omega_P^2}{\omega(\omega^2 + \gamma^2)} \tag{2.4}$$

This permittivity equation can be used to model the permittivity of InSb plasmonic nanostructures in the THz range [16, 22].

In a similar way to deriving the permittivity, the permeability of a material can also be derived using the Lorentz model by considering the magnetization of a magnetic dipole. This will give an equation similar to Equation 2.2. The Lorentz model predicts the relative permeability in Equation 2.5.

$$\mu_r(\omega) = 1 + \sum_{k=1}^{M} \frac{\omega_{pmk}^2}{\omega_{0mk}^2 - \omega^2 - i\gamma_{mk}\omega} \tag{2.5}$$

In this equation, M is the number of oscillators (dipoles), ω_{pmk} is the plasma frequency for the magnetic dipole, ω_{0mk} is the resonant frequency, and γ_{mk} is damping coefficient for the kth oscillator.

However, for modelling InSb we will consider only nonmagnetic media and thus it is assumed that the relative permeability is unity. This assumption is often used when modelling the permeability of metals and dielectrics [21].

2.1.3 Surface Plasmons

A structure can be considered plasmonic if it meets the requirements to support surface plasmons. There are generally two classifications of surface plasmons. The first type is the surface plasmon polarition (SPP). SPPs are the collective oscillation of electrons that are coupled with an electromagnetic wave at a dielectric-metal interface [21, pg. 41-63] [21, pg. 21-52]. SPPs are longitudinal surface waves that propagate along the interface between a metal and a dielectric material.

To understand how SPPs exist we need to examine the EM fields at the dielectric-metal interface. We can find the electric and magnetic fields by applying Maxwell's equations to a planar interface between a dielectric and a metal.

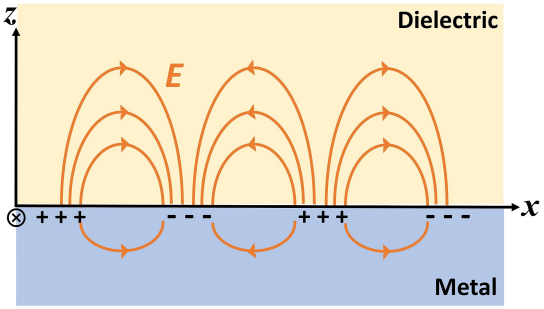


Figure 2.1: Diagram of the surface charge oscillations (electron oscillations) coupled with electric field oscillations in a SPP propagating along a dielectric-metal interface. Figure sourced from [1].

SPPs are longitudinal excitations, like a pressure wave. This means there is a component of the electric field that is parallel to the direction of travel. Imagine this pressure wave travelling along the dielectric-metal interface shown in Figure 2.1. The electric field, E_j, and the magnetic field, H_j are given by equations 2.6 and 2.7 [21]. Positions with $z > 0$ will be inside the dielectric medium and $j = d$ for the dielectric. While for $z < 0$, in the metallic medium, $j = m$ [21]. $k_{x,j}$ and $k_{z,j}$ are the x and z components of the wave vector in the respective medium. $E_{x,j}$ and $E_{z,j}$ are the x and z amplitudes of the electric field. Finally, $H_{y,j}$ is the amplitude of the magnetic field which only has a component along the y direction, into the page. The interface exists along the plane $z = 0$ and the wave is propagating along the $\hat{x}$ vector with a frequency of ω. The additional electric field component along the $\hat{z}$ vector results from the surface charge oscillation to which this EM field is coupled.

These two coupled components exist on opposite sides of the interface. On the dielectric side of the interface exists the evanescent EM wave which only penetrates a very short distance into the dielectric medium. On the metallic side of the interface the SPP exists as a surface charge oscillation which again only penetrates a very short distance into the metallic medium.

$$\mathbf{E}_j = (E_{x,j}\hat{x} + 0\hat{y} + E_{z,j}\hat{z})e^{i(k_{x,j}x + k_{z,j}z - \omega t)} \tag{2.6}$$

$$\mathbf{H}_j = (0\hat{x} + H_{y,j}\hat{y} + 0\hat{z})e^{i(k_{x,j}x + k_{z,j}z - \omega t)} \tag{2.7}$$

The solutions given above are both transverse magnetic (TM) polarized waves because

SPPs can only exist as TM polarized modes. This is a result from the continuity conditions for E_y and H_x at the interface [2]. For a proof of this see Appendix A.

The dispersion relation related to these waves is given in Equation 2.8, where $k_x = k_{x,d}$ and ε_d and ε_m are the permittivities of the dielectric and metallic mediums [21].

$$k_x = \sqrt{\frac{\varepsilon_d \varepsilon_m}{\varepsilon_d + \varepsilon_m}} \frac{\omega}{c} \tag{2.8}$$

For SPPs to propagate this value under the square root must be positive which places unique restrictions on the values of ε_d and ε_m.

For the real parts of the permittivities, this condition is $\varepsilon'_m(\omega) < -\varepsilon'_d(\omega)$ [21]. This condition means that the permittivities of the two materials must have opposite sign and the magnitude of the negative permittivity must be greater than that of the positive one. When designing plasmonic devices this condition limits which materials can be used for the metal-dielectric interface. One more condition limits the excitation of SPPs. SPPs can not be excited directly by electromagnetic waves in free space [2].

Figure 2.2 shows the dispersion relation for SPPs (curved lines) in the ideal case where the frequency dependent permittivity of the metal ε_m is given by Equation 2.1 but with no damping (i.e. $\gamma = 0$) [2]. The light lines (straight lines) correspond to $k = \frac{\omega}{c}\sqrt{\varepsilon_d}$ and it is the boundary for which electromagnetic waves can exist in the dielectric medium [21]. Only with frequencies and wave vectors to the left of the light line can electromagnetic waves exist inside the dielectric material. The curve for SPPs always remains below and to the right of the light line for the medium that is at the interface.

At lower frequencies, the SPP dispersion curve forever approaches the light line until $\omega = 0$ and at higher frequencies it asymptotically approaches $\omega = \frac{\omega_p}{\sqrt{1+\varepsilon_d}}$ [21]. For any incident light on the dielectric-metal interface there will always be a gap between the wave vector of the incident light, k_{in}, and that of the SPP, k_{sp}. This means there is a momentum mismatch (or phase mismatch) of G between the incident light and the SPP [2, 21]. G increases with frequency.

The only way to excite an SPP is to match the wave vectors along the propagation direction $\hat{x}$. So, an SPP can never be directly coupled to the electromagnetic wave that is incident on the dielectric-metal interface [21].

If it is impossible to have an electromagnetic wave in the dielectric medium with a wave vector x component large enough, then how can one excite an SPP at the dielectric-metal

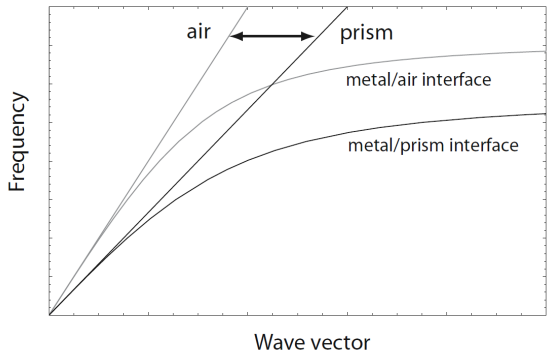


Figure 2.2: A plot of the dispersion relations of the possible SPPs that could be excited at the interfaces of prism, metal, and air in the prism coupling technique. The straight lines are the light lines in the labelled material. Figure sourced from [2, pg. 43]. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer. Plasmonics: Fundamentals and Applications by Stefan A. Maier © 2007.

interface? There are a variety of ways to overcome this momentum mismatch and excite SPPs. One such method is prism coupling.

In prism coupling, the idea is to introduce a third medium with a permittivity ε_p greater than ε_d [2]. Figure 2.3 illustrates this system. If we consider the dielectric medium to be air, $\varepsilon_d = \varepsilon_{air} = 1$, and $\varepsilon_p > 1$ then we get a new light line for the electromagnetic waves in the prism, $k = \frac{\omega}{c}\sqrt{\varepsilon_p}\sin\theta_{in}$ where θ_{in} is the incident angle of the light beam at the metal-prism interface [2]. Figure 2.2 shows these dispersion relations and now we can see that the momentum of the incident light from the prism can match that of the SPPs at the air-metal interface because the lines intersect.

For a range of incident angles SPPs will be able to be excited by the light that penetrates all the way from the prism and through the metal film to the metal-air interface. This is only one of a number of ways to excite SPPs but all of the methods use some technique to match the x component of the momentum of the incident light to that of the SPP.

The second type of surface plasmons are localized surface plasmons (LSPs) which do not propagate [2]. LSPs can occur in nanostructures that have dimensions smaller than the wavelength of the incident light, so-called sub-wavelength structures such as nanoparticles. Like SPPs, LSPs also result from the coupling of electromagnetic waves with electron oscillations. Instead of propagating along an interface, there is an effective restoration force due to the sub-wavelength geometry of the nanostructure and the polarization caused by the displacement of electrons from their equilibrium positions [2, 21]. This allows for

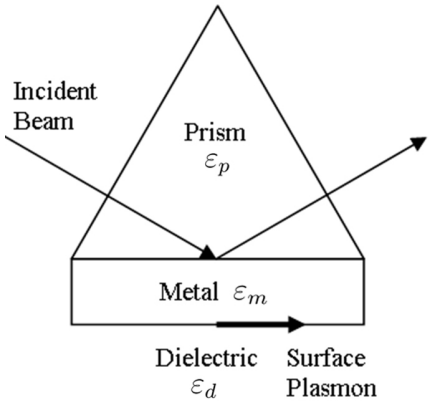


Figure 2.3: The schematic of a simple prism coupling set up to excite SPPs. Figure adapted from [3].

specific resonant oscillations, called localized surface plasmon resonances (LSPRs), where electromagnetic fields are greatly amplified in a small volume on both sides of the metal-dielectric interface.

LSPRs can be excited by direct coupling with an incident electromagnetic wave and there is no need for the phase-matching techniques that are used to excite SPPs [2, 21]. Figure 2.4 shows a schematic of two nanoparticles subjected to an electric field. The electron clouds are displaced by the electric field and will oscillate while the restoring force caused by the induced polarization tries to return them to equilibrium.

When the frequency of the electric field approaches the natural frequency of the dipole an LSPR occurs and the local fields inside and outside of the nanoparticle are greatly enhanced. At the LSPR frequency there is also a strong absorption. The properties of the LSPR are dependent on the shape, size, and material of the nanostructure [21].

LSPs are not limited only to nanoparticles but can appear anywhere within a structure where the dimensions are sub-wavelength such as at corners and edges of metallic nanostructures [2]. No derivation of the fields of an LSPR will be given here because they are highly dependent on the geometry of the nanostructure and are best analyzed through simulations.

Both types of surface plasmons are useful for controlling the propagation of light by localizing electromagnetic waves. SPPs localize light along one dimension ($\hat{z}$ in the diagrams above) to effectively create two-dimensional electromagnetic waves that propagate along the metal-dielectric surface [2]. While LSPs will strongly enhance fields which then decay in a short distance from the interface.

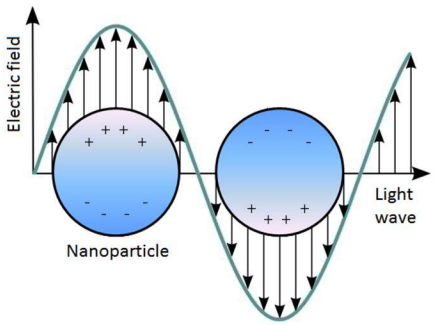


Figure 2.4: A simplified visual representation of how an oscillating electric field will interact with a metal nanoparticle to excite an LSP. Figure sourced from [4].

2.1.4 Indium Antimonide

Continuing towards this chapter's goal of simulating the interactions of nanostructures with THz radiation, more needs to be said about InSb. InSb is a III-V semiconductor compound that has a small band gap (E_g) , high electron mobility (μ_e) and small electron effective mass (m_e^*) [16, 23]. InSb is a good candidate for THz optoelectronic devices. This is because, when subjected to light at THz frequencies, InSb acts characteristically like a metal material would at visible frequencies [16, 22, 24, 25]. That is to say, InSb has a permittivity that can be modelled by the modified Drude-Lorentz model (Equation 2.4) with a plasma frequency in the THz range and thus can support surface plasmons as outlined in the previous sections.

InSb is the best semiconductor material to choose for THz plasmonics because no other semiconductor material intrinsically supports surface plasmons in the THz frequencies. Other semiconductor materials would need to be heavily doped to support THz surface plasmons.

Metals can support SPPs at visible light frequencies but they are not a good choice for THz frequency SPPs. Metals will have a plasma frequency far above the THz region and thus will have a negative permittivity at THz frequencies. This means that they will actually be capable of supporting SPPs. However, due to the fact that $\omega << \omega_P$ there will be a long SPP decay length into the both the metal and dielectric material [26]. Many applications of SPPs require high field localization and large enhancement of the field. Because the SPPs will not be within as small of a volume, metals are not practical materials to choose for low frequency (THz) plasmonic applications [2].

Quantity	Value	
ε_∞	15.68	[22, 24, 25, 27, 28]
m_e^*	$0.014m_0$	[24, 25]
γ	1.79×10^{12} rad/s	[22, 28]
N	$\sim 2 \times 10^{16} cm^{-3}$	[16, 22, 28]
ω_P	$\sim 1.52 \times 10^{13}$ rad/s, ~ 2.41 THz	[22, 28]
μ_e	$\sim 7.7 \times 10^4 cm^2/Vs$	[27]
E_g	0.174 eV	[23, 27, 28]

Table 2.1: Typical values for parameters used in modelling InSb at 300K.

InSb will have a plasma frequency within the THz range so it is a better choice. Better yet, because InSb is a semiconductor its plasma frequency is tunable by means of carrier injection. Equations 2.9 and 2.10 are commonly used to model the permittivity and the plasma frequency of InSb, respectively [16,22,24,25]. Equation 2.9 appears slightly different from the earlier Equation 2.4 but it is actually the same. In experimental work, sometimes the square of the plasma frequency is normalized by the high frequency permittivity ε_∞ which allows it to be factored out in equation 2.9. The equation is otherwise the exact same. Returning to equation 2.10, the plasma frequency depends on N, which is the carrier concentration, or electron density of the material. ε_0 is the permittivity of free space, e is the elementary charge, and m_e^* is the electron effective mass. Typical values for these parameters are given in table 2.1.

$$\varepsilon_{InSb}(\omega) = \varepsilon_\infty \left(1 - \frac{\omega_P^2}{\omega^2 + i\omega\gamma}\right) \tag{2.9}$$

$$\omega_P = \sqrt{\frac{Ne^2}{\varepsilon_0 \varepsilon_\infty m_e^*}} \tag{2.10}$$

Semiconductor materials offer a few different ways to actually control and tune the carrier concentration N. First of all, the number of free electrons in a semiconductor can be controlled through the doping process [24, 25, 27, 29]. In the fabrication of a plasmonic semiconductor structure the carrier concentration can be chosen by controlling the amount of impurities (dopants) introduced into the structure. There are a variety of fabrication methods that allow control over the doping of a semiconductor. One such method is molecular beam epitaxy which will be discussed in Chapter 3. Controlling the doping

level will set an equilibrium carrier concentration in the semiconductor but there are also methods to dynamically tune the carrier concentration.

Carriers can be introduced through temperature control because at higher temperatures, electrons in the valence band will be excited into the conduction band. Another method is to directly inject carriers by applying a voltage across the material. The third method for dynamically controlling carrier concentration is to generate carriers through photogeneration [16, 22–25]. By subjecting the material to a light source (this will be an additional light source to the THz radiation at a different frequency) electrons can absorb photons and become excited into the conduction band.

There are still a few final reasons why semiconductors in general are a good choice for THz plasmonic materials. InSb, under THz illumination, acts like a metal would under visible light. Because of this, we can use our understanding of plasmonics with metals, which have already been extensively studied, to adapt our ideas to THz frequencies by using a semiconductor in place of a metal. We can also add more functionality to these devices by taking advantage of the semiconductor's tunable permittivity. Furthermore, semiconductor fabrication has been widely used for industrial and research purposes for a long time. While there is always room to improve, the technology has already matured past the point that is needed to reliably produce semiconductor nanostructures on the sub-wavelength scales that are needed to produce THz plasmonic devices.

As a final note, to demonstrate how the permittivity of InSb can be tuned, one can consider how changing the temperature of the material would affect ω_p. Equation 2.10 shows that the plasma frequency is dependent upon the carrier concentration and the electron effective mass. Both of these quantities are themselves temperature dependent. Their temperature dependencies are given in the empirically determined Equations 2.11 and 2.12 [30]. In these equations the temperature is given in units of K and k_B is the Boltzmann constant in units of eV/K. Equation 2.11 is not directly given in [30] but it can be derived from the relation between the temperature dependencies of the band gap energy and the electron effective mass.

$$m_e^*(T) = 1.293 \times 10^{-32}(1 - 3.941 \times 10^{-4}T) \quad [Kg] \tag{2.11}$$

$$\begin{aligned} N(T) = 2.9 \times 10^{11}(2400 - T)^{3/4}(1 + 2.7 \times 10^{-4}T)T^{3/2} \\ \times exp\left(\frac{-(0.129 - 1.5 \times 10^{-4}T)}{k_B T}\right) \quad [cm^{-3}] \end{aligned} \tag{2.12}$$

Comparing to other works, when using Equations, 2.10, 2.11 and 2.12 and $T = 300$ K, the plasma frequency found is $\omega_p = 1.78 \times 10^{13}$ rad/s. This is only a 17% difference to the value listed in Table 2.1. With an increase of only 15 degrees to $T = 315$ K, the plasma frequency is blueshifted to 2.04×10^{13} rad/s, an increase of 15%!

2.2 Methods

With a better understanding of the physics of surface plasmons and an appropriate model for the permittivity of InSb, the next stage is to use this to assemble a simulation of these effects. In this research, the commercial software COMSOL Multiphysics has been used to develop a three dimensional model of an InSb nanostructure and simulate its response to EM radiation with a range of THz frequencies.

2.2.1 COMSOL Multiphysics

COMSOL Multiphysics is a simulation environment that is capable of simulating not only electromagnetics but also many other types of physics including fluid dynamics, structural mechanics, heat transfer, etc. The environment provided by COMSOL uniquely allows you to couple these different physics into one simulation so that results can show multiple characteristics of the system while also revealing the interplay between different physics taking place in the system.

2.2.2 The Simulated Model

Simulations of multiple types of physical effects can be time consuming, resource heavy, and sometimes needlessly complex. For the purpose of revealing the effects of surface plasmons in InSb nanostructures, only a single COMSOL module is essential, the Wave Optics Module. The Wave Optics Module extends the ability of COMSOL Multiphysics to simulate the electromagnetic effects of systems such as optical fibers, photonic waveguides and crystals, as well as other photonic devices [31]. It has also been extensively used for studying the plasmonic effects of InSb [16, 22–24, 27, 32].

To simulate the plasmonic response of an InSb nanostructure, a study was conducted in the frequency domain rather than in the time domain. Electromagnetic waves vary in time as a sinusoidal wave and we can assume that the solutions for the waves' interactions

Name	Value	Description
t	$1.5\mu m$	InSb layer thickness
w	$4.0\mu m$	Width of lines at base
g	$2.0\mu m$	Width of gap at base
ϕ	$80°$	Angle of trapezoid at base
s	$150\mu m$	Substrate (GaAs) thickness
a	$2.5\mu m$	Height of air block

Table 2.2: Parameters of the simulated geometry.

will also vary sinusoidally. This makes the problem time harmonic and allows for the system to be solved as stationary in the frequency domain [31, pg. 14]. The solutions will have complex values containing the amplitude and phase of the field. This is much more efficient and less expensive computationally than using the time domain. By using Fourier analysis to be able to include any signal that is periodic this can be a powerful tool for many systems.

As a proof of concept, a simple geometry was chosen with arbitrary yet, sub-wavelength dimensions. The parameters of this geometry are given in Table 2.2 and a 2D projection of this 3D geometry is shown in Figure 2.5. The InSb takes the shape of a trapezoidal prism that, in the simulation, extends infinitely in both directions along the x axis. The full simulated structure will be an array of these prisms with the structure repeated along the y axis, again infinitely in both directions. In Figure 2.5 there are three domains with different colours, indicating the material occupying that volume. To obtain realistic results from the simulation you could not have a detached nanostructure floating in the void of space unless you were conducting your experiment aboard the international space station. This is why the block of air and block of Gallium Arsenide (GaAs) were included. The GaAs acts as the substrate which supports the weight of the InSb, holding the prisms in place so that they remain evenly spaced. The air block fills the voids between the InSb prisms giving volume that can be used to simulate the near-field patterns of the nanostructure. The trapezoidal face was chosen instead of a rectangle because it is unrealistic to achieve a perfectly vertical sidewall from the nanofabrication etching procedure.

After defining the geometry a few other simulation design decisions need to be made. First of all, the boundary conditions must be established. As mentioned, the intended simulated structure extends infinitely in both directions of the xy plane. This is achieved in the simulation by using periodic boundary conditions on both pairs of the xz and yz faces of the structure. The entire geometry is made of repetitions of the cell in Figure 2.5. Thus each boundary of the cell is effectively connected to another cell, even though

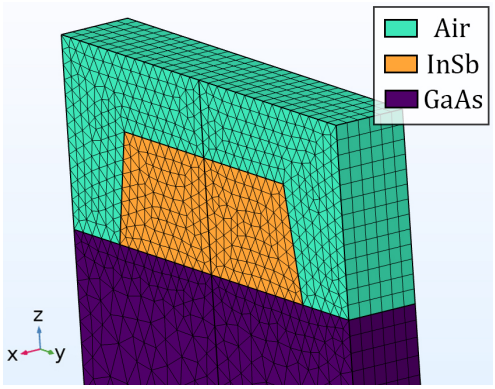


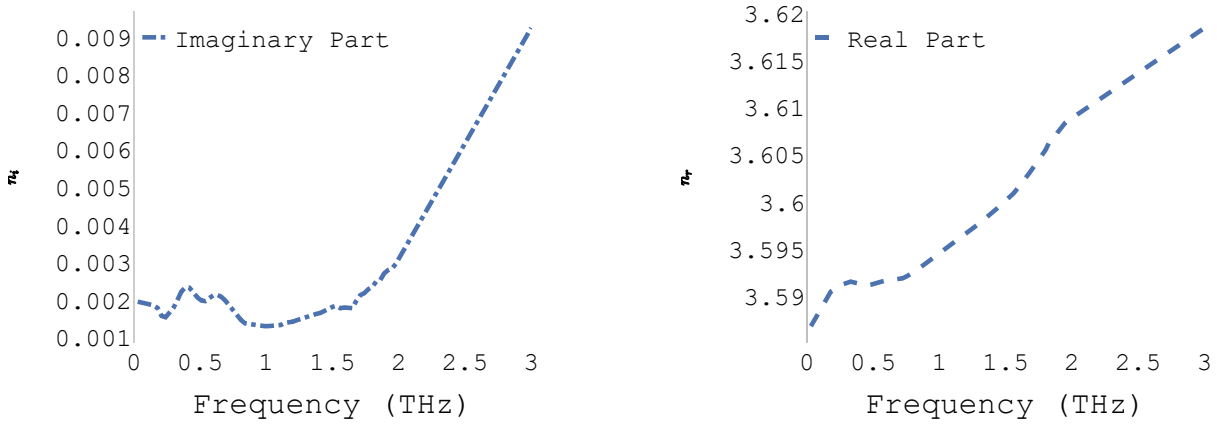
Figure 2.5: A stereoscopic projection of the 3D model used in the simulation. There are three domains containing different materials. The mesh used in the simulation is displayed on the surface.

only a single cell needs to be simulated. The remaining two boundaries on top of the air block and below the GaAs substrate are defined as ports. This is where EM waves can enter and exit the system. The top port, above the air block, is then set to emit a plane wave travelling towards the InSb layer along a path parallel to the z axis. At this port, the polarization of the EM wave can be set, and both transverse electric (TE) and TM modes will be used. The bottom port is where energy escapes the system and measurements of transmission are made.

EM radiation travelling from the entry port to the exit port will pass through all three different materials and it will feel a different effect from each of them. In the air domain, the refractive index has been set to 1 and therefore EM radiation travels through unimpeded by any absorption as we would expect over a short distance. Once entering the InSb domain the electromagnetic radiation is affected by the permittivity of the InSb as calculated by the Drude-Lorentz model in the earlier section 2.1.2. Finally, any radiation that is transmitted through the InSb layer will interact with the GaAs.

To get realistic values for the refraction index for GaAs, measurements were taken from [5]. Measurements for both the real refractive index and absorption coefficient, which gives the imaginary part of the refractive index, are used. The measurements from this source only included frequencies from 0.2 THz to 2 THz. The target range of frequencies of this simulation however is from 0 to 3 THz and so the data was linearly extrapolated anywhere beyond the range of measurements. The figures 2.6a and 2.6b show the plots of this data.

Even with the geometry defined, to simulate a 3D model such as this, the geometry



(a) (b)

Figure 2.6: Plots of the imaginary (a) and real (b) parts of the refractive index for GaAs. Anywhere outside of the range of 0.2 THz to 2.0 THz was linearly extrapolated. The data for these plots was sourced from [5].

needs to be meshed. Dividing the geometry up into many smaller constituent geometrical shapes gives a finite number of smaller elements. This will be the only places where the solution will actually be calculated. This is called the Finite Element Method. This is only an approximation to the real solution. However, by decreasing the element size, and therefore increasing the number of elements, this approximation can very closely match the real solution. The increased number of elements comes at a cost of computational time. So, it is important to find a balance between element size and computational time to get a reasonable result in a reasonable amount of time. One method employed here to improve accuracy without a significant increase in the number of elements is to use a variable element size. In Figure 2.5 the mesh that was used can be seen as a pattern of triangles on the face of the geometry. To reduce the number of elements without significantly affecting the accuracy of the simulation the smallest size elements of the mesh are only in the InSb and in the area in its immediate vicinity. This is where the biggest gradient will be and so the most number of elements should be here to capture the correct behaviour.

2.3 Results

In this section, the results from the simulation will be presented and discussed. The simulation was run under four different configurations to study multiple effects on the THz

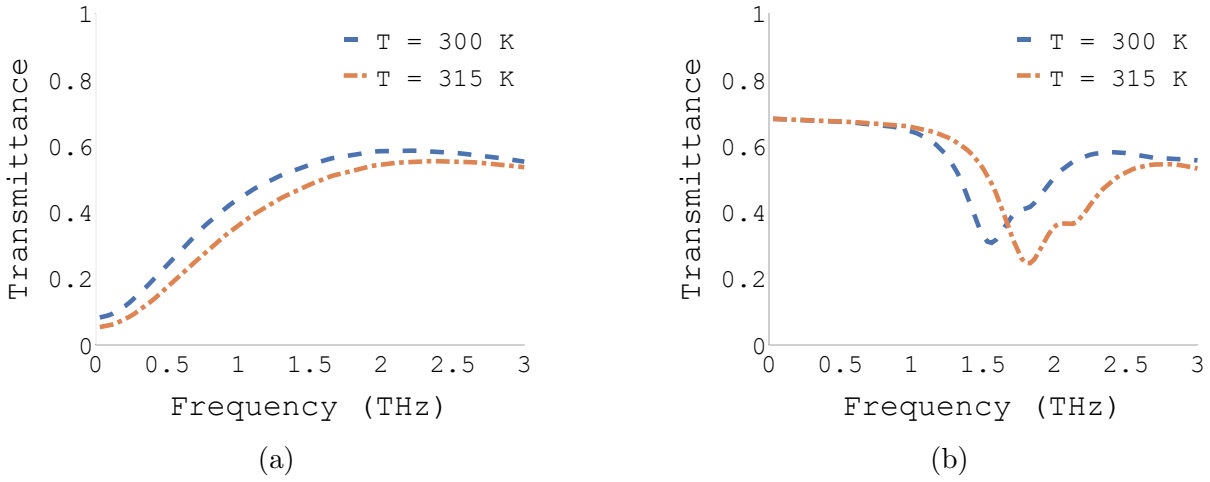
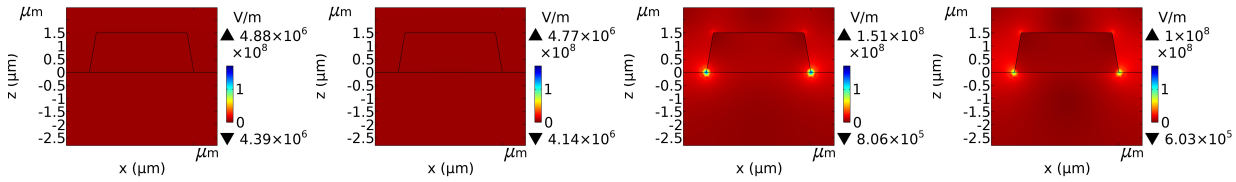


Figure 2.7: Two plots of the total transmittance through the simulated nanostructure. (a) The total transmittance when the incident radiation is TE Polarized at different temperatures of 300 K and 315 K. (b) The total transmittance when the incident radiation is TM Polarized at different temperatures of 300 K and 315 K.

plasmonic response of InSb. The simulation was run with two different temperatures for the InSb at 300 K and at 315K. As shown earlier in Section 2.1.4, this corresponds to two different values for the plasma frequency, ω_p, in the Drude-Lorentz model and so we should expect different results between them. The two remaining configurations changed the polarization of the light from TM to TE. In the context of the geometry in Figure 2.5, TM polarization corresponds to the polarization where the electric field is oscillating along the y axis. TE polarization has the electric field oscillating along the x axis. Each of these simulations were tested with a range of frequencies from 0 THz to 3 THz.

Beginning with Figures 2.7a and 2.7b there is a clear difference in the response behaviour to the two polarizations of light. Figure 2.7a shows that the transmittance for the TE polarization at both simulated temperatures are similar. There is a clear rise in transmittance with frequency but no strong peaks showing any kind of resonance.

For each of these plots, the overall height of the curve beyond 1.75 THz should not be depended upon. This is because, beyond this point, even the authors of [5] were not confident in their data for the refraction index of GaAs and then, for this simulation, the data was extrapolated further. That being said, the trend of the behaviour is still valid because the overall effect of the GaAs substrate is just to attenuate the amplitude of the field further.



(a) (b) (c) (d)

Figure 2.8: Four 2D plots of the amplitude of the electric field across a cross-section of the nanostructure when the incident light frequency is 1.56 THz with different polarizations and InSb temperatures. The colour indicates the amplitude according to the colour bar. The maximum and minimum values for the electric field amplitude is displayed above and below the colour bar, respectively. (a) TE Polarized at 300 K. (b) TE Polarized at 315 K. (c) TM Polarized at 300 K. (d) TM Polarized at 315 K.

With the TM modes in Figure 2.7b, there is an interesting dip in transmittance for both temperatures of InSb. At $T = 300$ K this minimum corresponds to a frequency of 1.56 THz and at $T = 315$ K it is at 1.83 THz. Raising the temperature of the InSb had a clear effect of increasing the resonance frequency in the transmittance. These dips in transmittance also correspond to an increase in absorption, indicating that this may be the result of a surface plasmon excitation.

Strong absorption and highly localized field enhancement are two telltale indicators of a LSPR. Figures 2.8c and 2.8d exhibit the latter indicator, giving more evidence that this loss of transmittance is the result of an LSPR. Figure 2.8 shows the amplitude of the electric field across a cross-section of the nanostructure when the frequency is 1.56 THz with two different polarizations and at two temperatures. Again, we see no action from the TE polarizations. From the TM polarizations we see two highly localized field enhancements near the sharpest corners of the nanostructure. The field enhancement is clearly stronger for the $T = 300$ K simulation. This is because increasing the temperature blueshifted the resonance frequency, just like it did to the plasma frequency and now, at 1.56 THz we are slightly off-resonance. At 1.56 THz this increase of 15 K dropped the maximum amplitude to 2/3 of its original value.

In conclusion, this chapter has clearly demonstrated that by modelling the permittivity of InSb with the Drude-Lorentz model, the effects of surface plasmons can be revealed for THz frequency excitations. Through simulations using the Finite Element Method, the localized field enhancement of the LSPR can be mapped in the near-field region. InSb is most definitely a good candidate for actively tunable plasmonic devices, as can be seen from the changed resonance frequency when the temperature of the material is increased.

This simple proof of concept nanostructure demonstrates the potential for InSb to be used for a number of applications. One can imagine this nanostructure, or one very similar, being used as a filter for THz light that can be actively switched to allow different frequencies of light through. There may also be a way to exploit this polarization dependence so that only the desired frequency of light with the desired polarization is allowed to pass. With such a strong dependence on the geometry of the nanostructure itself, it is hard to say what devices will come from this, but it is sure to be interesting.

Chapter 3

Molecular Beam Epitaxy

Chapter 2 introduced the concepts of surface plasmons and established the Drude-Lorentz model. This was used to predict the THz plasmonic response of an array of InSb trapezoidal prisms that comprise a sub-wavelength nanostructure. The topic of chapter 3 will be the nanofabrication technique, molecular beam epitaxy (MBE). MBE is a crystal growth method that is used here to create the InSb crystal from which the final sub-wavelength nanostructure will be etched.

3.1 Theory

3.1.1 MBE Systems

This nanofabrication process begins with deposition. To develop a material from scratch, we need a way to amass molecules of its constituent elements. This can be done through vapour deposition onto a semiconductor wafer. The semiconductor wafer is a crystal itself that was mass produced using another crystal growth technique called vertical gradient freeze (VGF). For more information on this technique refer to [33, pg.27]. This wafer acts as the base for our deposition and is thus referred to as the substrate.

Vapour deposition techniques can be grouped into two different classifications, either chemical vapour deposition (CVD) or physical vapour deposition (PVD) [34, pg. 55]. The main distinction between the two types of techniques is the physical state in which the starting materials begin. In PVD, the starting materials are in a solid state which is then evaporated and then deposited on the substrate. In CVD, the starting materials are in

vapour or liquid state and, through chemical reactions on the surface of the substrate, atoms are deposited onto the substrate [34]. MBE is an example of a PVD technique. For information on other CVD and PVD techniques readers are referred to [34, pg. 55 - 85].

MBE is a method that can deposit thin layers of materials onto a substrate with high precision control over the layer thickness and chemical composition. A schematic of a typical MBE chamber is given in Figure 3.1. MBE is performed in an ultra-high vacuum environment to prevent molecules of any unwanted materials from depositing onto the substrate [35, pg.2]. To maintain this vacuum throughout, there are two or three chambers each of which has at least one pumping system attached to it [36]. These pumping systems can be ion pumps or cryogenic pumps. Additionally the growth chamber walls are cooled by a liquid nitrogen shroud so that anything that isn't pumped out will be captured on the walls of the chamber. The main growth chamber is only ever brought to atmospheric pressure during maintenance and instead wafers are exchanged through a separate load lock chamber. Background pressures in the growth chamber can be as low as 10^{-11} Torr which is 14 orders of magnitude below atmospheric pressure [36].

During growth, the substrate is placed upside down in the center of the growth chamber. There is a number of components around the underside of the chamber called effusion cells [35, pg. 4-7]. The effusion cells are enclosed in the same vacuum with a mechanical shutter to block or unblock the line of sight between the material in the cell and the substrate. There is an effusion cell for each chemical element available in the system. The material within each cell is extremely pure. Each cell contains an amount of a single different chemical element. These different elements can be heated to vaporize them at a rate that depends on the temperature to which they are heated [35]. The shutters can then be used to release controlled amounts of vapour in a beam of molecules directed towards the substrate. The molecules adsorb to the surface of the substrate and migrate along the surface until they reach a vacant lattice site [35]. This results in a thin layer being grown atomically layer-by-layer.

Controlling the shutters and the temperatures of the cells determines the beam fluxes. Complex chemical compositions can be grown with multiple constituent elements. This also allows for the intentional introduction of impurity atoms into the material through doping. The amount of each element reaching the substrate can be changed rapidly resulting in complex chemical composition profile along the growth direction that can change just as rapidly [6, pg. 112]. These growth conditions as well as the substrate temperature will all play a role in determining the final product from a growth. To make certain that the whole surface is uniformly exposed to the same amount of each beam, the substrate must also be rotated.

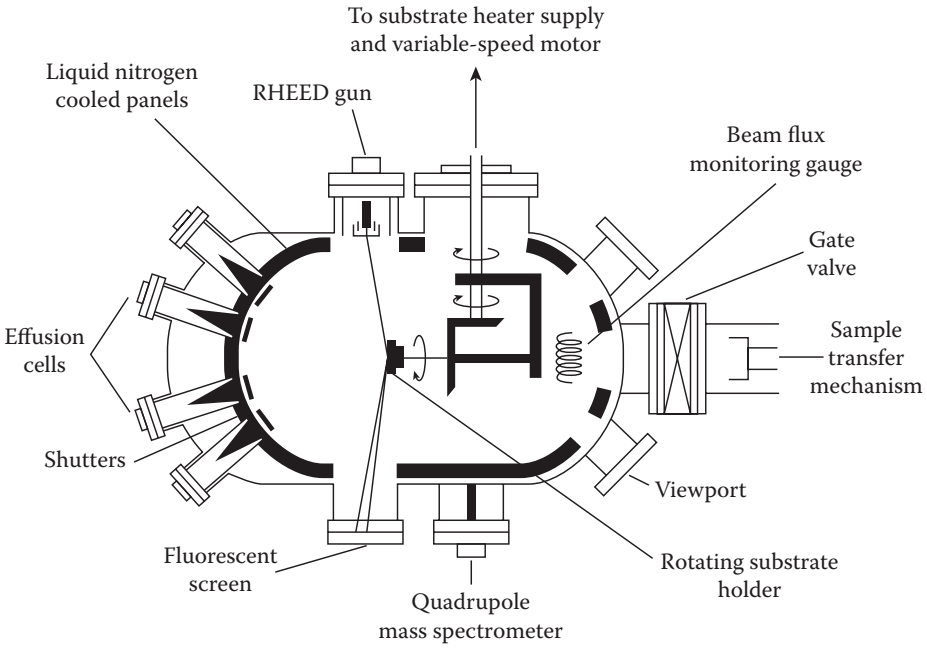


Figure 3.1: A schematic of a typical MBE chamber equipped for RHEED analysis. Figure sourced from [6, pg. 113].

The high costs required for the equipment and assembly of an MBE system as well as the regular maintenance and cleaning procedures are a tall barrier for researchers entering this field. However, these drawbacks are counterbalanced by a number of advantages that make MBE the superior growth method for a number of different applications. The ability to grow all kinds of semiconductor materials and the alloys from combining them gives MBE great flexibility. While not the fastest deposition method, MBE can be controlled to have growth rates higher than 1 μm per hour [35, pg.3]. In addition to this, MBE provides unique methods of in situ (during growth) observations of the material layers being deposited [6, pg. 112]. The ability to monitor the growth while it is under way is invaluable for controlling and optimizing the process.

Characterization methods such as reflection high energy electron diffraction (reflection high energy electron diffraction (RHEED)), mass spectrometry, absorption spectrometry and reflectance are used to monitor temperatures, surface reconstruction, surface chemistry, and growth rates.

3.1.2 Reflection High Energy Electron Diffraction

Reflection high energy electron diffraction uses high energy (up to 100 keV) electron beams at a shallow angle with respect to the surface of 1° to 2°. The electon beam probes the surface of the crystal during MBE growth. In Figure 3.1 the typical installation of an electron gun and fluorescent screen on an MBE growth chamber is depicted. From the diffraction pattern recorded on the fluorescent screen opposite to the electron source, many characteristics of the growth can be ascertained. From RHEED analysis, growth rate, growth mode, surface roughness, and surface reconstruction can all be determined from in-situ measurements.

The underlying physics that rule this interaction are similar to the physics of x-ray diffraction (XRD) and will be explained in a different context later in Chapter 5. First of all, the behaviour of the electrons will be similar to waves of EM light but with a de Broglie wavelength

$$\lambda = \frac{hc}{E}$$

determined by the energy E of the beam where h is the Planck constant and c is the speed of light [6, pg. 393 -394]. As EM radiation would, the electrons will only show a signal when both the incident beam and diffracted beam are at an angle that meets the condition for constructive interference [6, 7, 35]. This is why, across the fluorescent screen, the signal will vary in intensity creating a diffraction pattern. The electrons however only penetrate

a shallow depth into the surface of the crystal [6, 7, 35]. So, when atomically smooth, the resulting diffraction pattern corresponds only to an effectively 2D lattice from the top few atomic layers of the crystal.

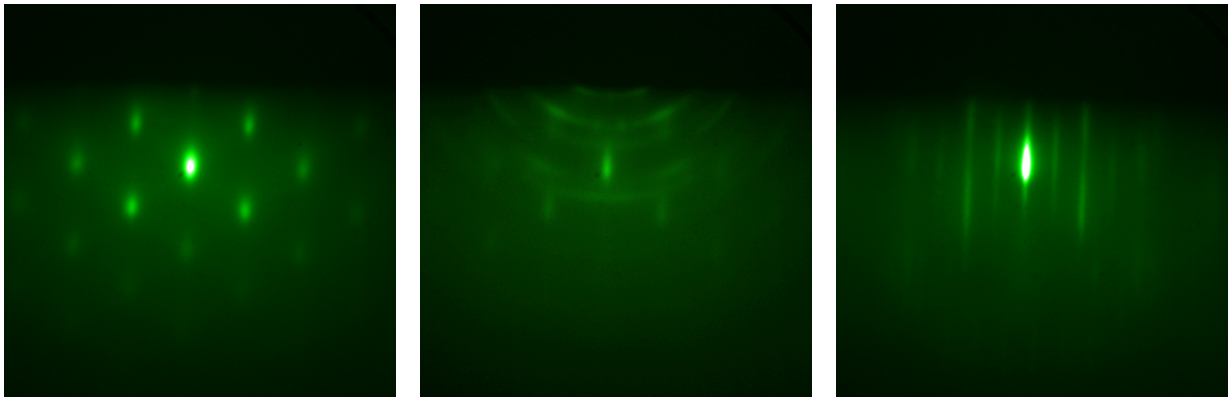
The diffraction pattern is a projection of the reciprocal lattice of the crystal [7, pg. 393-394]. The 2D lattice in direct space will correspond to a set of rods in reciprocal space. Constructive interference occurs at points where these rods, in reciprocal space, intersect with the Ewald Sphere. The position and radius of this sphere depend on the angle of incidence and wavelength of the incoming electron beam, respectively. The collection of rods intersect the Ewald Sphere at many points distributed along an arc. However, the diffraction pattern on the fluorescent screen shows streaks instead of dots. In real growths, the surface reconstruction is not perfect and this gives some width to the rods in reciprocal space. So, the dots are spread and appear as streaks.

Real crystal lattices have defects as well as thermal vibrations that give some width to the features in the reciprocal lattice and thus the patterns are rods and dots [35, pg.13]. The type of RHEED pattern that is shown reveals whether the growth mode is 2D (i.e. layer-by-layer) or 3D (i.e. islands form and grow in all dimensions) [6, pg. 490-493].

With the short wavelength of the high energy electrons, multiple rods will appear in the diffraction pattern as in Figure 3.2c [6, 7, 35]. However, when the surface is not atomically smooth the electrons will penetrate islands or other three dimensional structures on the surface. In this case, the diffraction pattern will correspond to a 3D structure. Instead of a set of rods, the reciprocal lattice of a 3D structure is a series of points and so the projection of it appears as an ordered arrangement of dots on the fluorescent screen, as in Figure 3.2a [6, 7, 35]. There is also a third case where the sample being grown is polycrystalline. In this case, the surface is comprised of many differently oriented smaller crystals and the diffraction condition is met along only specific angles between the incident beam and the surface but with every possible amount of tilt in the opposite axis because of the random orientation of these smaller crystals. This ring pattern is shown in Figure 3.2b.

By analyzing the diffraction patterns, the periodicity along a single crystallographic direction can discerned from the periodicity of the pattern. If measurements are taken at two different azimuthal directions, then periodicity of the unit cell that makes up the surface lattice can be determined [6, 7, 35]. The periodicity along the dimensions identifies which surface reconstruction is present.

Surface reconstruction is the rearrangement of atoms at the surface of a crystal [7, pg. 132-133]. This happens because, at the surface, the crystal lattice is terminated on one side and so the forces from neighbouring atoms that would normally keep the lattice in shape are absent. Even with their neighbours missing, surface atoms still 'want' to complete the



(a) (b) (c)

Figure 3.2: RHEED images (a) during the initial stages of the growth, (b) Shows an intermediate stage where rings form indicting a polycrystalline surface, and (c) after the MBE growth of InSb is completed.

same number of bonds to atoms. This imbalance changes the equilibrium arrangement of atoms near the surface [7]. Many different surface reconstructions may be possible and are dependent upon the growth conditions [35, pg. 14]. Different surface reconstructions can be associated with different surface stoichiometries which may affect how atoms are incorporated into the crystal [35].

The growth rate can also be determined by measuring the RHEED pattern intensity over time [6, 7]. During epitaxial growth the intensity of the diffraction pattern oscillates. As a new layer begins to grow, the surface of the substrate becomes rough which corresponds to a decrease in the intensity measured. Incidentally, this also gives an indication to the surface roughness. As growth continues, a new layer is completed and the intensity increases, peaking when the surface is atomically smooth again and a layer has been completed. Thus, the period of the RHEED oscillations corresponds to a single atomic layer growth [6, 7].

3.1.3 Other In-situ Characterization

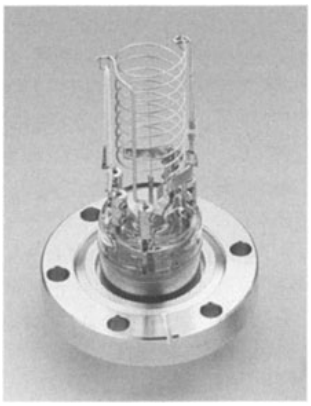
RHEED is one of the most useful and reliable in-situ characterization techniques but one of the great advantages of MBE is that it allows for many other in-situ techniques to be used simultaneously. Techniques such as pyrometry or band edge thermometry (BET) can be used to monitor the substrate temperature by analyzing reflected or emitted light from the

substrate surface [35, pg. 230-231]. Pyrometry works by measuring the absolute intensity of near infrared radiation emitted from the substrate surface. Based on Planck's law the measured intensity can be related to the surface temperature. This method however is influenced by light emission from other sources inside of the growth chamber such as the heated effusion cells [35, pg. 230-231]. This can lead to measurements with large error. A method that usually measures the temperature of the substrate more accurately is diffuse reflectance spectroscopy. Infrared light is transmitted through the substrate and collected from the backside of the wafer. The substrate material will have a bandgap absorption edge in this infrared spectrum and the positon of this absorption edge will depend on the temperature of the substrate [35, pg. 230-231]. Comparing the incident and output spectra to a precalibrated spectrum for the same type of wafer will reveal the temperature of the material in a method called band edge thermometry. BET only works if the substrate is thin enough to measure the transmitted infrared signal and so it only works for thin grown layers or for the intial part of a longer growth [35, pg. 230-231]. The two methods have their strengths and weaknesses and both methods complement each other.

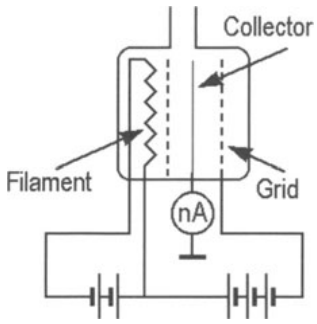
Other types of in-situ characterization include, vacuum ion gauges, beam flux monitor. and residual gas analysis (RGA). Ion gauges are the typical method for measuring vacuum pressures below 10^{-3} Torr [7, pg. 28-29]. Figure 3.3 shows the design of a Bayard-Alpert gauge. A typical hot-filament, Bayard-Alpert, ion gauge accelerates electrons from the filament towards a positvely charged grid separated by a short distance from the filament. Atoms and molecules that cross paths with the electrons will be ionized and charged. The ions are attracted to the nearby ground collector wire and a current can be measured. The collector current measured by the gauge can be converted to a pressure measurement [7, pg. 28-29]. The beam flux monitor is a specially configured ion gauge that can be used for pre- and post-growth calibrations. The beam flux monitor can be positoned in the center of the growth chamber where the substrate would be during growth and it is used to measure the flux from the effusion cells during calibrations. This is used to determine which temperatures to set each cell to, to get the desired flux for a specific growth.

The RGA is a quadrupole mass spectrometer that is used to measure the composition of the gas inside of the growth chamber [7, pg. 29-30]. RGA can be used to check the state of the system's vacuum. Leaks can be found by intentionally blowing helium gas into seams between components of the system. If there is a leak, the helium signal measured by the analyser will sharply increase above the normally very low noise floor. During growths this analyser can also be used to monitor the processes at the substrate surface. For example, if the substrate is heated such that the top layer on the surface is evaporated an increase will appear in the mass spectrometer at a mass corresponding to the evaporated elements.

The list of characterization techniques compatible in-situ with MBE is extensive and



(a)



(b)

Figure 3.3: (a) A photograph of a Bayard-Alpert Ion Gauge. (b) A schematic of the Bayard-Alpert Ion Gauge. Figures sourced from [7, pg. 29].

many techniques have been left unmentioned. Hopefully, this section has made clear just how versatile MBE can be and why it is worth the high costs and maintenance to establish a productive system.

3.2 Methods

3.2.1 Growing InSb

Taking the first step towards fabricating a nanostructure like the modelled one in Section 2.2.2, an InSb layer is grown using MBE. Growing InSb directly on GaAs is not a trivial task, even for MBE. The lattice constant of a crystal is the distance between unit cells in the crystal lattice. Each crystal has its own lattice constant. A larger lattice constant would mean that the space between atoms in the lattice is larger. The lattice constant of InSb is 14.6% larger than the lattice constant of GaAs [37]. This indicates that there is a large lattice mismatch between the two crystal lattices. This high lattice mismatch results in a 3D growth mode (Stranski-Krastanov mode) where islands form on the surface [35, pg. 26]. This means the surface is not smooth and therefore it becomes difficult to grow a layer that is free of defects. Fully aware of this challenge the decision was made to proceed with the plan to grow InSb directly on GaAs.

A 1/4 segment of a 3" semi-insulating GaAs (100) wafer was loaded into the MBE

system to be the substrate for the growth. Throughout the initial stages, the growth was monitored using RHEED. The growth starts with annealing of the substrate, heating the sample to $\sim 800°C$, to desorb any oxide from the surface. The next step is to grow GaAs buffer layer to ensure that the surface that the InSb will be deposited on is perfectly clean and atomically smooth. The chamber is first allowed to "soak" in As and then the Ga is released and the GaAs buffer layer is grown at a temperature of around $600°C$. This buffer layer is grown in two steps of 100 nm. At this point the RHEED pattern was streaky, similar to Figure 3.2c, indicating that the substrate surface was now atomically smooth and ready for the deposition of InSb.

Next, the As source is closed and then the Sb source is opened to switch the chamber to a Sb soak. Finally, the InSb layer can be grown. For the growth of the InSb layer a Sb/In ratio of ~ 1.7 is used with a target growth rate of 2 Å/s. The first attempt to grow the InSb layer was not successful. The substrate temperature just before the InSb layer growth was $450°C$ according to BET. The moment the In effusion cell shutter was opened the BET reading jumped to $500°C$ and the RHEED pattern quickly turned to a spotty pattern like Figure 3.2a. This indicated that the surface was getting rougher and the growth mode was 3D. The initial 3D growth mode is to be expected with such a large lattice mismatch. But what was not expected was the RHEED signal disappeared entirely just moments after the start of the growth. In a quick response the shutters were closed and the growth was halted. This is a perfect example of how advantageous the in-situ characterization during MBE can be. Without the RHEED images the growth may have been allowed to proceed to the end for it only to be found out that the grown film was not crystalline but amorphous. Instead, the process was stopped and the substrate was $600°C$ and the deposited InSb was evaporated from the surface. The desorption of In and Sb atoms created a clear signal in the RGA. To ensure the growing surface was again clean and atomically smooth another GaAs buffer layer was grown.

From the first attempt it was learned that the substrate temperature of $450°C$ was too high for growing InSb. From the resources of [38, 39] it was learned that InSb can be grown directly on GaAs (100) substrates if the initial growth temperature is very low. After verifying, using RHEED, that the substrate surface was again atomically smooth the growth was attempted again. This time however, as a precaution a thin 10 nm layer of aluminum arsenide (AlAs) was grown on top of the GaAs. This serves two purposes, (i) should the growth of InSb fail again the AlAs has a lower temperature of vaporization and should allow easier recovery of the atomically smooth GaAs surface and (ii) if the AlAs layer remains all the way through to the nanofabrication process it can be used later as a lift off layer to separate the InSb nanostructure from the GaAs substrate. This allows for the the InSb nanostructure to be transferred to a new substrate.

Description	Material	Thickness (nm)
Substrate	SI-GaAs (100)	6.25E+05
Buffer Layer 1	GaAs	100
Buffer Layer 2	GaAs	100
Lift Off Layer	AlAs	10
LT InSb Layer	InSb	20
HT InSb Layer	InSb	1500

Figure 3.4: A schematic representation of the layers grown by MBE.

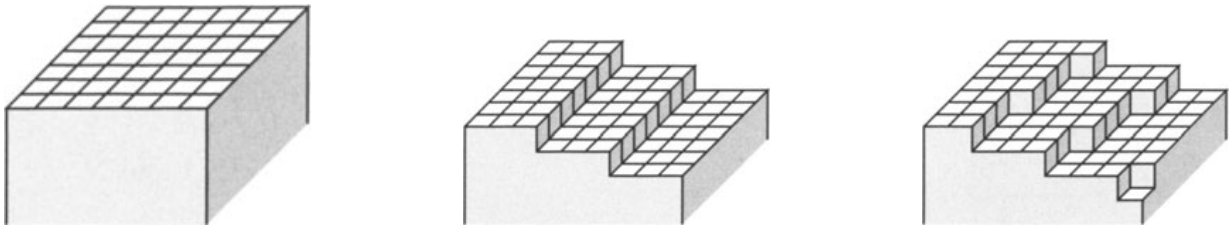
The InSb can continue to be grown on the relatively inexpensive and accessible GaAs substrates but then benefit from the properties of a new substrate. Other substrates that have the desired properties for a device may be expensive, or hard to acquire, or they could just be incompatible with MBE. With this technique the InSb could transferred to a substrate that is flexible or transparent to THz radiation.

For example, a substrate that is more transparent to THz radiation than GaAs (refer to Figure 2.6a) would be advantageous for any THz device that depends upon the transmission of THz radiation through this nanostructure.

The final growth structure is shown in Figure 3.4 with grown layers underneath the substrate. In Section 3.1.1, it was explained that the substrate is mounted upside down in the growth chamber with the sources underneath it. The final growth layers show the two-stage GaAs buffer layer with a total thickness of 200 nm followed by the 10 nm AlAs lift off layer. The next two layers are the low temperature (LT) InSb and high temperature (HT) InSb layers which were grown at $200°C$ and $350°C$ respectively. The LT layer is only 20 nm thick, while the HT layer is 1500 nm thick.

Monitoring RHEED during this second attempt, again initially showed a dotted pattern for 3D growth but instead of the RHEED signal fading out, when the HT layer began growing, the pattern transitioned from a set of dots to a set of rings as in 3.2b. This is a significant improvement because now the surface is forming a crystal, albeit a crystal with a high amount of defects, but still at least polycrystalline. After a short time of growing the ring pattern eventually transitions to a streaky pattern like Figure 3.2c and crystalline InSb has been grown directly on top of a GaAs substrate.

With the growth procedure for InSb developed, this growth was repeated three more times except on full 3" wafers this time. The same structure as in 3.4 was used for all three of these growths but they differed in the substrate misorientation. In the manufacturing of



(a) (b) (c)

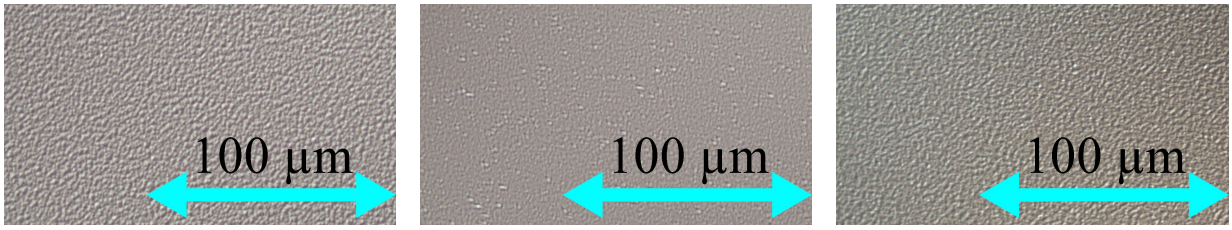
Figure 3.5: Three images of representative crystal surfaces. (a) A crystal surface cleaved with no misorientation to the (100) planes. (b) When the surface is cleaved with a misorientation from the (100) planes along a single crystallographic direction. (c) When the surface is cleaved with a misorientation from the (100) planes along two crystallographic directions. Figure sourced from [7, pg. 233].

GaAs substrate wafers, if the angle at which the bulk crystal is cleaved, when slicing it into a number of wafers, is off-normal to the (100) planes this is called a misorientation or it is sometimes referred to as the offcut angle. Figure 3.5 shows three cases of misorientation. When there is no misorientation, then the surface is atomically smooth as in Figure 3.5a. In the case of a misorientation in a direction that is parallel to one of the crystal lattice directions, a vicinial surface with steps is seen like in Figure 3.5b. In general, no substrate is guaranteed to be perfect with no misorientation and there will usually be some amount of misorientation with a surface like Figure 3.5c which occurs when the misorienation is along a direction that has a component in two directions along the crystal [7, pg. 232-233].

The three growths on these wafers were assigned the numbers of G0845, G0846, and G0859. G0845 had no intentional misorientation with respect to the (100) planes. G0846 had a nominal 2° offcut and misorientation in the direction towards the [010] crystallographic direction. Since this is parallel to one of the crystal directions, as long as there isn't a significant unintentional misorientation in the other direction, then this substrate's surface looks something like Figure 3.5b. G0859 had a nominal 0.55° misorientation in the direction towards the [0-11] crystallographic direction and so the surface of this substrate should look like Figure 3.5c.

3.3 Results

The next three chapters of this thesis will be about a variety of different characterization methods and nanofabrication methods used on samples that were cleaved from these three



(a) (b) (c)

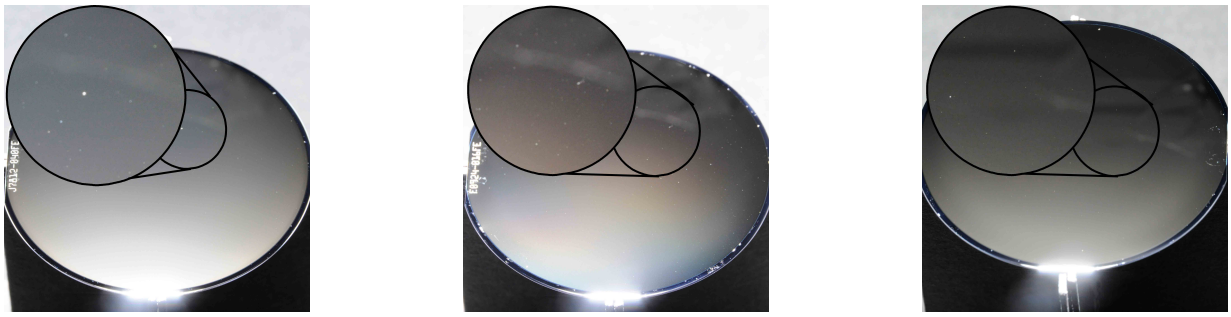
Figure 3.6: Three photographs from Nomarski interference contrast microscopy of the surfaces of all three wafers, G0845, G0846, and G0859. Each photo is taken with a 40 X lens. (a) Surface of G0845 (0° misorientation). (b) Surface of G0846 (2° misorientation). (c) Surface of G0859 (0.55° misorientation)

wafers. However some preliminary results from immediately after the growth can be shown here.

Figure 3.6 shows the surfaces of all three wafers as seen at 40 X zoom using Normarski interference contrast microscopy. The Nomarski images allow for a brief qualitative analysis of the surfaces. The images show that all of the samples do have a textured surface and that the surfaces appear uniform across the wafers. From an initial look, it appears that G0846, in Figure 3.6b has fewer hills on its surface than the other two samples. These surfaces will be explored further with different characterization methods in the following chapters.

Another quick method to study the surface is to take a picture of the light reflecting off of it's surface from the flash of a camera. With a bright enough light and the correct camera settings, defects in the surface of the wafer can be seen from the reflectance. The images in Figure 3.7 demonstrate this technique on all three wafers. This again gives a qualitative result revealing the overall density of these surface defects. Each of the reflectance images shows some defects across the wafers.

From the microscope and from the reflectance, it can be seen that none of the wafers are without their flaws, but just how flawed are they? To figure this out, more advanced characterization techniques will be used to study both the surfaces and the material underneath.



(a) (b) (c)

Figure 3.7: Three photographs showing the reflectance of the surfaces of all three wafers, G0845, G0846, and G0859. Each photo is taken with the same exposure time and along the same crystallographic orientation. (a) Surface of G0845 (0° misorientation). (b) Surface of G0846 (2° misorientation). (c) Surface of G0859 (0.55° misorientation). The primary flat is on the opposite side in (c) because it is the only EJ cut wafer of the three.

Chapter 4

Atomic Force Microscopy

The surfaces of the InSb crystals grown through MBE mentioned in Section 3.2.1 were studied by use of atomic force microscopy (AFM). This chapter will be a brief review of this surface study and its results.

4.1 Theory

Atomic force microscopy is a method of measuring the topography of a sample surface and it is capable of capturing vertical features with resolutions greater than 1 Å. The images produced by AFM are not from measurements of any kind of wave, EM radiation or otherwise. The images are instead produced from the interaction forces between a probe tip and the sample surface while the probe is scanned along the surface in an array of lines. The scanning probe has a tip that is just a few micrometers long with an atomically small point at the end. This tip is attached to a cantilever that is a couple hundred micrometers long [7, pg. 164-168]. To scan the surface, the probe is mounted into a position control system within the atomic force microscope (AFM). The position of the probe is measured and controlled with extremely high precision through the use of piezoelectric transducers.

Piezoelectric ceramics can expand or contract based on applied electric voltage. Piezoelectrics allow for precise control of motion in the range of less than 1 Å and up to a few μm [7, pg. 159]. This is how the AFM achieves such high resolution scans of atomic surfaces.

When a sample is brought near the probe, or vice versa, the probe tip is either attracted or repelled by the atomic interaction forces between the atoms of both materials [7, pg.

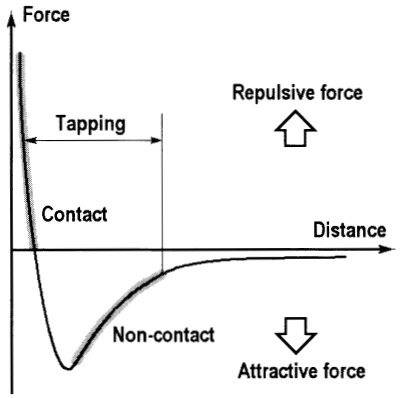


Figure 4.1: A graph of the force response curve from the combination of the long range attractive van der Waals forces and the short range repulsive Coulomb interaction forces between two atoms. Figure sourced from [7, pg. 166].

164-168]. This behaviour can be modelled by the combination of the long range van der Waals attraction forces and the short range electrostatic repulsion force in the interaction between two atoms [40]. The van der Waals forces are weak but act over a relatively long distance compared to the electrostatic forces. The combination of these forces gives a force profile similar to the one shown in Figure 4.1.

At far separation distances, greater than a few hundred Å, the force has weakened to the point of disappearing entirely. Once close enough, there will be a weak attraction force that increases as the distance decreases. When the probe and the sample are brought into close contact there is an electrostatic repulsion between the two as the electrons surrounding each atom repel each other [7, pg. 166].

Three of the basic operation modes for AFM are: non-contact, contact, and tapping mode. The main distinction between the three modes is which range of separation distance the probe is kept in during the device operation [7, pg. 164-168]. Contact mode corresponds to the left side of the curve in Figure 4.1 where the separation distance is shorter, around a few Å, and so repulsive forces dominate. The probe is kept in close proximity to the surface and the repulsive forces deflect the probe by slightly deforming the cantilever. If the cantilever is too stiff then the probe tip may scratch the surface.

As the probe is scanned across the surface, the cantilever bends, following the topography of the surface. The deflection of the probe tip is measured using one of many methods. The most conventional method is to reflect a laser off of the backside of the cantilever and measure the deflection by measuring how much the reflected beam spot moves. When

operating in constant-height mode, the deflection of the probe can be used to measure topographic features by calculating the force required to deflect the probe.

The preferred method however is usually constant-force mode, where the AFM uses this measured deflection in a feedback loop. The deflection of the probe is used to decide how to adjust the height of the probe to maintain a constant force, which then also affects the deflection and the loop continues [7, pg. 164-168]. This feedback loop keeps the deflection of the probe constant and therefore at constant force of interaction. In constant-force mode, the height of the surface along each scan line can be taken simply as the height of the probe along the scan line.

Non-contact mode uses the weakly attractive forces that occur when the sample is kept at a further distance from the probe (more than 10 Å) [7, pg. 164-168]. In this mode, the cantilever is vibrated at frequencies typically from 100 kHz to 400 kHz which is at or near the cantilever's resonance frequency. The amplitude of these oscillations is on the order of 10s of Å. When the probe is interacting with the sample surface, the resonance frequency of the cantilever is changed. Again using a feedback loop with the measured deflection, the height of the probe is adjusted to keep the resonance frequency of the oscillations constant. This is called constant-gradient mode because the tip is effectively tracing lines of constant gradient in force as the probe is scanned along the surface [7, pg. 164-168].

The final mode, tapping mode, is a combination of the previous two modes. Tapping mode is essentially the same as non-contact mode however the sample is brought closer to the probe at the lowest point in the oscillation. At this point, the tip approaches the surface, hardly tapping it [7, pg. 164-168]. This mode has higher resolution like contact mode but it removes much of the risk of scratching the surface because the tip only briefly touches the surface and so it is not dragged across it [40]. This is is what has made tapping mode the most often chosen method for AFM [40].

Each mode has their benefits and drawbacks, depending on the sample being characterized. For a long time AFM has been an important tool in the characterization of materials since its conception in 1986 [41], and the technique has been developed well beyond the principles described here. The technique used in this research uses a different mode called PeakForce Tapping along with ScanAsyst for automatic optimization of the feedback loop. PeakForce Tapping can be considered similar to regular tapping mode but the frequency of the cantilever oscillations is intentionally below the resonance frequency. The details behind the operation of this mode can be found in the documentation provided here [40] by the company, Bruker, that developed these techniques.

Using AFM images, an analysis can be done to determine the facet angles of the pyramidal mountains, called hillocks, that form on the surface of a wafer. To understand the

formation of these hillocks we can invoke the terrace ledge kink (TLK) model of crystal growth [7, pg. 229]. In the TLK model, dislocations can occur within the crystal lattice either due to point defects such as vacancies or interstitial atoms. These dislocations cause a shift of all of the neighbouring atoms along what is called the dislocation line. This line extends until it either reaches another dislocation or until it reaches the end of the crystal. At the point where the dislocation emerges, screw dislocations on the surface can form.

A screw dislocation emerges from the surface as a misaligned step where one end of the step ends at a layer below the other end of the step, forming a gradual slope. When growth occurs around this step, deposited adatoms find it energetically favourable to join the lattice at this ledge created by the screw dislocation. However, because this ledge does not extend along the entire length of the surface, growth in this way will effectively create another partial ledge facing the perpendicular direction. As the growth continues, ledges keep forming in perpendicular directions around the point of emergence of the screw dislocation. This results in a spiral growth pattern creating surface features like Figure 4.2a and these are referred to as hillocks.

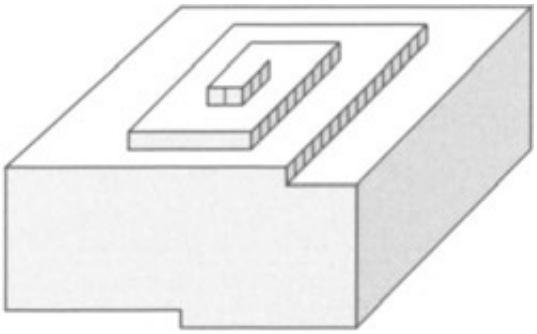
4.2 Methods

4.2.1 Surface Roughness

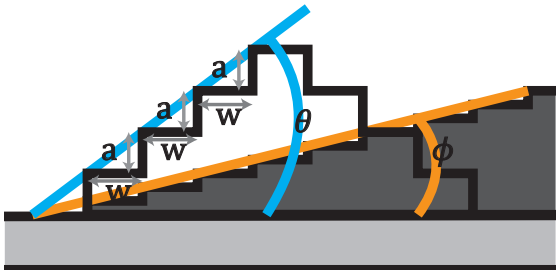
All three of the growths from Section 3.2.1 had their surfaces imaged by AFM scans to examine the difference in their topography. The equipment used to perform the AFM scans was a Bruker Icon/Fastscan AFM and the mode used was the aforementioned PeakForce Tapping mode. Comparisons of these scans should reveal how the surface topography changes as the misorientation of the substrate surface changes.

One way to quantify this dependence is by measuring the roughness of the surface. The roughness of the surface is determined using the Bruker NanoScope Analysis software. The built-in tool provides a quick measurement of the roughness of any region from an AFM scan. Typically the roughness is reported either as the average roughness, R_a or the root mean square (RMS) roughness, R_q. Here R_q will be reported. The RMS roughness is calculated by Equation 4.1 [42].

$$R_q = \sqrt{\frac{\sum (z_i)^2}{N}} \tag{4.1}$$



(a) The formation of a hillock from a screw dislocation. Figure sourced from [7, pg. 245]. Used with permission of Springer, from Surface science : an introduction, Oura, K., © 2003; permission conveyed through Copyright Clearance Center, Inc.



(b) The facet angle, θ, of a hillock and the misorientation, ϕ, of the substrate surface.

Figure 4.2

z_i is the measured height by the AFM and N is the number of measured points considered in the calculation. A proper measurement of roughness however should not be affected by things such as plane tilt or bowing of the surface. Roughness should be a measure of the features on the surface and not of the overall curvature of the surface, nor should it depend on how the surface is tilted due to the way the sample is mounted or otherwise.

To correct for this, plane fitting or flattening of the surface should be done before roughness calculation [42]. Thankfully, NanoScope Analysis has a tool for this as well. So, each image was flattened. This is done by calculating a least-squares fit to a second degree polynomial, $z = a + bx + cx^2$, for each scan line and subtracting it from the measured values [42].

RMS roughness is a quick and easy way to describe surface roughness but, it is not the answer to every question. When reducing all of the information contained in hundreds of thousands of pixels into one single number, many characteristics of the surface may be neglected. RMS roughness for example, can not distinguish between surfaces that have different frequency profiles. This means two surfaces can have the same RMS roughness if they have bumps of the same height out-of-plane but the bumps are different sized in-plane [42].

4.2.2 Effects of Varying Offcuts

In complement to the study completed by Peyton Shi in [43, pg. 106-113], a similar study of the effect of substrate offcut on surface topology was completed. It begins with finding the natural facet angle of the hillocks on the surface.

The facet angle of a hillock is the angle measured as θ in Figure 4.2b that averages the slope of a hillock with respect to an atomically flat surface. It is a reasonable assumption that the steps of the hillock each increase the height from the flat atomic surface by one monolayer, i.e. each step has the height of 1/2 the lattice constant a. The width of each terrace is w and its value will depend on the material being grown, as well as its growth conditions such as substrate temperature [43, pg. 101-106]. All of the hillocks grown should have the same natural facet angle, and this can be used to our advantage. It should however, be noted that the facet angle in different crystallographic directions does not need to be equal and a hillock may have a steeper slope along one direction than it does along the other.

To determine the facet angle, the terrace width must be determined. The simplest way to determine this is to draw a line that intersects the steps at a 90° angle in the AFM image. The average terrace width is then calculated as $w = l/n$ where l is the length of the line and n is the number of periods in surface height (steps) along this line. The facet angle is then calculated from $\theta = \tan^{-1}(a/(2w))$.

This natural facet angle should be the same across all three growths of G0845, G0846, and G0859 because they were all grown with the same growth conditions. However, the substrate surface misorientation of the final two growths will obscure this measurement. For example if the surface is misoriented by the angle ϕ as in Figure 4.2b then the hillock as observed along the tilted surface will appear to be partially buried under the surface.

In Figure 4.2b, when the surface is misoriented, the growth will occur on a vicinial surface, as mentioned in Section 3.2.1. The steps already present on the substrate surface suppress one side of the hillock. For the hillock to grow with its natural facet angle, one side of the hillock must have fewer steps to compensate for the tilted base of the hillock, due to misorientation. If the AFM image is corrected for this tilt, then the hillocks will appear as if they have been tilted instead. So, to grow a crystal that is free of hillocks, the misorientation of the surface should be greater than or equal to this natural facet angle. Further issues, such as step bunching, arise if the misorientation is significantly greater than this natural facet angle and the surface becomes rough again [43, pg. 107].

Since the amount of the hillock that is hidden depends on the misorientation of the sample, this provides a method of measuring the misorientation of the surface by measuring

the lengths of sidewalls of the hillocks. If we maintain that a positive sidewall of a hillock is the side that is fully uncovered by any misorientation that is less than the facet angle and a negative sidewall is thus the side that appears shorter, then the misorientation of the surface can be found from any hillock as Equation 4.2. ϕ_i is the misorientation at the sample point i while S_+ and S_- are the positive and negative sidewall lengths. The total uncovered length of the hillock is $L \neq S_+ + S_-$ because the width of the top terrace should be counted twice, once going up and once going down. This allows the misorientation to be calculated with only two measurements, the positive sidewall and the total length. It should be noted that this method of determining the offcut is affected by changing the terrace width w.

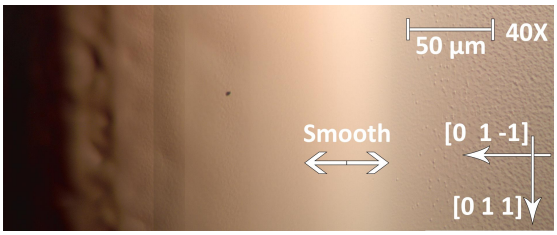
$$tan(\phi_i) = \frac{S_+ - S_-}{L} \times \frac{a}{2w} = \frac{a}{2wL}(2S_+ - L - w) \tag{4.2}$$

There is one last trick. The effects of changing the misorientation for a number of different misorientaions can actually be seen on only a single wafer. The polishing procedures that are standard in the production of common stubstrate wafers actually leaves a curved edge around the entire circumference of the substrate. This effectively gives a gradual increasing misorientation near the edges of the wafer. The shape in Figure 4.3b is an exaggeration of this effect. In Figure 4.3a, a composite of Nomarski images with varying focus show the bowing of the surface near the edge. The composite image technique was used to keep the surface in focus even as it dropped away from the microscope nearer to the edge.

Even though the entire edge of the substrate wafer will have this curved edge, the substrate holder that carries the substrate around the MBE system covers almost the entire edge and no growth occurs on these parts. Without designing a special substrate holder, the only segments along the edge of the wafer that are exposed to the molecular beams are at the primary and secondary flats of the wafer. The wafer flats are key in keeping track of the crystallographic orientation of the wafer, they just also happen to be useful for studying the effects of misorientation.

4.3 Results

Figure 4.4 shows AFM 10 μm by 10 μm scans of three growths of InSb on substrates with misorientations of 0° (4.4a), 2° (4.4b), and 0.55° (4.4c). From these images the RMS roughness was calculated by using the flattening and roughness tools as described



(a)

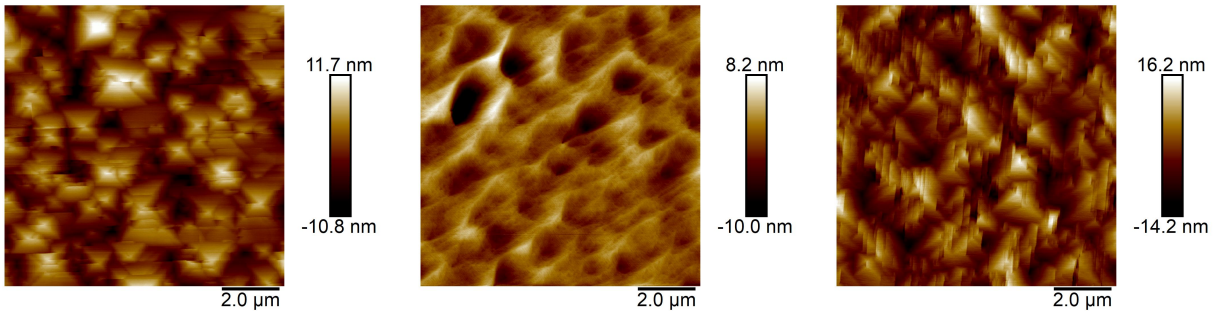
(b)

Figure 4.3: (a) A composite of images from a Nomarski microscope showing the bowing of the G0845 substrate surface at the primary flat. The vertical lines that seem to bisect the entire image are the result of combining multiple images that had slightly different exposures and are not a feature of the surface. (b) An illustration emphasizing the rounded edge of a one side polished substrate wafer.

above. The lowest roughness measured was from G0846 (2° misorientation) with an RMS roughness of 2.28 nm. In the other two growths, the RMS roughness was measured to be 3.24 nm and 4.27 nm for G0845 (0°) and G0859 (0.55°), respectively.

By visual inspection, G0846 does appear to have the smoothest surface and G0859 does appear rougher than G0845. In the case of G0846, the surface is smoother than the surface of the nominally 0° misorientation growth, G0845, and this makes sense because of the suppression of hillocks mentioned above.

On the other hand, G0859 has a small misorientation that should slighty suppress the hillocks and make the surface smoother than the 0° misorientation wafer. But this is not



(a) (b) (c)

Figure 4.4: 10 μm by 10 μm AFM scans of (a) G0845, (b) G0846, and (c) G0859

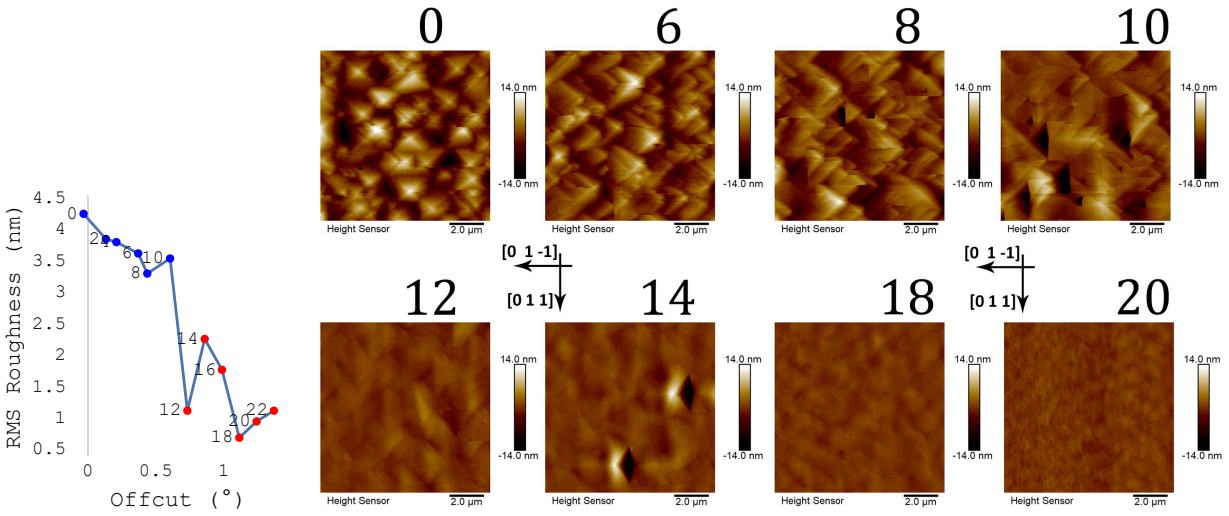
the case. To be discussed later in Chapter 5, an issue was discovered with the G0859 substrate when examining the sample with XRD. Further evidence that something strange has happened is that another AFM scan of a different region of the wafer gives an RMS roughness less than the roughness of the G0845 wafer. So, the surface of G0859 is not homogeneous.

Studying G0845, at the center of the wafer, the natural facet angle of the hillocks can be found. Counting the number of steps along a line parallel to the [0 1 -1] direction in $1\mu m$ by $1\mu m$ AFM scans gives an average terrace width of 34.47 nm. The lattice parameter for InSb can be determined by XRD. In the case of G0845, XRD shows the InSb layer is fully relaxed, meaning that the lattice parameter for bulk InSb at room temperature can be used. This lattice parameter is 6.479 Å [44, pg. 156], giving a hillock facet angle of 0.54°.

So, the optimal offcut for obtaining an atomically smooth surface will give a misorientation of at least 0.54° towards the [0 1 -1] direction. To truly get the best offcut you must also account for the hillock facet angle in the [0 1 1] direction. However, as mentioned earlier, the only part of polished edge that is exposed during growth is the substrate flats. The [0 1 1] direction corresponds to the secondary flat of the wafer. Before this study of misorientations, the wafer had been cleaved into quarters, and so an examination of the secondary flat, which is already shorter than the primary flat, was unfeasible. The offcut of G0859 is 0.55° towards the [0 -1 1] direction. The facet angle will be identical in the [0 1 -1] and [0 -1 1] directions. So, this almost exactly matches the optimal offcut that is predicted from the natural facet angle. But from the high density of defects showing in image 4.4c it is difficult to say whether this improved roughness or not.

To verify this hypothesis, a series of $10\mu m$ by $10\mu m$ and $1\mu m$ by $1\mu m$ AFM scans were made on the primary flat of the G0845 wafer. Figure 4.5b shows a selection of these $10\mu m$ by $10\mu m$ images. Each image is numbered and larger numbered images are closer to the edge of the wafer and thus have an increased effective offcut along the [0 1 -1] direction. In the first image, the surface appears rough and is overcrowded by hillocks. Hillocks in image 0 appear to have even length positive and negative sidewall lengths. However by image 6 the hillocks now appear tilted since the negative sidewalls have become significantly shorter than the positive sidewalls. This trend continues until in image 12, the hillocks have disappeared entirely and the surface appears very smooth.

The RMS roughness for each image is plotted versus the effective offcut in Figure 4.5a. Since the method of determining the effective offcut from Equation 4.2 depends on the appearance of the hillocks, it may be somewhat subjective. The hillocks in each image are overlapping and of various sizes. This can make it difficult to measure the length of the



(a) (b)

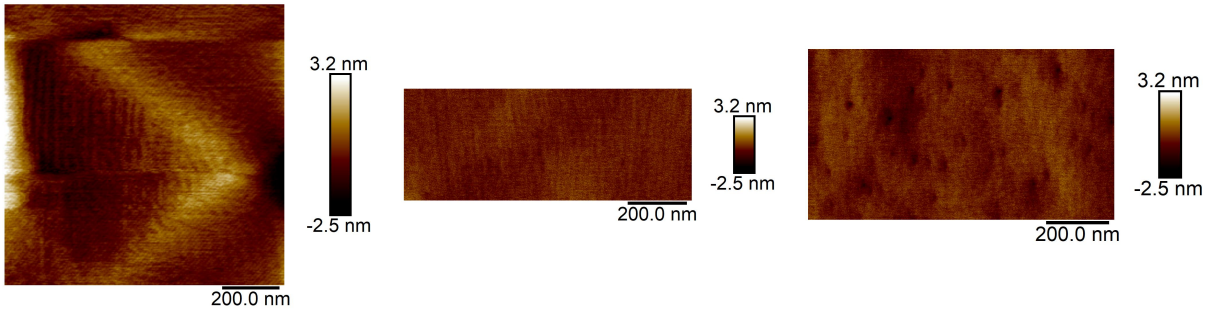
Figure 4.5: (a) A plot of the RMS roughness as the effective offcut changes. (b) An arrangement of AFM scans proceeding closer and closer to the substrate edge.

positive and negative sidewalls. Care was taken in choosing larger hillocks that appeared relatively unobstructed.

Additionally, this method can't be used once the hillocks have disappeared. So, for images closer to the wafer edge than image 10, the effective offcut was calculated by linearly extrapolating from the offcuts of the previous images and these data points have red markers. The accuracy of this is questionable but, since images were taken in somewhat regular intervals approaching the edge of the wafer, it will give a rough estimate.

Overall, these measurements of offcuts should be considered to have a high uncertainty. However, it is a good sign that the offcuts measured for images 10 and 12 are 0.60° and 0.73° which is just above the natural facet angle and it is where the transition to a smoother surface is seen.

The material grown here is InSb on GaAs with no buffer layers and unique growth conditions. This will have different growth characteristics to the aluminum indium antimonide (AlInSb) material which was used in the complementary study [43, pg. 106-113]. Nevertheless, the results of both studies do show a similar trend in the increased suppression of hillocks as the offcut increases. In contrast to [43], Figure 4.5a does not show a significant increase in surface roughness, even once the offcut angle is significantly greater than the



(a) (b) (c)

Figure 4.6: Three tighter view, 1 μm across, AFM scans of the surface of G0845. (a) Image number 1 showing a hillock at a 0° offcut with RMS roughness 0.727 nm (b) Image number 15 with an effective offcut of 0.86° and RMS roughness 0.258 nm. (c) Image number 21 with an effective offcut of 1.24° and RMS roughness 0.351 nm.

natural facet angle. This could suggest that with these growth conditions there may be a wider range of offcut angles that do produce an atomically smooth surface. On the other hand, this may also be the result of poor measurement of the offcut angle and Image 22 may not really be that far from the optimal offcut angle anyways.

The latter case seems more likely when studying the Nomarski image of Figure 4.3a. It can be seen here that the surface does have an increased roughness, once outside of the range indicated by the arrows in the figure. Roughness does appear to increase significantly as the edge of the substrate (left side) is approached.

Some of the late stage surface morphologies can be seen in Images 15 (4.6b) and 21 (4.6c) in Figure 4.6. While both show smooth surfaces, Image 21 does begin to develop depressed bumps in it surface and it does become a little bit rougher. Another point to consider is that the substrate polishing technique was not developed for the purpose of studying substrate offcuts. At the very edge of the substrate we can not guarantee that the roughness comes from step bunching or instability due to the large offcut. It may also be a result of the mechanical polishing of the substrate edge.

Chapter 5

X-Ray Diffraction

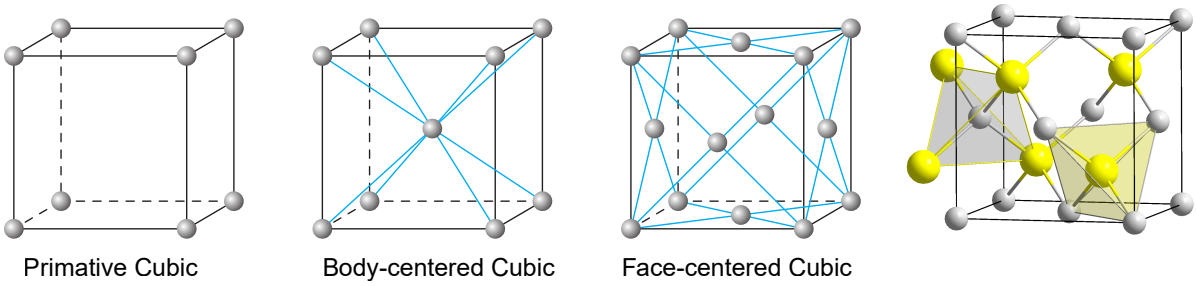
While AFM studies probe the surface of a sample, giving images with height data that can be directly interpreted as the topographical surface, XRD, by contrast, is used to study deeper below the surface revealing characteristics of bulk materials. XRD studies can be carried out in-situ, or ex-situ (during or after growth) in much the same way as RHEED and other diffraction based methods can. The main distinction between RHEED and XRD is that RHEED uses the wave nature of electrons and XRD uses the wave nature of photons.

Where high energy electrons may only penetrate as far a 10 Å, X-rays can penetrate up to mm into a sample [33, pg. 405]. In most XRD studies, the influence of the surface is about five orders of magnitude weaker than that of the bulk crystal [7, pg. 66]. In special configurations XRD can be used for surface analysis, but, in this chapter, only the bulk analysis will be considered.

XRD and crystallography in general are topics large enough to deserve multiple textbooks dedicated to them alone and so, in this relatively brief dissertation, only the fundamental material necessary to comprehend the results of this research will be included. For a more thorough examination of these topics readers are referred to [45] and [46].

5.1 Theory

In Section 3.1.2, the reason for the appearance of the electron diffraction pattern was vaguely said to be due to constructive interference. However, the condition for constructive interference was never mentioned. It will be explained here.



(a) (b)

Figure 5.1: (a) The unit cell of cubic Bravais lattice types. Figure adapted from [8]. (b) Zincblende crystal structure. Indium is yellow and antimony is grey. Figure sourced from [9].

5.1.1 Crystal Structure

Crystals are solids made up of an ordered arrangement of atoms or molecules that, together, form a periodic structure that corresponds to a lattice type [45, pg. (1-26)]. In three dimensions there are only 14 possible types of a lattice. These are called the Bravais lattices [45, pg. (1-22)-(1-23)]. The lattices are arrays of points that have repeating unit cells that fill all space. Any lattice that can be imagined in three dimensions can be reduced to one of the Bravais lattices. Only three of these lattice types have cubic symmetry and the unit cells for these are shown in Figure 5.1a.

The crystal structure of InSb is the zincblende structure, Figure 5.1b [44, pg. 154]. In the zincblende structure one face-centered cubic lattice of indium and one of antimony interpenetrate each other. This creates tetrahedral shapes where each atom, indium or antimony, is bonded with four of the opposite type of atom.

5.1.2 Miller Indices

To study crystal structure only the unit cell needs to be considered. To distinguish between different crystallographic directions and planes, the Miller index notation is standard in crystallography [45, pg. (1-2)-(1-4)]. Figure 5.2a shows a cubic unit cell. Since the unit cell of the InSb crystal structure is cubic, only cubic unit cells will be considered here. In general, the unit cell does not need to be cubic but, for simplicity, here it will be. In

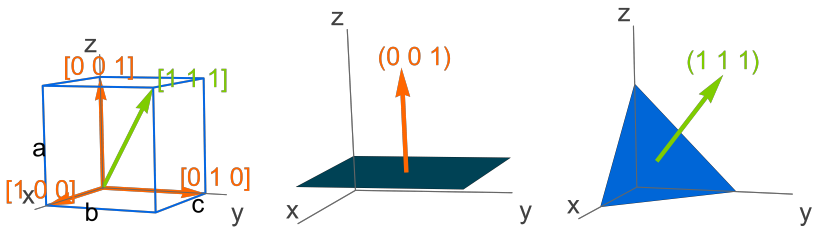


Figure 5.2: Three graphical representations of crystal structure. (a) The unit cell of a cubic crystal structure. The arrows indicate different crystallographic directions. (b) and (c) the (001) plane and the (111) plane, respectively, and their respective normal vectors.

general, the unit cell is usually a parallelepiped shape. Figure 5.2a shows there are three lengths of the unit cell, a, b, and c. These are referred to as the lattice parameters.

Choose basis vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ such that $\mathbf{a} = [001]$, $\mathbf{b} = [010]$, and $\mathbf{c} = [100]$. The length of each basis vector is a, b, and c, respectively and the directions are along the edges of the unit cell. Since we are only considering a cubic cell here $b = c = a$ and the basis vectors are orthogonal as in Figure 5.2a.

Crystallographic directions are indexed by $[hkl] = h\mathbf{a} + k\mathbf{b} + l\mathbf{c}$ while the notation for planes is with round brackets, (hkl) [45, pg. (1-2)-(1-4)]. In cubic lattices, not in general, the direction perpendicular to a plane (hkl) is given by the same indices [hkl]. Figure 5.2a shows two crystallographic directions [001] and [111]. Figures 5.2b and 5.2c show the corresponding planes. It is important to note that although only one plane is shown in the figure, a crystal contains many planes separated by distance d_{hkl} that all correspond to (hkl).

The indices h, k, and l are almost always integers, except for in a few rare cases. However, they can also take on negative values, and in this case the standard notation is to mark the number with a bar like so, $[\bar{1}\bar{1}1] = [-1\ -11]$ [45, pg. (1-2)-(1-4)].

Families of equivalent planes, again in cubic lattices only, are permutations of the same indices, positive or negative. The family of planes is denoted by curly brackets {hkl}. Each plane in the family is equivalent in terms of the symmetry of the crystal. The same is true for a set of directions, indicated by angled brackets, ⟨hkl⟩. For example, the {111} family contains eight planes of (111), $(\bar{1}11)$, $(1\bar{1}1)$, ...

5.1.3 Reciprocal Lattice

The lattice in the previous section is referred to as the direct space lattice because it is the physical lattice that corresponds to the crystal structure of real materials. The reciprocal lattice on the other hand is a hypothetical tool that is used to understand and interpret results from experiments with crystals.

The reciprocal lattice is defined by another set of basis vectors called $\mathbf{a}^*$, $\mathbf{b}^*$, and $\mathbf{c}^*$ [45, pg. (2-2)-(2-4)]. In the case of a cubic lattice in direct space, the reciprocal lattice vectors are parallel to their direct space counterparts. The length of these basis vectors is inversely proportional to the length of the direct space basis vectors such that, $|\mathbf{a}^*| = 2\pi/a$, $|\mathbf{b}^*| = 2\pi/b$, and $|\mathbf{c}^*| = 2\pi/c$.

We can define a reciprocal lattice vector $\mathbf{Q}_{hkl} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$. Each point in the reciprocal lattice is denoted by hkl, without any brackets this time [45, pg. (2-2)-(2-4)]. $\mathbf{Q}_{hkl}$ is the vector from the origin to each point hkl. The magnitude of $\mathbf{Q}_{hkl}$ is $2\pi/d_{hkl}$ where d_{hkl} is the distance between the (hkl) planes in direct space. For cubic lattices the distance corresponds to

$$\frac{1}{d_{hkl}^2} = \frac{h^2 + k^2 + l^2}{a^2} \tag{5.1}$$

So, higher indexed planes are closer together in direct space, and correspond to points further apart in reciprocal space. The reciprocal lattice is just another array of points. In reciprocal space, every point at the tip of $\mathbf{Q}_{hkl}$ corresponds to a set of planes, (hkl), in direct space.

5.1.4 Diffraction

Finally we are ready to talk about diffraction. The discussion in this section applies to anything wave-natured, energy or matter, such as EM waves or high energy electrons as long as the wavelength is comparable to the spacings of the crystal lattice.

Bragg's law treats crystal planes as if they were mirrors [45, pg. (3-3)-(3-6)]. The incident wave-vector is $\mathbf{k}$ at an angle θ with respect to the surface and the reflected wave-vector is $\mathbf{k}'$. For a reflection from a mirror, the reflected angle is equal to the incident angle θ. Note that this means the angle between the incident and reflected rays is 2θ.

The scattering vector is defined as $\mathbf{\Delta k} = \mathbf{k}' - \mathbf{k}$. If scattering is elastic then energy is conserved and $|\mathbf{k}'| = |\mathbf{k}| = 2\pi/\lambda$ where λ is the wavelength of the electrons or x-rays. The

Bragg angle is θ_B and it is the incident angle where intensity peaks are seen in diffraction patterns.When the incidence angle is equal to the Bragg angle, the tip of the scattering vector overlaps a reciprocal lattice point. Equation 5.2 is the Laue condition [45, pg. (3-3)-(3-6)].

$$\mathbf{\Delta k} = \mathbf{k}' - \mathbf{k} = \mathbf{Q}_{hkl} \tag{5.2}$$

By equating the magnitude of the scattering vector and the magnitude of the reciprocal latice vector you will arrive at Bragg's Law, Equation 5.3 [45, pg. (3-3)-(3-6)].

$$n\lambda = 2d \sin\theta \tag{5.3}$$

n is a positive integer and it represents the order of diffraction and d corresponds to the separation distance between one set of (hkl) planes. Higher order diffractions correspond to the periodic sidebands that are usually seen around a central maximum of intensity in a diffraction pattern.

Instead of indexing by the integer n, crystallographers use the concept of the reciprocal, lattice treating higher order diffraction peaks as having reflected from hypothetical planes that have higher indexes (hkl) [45, pg. (3-3)-(3-6)].

So, for each set of planes in the direct space lattice, there will be a corresponding reciprocal lattice point and a unique θ_B to capture it. To capture a peak intensity, the angle of the incident wave-vector (θ) must be at the Bragg angle (θ_B), with respect to the plane. The position of the detector must also be at θ_B with respect to the plane. That means that the angle between the incident and diffracted beams will be twice the Bragg angle (2θ). This will satisfy the Laue condition in Equation 5.2.

However for real crystals the Laue condition is not the only requirement to see a diffraction peak. There are many other factors that determine the strength or existence of a peak. Atomic form factors, crystal structure factors, geometry factors, texture factors, polarization factors and absorption factors all depend on either aspects of the crystalline material being examined or aspects of the experimental setup.

This limited discussion of the background will end here but it should be understood that there is a lot of complexity to XRD that is omitted here.

5.2 Methods

5.2.1 High-resolution X-ray Diffraction

High-resolution x-ray diffraction (HRXRD) is exactly that, XRD but with very high angular resolution. The reciprocal lattice of a perfect ideal crystal would have points with infinitesimal width. Real crystal measurements however suffer from instrumental effects as well as crystal defects, thermal vibrations, curvature, etc. that all contribute to increasing the width of these points, effectively spreading the intensity in reciprocal space.

If the x-ray source remains stationary, the reciprocal space can be probed by measuring the intensity of x-rays while rotating the sample through an angle θ and the detector through 2θ. As θ approaches θ_B the intensity rises and peaks before dropping again, making a Gaussian peak shape. These are commonly referred to as Bragg peaks.

In general, the incident angle of the x-ray can be changed separately from the angle at which diffracted x-rays are measured by the detector. This is why the axis controlling the incident angle is commonly referred to as ω instead of θ. Since each axis can be controlled independently, different types of scans can be made. For example, when both the ω and 2θ axes are moved together such that $2 \times \omega$ is always equal to 2θ, the scan is called symmetric. This type of scan is also called a theta-two-theta scan or an omega-two-theta (ω-2θ) scan.

In HRXRD, the sample is loaded into a diffractometer that allows the sample to be rotated into nearly any orientation in 3D space. Figure 5.3 shows a schematic of the relevant angles in a four circle diffractometer. There are four relevant angles, ω, 2θ, χ, and ϕ.

ω determines the angle of incidence of the x-ray onto the sample while 2θ determines the angle of measurement which is the angle between the incident and diffracted x-rays. χ allows for the sample to be tilted around the axis that is perpendicular to the ω axis but in the plane of the sample holder surface. Finally, ϕ is the angle that rotates the entire sample holder around the axis that is normal to the holder surface.

With high quality single crystal specimens such as those produced by MBE the effects that are broadening these peaks are reduced but not eliminated as the crystal quality approaches perfection. The features of one of these so-called Bragg peaks can be on the order of arcseconds. An arcsecond is 1/3600th of a degree and denoted by 1" = 1 arcsecond. Needless to say, making measurements with such high angular resolution is quite a feat of technology yet advancements in this field have made it commonplace.

What's not included in Figure 5.3 is the multitude of x-ray optics that is required to achieve these high resolutions. The best results for HRXRD are achieved when the incident

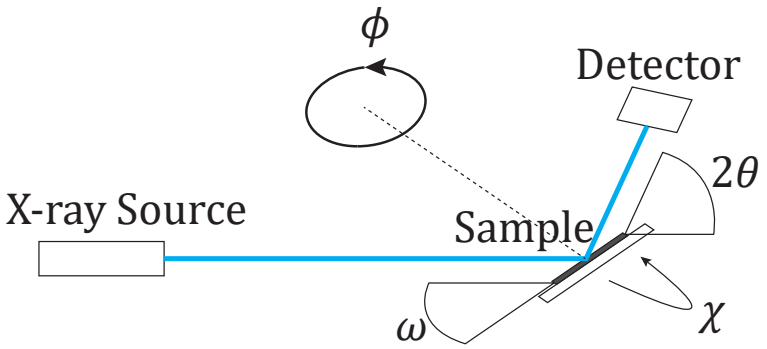


Figure 5.3: A schematic of the diffractometer and the angles that relate to the system in a four circle geometry.

beam has a low divergence and is monochromatic. That is to say the x-ray wave should be a plane wave of a single wavelength. This will ensure that each incident ray is incident at the same angle and has the same wavelength so that the only Bragg peaks seen are the ones selected by the diffractometer motor positions.

Additionally, the beam can be conditioned after reflecting as well to measure different types of results. A double-axis scan is one that has an unobstructed detector measuring x-ray reflections. That means all of the light that is reflected from the sample is captured by the detector as long as it falls within the aperture of the detector [46, pg. 193-194]. This is the simpler option but it means that reflections that are slightly off of the Bragg angle will be integrated into the intensity measurement. This is not always a bad thing depending on what characteristics of the crystal you are looking to measure. However, it does ultimately lower the resolution of the scan and sharper peaks may be obscured [46, pg. 193-194]. Triple-axis scan configurations add an additional crystal between the sample and the detector. This crystal has a channel cut into it so that diffracted x-rays only reach the detector if the angle of diffraction is within a small range of values selected by the crystal. This crystal is called an analyzer and it is on a rotating axis that can be scanned, changing precisely which angle of reflected light is selected to reach the detector [46, pg. 193-194]. This small acceptance angle allows for the effects of defects or curvature to be removed from the intensity scan so that they may be distinguished from reflections that would correspond to nearly the exact Bragg angle [46, pg. 193-194].

5.2.2 Experimental Setups

XRD measurements were taken in two different laboratories using two different experimental setups. The first experimental setup was done in-house with a Bruker QC3 system. This system uses an x-ray tube as its source. X-ray tubes generate x-rays by accelerating high energy electrons at a metal target. If the energy of the collision is great enough, an inner shell electron can be excited to a higher state. This is an unstable state and when a subsequent downward transition occurs to fill the vacancy, the excess energy is released as a photon with an x-ray wavelength [47, pg. 44-45]. X-ray tubes are a practical source of high intensity x-ray beams with precise control of wavelength appropriate for use in XRD. The x-ray tube used is equipped with a copper (Cu) target. The QC3 system is also equipped with a multilayer mirror and channel-cut germanium (004) crystal to condition the incident x-ray beam. The result is a quasi-parallel beam of Cu Kα x-rays with beam divergence less than 10" and the x-ray wavelength is 1.54 Å. This system is also capable of being operated in double-axis or triple-axis configuration. When in triple-axis configuration the two-bounce channel cut silicon (022) analyzer crystal limits the detector acceptance angle to just 12". Finally, the detector is a scintillation detector that, in double axis configuration, has an adjustable size slit. X-rays striking anywhere on the entire area of the scintillation detector are recorded as a single data point.

The other experimental setup used synchrotron radiation from the BXDS-IVU beamline at the Canadian Light Source (CLS). Synchrotrons generate EM waves through bremsstrahlung radiation by accelerating electrons using magnetic fields [47, pg. 108-109]. There is great control over the path of the electrons and they can be made to emit EM waves with a wide range of energies, including x-ray energies. In these experiments the beamline was configured to emit x-rays with a wavelength of 1.24 Å. The setup of a synchrotron beamline is more complex and goes outside the scope of this research. However, the incident beam assuredly was conditioned to be monochromatic with a low divergence comparable to that of the QC3 system.

Diffraction measurements at the CLS have a few advantages over the ones done in-house with QC3. First of all, the synchrotron is a significantly brighter source of x-rays, meaning weaker Bragg peaks can be detected. Although in this case it was not used, synchrotrons also allow for the energy of the x-ray to be changed and this can be useful for resonance anomalous x-ray scattering experiments [45, pg. (3-51)-(3-53)]. Finally, the diffractometer at CLS was equipped with a 2D array detector that allows x-ray intensities to be measured at discrete points in a 2D area. This is significant because the images recorded by the detector show the shape of the Bragg peaks as a 2D projection of the reciprocal lattice. Furthermore, this means that by combining all of the 2D projections

through a regular ω-2θ scan, a 3D map of the reciprocal space around the Bragg peak can be made. This is fantastic because the full 3D reciprocal space map (RSM) contains a great deal of information about the crystal lattice that without an array detector would take a significantly long time to collect.

5.2.3 Reciprocal Space Maps

At the CLS, multiple ω-2θ scans were taken of the (004) Bragg peaks for each of the samples from all three growths G0845, G0846, and G0859. These scans were completed in 12 different azimuthal positions by rotating the ϕ motor 30° between each successive scan. The scans were made at 12 different azimuths spaced by 30° in ϕ such that $phi = 0°$ corresponds to a plane of incidence parallel to the [0 1 1] direction.

These additional positions should reveal anisotropic features in the shape of the Bragg reflection. 3D RSMs were made for each individual scan. The method for creating these RSMs was derived from the similar method given in [48]. Each scan analyzed here was relatively short and coarse, as far as HRXRD can go, yet a single ω-2θ scan still had over 18 million data points. To handle the calculations involved in analyzing this large data set Wolfram Mathematica was used. The first attempts to generate these 3D RSMs failed, causing a computer, even with high specifications, to lock up entirely and crash. A great deal of work went into learning about Wolfram code optimization to be able produce any maps at all. Still, the calculation can take upwards of an hour depending on the specifications of the computer and data collected in the scan. Furthermore, although all points have successfully been mapped to reciprocal space, visualizing a data set with this many data points has been another challenge on it's own. To reduce the number of data points, for the sake of generating plots within a reasonable amount of time, methods of data aggregation were explored, such as clustering and binning but the only method of visualizing the data that has produced acceptable figures is by taking a subset of the data points. A detailed description of how the real space measured values were mapped to reciprocal space is given in Appendix B.

Two types of figures have been produced from these 3D RSMs: 3D scatter plots of the data points from an individual scan, and polar plot figures combining data taken from each of the azimuthal scans. The 3D scatter plots show a subset of the data points from an individual scan. The intensity measured for each data point determines both the colour and opacity of the datapoint, so that the highest intensity points are most prominent. The other figures are 2D polar plots of reciprocal space that flatten part of a spherical shell of finite thickness. These can be thought of as narrow range, high resolution pole figures.

Proper pole figures are typically used in XRD for the identification of texture and defects as in [48]. Here, these high resolution polar plots are used to observe the shape of the Bragg peak and identify the presence of defects. More details about how these plots were made is also given in Appendix B.

5.2.4 Rocking Curves

Another type of scan that can be done in XRD is the rocking curve scan, otherwise known as an ω scan or a transverse scan [46, pg. 148]. In this type of scan, the sample and detector, (ω and 2θ) are first aligned on the Bragg peak at the point with the brightest signal, and then the sample (ω) is rocked without moving the detector.

In the triple-axis configuration, the rocking curve reveals information about the amount of tilt within a crystal. In double-axis configuration, the rocking curve contains information about both the tilt and the strain within a crystal. With only a double-axis scan the effects of tilt and strain cannot be separated from each other.

From this type of scan we can learn about dislocations, mosaicity, curvature, misorientation, and inhomogeneity [49, pg. 214]. A rocking curve maps reciprocal space but only along a circular arc and so this information is also contained within an RSM [46, pg. 148]. The measuring instrument as well as layer thickness create an intrinsic width to the peak seen in a rocking curve, even if the crystal is perfect. Furthermore, different planes of the crystal have different intrinsic peak widths.

When a crystal has a high dislocation density, the rocking curve broadening can be understood by picturing the structure as made up of many tiny crystal blocks. The collection of these blocks, each at slightly different tilts, is the mosaic of the crystal. As the dislocation density increases, the angles by which the blocks are tilted increases and so does the mosaicity. As a result of these tilts, the diffraction peak will be broadened in ω as intensities will now be measured at points further away from the Bragg angle [49, pg. 214-215]. In reciprocal space this can be seen as a broadening of the reciprocal lattice point along a direction parallel to the (hkl) plane.

A triple-axis rocking curve directly measures the distribution of the orientations of the mosaic blocks. This allows for a quantitative measurement of the dislocation density in an epitaxial film from the width of an x-ray rocking curve scan [49, pg. 214-215].

The significance of the dislocation density in this work is not so much related to the plasmonic properties of the InSb layer. The dislocation density will not affect the plasmonic properties of the InSb material in any significant way but it may play an important role in

nanofabrication. Dislocations have been known to affect etch rates and create anisotropy. So, for the fabrication of high quality nanostructures with well defined shapes it may be important to reduce the number of dislocations present in the layer.

Using the in-house QC3 setup, a long series of rocking curve scans were measured for each of the InSb growths, G0845, G0846, and G0859. The scans were completed for the (004) planes for both the InSb and GaAs peaks. Additionally, these scans were completed in both triple-axis and double-axis configurations as well as at different azimuthal angles (ϕ).

To quantitatively measure the broadening of the Bragg peak due to the mosaicity, defects, curvature, etc. the measurements are fit to a Gaussian peak and the full width at half maximum (FWHM) of each peak is measured. A slight modification is applied to the theory that was developed in [50] to calculate the corresponding dislocation density for each FWHM.

From [50], the measured FWHM, $\beta_m(hkl)$, is predicted by Equation 5.4. Each term in this equation corresponds to the broadening due to various effects.

$$\beta_m^2(hkl) = \beta_0^2(hkl) + \beta_d^2(hkl) + \beta_\alpha^2(hkl) + \beta_\varepsilon^2(hkl) + \beta_L^2(hkl) + \beta_r^2(hkl) \tag{5.4}$$

As mentioned crystals have their own intrinsic width. $\beta_0(hkl)$ is the intrinsic width of the sample crystal while $\beta_d(hkl)$ is the intrinsic width of the diffractometer's beam conditioning crystal. $\beta_\alpha(hkl)$ is the broadening due to the rotations caused by disloctions, i.e. the tilts of the mosaic blocks. $\beta_\varepsilon(hkl)$ is the broadening due to strain around the dislocations [50]. Strain ε, simply put, is a stretch or compression of the crystal structure and thus the unit cell, causing a change in lattice constant. The strain is induced largely by stresses from lattice mismatch and dislocations form to reduce stress in the layer [49, pg. 203-211]. $\beta_L(hkl)$ is the rocking curve broadening due to the size of the crystal. Finally, $\beta_r(hkl)$ is the broadening due to the curvature of the sample [50].

There are reasons for why most of these effects can be neglected in a typical rocking curve scan. In general, broadening intrinsic to a crystal is less than 10" which is at or below the beam divergence of QC3 and will not show a significant effect [50]. This is well justified in the case of (004) Bragg peaks for the all of the crystals involved in the QC3 setup. $\beta_L(hkl)$ is inversely proportional to the thickness of the layer and can be neglected for layers thicker than $1\mu m$ [50]. Broadening due to curvature ($\beta_r(hkl)$) can be neglected because InSb is so highly mismatched with GaAs. In heteroepitaxial growths, curvature may be induced in both the substrate and epitaxial layer by the strain that forces lattice constants to match. However, in this case the mismatch is so great that the InSb layer

would be nearly fully relaxed after the first few monolayers are grown and thus there would be very little strain still to cause curvature.

This leaves us with only two factors to account for the broadening of the FWHM of the scan. The measured FWHM is thus compared to Equation 5.5.

$$\beta_m^2(hkl) = \beta_\alpha^2(hkl) + \beta_\varepsilon^2(hkl) \tag{5.5}$$

These two factors are then determined by Equations 5.6 and 5.7 [50].

$$\beta_\alpha^2(hkl) = 2\pi \ln 2b^2 D \tag{5.6}$$

$$\beta_\varepsilon^2(hkl) = 0.160 b^2 D |\ln\left(2 \times 10^{-7} cm\sqrt{D}\right)| \tan^2\theta \tag{5.7}$$

D is the threading dislocation density (TDD), input as $[cm^{-2}]$, θ is the Bragg angle and b is the length of the Burgers vector for these dislocations. For a zincblende crystal structure on the (001) surface, all dislocations have a Burgers vector of the form $\frac{1}{2}a\langle 110\rangle$ where a is the cubic lattice constant [51]. For InSb $b = 4.581\mathring{A}$. Equation 5.7 assumes that all dislocations are 60° type dislocations. The angle between the between the burgers vector and the line of the dislocation specifies what character the dislocation has. 60° type dislocations have a mixed character of an edge and a screw dislocation [51].

In [50], a method for measuring dislocation densities is used. This method requires, at minimum, the measurement of three rocking curves for three different Bragg angles. Instead, a simplified version of this method is proposed here to estimate TDD from only a single triple-axis rocking curve scan.

The key difference here is that the rocking curve scans are completed in the triple-axis configuration with only a very small acceptance angle. When in the triple-axis configuration the scan is not sensitive to the effects of strain, i.e. the only significant source of peak broadening is due to rotations or tilts of the mosaic blocks. In this case, Equation 5.5 can be simplified even further and the measured FWHM can be taken as exactly the broadening due to rotations, $\beta_m^2(hkl) = \beta_\alpha^2(hkl)$. An estimate for theTDD can then be made from Equation 5.6. To check that this method is self-consistent, measurements are also taken in double-axis configuration. $\beta_\varepsilon^2(hkl)$ is found by subtracting the measured broadening in the triple axis scan $\beta_\alpha^2(hkl)$ from the measured broadening in the double axis scan $\beta_m^2(hkl)$. Solving Equation 5.7 numerically for D another estimate for the TDD is made.

Furthermore, rocking curve scans at different azimuthal positions ϕ can reveal the anisotropy in the rocking curve width. For each of the wafers, G0845, G0846, and G0859, the same azimuthal positions were used as in the 3D RSMs. The rocking curves were made at 12 different azimuths spaced by 30° in ϕ such that $phi = 0°$ corresponds to a plane of incidence parallel to the [0 1 1] direction.

5.3 Results

The first result to discuss is the result from a single ω-2θ scan. As a standard procedure after completing a growth, all three of the growths, G0845, G0846, and G0859 were characterized by an ω-2θ scan. This scan covered the range all the way from the substrate Bragg peak, to the InSb Bragg peak, which is separated by $\sim 4.6°$. The RADS software made by Bruker was used to model and fit the scan. The results of these scans confirmed that InSb layers were grown on the GaAs with the specifications that were laid out in Section 3.2.1.

Before continuing any further, the issue with G0859 that was mentioned in Section 4.3 should be addressed. While collecting images for the 3D RSMs for G0859 an issue was discovered that was not identified in the QC3 ω-2θ scans. When the diffractometer was set in position for the 004 Bragg peaks the image captured by the detector revealed two independent peaks, shown in Figure 5.4. This same result was found for both the InSb layer as well as the GaAs substrate. This reveals that both layers show reflections from at least two different crystal grains within the material. The conclusion was that, even before any growth occurred, the substrate was of poor crystalline quality. Therefore, the results of the analysis of G0859 do not show, independently, the effects of the substrate misorientation. No more analysis of this sample will be presented here.

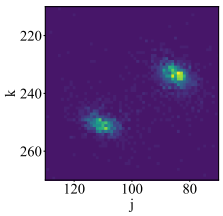


Figure 5.4: A cropped selection of the image captured by the x-ray detector for the InSb 004 Bragg peak in G0859. An anomalous second reflection appears separated from the first.

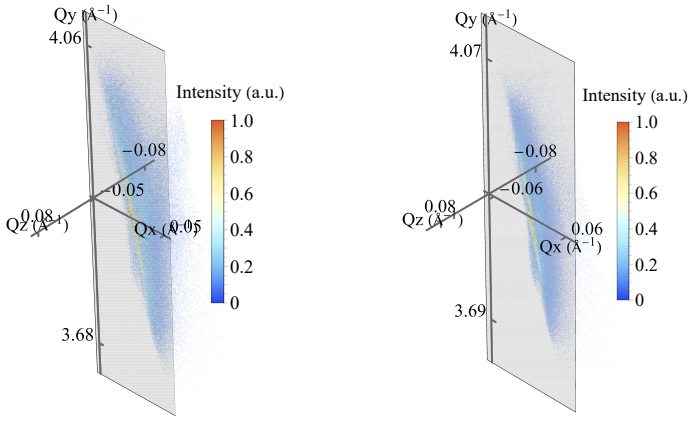


Figure 5.5: Two 3D scatter plots of the RSMs of the InSb 004 Bragg peaks in (a) sample G0845 and (b) sample G0846. The data is sliced by the xy plane to show the inner structure.

5.3.1 3D Reciprocal Space Maps

Each ω-2θ scan performed at the CLS was mapped to a 3D volume in reciprocal space through the method described above and in Appendix B. To visualize the 3D RSMs in the most direct way, a scatter plot for each one was made on a set of 3D axes. Figure 5.5 shows the 3D RSMs for G0845 and G0846 at $\phi = 90°$ around the Bragg peak for the (004) planes. To reveal the inner structure of the Bragg peak, a plane, normal to the z axis, cuts through the highest intensity region. Qualitatively each of the plots look very similar, even at different orientations of ϕ. The 3D plots are useful for verifying that the 3D mapping process proceeded correctly and for observing the larger shape of the Bragg peak.

For example, in each 3D plot the shape of the Bragg peak was seen as a streak broadened along the Qy axis, which would correspond to the [001] crystallographic direction, normal to the (004) planes. From Equation 5.1 the spacing of the (004) planes should be $d_{004} = 1.62\mathring{A}$. From the 3D RSMs for G0845 and G0846 the maximum intensity point occurs at $(Qx, Qy, Qz) = (0.00, 3.87, 0.00)$. The magnitude of the reciprocal space vector is $|\mathbf{Q}_{hkl}| = 2\pi/d_{hkl}$. The corresponding measurements for d_{004}, based on the highest intensity point from G0845 and G0846 both match the theoretical predicted point for the (004) reflection with less than 0.3% difference.

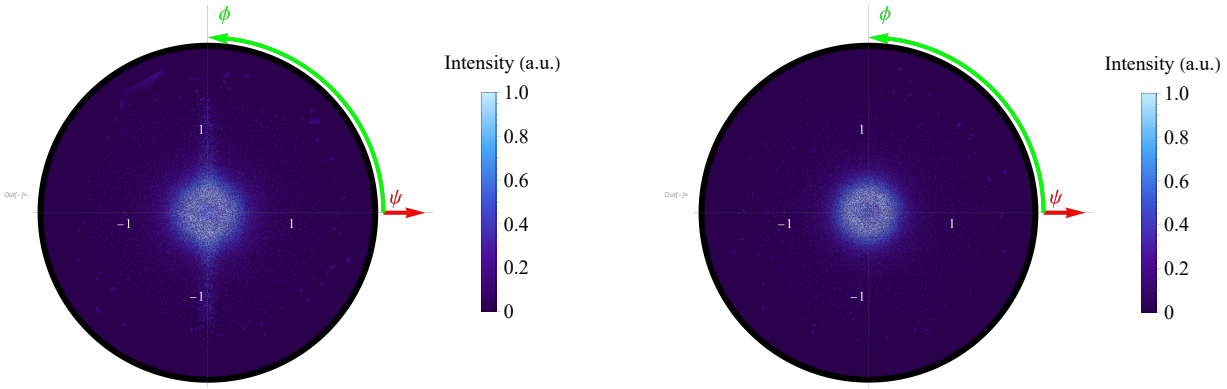
However, these Bragg peaks are also broadened along all three dimensions, but mostly along Qy. This shows there are reflections from parts of the crystal where d is not exactly

corresponding to d_{004}. This is an expected result. Real crystals have dislocations and internal strains that determine the distribution of d spacing as well as other effects that spread the intensity of the Bragg peak. The result that was not expected, however, is that the streaks are not parallel to the [001] direction. The Bragg peak actually appears slanted in the cut plane along the Qx and Qy directions. Broadening of the Bragg peak along the Qx axis might correspond to tilts in the mosaic of the crystal. But the distribution of tilts should be random and thus the broadening should be symmetric across a line parallel to the Qy axis. Furthermore this same tilt appears in the 3D RSM for the GaAs substrate 004 peak as well as every other scan. But we know that this substrate should have no tilt, or very little. The conclusion then is that this tilt is instrumental in nature. The exact cause of this is still under investigation.

It is difficult to present a static figure of a 3D plot in a way that appropriately represents all of the information about reciprocal space that is contained within the RSM. For that reason it would be beneficial to show the RSM as a projection along a surface in a 2D plot. Pole figures are usually used to show a RSM for a large area covering multiple Bragg peaks. They are not used for the smaller areas covered with high-resolution scans in HRXRD. In this research, the concept of pole figures is applied in a novel way to try to examine the tilt and strains caused by dislocations in the crystal. A pole figure is a polar plot of a region of reciprocal space. The surface of a hemisphere in reciprocal space would be mapped to a 2D circle in the polar plot. In the spherical coordinate system this hemisphere is described by a radial distance r, azimuthal angle, ϕ, and polar angle, ψ. The pole figure is made by plotting in the 2D polar coordinate system $(r, \theta) = (\psi, \phi)$. The angle ψ is plotted as the radius. In Figure 5.6, the angle ϕ is plotted as the angular position with the green arrow indicating the $\phi = 0^\circ$ position and the direction ϕ is measured in. This effectively gives a 2D projection of the hemisphere onto a circle.

At the $\phi = 0^\circ$ position the incident x-ray beam is travelling in the [011] crystallographic direction.The pseudo pole figures given in this dissertation are similar to regular pole figures except that measurements are taken only for very small polar angles $< 2^\circ$ and the surface has some finite thickness. This means that the circular projections seen in Figure 5.6 are projections of a spherical cap in reciprocal space rather than projections of the whole hemisphere. Specific details regarding the creation of these plots is given in Appendix B.

The pole figures for the InSb 004 Bragg peaks of G0845 and G0846 are given in Figure 5.6a and Figure 5.6b, respectively. The two plots are distinguishable by the faint vertical line that passes through the central spot in the G0845 plot. This line appears only in the $\phi = 90^\circ$ and $\phi = 270^\circ$ scans. In these azimuthal positions the x-ray beam is parallel to the $[01\bar{1}]$ direction. Also, note that the radius of the central spot is slightly larger for G0845. These two effects are thought to be the result of a higher density of dislocations in the



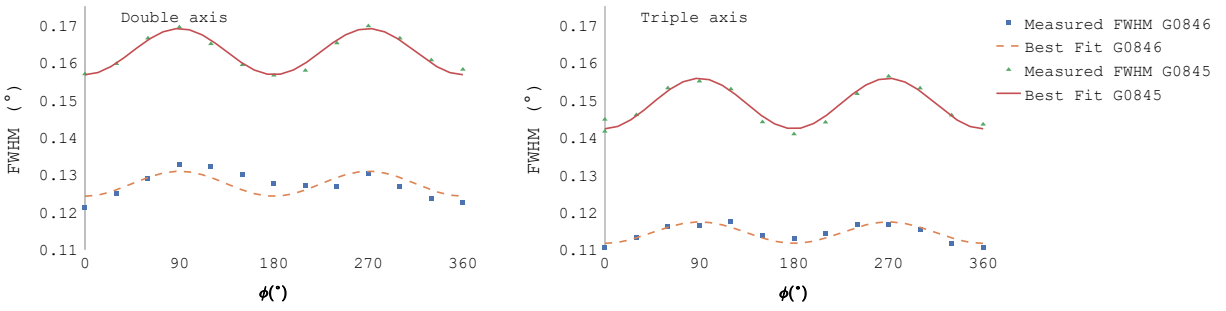
(a) (b)

Figure 5.6: Two polar plots of the 3D RSMs of the InSb 004 Bragg peaks in (a) sample G0845 and (b) sample G0846. ψ is measured as the radius and ϕ is the angle measured counter-clockwise from the positive x axis.

sample that was grown with 0° misorientation, G0845. With a higher dislocation density the Bragg peak is expected to be broader. It will be shown later that these results are in agreement with other measurements made to measure the dislocation densities.

There is a ring shaped central bright spot in both images that appears slightly dim in the very center of the image. This is because these plots were made by selecting, in spherical coordinates, only the data points that have a radius (Q_r) close to the nominal radius for the InSb 004 planes (Q_{004}). This shows that, at the nominal radius, the brightest spot in the Bragg peaks for both samples does not correspond to a tilt of zero. This is a result of the instrumental effect that made the Bragg peaks appear tilted in Figures 5.5.

The significance of creating these pole figure plots is emphasized by the fact that we can distinguish two types of effects that both influence the broadening of the Bragg peak. Furthermore, I postulate that creating figures like these can be valuable for observing the effects of strain separately from the tilt. First of all, the effect of strain would be to change the d_{hkl} either by compressing or stretching the unit cell of the lattice. These plots however are created by selecting data points that only correspond to $Q_{hkl} \pm \delta$ where δ is the tolerance in deviation from Q_{hkl}. This means that, with an appropriately low tolerance, the data shown in these plots is completely independent of any effects of strain on the shape and size of the Bragg peak. Two or more plots made with different central values selected for Q_{hkl} could be compared to study the effect of strain alone. Conversely, studying a single plot alone would reveal the tilts present in the sample for a single amount of strain. The



(a) (b)

Figure 5.7: Two plots of the FWHMs measured versus the azimuthal angle ϕ for (a) double axis measurements and (b) triple axis measurements. Each plot shows the measurements for both samples G0845 and G0846.

intensity corresponding to different amounts of tilt can be read directly from the plot as the radius ψ.

5.3.2 Rocking Curves

The FWHM for each rocking curve scan was measured by fitting the raw intensity vs. ω data points to a Gaussian function. For specific details on how this was done, Appendix D has a brief explanation. Figure 5.7 shows the FWHMs values that were measured for each of the rocking curve scans made at different azimuthal angles. Figure 5.7a shows the FWHMs for the double axis scan configuration and Figure 5.7b shows the results for the triple axis configuration. Across all of the azimuthal positions from 0° to 360° it can be seen that the FWHMs fluctuate.

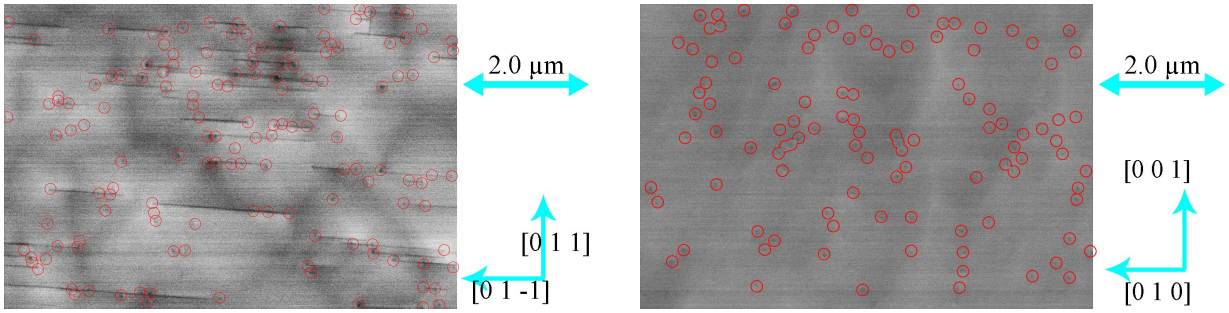
Crystallographic directions separated by 180° in ϕ should show equivalent measurements in XRD. This is why two peaks are seen for each plot in Figure 5.7. The measured FWHM from each triple axis scan was used to estimate the TDD using Equation 5.6. Since the FWHM changes with ϕ, the TDD estimate follows the same trend. Table 5.1 shows the measured FWHMs and TDDs averaged across equivalent orientations separated by 180° in ϕ . Table 5.2 shows the estimates from the double axis analysis using Equation 5.7. From the figures and the tables it can be seen that the G0846 growth with a 2° misorientation always has a lower FWHM and therefore a lower TDD and the variations with ϕ are smaller too. This all points to the conclusion that performing the growth on a substrate with misorientation improved the quality of the grown crystal.

Sample	G0845	G0845	G0846	G0846
$\phi(^\circ)$	FWHM ($^\circ$)	TDD (cm^{-2})	FWHM ($^\circ$)	TDD (cm^{-2})
0	0.141	6.647×10^8	0.112	4.170×10^8
30	0.145	7.008×10^8	0.114	4.312×10^8
60	0.152	7.740×10^8	0.117	4.523×10^8
90	0.156	8.072×10^8	0.117	4.534×10^8
120	0.153	7.801×10^8	0.116	4.516×10^8
150	0.145	7.007×10^8	0.113	4.233×10^8

Table 5.1: A table showing the measured FWHMs and TDDs in the triple axis configuration for both samples G0845 and G0846. The values are averaged between the measurements for azimuthal angles that are separated by 180° in ϕ.

Sample	G0845	G0845	G0846	G0846
$\phi(^\circ)$	FWHM ($^\circ$)	TDD (cm^{-2})	FWHM ($^\circ$)	TDD (cm^{-2})
0	0.157	3.008×10^9	0.125	1.984×10^9
30	0.166	2.932×10^9	0.126	1.931×10^9
60	0.170	3.135×10^9	0.128	1.822×10^9
90	0.166	2.762×10^9	0.131	2.474×10^9
120	0.160	3.163×10^9	0.130	2.148×10^9
150	0.157	3.141×10^9	0.127	2.255×10^9

Table 5.2: A table showing the measured FWHMs and TDDs in the double axis configuration for both samples G0845 and G0846. The values are averaged between the measurements for azimuthal angles that are separated by 180° in ϕ.



(a) (b)

Figure 5.8: SEM ECCI images for (a) sample G0845 and (b) sample G0846. Small dots of increased contrast caused by TDDs are circled in red. The corresponding crystallographic directions are indicated by the blue arrows on the right.

Comparing the double axis TDD measurements to the triple axis TDD measurements along the same directions reveals there is some inconsistency in the methods of estimating TDD. It appears as though there is some additional broadening in the double axis scans that was unexpected. This results in a TDD measurement that is more than four times greater than the values predicted in the triple axis scans. Equation 5.4 included the effects of many different factors in the broadening of the Bragg peak. It was well justified to exclude all of the effects except for broadening due to strain or tilts. One possible scenario is that the triple axis TDD estimates are correct and therefore the reason the double axis measurements are incorrect is because there is some additional strain in the crystal that was not accounted for in this model.

To test this hypothesis another method for measuring the TDD independently from XRD is used. This method is electron channeling contrast imaging (ECCI). ECCI is used to create images that represent the surface and near surface region of the sample. The contrast in these images is produced by the tilt and strain induced around dislocations in the crystal structure. A brief description of how ECCI works is given in Appendix C.

Figure 5.8 shows some of the ECCI results for G0845 (Figure 5.8a) and G0846 (Figure 5.8b). Notice that the horizontal and vertical directions correspond to different crystallographic directions in these images. This is because the 2° misorientation of the G0846 surface meant that, to record a good contrast in ECCI, the sample had to be oriented and tilted correctly to account for the tilt of the crystallographic planes. Meanwhile, the G0845 sample which was on-orientation did not need to be tilted.

In the images, red circles are drawn around any small dot where an increase or decrease

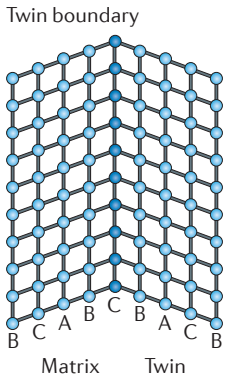


Figure 5.9: A schematic representation of the atomic structure at a twin boundary. Figure sourced from [10]. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer, Nature Reviews Materials. Stabilizing nanostructures in metals using grain and twin boundary architectures, Lu, K. © 2016.

in contrast is seen. These same images are shown without the circles in Appendix C. The dots are found through visual inspection and so the results are somewhat subjective. Some dots appear only with very faint contrast and it can be difficult to determine when to count another one. Making identifying the dots even more difficult, in the results for G0845 there is an additional effect. In addition to the dots, there is also a high density of lines parallel to the $[01\bar{1}]$ direction and many of the dots coincide with these lines. Notice how there are no lines in the image for G0846.

These lines are thought to be twin boundaries formed in the crystal. These same lines can be seen in the AFM images of Section 4.3. A particularly clear image of one is shown in Figure 4.6a splitting the hillock in two. On either side of a twin boundary the crystal structure is the same however in a different orientation. Along the twin boundary line the two crystal grains may appear as if reflected across the boundary as in Figure 5.9.

The dots however correspond to threading dislocations (TDs). Each dot indicates another TD and by counting the number of dots in an area another estimate for the TDD can be made, separately from the XRD measurements. The difficulty of identifying relatively low contrast dots means that there will be a high uncertainty to these measurements and likely some personal bias. The strategy chosen for counting dots was to over count rather than under count. So, practically any small deviation in contrast from the white noise background would be counted as a TD. Therefore, the TDD measurements obtained from these images can be seen as an upper limit.

For G0845 a TDD of $6.565 \times 10^8 cm^{-2}$ was measured by counting 145 dots in a $2.209 \times 10^{-7} cm^2$ area. The same sized area for G0846 only showed 107 dots giving a TDD of

$4.844 \times 10^8 cm^{-2}$. These estimates much more closely match the results from the triple axis analysis rather than the double axis results.

The XRD TDD estimate for G0845 has only at minimum a 1.25% difference from the ECCI measurement when $\phi = 0°$ and a maximum of 23.0% difference at $\phi = 90°$. The fact that the TDD estimate matches much more closely at $\phi = 0°$ makes sense. $\phi = 0°$ corresponds to an incident x-ray beam along the [011] direction. This is perpendicular to the direction of the twin boundaries identified in Figure 5.8a. It would make sense that, if these grain boundaries are significantly contributing to the broadening of the Bragg peak, that they would only contribute significantly to broadening in the $[01\bar{1}]$ direction parallel to its widest dimension.

Photons travelling through this grain boundary will have more chances to be scattered by it if they travel parallel to the longer dimension in the $[01\bar{1}]$ rather than when they traverse this boundary across the much shorter dimension in the [011] direction. Therefore, when $\phi = 0°$ the FWHM measurement in triple axis XRD is solely from the TDD and not influenced by the twin boundaries. With this new information, returning to the pole figures, Figures 5.6a and 5.6b, it seems certain that the vertical line present in the pole figure for G0845 must be a direct result of scattering from the twin boundaries.

The twin boundaries definitely should contribute to the overall anisotropic broadening of the Bragg peak however twinning is not present in the G0846 sample and yet the FWHM does still depend on ϕ, although the variations are significantly smaller. The TDD estimate from ECCI for G0846 has at most 13.9% difference to the triple axis estimate at $\phi = 0°$ and at minimum 6.40% difference when $\phi = 90°$.

Returning to the hypothesis, the low percent difference between ECCI and triple axis XRD TDD estimates does indicate that Equation 5.6 successfully predicts the the influence of the tilt of mosaic blocks on the Bragg peak width except that it does not explain the weak dependence on ϕ. Also, Equation 5.7 is not able to predict all of the influence that strain has on the peak width, at least in these samples. There is still some unknown source of additional strain in both of these samples. The answer to this question likely lies somewhere in the 3D RSM it will just take some searching to find it.

This dependence on ϕ may be related to an unequal distribution of dislocations on different slip systems within the crystal structure as in [52–54]. In fact, a very similar variation of the FWHM was seen in [52] which showed minimum and maximum values in the same azimuthal positions that they were found in here.

For a related III-V semiconductor structure, GaInAs/AlInAs buffers which are reported in [54], this anisotropy is explained by the mechanisms of strain relaxation that occurs during crystal growth. First some things must be understood about how dislocations can

move during crystal growth. In zincblende crystals, active dislocations can slip on glide planes that belong to the {111} family of planes [51]. A slip system is defined by a glide plane and a Burgers vector. The Burgers vector is contained within the glide plane and dislocations can glide in the direction of this vector. The 60° dislocations modelled in Section 5.2.4 form $\frac{1}{2}a\langle 110\rangle\{111\}$ slip systems. Each of the planes in the {111} family is exactly parallel to one of the other planes. For example, $(\bar{1}\bar{1}1)$ is equivalent to $(11\bar{1})$. Then, slip systems can be made with only four different planes. With a (001) surface orientation, four {111} planes intersect this surface [51]. The slip systems can have Burgers vectors in three different $\langle 110\rangle$ directions within these planes. The (111) plane is shown in Figure 5.2c and each of the three sides of the triangle point in these directions. Considering all four of the glide planes and noting which directions are equivalent, the final number is eight slip systems that can intersect the (001) plane. These are the slip systems whose dislocations can cause tilts and strains in the (001) planes.

To help visualize this, Figure 5.10 shows a pyramid with four triangular faces. The triangular faces of the pyramid are the four {111} planes and the pyramid is intersecting the (001) plane. The four edges of the pyramid that intersect the (001) plane are along the directions of the $\langle 110\rangle$ Burgers vectors. Each Burgers vector is parallel to two different planes and so there are eight possible combinations of $\langle 110\rangle$ Burgers vectors with {111} planes when considering the intersection with the (001) oriented surface of a crystal.

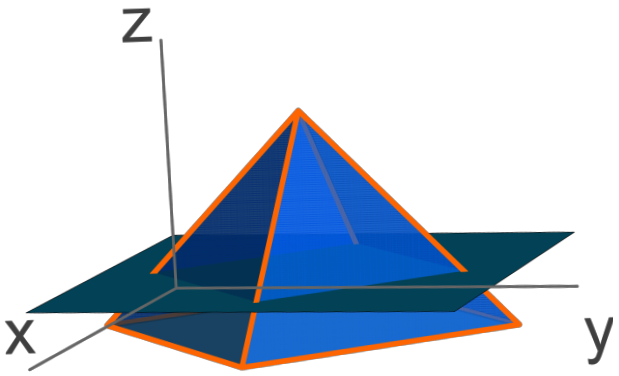


Figure 5.10: A graphical representation of the active slip systems that intersect the (001) plane in a zincblende crystal structure. The triangular faces of the pyramid represent the {111} glide planes and the edges of these faces that intersect the (001) plane are along the $\langle 110\rangle$ Burgers vectors.

The $[1\bar{1}0]$ direction (this is in the (001) plane) is parallel to two of these planes while [110] is parallel to the other two planes. The dislocations that are in the planes parallel to

the $[1\bar{1}0]$ direction are called α type dislocations. Likewise, the dislocations in the planes parallel to the [110] direction are called β type dislocations [54]. An unequal amount of α and β dislocations would explain why the TDD measurement is different along the $[1\bar{1}0]$ and [110] directions.

These dislocations are formed because of the large lattice mismatch between the InSb layer and GaAs substrate. The InSb will be subjected to a high amount of strain. After a certain critical thickness is reached it becomes energetically favourable for dislocations to form which allows the strain to relax. Initially, there will be strain in both the $[1\bar{1}0]$ and [110] directions. The strain along either direction can be relaxed when half loops are nucleated [54]. A half loop dislocation line is one that connects two edge dislocations on the same plane by threading into the crystal and then looping back around to the starting plane. As crystal growth continues the half loop would glide and expand creating a section of misfit in the crystal [54]. If the glide of this threading component is impeded then as the crystal becomes thicker, the dislocation line may continue to thread throughout the entire crystal, and this is called a threading dislocation.

The α and β type dislocations may have unequal distributions because they have different nucleation energies and also different glide velocities [54]. This inequality is due to the differences in atomic structure of indium and antimony atoms [54]. A lower nucleation density and a higher glide velocity will lead to an increase in TDD.

Another factor that influences the ratio of α dislocations to β dislocations is the substrate misorientation. Changing the orientation of the surface of the crystal with respect to the {111} glide planes will also change the angle between the planes and the direction of stress caused by edge dislocations [55].

This will either increase or decrease the resolved shear stress (RSS) for each slip system. Slip systems with higher RSS will have a lower nucleation energy and higher glide velocities [55]. However, because the offcut of G0846 is towards the [010] direction, which has equal components along the $[1\bar{1}0]$ and [110] directions, the change in RSS is equal for both α and β slip systems and there is no net effect on the ratio of TDD in the two systems. Therefore, from this misorientation we would not expect to see a change in the amplitude of the FWHM oscillations.

Simply by comparing the top and bottom lines in Figure 5.7b, we do see significantly smaller variations in FWHM for G0846. The same can be said about the double axis scans in Figure 5.7a. However, there seems to be some erroneous data in the double axis measurements for G0846 which reduces confidence in this result. Regardless, this means that the misorientation reduced the β TDD more than it did the α TDD suggesting that the offcut was not exactly along the [0 1 0] direction.

In both Table 5.1 and Table 5.2 the maximum TDD measurements are seen for orientations of $\phi = 90°$. This is when the x-ray beam is parallel to the $[01\bar{1}]$ direction and so it is parallel to the α glide planes. The α TDs that penetrate into these planes will have dislocation lines perpendicular to the direction of travelling x-rays. Just as with the twin boundary, little to no effect will be seen in the broadening of the Bragg peak from these dislocations. On the other hand the β dislocations will have components along the direction of travel and so they will significantly broaden the Bragg peak. This is how the FWHM depends on the azimuthal angle. These results indicate that the β TDD is larger than the α TDD.

Chapter 6

Nanofabrication

In this chapter we embark on a quest to fabricate InSb plasmonic nanostructures. The goal is to establish a sequence of nanofabrication procedures that will take epitaxial layers of InSb, grown by MBE with the methods described in Chapter 3, and transform them into a designed nanostructure for use in THz plasmonic applications. This chapter summarizes the early stage of this research. The nanostructure modelled in Chapter 2 is used as the blueprint for the design of our nanostructure.

6.1 Theory

Nanofabrication encompasses a multitude of different techniques, each of which involves processes governed by their own unique chemical and physical laws. The whole theory behind all of these techniques can not be described within this dissertation. Instead, the focus of this section will be only on the techniques that were made use of and only the main principles of the theory will be described. Regardless, a thorough understanding of the complexities is not needed to make use of the various techniques. In nanofabrication, firsthand experience is key to producing positive results.

The nanostructures are fabricated through ultraviolet (UV) photolithography and inductively coupled plasma (ICP) etching. The principles of each process step are described in the below sections.

6.1.1 Sample Cleaning

The first step in the nanofabrication process is to clean the sample. This is done to remove any dust or dirt that may be present on the surface of the sample. The samples, by household standards, are never dirty and appear clean to the naked eye. However, on the nanoscale, dust particles too small to be seen by the naked eye can significantly affect processes on the surface of the sample. Samples that are removed from the MBE system are always handled and stored in clean environments. Nevertheless, all samples will inevitably collect some dirt over time and this is why it is best to clean them right before any nanofabrication processing.

A variety of cleaning processes have been developed for different use cases with varying strengths and ability to clean. The most common type of cleaning process combines, in separate steps, the use of acetone, methanol, and isopropyl alcohol (IPA) [34, pg. 139]. This process is referred to as acetone-methanol-isopropyl alcohol (AMI) cleaning and it is most effective at removing organic contaminations. The first step is to either spray or soak the sample in acetone which softens and dissolves many organic residues. The second step is to clean the substrate with methanol before the acetone dries because acetone leaves behind an organic film. The third step uses IPA to remove the methanol. Different laboratories may have different procedures for sample cleaning. In the work presented here, sample cleaning is done using acetone and IPA only and this is sufficient in most cases.

The final step in this cleaning process is to blow dry the sample. It is generally bad to let any liquid dry by evaporation because, no matter how clean the liquid is, there is always some contaminants present. The evaporating liquid would then leave these contaminations behind [34, pg. 139]. So, the sample is blown dry using, clean and dry nitrogen gas. Nitrogen gas is also used after any process where dust may be left behind. It is an inert Noble gas and so it is very useful in cleaning samples.

6.1.2 Spin Coating

Spin coating is the method used to create a thin film of photoresist on the cleaned sample surface. Photoresists are UV sensitive organic polymers that are permanently altered when exposed to UV light. To use photoresists, the polymers are dissolved into liquid solvents which are then dispensed over top of the sample [34, pg. 142]. The sample itself, is held onto a rotating stage by vacuum suction.

Initially the sample is spun slowly with a lower acceleration during what is called the spread stage. This initial spin allows the fluid to spread across the entire sample surface in

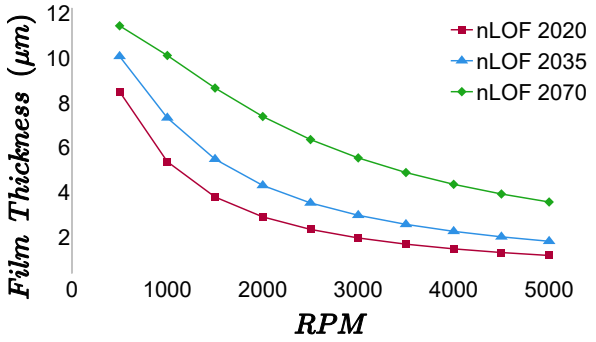


Figure 6.1: A plot of the spin curves for the three nLof 2000 series photoresists. Final film thickness values correspond to a sample spun for 60 seconds at the indicated rotational velocity.

a path that spirals outward from the center. The angular velocity and acceleration need to be chosen carefully so that the entire surface is covered without all of the fluid being thrown off. Increasing the angular acceleration increases the tangential force on the fluid. When the angular acceleration is too low some of the sample surface may not get covered but when it is too high the fluid is separated from the surface [34, pg. 143].

The next stage is the thinning stage that determines what the final thickness of the film will be. The angular velocity reaches a target value which is selected based on spin curves provided in the material datasheet. Approximate spin curves for three different photoresists are show in Figure 6.1. As the sample spins, the film thins out over time reaching an asymptotic value. The final angular velocity of the sample is what determines the film thickness.

Equation 6.1 can be used to approximate the final film thickness for a photoresist [34, pg. 146]. h is the final thickness of the film, ρ is the density of the fluid, η is the viscosity of the fluid, t is time, ω is the angular frequency, and h_0 is the initial film thickness after the spread stage. Equation 6.1 is used to plot three spin curves for the AZ nLOF 2000 series photoresists. The values for ρ and η are taken from the materials safety datasheets and t is chosen to be 60 seconds while the initial thickness is $12\mu m$. It only takes about 30 seconds for the film thickness to reach a steady-state [34, pg. 147].

$$h(\omega) = \frac{h_0}{\sqrt{1 + \frac{4\rho\omega^2 h_0^2 t}{3\eta}}} \tag{6.1}$$

With higher spin speeds and lower fluid viscosities a thinner film will be produced. The AZ nLOF 2000 series photoresists are all identical except with different concentrations of solvents, which influences the viscosity of the fluid. This allows a wide range of film thicknesses to be selected.

During the spin, the airflow over the sample will affect how the fluid evaporates. To get uniform evaporation it is important that a laminar air flow is maintained over the sample. Once the thinning stage is complete the sample is moved to a hotplate. The temperature of the sample is raised to above 100° and so the photoresist is baked. The purpose of this bake is to remove as much of the solvents from the film as possible. This is called a soft bake [34, pg. 151]. The hotplate increases the evaporation rate of the solvents in the fluid. After about 90 seconds, depending on film thickness, the photoresist will be stabilized and ready for UV exposure.

6.1.3 UV Photolithography

Photolithography is a type of lithography that exposes photosensitive materials to light creating patterns on the surface of the sample that will later define the shape of the nanostructure. Photoresists, such as the AZ nLOF 2000 series are designed to be particularly sensitive to UV light frequencies. The light comes from either a laser, exposing only one small spot at a time, or a wide light beam is passed through a photomask [34, pg. 153]. The photomask works as a stencil, exposing the photoresist everywhere except for where the photomask blocks the light.

Photoresists have three main components: solvent, resin, and photoactive compound (PAC). The purpose of the solvent is for dispensing the photoresist and spreading it into a thin layer. After the soft bake, the solvents are gone and only the resin and PAC remain [34, pg. 163-164]. The bulk of the photoresist layer is made of the resin. After the pattern has been exposed, the sample and photoresist would be submerged into a photoresist developer. The photoresist developer is a weak alkaline solvent that can dissolve the resin material. In a typical photoresist the PAC is bound to the resin which creates a material that is less soluable than the resin would be on it's own. When the PAC is exposed to UV light it releases from the resin and reacts with water to create a photoacid. The photoacid now increases the soluability of the resin and the photoresist devloper can dissolve it away. This means that areas that are exposed to the UV light are more soluable and thus are removed from the sample by the photoresist developer. This is called a positive-tone photoresist [34, pg. 164]. This is how patterns can be defined in the photoresist layer.

Alternatively, there are also negative-tone photoresists that become less soluable when

exposed to UV light. The AZ nLOF 2000 series photoresists are negative-tone type. The chemistry regarding negative-tone photoresists is very different from positive-tone photoresists [34, pg. 181]. In negative photoresists, the solubility of the resin is reduced when UV exposure causes a cross-linking effect extending the length of the polymer chains in the photoresist. Negative photoresists have a few advantages over positive photoresists. They have better adhesion to the surface sample and have a higher structural integrity and lower erosion rates with plasma etching [34, pg. 163-164].

Additionally, it may be beneficial to switch to a negative-tone photoresist depending on the dimensions of the pattern that needs to be exposed. Switching tones of photoresist means the exposed and unexposed areas of the pattern are inverted. If the areas that would be exposed in a positive resist have smaller dimensions than the areas that wouldn't be exposed, switching to a negative tone photoresist means the same pattern can be made by exposing the features with longer dimensions. Smaller dimensions are typically harder to expose with good resolution.

After exposing the photoresist to UV light, the resin begins to absorb light in a wider range of UV frequencies resulting in a poor selectivity. To overcome this issue, photoresists have been developed with chemical amplification processes. This means that a lower dose exposure can be used to affect the solubility of more resin molecules. The chemical processes are thermally activated and so these types of photoresists require a post-exposure bake (PEB). This also means that the solubility of exposed areas will depend somewhat on the PEB temperature. The AZ nLOF 2000 photoresists are examples of chemically amplified photoresists.

6.1.4 Inductively Coupled Plasma Etching

To transfer the pattern from the photoresist into the InSb layer, the sample will need to be etched. In an etch process, the photoresist layer acts as a chemical and physical barrier protecting the InSb layer underneath from being etched. In a developed photoresist, the gaps in the pattern will leave gaps in the photoresist layer so that parts of the underlying InSb layer are uncovered. The uncovered parts of the layer is where etching will occur. This is how the pattern of the photoresist is transferred to the InSb layer.

Inductively coupled plasma (ICP) etching removes material from the sample by exposing it to reactive gasses [34, pg. 227]. Reactions take place inside of a vacuum chamber where gas flow is controlled. A plasma is generated inside of the chamber by ionizing the gasses inside. To ionize the gasses, a time-varying current is passed through a coil wrapped around the outside of the chamber. The corresponding magnetic field generated inside of

the chamber also varies with time. The gasses inside are then ionized by energy supplied through this magnetic field [34, pg. 227].

The ionized gasses (plasma) will be comprised of charged ions, neutral excited free radicals (for example, O_3), and the remaining unexcited molecules. Etching proceeds through two mechanisms: (1) chemical reactions occur between free radicals and the surface of the sample to create volatile products which easily desorb from the surface and are removed from the system, (2) ions bombard the surface causing physical sputtering in high energy collisions which eject atoms from the surface [34, pg. 224]. Both etching mechanisms work to remove material from the sample on their own but when they are both combined the etch is significantly faster than the sum of both individually. The high energy collisions of the ions with the surface in mechanism (2) will not only cause sputtering but they will also lower the activation energy of the chemical reactions that take place in mechanism (1) [34, pg. 225].

Furthermore, the ions are accelerated toward the material along a straight path and so they mostly speed up reactions along horizontal surfaces where most collisions occur. This is why plasma etching is the preferred method in many nanofabrication procedures [34, pg. 222]. It creates better vertical profiles and straighter edges in the patterns than wet chemical etching would.

Generating the plasma using a magnetic field rather than an electric field allows the motion of the ions in the plasma to be controlled completely independently from the amount of ionization. This can be done by creating a second electric field by applying a negative bias to the sample [34, pg. 227]. This is referred to as the DC bias. In this case, the ions will be accelerated towards the substrate causing ion bombardment. Because of the many high energy collisions, the sample must also be cooled by helium to maintain a constant temperature, otherwise etching may be inconsistent [34, pg. 223]. This is important if you want to have control over how much sputtering occurs but still be able to increase gas pressure for a faster etch rate.

There are many choices for etchant gasses and the best selection will depend on the chemistry between the etched sample and the gasses. For etching InSb the gasses CH_4 and H_2 are appropriate choices. Once etching is complete, the photoresist is dissolved using a photoresist remover solvent and the remaining material on the substrate will be your patterned nanostructure.

6.2 Methods

6.2.1 Cleanroom

The entire nanofabrication process documented here was kept in a cleanroom environment. Cleanrooms are special laboratories which are designed to have constant air filtration to reduce the number of particles in the air. This reduces the chance of sample contamination from stray dust particles. A typical office space is around class 100 000 which means there are 100 000 particles per ft^3 that are larger in dimension than $0.3\mu m$. Humans are the greatest source of contamination in cleanroom environments. People constantly shed skin and drop hairs. Our skin has natural oils that are left behind on everything we touch. Other sources of contamination are many commonly used items such as paper, pencils and even clothing. For these reasons, entry into a cleanroom environment is restricted only to those who understand and follow the proper cleanroom gowning and behaviour protocols. Every time that someone enters a cleanroom they have gown themselves from head to toe in specially made outfits to reduce the number of particles that they are outputting into the room. During the nanofabrication process, the samples are kept within a class 100 cleanroom environment.

6.2.2 Procedure

The wafers of the MBE growths in Chapter 3 were 3" in diameter and so each one had enough material to be used many times in nanofabrication. Whenever a new sample was needed for nanofabrication a small rectangular section would be cleaved from the wafer. The samples were approximately 1.5 cm by 1.3 cm.

The procedure that was developed for making nanostructures in InSb can be broken down into five parts that are followed in order.

1. Sample cleaning
2. Photoresist spincoating
3. Photoresist exposure
4. Photoresist development
5. ICP etching

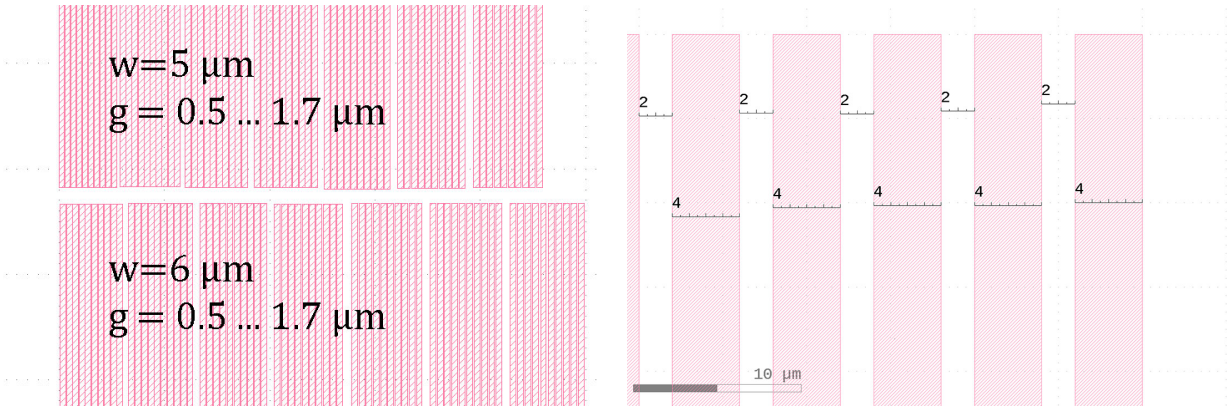
Sample cleaning was straightforward. New samples were always cleaned before anything else. They were cleaned using the techniques described above for AMI cleaning except that the methanol step was skipped. Before proceeding, the cleanliness of each sample was assessed with an optical microscope to ensure no residues were left behind.

Next, the samples were spincoated with AZ nLOF 2035 photoresist. The target film thickness was just over $2\mu m$ because this would be thick enough to survive the ICP etching procedure at the end. For the spread stage an initial acceleration of 1000 rotations per minute (RPM)/s was chosen and the speed was held at 1000 RPM for 1 second. Looking at Figure 6.1, a final speed of 4000 RPM after 60 seconds of spinning should achieve a film thickness of a little more than $2\mu m$. So, for the thinning stage the rotation speed was ramped up from 1000 RPM to 4000 RPM at an acceleration of 4000 RPM/s. Combinations of spread speeds and accelerations were tested and these were the values with the best results. Accelerations of 500 RPM/s left too much of the sample uncovered while accelerations over 1000 RPM/s showed inhomogeneous shapes, likely due to fluid slipping and being thrown off of the surface. After a sample was spin coated it was then soft baked for 90 seconds at 115 $^{\circ}C$.

In the third step, patterns were exposed into the photoresist layers using the Heidelberg MLA150 Maskless Aligner. This tool uses a 375 nm UV laser and precision position control to faithfully draw out patterns in photoresists with no need for a photomask. The patterns can be drawn arbitrarily using a software called KLayout. Some of the patterns that were used are shown in Figure 6.2. The pattern in Figure 6.2b is the desired pattern that we would like to etch into the InSb. The dimensions of this pattern match those in Table 2.2 where $w = 4\mu m$ is the width of the lines and $g = 2\mu m$ is the width of the gaps. This pattern can be seen as a view of the InSb lines from above.

There are number of variables to consider in how the photoresist is exposed. First there is the dose of a pattern, which is a measure of the energy density of light that is applied to each spot on the photoresist. Second, there is the defocus variable. This variable sets the height of the laser focus point relative to the sample surface. Finally, the dimensions drawn in the pattern file do not always come out to be what they should be. To get the correct dimension to develop in the photoresist, the correct value for each of these variables must be set.

To find these varabiales many combinations of all of them were used to develop small patterns on a sample at once. Figure 6.2a shows an array of patterns all with the same shape but with different dimensions for the line width and gap width. Each element of this array can be exposed in sequences covering a range of dose and defocus values. After the pattern is developed the different patterns are searched through with a microscope or,



(a) (b)

Figure 6.2: Some examples of the UV exposure patterns that are used to shape photoresists. Red areas indicate where exposure will occur. (a) A pattern of an array of lines with different line widths and gap widths. (b) A pattern with line width of $4\mu m$ and gap width of $2\mu m$ which is the pattern used in the final nanofabrication process.

for more accurate measurements, a scanning electron microscope (SEM) is used.

The values that produced a pattern with the correct dimensions were a dose of $150mJ/cm^2$ and a defocus of 0. With these settings the dimensions of the developed pattern very closely matched the dimensions of the input pattern file.

AZ nLOF2035 is a chemically amplified photoresist so after exposure it needs to be baked. Each sample that was exposed was then immediately baked at 115 $^{\circ}C$ for 60 seconds. The temperatures of the soft bake and PEB do not need to be the same as each other but there is not a lot of time between the soft bake step and the PEB step and so it was decided just to use the same temperature for both so that the hotplate is not still adjusting between temperatures during the bake.

After the PEB, the photoresist needs to be developed. The developing procedure is fairly simple. Holding the sample with tweezers and without letting go, the sample is submerged into a beaker filled with the photoresist developer. For the AZ nLOF 2035 photoresist, the appropriate developer is MIF 300. While the sample is in the developer, small motions are made with the tweezers to adjetate the dissolving material. After a precisely measured amount of time the sample is quickly moved into a beaker of deionized water to halt the development process. The sample is again, adjetated while in the water and it is kept in the water for 60 seconds. If the sample is just left in the developer then it will continue to dissolve away the photoresist. That means that the development time

is actually another factor that affects the final dimensions of the patterned photoresist. Again, to find the right development time a number of combinations need to be tried. After doing so it was found that 110 seconds of developing time was ideal for producing the dimensions of the pattern.

After the patterns were developed they were all checked under an optical microscope to confirm that photoresist developed correctly without losing adhesion or breaking the pattern somewhere. Before the next step, the height of the photoresist was measured using a profilometer. This measurement is made to determine how far the InSb surface is away from the top of the photoresist. This same measurement will be made again after etching to give an approximate measurement of the etch rate.

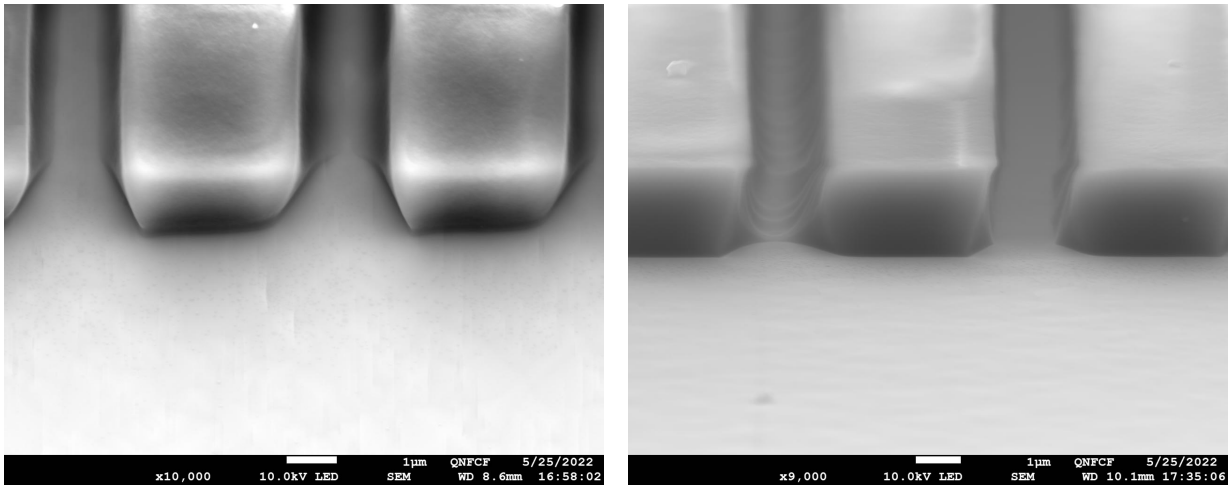
This is only approximate because, in ICP etching the photoresist will also be etched away as well. However, the selectivity of the etch recipe that is chosen is greater than 17:1 meaning the etch rate of the InSb layer will be more than 17 times faster than that for the photoresist. If the amount of time spent etching is small than the difference in the height of the photoresist will not be very much and this is a good approximation for the etch rate.

Finally, the sample is ready for etching. The machine used for etching is the Oxford Instruments Plasmalab System 100. At the Quantum-Nano Fabrication and Characterization Facility (QNFCF) users of this system are not allowed to change any variables of the etch recipe. It would be easy for someone who is not experienced with the machine to make a mistake and damage the machine. The only parameter that can be changed is the amount of time the etch runs for. This is okay, because a properly configured etch should be etching at a constant rate and so if the etch rate is known, then the recipe can just be extended long enough to etch the desired amount.

The etch recipe that was used was obtained from Oxford Instruments process engineers. It was recommended to use CH_4 and H_2 gasses in a 3:1 ratio. From the Oxford instruments benchmark tests, this recipe should etch InSb at a rate of 42 nm/min. This is a good starting point but it should not be assumed that the etch rate will be the same here.

To measure the etch rate, the height difference between the photoresist layer and the InSb surface is measured before and after a 10 minute long etch. Then with this estimate for the etch rate, another sample, prepared identically to the first will be etched for as long as it would take to go all the way through the InSb layer.

The very last step after etching is to remove the remaining photoresist. Specific solvents are made to remove photoresists, whether they have been exposed or not. The AZ 400T photoresist remover is appropriate for AZ nLOF 2035.



(a) (b)

Figure 6.3: Two SEM images of a developed photoresist pattern that had a gap width of $1.3\mu m$. Images (a) and (b) are in adjacent locations on the pattern. (b) shows that the photoresist was not properly developed in some areas.

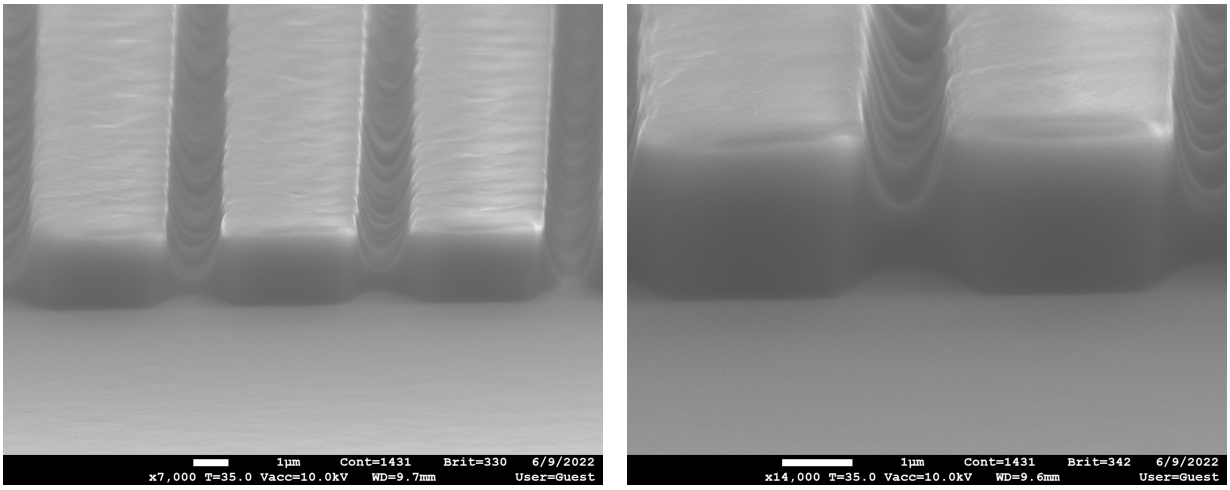
6.3 Results

6.3.1 Photoresist Patterns

In this section some of the results from photoresist development are shared. The target nanostructure mentioned above is an array of $4\mu m$ wide lines with $2\mu m$ gaps in between. This is only after changing targets. Originally, the target was for gaps of $1\mu m$ width. Although the specifications of the MLA150 say that the minimum feature size it can produce is $0.6\mu m$. The closest that was achieved was $1.3\mu m$ but even in this case the pattern was not fully developed. Figure 6.3 shows two SEM images of this pattern in different locations. Figure 6.3a shows that in some places the pattern was developed quite well, but, in Figure 6.3b, at a nearby location the photoresist does not show sharp lines and is not as well develeoped.

Figure 6.4 shows some SEM images of the poorest results for the photoresist pattern when the goal was still to get a $1\mu m$ gap width.

It is clear to see that the gaps in between the lines in Figure 6.4 are not developing well. Furthermore, the developed pattern shows dimensions quite far, relatively speaking, from what the pattern file was. Any time that a pattern file with $1\mu m$ gap widths was



(a) (b)

Figure 6.4: Two SEM images of different developed photoresist patterns that correspond to a pattern file that has a gap width of $0.5\mu m$. Both images show that many of the gaps were not developed fully. Where gap widths were measureable the widths were larger than $1\mu m$.

used, the resulting gap was always significantly more than that, usually close to $2\mu m$. To try to correct for this the dimension of the gap was reduced in the pattern file to $0.5\mu m$. Yet the developed pattern in Figure 6.4a had gap sizes of $1.5\mu m$ (if you measured it where the gap is actually clear) when the pattern had $0.5\mu m$. Figure 6.5 shows what happens when the exact same fabrication procedure is used, with the same exposure parameters. The only difference between the two is that the gap width in the pattern file for Figure 6.5 is $2\mu m$.

When the target gap size was increased to $2\mu m$, the pattern file and the developed pattern matched almost perfectly. In the pattern file the lines had widths of $4.0\mu m$ with gaps of $2.0\mu m$. And then the developed pattern had line widths of $4.1\mu m$ with gaps of $2.0\mu m$. Clearly, the goal of $1\mu m$ gaps was too ambitious and it caused mostly all of the trouble in achieving a well developed photoresist pattern with correct dimensions.

6.3.2 Etching

This part of the research is the most recent, and etching the InSb sample has only been attempted once so far. Although two etches were completed, the first was just for calibration

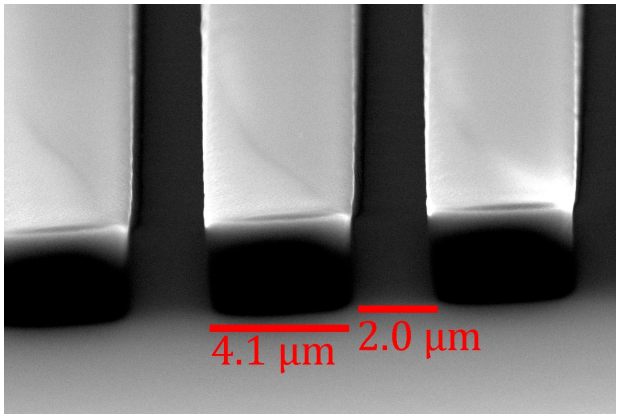
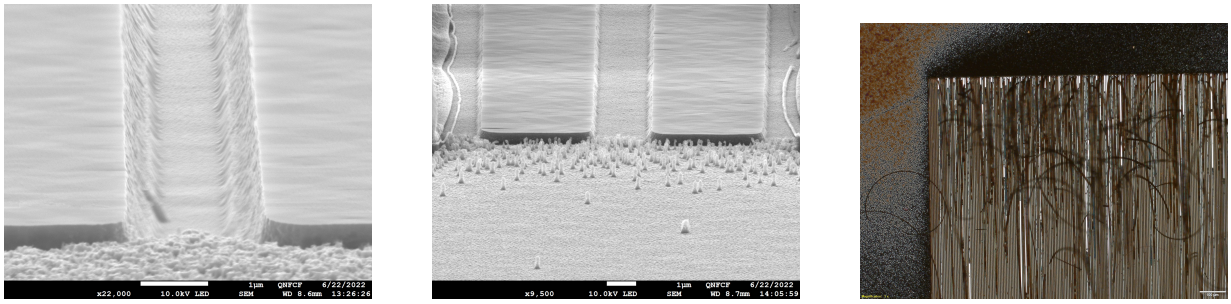


Figure 6.5: An SEM image of a developed photoresist pattern that almost perfectly matched its corresponding pattern file. The pattern file has a gap width of $2.0\mu m$ and a line width of $4.0\mu m$. All of the gaps were clearly well developed with sharp edges. The line width was only $0.1\mu m$ different from the pattern file which is an acceptable amount of deviation.

of the second etch. Two samples prepared identically to the way of the sample in Figure 6.5 were etched. So, these samples have photoresist layers with the pattern of $4\mu m$ lines and $2\mu m$ gaps. The first sample, G0845FC, was etched only for 10 minutes. From the the profilometer measurements, before and after the height of the photoresist layer went from $2.69\mu m$ to $3.34\mu m$, a change of 650 nm in 10 minutes. So, the estimated etch rate is 65 nm/min. A more accurate way to find the etch depth is to instead, strip the photoresist and then measure the height differences with the profilometer. In this case, we saw that only 438 nm had been etched in 10 minutes. This means the etch rate was very close to the Oxford benchmark of 42 nm/min. So using this as our etch rate it was decided to etch the second sample, G0845FB, for 44 minutes, to ensure that the etch goes all the way through the InSb layer. The samples were examined under an optical microscope and with SEM. Figures 6.6 and 6.7 show the results of these two etches

Obviously neither etch went the way that we wanted it to. Both the 10 minute and 44 minute etches have quite a bit of damage. It can be seen that the pattern is broken in many places on the samples. In the optical image of G0845FB, Figure 6.6c, you can see that in some places the photoresist collapsed and fell. Because the photoresist was stripped off of the G0845FC, to measure the etch rate, we can't be sure if the photoresist would have collapsed in that sample as well.

There is a significant amount of damage across the pattern, but some areas do look



(a) (b) (c)

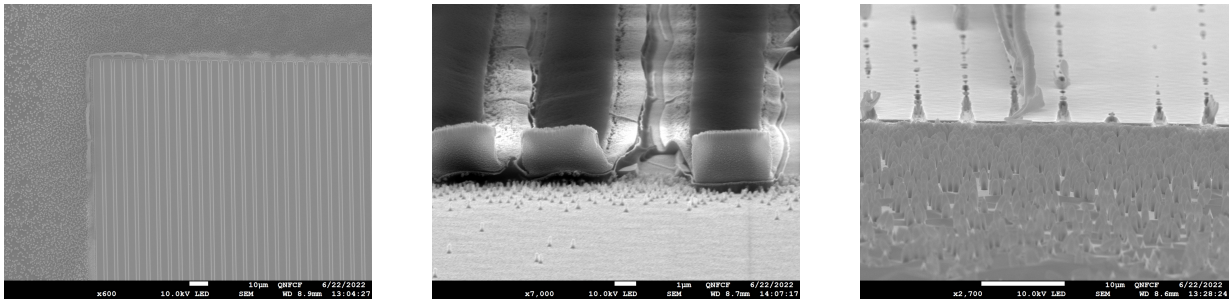
Figure 6.6: Three SEM images of the etching results. Images (a) and (c) correspond to the 44 minute etched sample. (a) shows a close up view of the gap between the etched lines and (c) shows a large area where the etched pattern was most homogeneous. Image (b) corresponds to the 10 minute etched sample and shows a close view of the gap between two lines.

okay. Some of the damage could just be a result of a poor quality spin coating of the photoresist. The size of the entire sample was just 1cm x 1cm and since it is so small and square, the effects from the edges and corners are still significant even in the center of the sample and the photoresist layer will not be completely uniform. If the photoresist layer was made stronger by hard baking it, this might prevent a lot of the damage.

Figure 6.7a shows the largest region that is in "good" shape and it is 7mm x 0.4mm on just one side of the sample. Figure 6.6a shows a close-up of the gap between two lines. You can immediately see there is something strange here. The gap between the lines has been etched further down than the area outside of the pattern. The sidewalls of the InSb lines look very rough and so does the area outside of the pattern.

The roughness outside of the pattern appears to be caused by a strange buildup of little pillars around the pattern, Figure 6.7c. The same pillars can be seen in Figure 6.6b for G0845FC which was only etched for 10 minutes. The density of these pillars is much higher after a longer period of etching so perhaps they are actually the result of some deposition or perhaps migration of atoms from the InSb. It also appears that this buildup may be preventing further etching. By comparing Figures 6.6a and 6.6b, you can see that the etch depths are about the same. Although the true depth of the G0845FB etch is obscured by the pillars.

This build up could be a fluoropolymer material left behind by reactions on the surface. This may be the result of contamination from previous etches performed in the same chamber since the equipment is shared amongst many researchers. One possible solution



(a) (b) (c)

Figure 6.7: Three images are shown that were taken after etching was complete. Images (a) and (c) correspond to the 44 minute etched sample. Image (a) is taken with an optical microscope immediately after etching was complete. The dark tangled lines are parts of the photoresist pattern that separated from the sample during etching. Image (c) is a SEM image that shows a large area of densely packed "pillars" that formed around the entire edge of the pattern. Image (b) corresponds to the 10 minute etched sample and is a SEM image of the photoresist separating from the sample.

to this issue would be to apply a correction to the DC bias used in the etch recipe. If the DC bias is too low then ions will not be accelerated with enough energy to assist in surface reactions and sputtering. This could be why material was left behind on the surface.

The only way to test this would be to repeat the etching again. Overall, it is not a bad result. The fact that you can start to make out the InSb lines is a good sign we are on the right track.

Chapter 7

Conclusion

Chapters 2 through 6 combined, contain a wealth of knowledge about InSb nanostructures. From the initial stages of modelling the THz response, all the way to fabrication. This thesis will be a valuable resource for future projects seeking to develop InSb nanostructures.

Chapter 2 presented the Drude-Lorentz model as a way to model the interactions of InSb with THz light. The useful properties of InSb for exciting surface plasmons were noted. Then, through a computer simulation the behaviour of a simple InSb nanostructure was chracterized. Control of the InSb temperature was used to demonstrate how the character of the THz response can be dynamically tuned. From this chapter, we can conclude that InSb is a promising candidate for developing tunable THz plasmonic devices.

Chapter 3 showed the first steps towards producing a real InSb device. Single crystal epitaxial layers of InSb were successfully grown on three GaAs (001) substrates with different surface misorientations from the [001] direction. Using MBE as the deposition method proved to be a great decision, as the variety of in-situ characterization methods available, namely RHEED, revealed when and where the first growth attempt went wrong. This allowed for a quick recovery and a successful growth was achieved only on the second attempt.

Chapter 4 began the first discussion on the characterization of these three MBE grown layers. Using, AFM, the surface topography was measured for each growth and the RMS roughness was measured. It was seen that the sample with the 2° misorientation towards [010] had the smoothest surface. Additionally, the effects of a varying offcut toward the $[01\bar{1}]$ direction was studied by exploiting the curvature of the edge of a polished wafer. As the offcut increased, approaching the edge of the wafer, the roughness of the sample decreased until the hillocks on the surface completely vanished. The conclusion is that

for offcuts around an angle of 0.73° toward the $[01\bar{1}]$ direction will produce a significantly smoother surface because this is close to the natural facet angle of the hillocks that form when the surface is on-orientation to the [001] direction. Away from this value on either side, the surface will become rough.

Chapter 5 saw the characterization of these growths by XRD. Two different experimental setups were used. From the in-house setup, HRXRD rocking curve scans of the 004 InSb Bragg peaks were studied. Triple axis and double axis scans were both used to estimate the TDD from the FWHM of the rocking curve peaks. Variation of the FWHM was found when the sample was rotated through azimuthal angles. This effect was suspected to be caused by mechanisms of strain relaxation that are unequal in different directions. This resulted in a higher TDD of β type dislocations as compared to the α type ones and this effect is measureable with XRD.

A large disparity was found between the predicted TDDs from double and triple axis scans. Independent measurements of the TDD were made using images from ECCI. A strong agreement was found between the triple axis TDD measurement and the ECCI mesaurement. The conclusion from this is that there is an additional unknown source of strain that increases the measured FWHM in the double axis scans. This source has yet to be identified.

Also in Chapter 5, the synchrotron source at the CLS was used in another study of the azimuthal dependence of the Bragg peaks. Using a diffractometer equipped with an array detector, symmetric ω-2θ scans were used to map out 3D reciprocal space from a number of different orientations. In a way that was unseen before, figures similar to pole figures were generated. Using these figures, it should be possible to study the effect of strain and tilt separately. This idea still needs to be tested. Regardless of whether or not this is true, the figures are still useful as they show the shape of the Bragg peak nearest to the nominal d_{hkl} spacing. The pole figure generated for the on-orientation grown wafer, Figure 5.6a, showed a vertical feature. It was found that this feature is parallel to the direction of maximum FWHM from the rocking curves and it is parallel to the twin boundaries that were imaged by ECCI. Thus, the conclusion is that, in addition to the unbalanced α/β TDDs, the twin boundaries contribute significantly to the anisotropy of the FWHM.

Chapter 6 outlined the procedure for creating InSb nanostructures through the use of UV photolithography and ICP etching. The procedure was used in an attempt to create a nanostructure that had the same shape as the simulated model in Chapter 2. The results after etching were not an acceptable production of this nanostructure but they were promising since the intended shape did begin to form. From start to finish, this procedure has only be followed once so far. Only the final etch still needs to be optimized. Once the

InSb etch recipe is properly optimized this nanofabrication process can be easily adapted to develop any number of differently shaped nanostructures.

Returning to Chapter 1 where the goals for this project were declared, this project can definitely be called a success, although it is far from completed. Procedures were successfully created and deployed for the development, characterization, and study of InSb nanostructures as well as their THz response. A variety of equipment, code and software, were used and in the future they will be easier to access. This project will serve as a stepping stone for the pathway into the development of InSb THz plasmonic devices.

In future projects, more complex nanostructures will be designed for specific applications, such as enhancing the generation of THz FCs by improving mode-locking in THz QCLs. For example, an extension to the nanofabrication procedure is already planned to make use of the thin AlAs layer as a lift off layer. This would allow the InSb nanostructures to be grown on the relatively cheap and abundant GaAs substrates that are compatible with MBE but then be transferred to a different substrate which is more transparent to THz light and may not be MBE compatible. There is a lot of room still for optimization of the MBE growth recipe as well as in the fabrication procedure. Once the etch recipe has been optimized it will be interesting to see how the etch results compare between the on-orientation samples and the ones with intentional misorientation. Also, a deeper analysis of the XRD results should be performed to better understand the properties of the resulting layer from MBE growth. The cause for the unexcepted tilt in the 3D RSMs still needs to be identified and the same is true for the additional strain in the double axis scans.

References

[1] pp. Schematic of a SPP mode, where surface plasmon polaritons propagate along an interface between metal and dielectric., 2021. [Online]. Available: https://en.wikipedia.org/wiki/File:Propagating_surface_plasmon_polaritons.jpg

[2] S. A. Maier, *Plasmonics: Fundamentals and Applications*, 2007.

[3] pp. (a) Kretschmann and (b) Otto configuration of an Attenuated Total Reflection setup for coupling surface plasmons. In both cases, the surface plasmon propagates along the metal/dielectric interface, 2007. [Online]. Available: https://commons.wikimedia.org/wiki/File:Prism_Coupler.png

[4] J. Hammond, N. Bhalla, S. Rafiee, and P. Estrela, "Localized surface plasmon resonance as a biosensing platform for developing countries," *Biosensors*, vol. 4, no. 2, pp. 172–188, 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4264378,https://creativecommons.org/licenses/by/3.0/

[5] D. Grischkowsky, S. Keiding, M. Van Exter, and C. Fattinger, "Far-infrared time-domain spectroscopy with terahertz beams of dielectrics and semiconductors," *Journal of the Optical Society of America B*, vol. 7, no. 10, p. 2006, 1990.

[6] J. E. Ayers, T. Kujofsa, P. Rago, and J. E. Raphael, *Heteroepitaxy of Semiconductors. Theory, Growth, and Characterization.*, 2nd ed. CRC Press, 2017.

[7] *Surface Science*, ser. Advanced Texts in Physics, 2003.

[8] p. The 14 Bravais lattices, 2008. [Online]. Available: https://commons.wikimedia.org/wiki/File:Bravais_lattices.svg

[9] p. Crystal structure of ZnS (sphalerite) with coordination polyhedra, 2007. [Online]. Available: https://commons.wikimedia.org/wiki/File:Sphalerite_polyhedra.png

[10] K. Lu, "Stabilizing nanostructures in metals using grain and twin boundary architectures," *Nature Reviews Materials*, vol. 1, no. 5, p. 16019, 2016.

[11] A. Khalatpour, A. K. Paulsen, C. Deimert, Z. R. Wasilewski, and Q. Hu, "High-power portable terahertz laser systems," *Nature Photonics*, 2020. [Online]. Available: https://www.nature.com/articles/s41566-020-00707-5https://www.nature.com/articles/s41566-020-00707-5.pdf

[12] W. P. Winfree and E. I. Madaras, "Detection and characterization of flaws in sprayed on foam insulation with pulsed terahertz frequency electromagnetic waves," in *41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, Conference Proceedings.

[13] R. M. Woodward, B. E. Cole, V. P. Wallace, R. J. Pye, D. D. Arnone, E. H. Linfield, and M. Pepper, "Terahertz pulse imaging in reflection geometry of human skin cancer and skin tissue," *Phys Med Biol*, vol. 47, no. 21, pp. 3853–63, 2002. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/12452577https://iopscience.iop.org/article/10.1088/0031-9155/47/21/325

[14] B. Ferguson and X.-C. Zhang, "Materials for terahertz science and technology," *Nature Materials*, vol. 1, no. 1, pp. 26–33, 2002.

[15] F. P. Mezzapesa, L. Viti, L. Li, V. Pistore, S. Dhillon, A. G. Davies, E. H. Linfield, and M. S. Vitiello, "Chip-scale terahertz frequency combs through integrated intersubband polariton bleaching," *Laser & Photonics Reviews*, p. 2000575, 2021.

[16] L. Deng, J. Teng, H. Liu, Q. Y. Wu, J. Tang, X. Zhang, S. A. Maier, K. P. Lim, C. Y. Ngo, S. F. Yoon, and S. J. Chua, "Direct optical tuning of the terahertz plasmonic response of insb subwavelength gratings," *Advanced Optical Materials*, vol. 1, no. 2, pp. 128–132, 2013. [Online]. Available: https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/adom.201200032?download=true

[17] T. Yasui, S. Yokoyama, H. Inaba, K. Minoshima, T. Nagatsuma, and T. Araki, "Terahertz frequency metrology based on frequency comb," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 1, pp. 191–201, 2011.

[18] G. P. Williams, "Filling the thz gap—high power sources and applications," *Reports on Progress in Physics*, vol. 69, no. 2, pp. 301–326, 2006.

[19] P. Shumyatsky and R. R. Alfano, "Terahertz sources," *Journal of Biomedical Optics*, vol. 16, no. 3, p. 033001, 2011. [Online]. Available: https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics/volume-16/issue-3/033001/Terahertz-sources/10.1117/1.3554742.pdf

[20] R. Kohler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, "Terahertz semiconductor-heterostructure laser," *Nature*, vol. 417, no. 6885, pp. 156–9, 2002. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/12000955https://www.nature.com/articles/417156a.pdf

[21] Y. Li, *Plasmonic Optics: Theory and Applications*, 2017.

[22] S. M. Hanham, A. I. Fernandez-Dominguez, J. H. Teng, S. S. Ang, K. P. Lim, S. F. Yoon, C. Y. Ngo, N. Klein, J. B. Pendry, and S. A. Maier, "Broadband terahertz plasmonic response of touching insb disks," *Adv Mater*, vol. 24, no. 35, pp. OP226–30, 2012. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/22807039https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/adma.201202003?download=true

[23] S. Lin, K. Bhattarai, J. Zhou, and D. Talbayev, "Thin insb layers with metallic gratings: a novel platform for spectrally-selective thz plasmonic sensing," *Opt Express*, vol. 24, no. 17, pp. 19 448–57, 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/27557222https://www.osapublishing.org/oe/fulltext.cfm?uri=oe-24-17-19448&id=348585

[24] R. Y. Kanyang, F. Zhang, G. Q. Han, Y. Liu, Y. Shao, J. C. Zhang, and Y. Hao, "Rainbow trapping and releasing in insb graded subwavelength grooves by thermal tuning at the terahertz range," *Optical Materials Express*, vol. 8, no. 9, pp. 2954–2966, 2018. [Online]. Available: ⟨GotoISI⟩://WOS:000443314600053https://www.osapublishing.org/ome/fulltext.cfm?uri=ome-8-9-2954&id=396673

[25] X. Y. Liu, Y. Liu, C. Z. Fang, G. Q. Han, and Y. Hao, "Investigation of enhanced transmission and beaming effect through an insb subwavelength grating with a slit at the terahertz range," *Ieee Photonics Journal*, vol. 12, no. 1, pp. 1–12, 2020. [Online]. Available: ⟨GotoISI⟩://WOS:000525345600001https://ieeexplore.ieee.org/ielx7/4563994/8951187/08922753.pdf?tp=&arnumber=8922753&isnumber=8951187&ref=

[26] J. Gómez Rivas, M. Kuttge, H. Kurz, P. Haring Bolivar, and J. A. Sánchez-Gil, "Low-frequency active surface plasmon optics on semiconductors," *Applied Physics Letters*, vol. 88, no. 8, p. 082106, 2006. [Online]. Available: https://pure.tue.nl/ws/files/3408298/388714280289362.pdf

[27] T. Golestanizadeh, A. Zarifi, T. Jalali, J. R. Maack, and M. Wubs, "Hydrodynamic acoustic plasmon resonances in semiconductor nanowires and their dimers," *Journal of the Optical Society of America B-Optical Physics*, vol. 36, no. 10, pp. 2712–2720, 2019. [Online]. Available: ⟨GotoISI⟩://WOS:000488515300025https://www.osapublishing.org/josab/abstract.cfm?uri=josab-36-10-2712

[28] M. Moridsadat and S. Golmohammadi, "Field enhancement in a terahertz plasmonic structure and it's sensing applications," in *Electrical Engineering (ICEE), Iranian Conference on*. IEEE, Conference Proceedings, pp. 320–324. [Online]. Available: https://ieeexplore.ieee.org/document/8472628/

[29] M. Moridsadat, S. Golmohammadi, and H. Baghban, "Tunable multi-band plasmonic response of indium antimonide touching microrings in the terahertz range," *Appl Opt*, vol. 57, no. 16, pp. 4368–4375, 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/29877380https://www.osapublishing.org/ao/abstract.cfm?uri=ao-57-16-4368

[30] M. Oszwałldowski and M. Zimpel, "Temperature dependence of intrinsic carrier concentration and density of states effective mass of heavy holes in insb," *Journal of Physics and Chemistry of Solids*, vol. 49, no. 10, pp. 1179–1185, 1988.

[31] COMSOL, *Wave Optics Module User's Guide*. COMSOL, 2020.

[32] S. Lin, K. Bhattarai, J. Zhou, and D. Talbayev, "Giant thz surface plasmon polariton induced by high-index dielectric metasurface," *Sci Rep*, vol. 7, no. 1, p. 9876, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/28852139

[33] T. Nishinaga, *Handbook of Crystal Growth*. Elsevier, 2015, vol. 1.

[34] A. Sarangan and A. Sarangan, *Nanofabrication*, 2016.

[35] S. Franchi, *Chapter 1 - Molecular beam epitaxy: fundamentals, historical background and future prospects*. Oxford: Elsevier, 2013, pp. 1–46. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123878397000014

[36] M. Henini, "Molecular beam epitaxy: from research to manufacturing," *Thin Solid Films*, vol. 306, no. 2, pp. 331–337, 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0040609097002423?via%3Dihub

[37] S. Yamaguchi and M. Matsumoto, "Stress reduction and electric properties of insb thin films grown by metalorganic vapor phase epitaxy on sapphire substrates with an inas buffer layer," *Vacuum*, vol. 84, no. 11, pp. 1323–1326, 2010.

[38] K. Kanisawa, H. Yamaguchi, and Y. Hirayama, "Two-dimensional growth of insb thin films on gaas(111)a substrates," *Applied Physics Letters*, vol. 76, no. 5, pp. 589–591, 2000.

[39] M. Oszwaldowski, T. Berus, A. Borowska, R. Czajka, and M. Zimniak, "Growth of insb thin films on gaas(100) substrates by flash evaporation epitaxy," *physica status solidi (c)*, vol. 1, no. 2, pp. 351–354, 2004.

[40] S. B. Kaemmer, *PeakForce Tapping – An Introduction*. Bruker Corporation, 2015. [Online]. Available: https://blog.brukerafmprobes.com/application-notes-atomic-force-microscopy/peakforce-tapping-an-introduction/

[41] G. Binnig, C. F. Quate, and C. Gerber, "Atomic force microscope," *Physical Review Letters*, vol. 56, no. 9, pp. 930–933, 1986.

[42] Bruker, *Nanoscope Analysis User Manual*. Bruker Corporation, 2015.

[43] Y. Shi, "Molecular beam epitaxial growth of insb quantum well heterostructures for applications in topological quantum computing," 2021.

[44] O. Madelung, *Semiconductors: Data Handbook*, 2004.

[45] A. M. Glazer, *A Journey into Reciprocal Space (Second Edition)*, 2021.

[46] D. K. Bowen and B. K. Tanner, *X-Ray Metrology in Semiconductor Manufacturing*, 2018.

[47] C. A. MacDonald, *An Introduction to X-Ray Physics, Optics, and Applications*. Princeton University Press, 2017. [Online]. Available: https://app.knovel.com/hotlink/toc/id:kpIXPOA006/introduction-x-ray-physics/introduction-x-ray-physics

[48] S. Gaudet, K. De Keyser, S. Lambert-Milot, J. Jordan-Sweet, C. Detavernier, C. Lavoie, and P. Desjardins, "Three dimensional reciprocal space measurement by x-ray diffraction using linear and area detectors: Applications to texture and defects determination in oriented thin films and nanoprecipitates," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 31, no. 2, p. 021505, 2013.

[49] M. Lee, *X-Ray Diffraction for Materials Research*, 2017.

[50] J. E. Ayers, "The measurement of threading dislocation densities in semiconductor crystals by x-ray diffraction," *Journal of Crystal Growth*, vol. 135, no. 1, pp. 71–77, 1994. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0022024894907277

[51] J. E. Ayers, S. K. Ghandhi, and L. J. Schowalter, "Crystallographic tilting of heteroepitaxial layers," *Journal of Crystal Growth*, vol. 113, no. 3-4, pp. 430–440, 1991.

[52] M. Li, Y. Qiu, G. Liu, Y. Wang, B. Zhang, and L. Zhao, "Distribution of dislocations in gasb and insb epilayers grown on gaas (001) vicinal substrates," *Journal of Applied Physics*, vol. 105, no. 9, p. 094903, 2009.

[53] B. Yarlagadda, A. Rodriguez, P. Li, R. Velampati, J. F. Ocampo, E. N. Suarez, P. B. Rago, D. Shah, J. E. Ayers, and F. C. Jain, "X-ray characterization of dislocation density asymmetries in heteroepitaxial semiconductors," *Applied Physics Letters*, vol. 92, no. 20, 2008. [Online]. Available: https://aip.scitation.org/doi/pdf/10.1063/1.2936078

[54] Y. Sun, K. Li, J. Dong, X. Zeng, S. Yu, Y. Zhao, C. Zhao, and H. Yang, "The anisotropic distribution of dislocations and tilts in metamorphic gainas/alinas buffers grown on gaas substrates with miscut angles toward (111)a," *Journal of Alloys and Compounds*, vol. 597, pp. 45–49, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925838814002710?via%3Dihub

[55] R. S. Goldman, K. L. Kavanagh, H. H. Wieder, S. N. Ehrlich, and R. M. Feenstra, "Effects of gaas substrate misorientation on strain relaxation in inxga1xas films and multilayers," *Journal of Applied Physics*, vol. 83, no. 10, pp. 5137–5149, 1998. [Online]. Available: https://aip.scitation.org/doi/pdf/10.1063/1.367331

[56] T. J. Bruno, R. Deacon, J. A. Jansen, N. Magdefrau, E. Mueller, G. F. Vander Voort, and D. Yang, *ASM Handbook, Volume 10 - Materials Characterization (2019 edition)*. ASM International, 2019. [Online]. Available: https://app.knovel.com/hotlink/toc/id:kpASMHVM2B/asm-handbook-volume-10/asm-handbook-volume-10

APPENDICES

Appendix A

TE Polarization SPP Excitation

For TE polarized light incident on the dielectric-metal interface in Figure 2.1 the wave equations governing this system are Equations A.1 [2]. The solutions for the electric and magnetic fields are given by Equation A.2. Positions with $z > 0$ will be inside the dielectric medium and $j = d$. While for $z < 0$, in the metallic medium, $j = m$. k_0 is the wave vector in vacuum.

$$\begin{aligned} 0 &= \frac{\partial^2 E_y}{\partial z^2} + (k_0^2 \varepsilon - k_x^2) E_y \\ H_x &= i \frac{1}{\omega \mu_0} \frac{\partial E_y}{\partial z} \\ H_z &= \frac{k_x}{\omega \mu_0} E_y \end{aligned} \tag{A.1}$$

$$\begin{aligned} \mathbf{E}_j &= (0\hat{x} + E_{y,j}\hat{y} + 0\hat{z}) e^{i(k_{x,j}x + k_{z,j}z - \omega t)} \\ \mathbf{H}_j &= (H_{x,j}\hat{x} + 0\hat{y} + H_{z,j}\hat{z}) e^{i(k_{x,j}x + k_{z,j}z - \omega t)} \end{aligned} \tag{A.2}$$

Also note that both in the dielectric medium and in the metallic medium $k_{z,j}$ is an imaginary number such that the field is attenuated completely over a short distance into the respective medium. $k_{z,d} = -i\kappa_{z,d}$ and $k_{z,m} = i\kappa_{z,m}$ where both $\kappa_{z,j}$ are positive and real valued because of the directions of penetration. If we apply the conditions to make E_y continuous across both mediums at $z = 0$ we get Equations A.3.

$$\mathbf{E}(z=0) = \begin{cases} E_{y,d}\, e^{i(k_{x,d}x-\omega t)}\, \hat{y} & j=d \\ E_{y,m}\, e^{i(k_{x,m}x-\omega t)}\, \hat{y} & j=m \end{cases} \tag{A.3}$$

$$E_{y,d}e^{i(k_{x,d}x)} = E_{y,m}\, e^{i(k_{x,m}x)}$$

Therefore we get the conditions $E_{y,d} = E_{y,m}$ and $k_{x,d} = k_{x,m}$. Now we apply the continuity condition to H at $z = 0$.

$$\mathbf{H}(z=0) = \begin{cases} (H_{x,d}\hat{x} + H_{z,d}\hat{z})e^{i(k_{x,d}x-\omega t)} & j=d \\ (H_{x,m}\hat{x} + H_{z,m}\hat{z})e^{i(k_{x,m}x-\omega t)} & j=m \end{cases} \tag{A.4}$$

$$\begin{aligned} (H_{x,d}\hat{x} + H_{z,d}\hat{z})e^{i(k_{x,d}x)} &= (H_{x,m}\hat{x} + H_{z,m}\hat{z})e^{i(k_{x,m}x)} \\ H_{x,d}\hat{x} + H_{z,d}\hat{z} &= H_{x,m}\hat{x} + H_{z,m}\hat{z} \end{aligned}$$

Therefore we get the conditions $H_{x,d} = H_{x,m}$ and $H_{z,d} = H_{z,m}$. However from Equation A.1 we know that

$$\begin{aligned} \frac{\partial E_{y,j}}{\partial z} &= ik_{z,j}E_{y,j}e^{i(k_{x,j}x+k_{z,j}z-\omega t)} \\ H_{x,j} &= -\frac{E_{y,j}k_{z,j}}{\omega\mu_0}e^{i(k_{x,j}x+k_{z,j}z-\omega t)} \end{aligned} \tag{A.5}$$

$$H_{x,d} = -\frac{E_{y,d}k_{z,d}}{\omega\mu_0} = i\frac{E_{y,d}\kappa_{z,d}}{\omega\mu_0} \qquad H_{x,m} = -\frac{E_{y,m}k_{z,m}}{\omega\mu_0} = -i\frac{E_{y,m}\kappa_{z,m}}{\omega\mu_0} \tag{A.6}$$

Therefore using TE polarized waves requires that $\kappa_{z,d} = -\kappa_{z,m}$. This can't be true because both were defined as positive real numbers. In conclusion, TE modes cannot excite SPPs.

Appendix B

Reciprocal Space Maps

B.1 Mapping to reciprocal space

This appendix will summarize the methods used to create 3D RSMs and the figures derived from the data set. Figure B.1 shows the system that we will be constructing. The coordinate axes will be defined in the reference frame of the sample plane. First, place the sample plane for which we want to measure the diffraction peak, at the origin of Cartesian axes with the the y axis normal to the plane. Then place the detector at some distance $d_{detector}$ along the positive x axis with the normal of the detector plane parallel to the x axis. The wave of the incident x-ray $\mathbf{K}_{ijk}$ is travelling towards the sample plane in the -y and +x directions within the xy plane. It is incident at angle θ to the sample plane.

Note that the incident wave vector actually only depends on the index i because each pixel on a frame captured by the detector corresponds to the same incident angle.

A diffracted wave vector $\mathbf{K'}_{ijk}$ will point towards one of the pixels on the detector and the pixels are indexed as ijk. Because we assume the scattering is elastic we know that the magnitude of this vector must be $2\pi/\lambda$ where, λ is the wavelength, but we do not have the direction.

It is best for the detector to have an odd number of pixels in each dimension so that one pixel can be considered the center of the array. The detector has J pixels along the z axis and K pixels along the y axis, where J and K are odd numbered. If the center pixel is centered on the x axis, then the index j counts the pixels along the z axis. Starting from $j = 1$, a larger j corresponds to a pixel further in the -z direction. Likewise, $k = 1$ is the first pixel along the y axis and then larger k values correspond to pixels further in the -y direction.

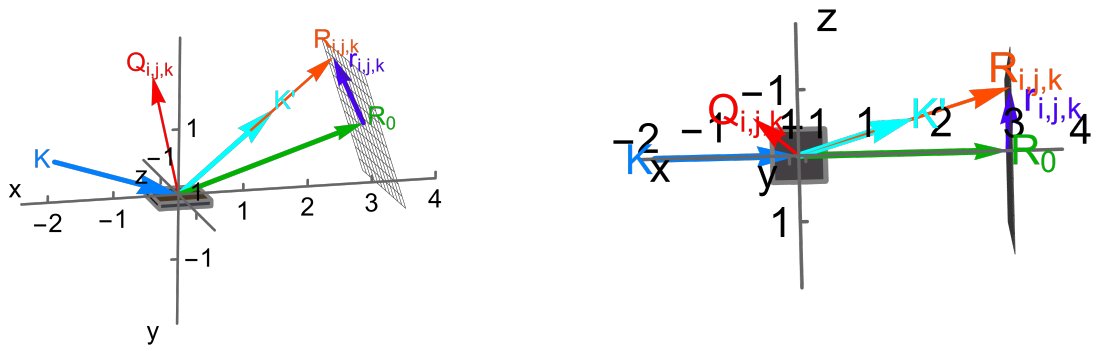


Figure B.1: 3D Schematic showing the relevant vectors in calculating a RSM. (a) Side angle view. (b) Top Down View.

The third index is used to index each image of the scan as the motors for ω and 2θ are stepped through a predetermined range. As mentioned earlier, in an ω-2θ scan the 2θ step size will be twice as large as the ω step size but in the reference frame of the sample plane, we only see the incident angles as equal to θ_i.

To include the effect of the motions of the sample and detector we define the rotation matrix $M(\theta_i)$ that, with the positive z axis out of the page, rotates any vector by θ_i in the clockwise direction around the z axis.

$$M(\theta_i) = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 \\ \sin\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

After rotating the ω motor, the incident wave vector is then given by

$$\mathbf{K}_{ijk} = \frac{2\pi}{\lambda} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot M(\theta_i)$$

Another vector $\mathbf{r}_{ijk}$ is in the detector plane and points from the center pixel of the detector, to any pixel ijk at any rotation θ_i. The detector rotates in the opposite direction to the incident x-ray and so $\mathbf{r}_{ijk}$ is

$$\mathbf{r}_{ijk} = S_p \begin{bmatrix} 0 \\ \frac{K}{2} - k + \frac{1}{2} \\ \frac{J}{2} - j + \frac{1}{2} \end{bmatrix} \cdot M(-\theta_i)$$

where S_p is the dimension of the square pixels. The addition of 1/2 to the y and z components makes sure that the vector points at the center of the pixel rather than the corner. $\mathbf{R}_0(\theta_i)$ is the vector from the origin to the center of the detector after it has rotated by θ_i.

$$\mathbf{R}_0(\theta_i) = d_{detector} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot M(-\theta_i)$$

Adding these two vectors together we get $\mathbf{R}_{ijk} = \mathbf{R}_0(\theta_i) + \mathbf{r}_{ijk}$ which points from the origin to any pixel ijk. $\mathbf{R}^{\star}_{ijk} = \frac{\mathbf{R}_{ijk}}{|\mathbf{R}_{ijk}|}$ is just a unit vector along this same direction.

Now that we have a unit vector in the direction from the origin to any pixel we find the diffracted wave vector to be

$$\mathbf{K}'_{ijk} = \frac{2\pi}{\lambda}\mathbf{R}^{\star}_{ijk}$$

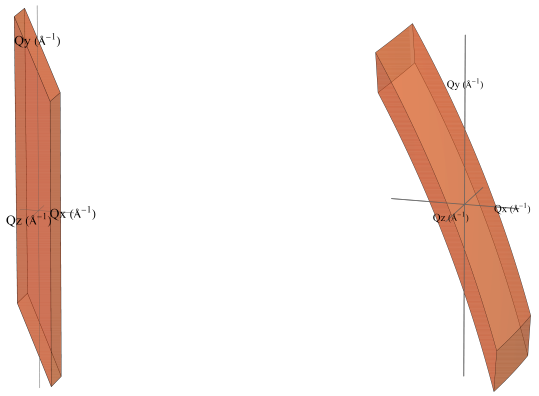
which points from the point of incidence to a pixel ijk.

Finally, we find, based on the positions of the sample and the detector, that each pixel measures the intensity of a point in reciprocal space corresponding to

$$\mathbf{Q}_{ijk} = \mathbf{K}'_{ijk} - \mathbf{K}_{ijk}$$

which is just the Laue condition. A large intensity should only be measured by a pixel ijk, corresponding to $\mathbf{Q}_{ijk}$, if this point meets the diffraction condition to a set of planes in a crystal. Images, taken at each step in an ω-2θ scan, contain the intensity measurements for every pixel in the array detector. Each pixel is indexed by ijk and then the corresponding point in reciprocal space is assigned the intensity of that pixel. As intensity measurements are arbitrarily dependent on the power of the x-ray source, each scan has its intensity values normalized between 0 and 1, while simultaneously converting to a logarithmic scale. The logarithmic scale is typical to reveal features in XRD intensity measurements.

Using Wolfram Mathematica to calculate each $\mathbf{Q}_{ijk}$ for all of the available indices, a volume of reciprocal space is mapped. A schematic representation of the volume is shown in Figure B.2. When the detector distance $d_{detector}$ is comparable to the pixel size S_p the shape



(a) (b)

Figure B.2: A schematic showing the shape of the region of reciprocal space that is sampled in an ω-2θ scan with an array detector.

is curved similar to that in Figure B.2b. But for lower angular acceptance, in HRXRD, the detector is kept much further away. This results in an approximately parallelepiped volume like in Figure B.2a.

B.2 3D Plots

The 3D plots in Section 5.3 are plots of data points taken directly from the 3D RSM. Using Wolfram Mathematica's built-in 3D plotting capabilities the data was plotted as individual points. The points are shown as circles of finite size with the colour and opacity determined by the intensity of the data point. Plotting millions of points in this way is taxing on computer resources and therefore it is not possible to plot every single data point on the 3D axes. A subset of 500 000 points is chosen by taking a pseudo random sample of the data set and plotting that instead. In the 3D plots the Bragg peak appears as a dense cloud of points. The outer layer of this cloud is made up of the lowest intensity points and so it obscures the high intensity points. To reveal the inside of the cloud, a cut plane is added to the figure, and points are only plotted on one side of the infinite plane.

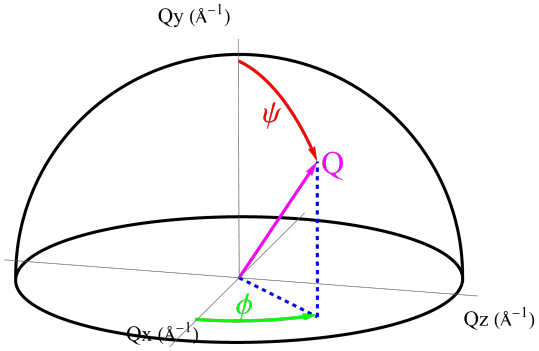


Figure B.3: Spherical Coordinates

B.3 2D Polar Plots

To create the polar plots in Section 5.3 the Cartesian coordinates (Q_x, Q_y, Q_z) of the 3D RSM are first converted into spherical coordinates (Q_r, Q_ϕ, Q_ψ) by using Equations B.1. The azimuthal angle, ϕ and the polar angle, ψ are defined in Figure B.3. The addition of ϕ_0 is used to correctly orient a scan in the sample reference frame for each of the different azimuthal positions in each ω-2θ scan.

$$\begin{aligned} Q_r &= \sqrt{Q_x^2 + Q_y^2 + Q_z^2} \\ Q_\phi &= \arctan\left(\frac{Q_z}{Q_x}\right) + \phi_0 \\ Q_\psi &= \arccos\left(\frac{Q_y}{\sqrt{Q_x^2 + Q_y^2 + Q_z^2}}\right) \end{aligned} \tag{B.1}$$

After converting each scan to spherical coordinates a subset of the data would be chosen. This is based on the radial coordinate Q_r. Pole figure measurements are typically taken only with a single radius along one sphere because the detector is kept stationary. In contrast, the data collected here corresponds to many different radii and so the data is first selected to be within a small range of a specific radius, $Q_{hkl} \pm \delta$. This effectively selects data along a spherical shell that has finite thickness of $2|\delta|$. Q_{hkl} is calculated from Equation 5.1 and $||\mathbf{Q}_{hkl}|| = 2\pi/d_{hkl}$. For the 004 reflection $Q_{004} = 3.879\mathring{A}^{-1}$. δ was chosen to be

$0.15\mathring{A}^{-1}$ because it selected about half of the values in the whole data set. $\delta = 0.15\mathring{A}^{-1}$ corresponds to Q_{004} within $\pm \sim 4\%$.

A polar plot was then created using only the Q_ϕ and Q_ψ coordinates. In the plot, the radius is Q_ψ while the angular position is Q_ϕ. This will flatten the spherical cap onto the circular surface of the polar plot. The measurements however, were not taken at continuous ϕ positions and so the plotted data does not fully cover a circular region. In the plot, the region covered by the measurements is actually made up of overlapping rectangles each rotated 30° from the next in the sequence. Even when most of the space is empty, the radius of the circle on which each plot is made actually covers the whole range over where measurements were taken, just not at every angle ϕ so there may be small gaps near the edges of the circle. This is how the polar plots were made in Figure 5.6.

Appendix C

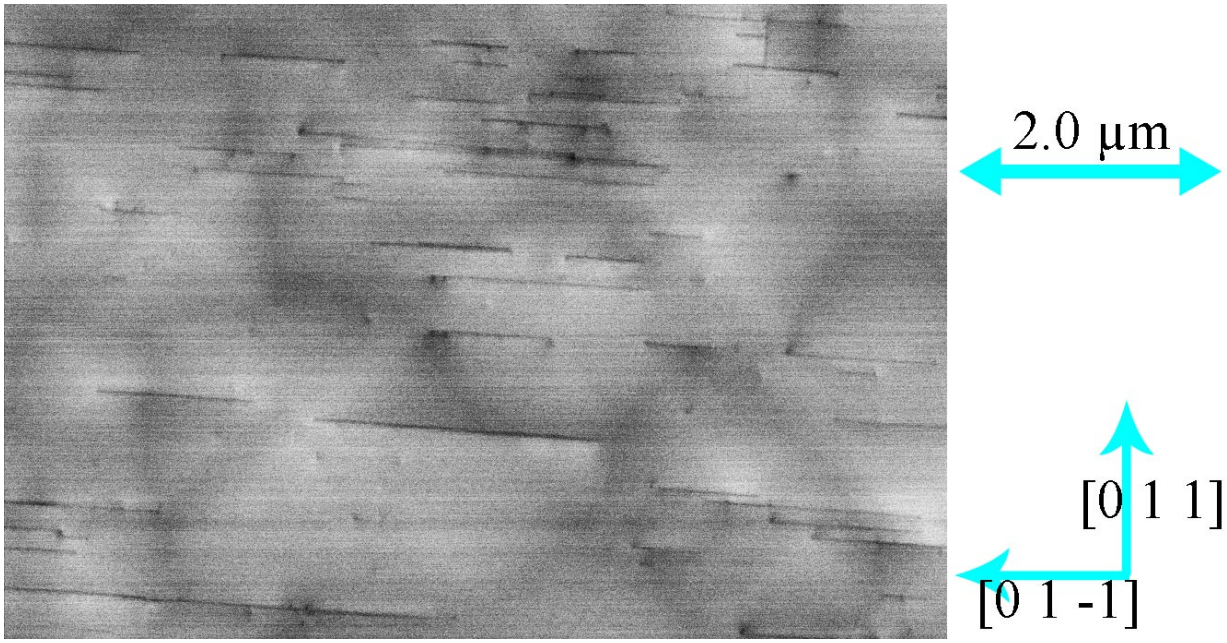
Electron Channelling Contrast Imaging

Electron channeling contrast imaging is a technique that uses backscattered electron signals acquired with a SEM. An electron beam is emitted towards the surface of the sample and the signal is measured as the intensity of the electrons that are scattered back out of the sample [56, pg. 564-565]. Images are collected by scanning the beam across the surface and using the electron signal to determine contrast in the images. The majority of the signal comes from electrons that were scattered without losing energy and so the electrons that are measured are the ones that only penetrated a very shallow depth into the sample. Thus, ECCI is mostly sensitive to features at and near the surface usually at depths of 10 to 100 nm [56, pg. 564-565].

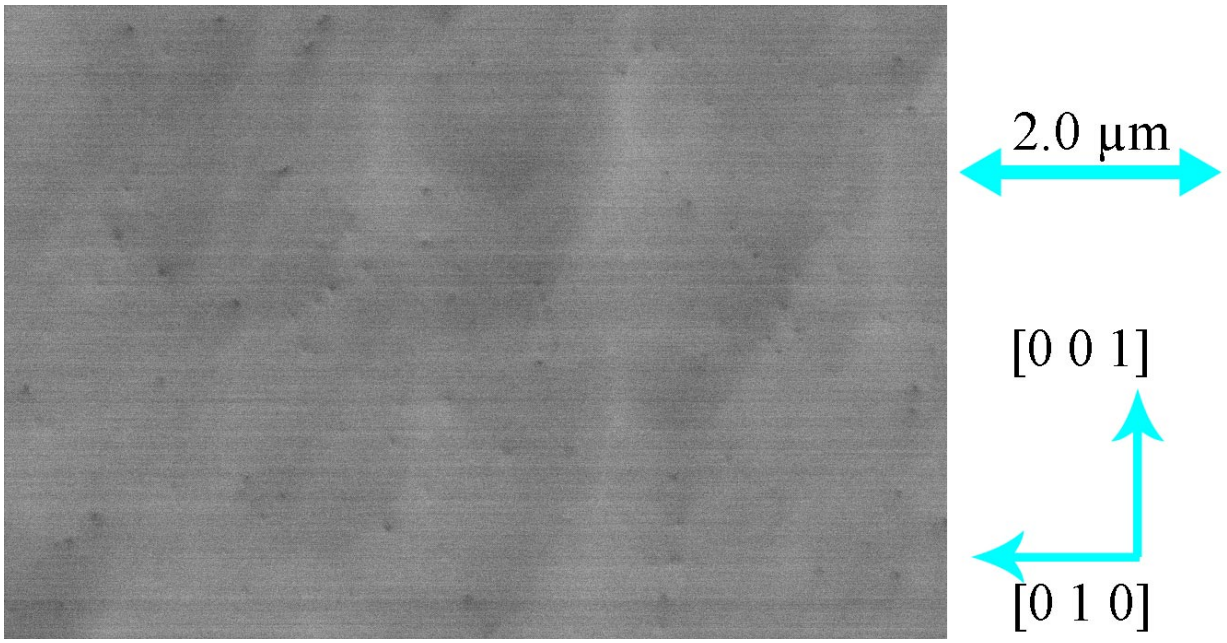
The wave nature of high energy electrons dictates the interaction with the surface in a similar way to x-rays in XRD. Bragg's law, Equation 5.3, applies to these interactions as well. The main distinction here however is that ECCI is sensitive to the vertical planes in the crystal structure rather than the horizontal planes [56, pg. 564-565]. For example, XRD signals are measured by diffraction peaks from planes such as the (001) planes where the [001] direction is normal to the sample surface as well as the plane. In contrast, ECCI signals measured from this sample would be sensitive to either the (100) or (010) planes where the [100] and [010] directions are in the plane of the sample surface. A perfect crystal would show a nonzero backscatter intensity that is constant at every point on the sample surface. Imperfections such as dislocations in the crystal structure causes slight tilts in crystallographic planes and localized strains are induced within the surrounding region.

For the best contrast image, the electron beam must be aligned to the Bragg angle θ_B

of a specific set of planes [56, pg. 564-565]. When a dislocation causes the planes to bend, a contrast in the signal is produced. The intensity can be either increased or decreased showing either bright or dark spots in the images depending on which direction the planes are tilted with respect to the incoming beam. The highest intensity is seen when the planes are most parallel to the electron beam and conversely, the minimum is seen when the planes are least parallel [56, pg. 564-565]. The beam can be aligned to either $+\theta_B$ or $-\theta_B$ and switching between positions will invert the contrast of the image. This is how ECCI is used to reveal the presence of dislocations in the sample.



(a)



(b)

Figure C.1: SEM ECCI images for (a) sample G0845 and (b) sample G0846. The corresponding crystallographic directions are indicated by the blue arrows on the right.

Appendix D

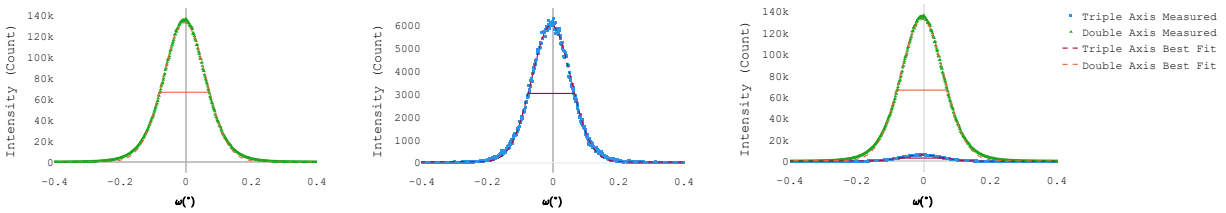
Rocking Curve Fitting

This section will explain the process used to measure width of data collected in rocking curve scans with the QC3 setup. In the rocking curve scan, the intensity of the x-ray beam is measured as the ω motor is stepped through an angular range. The range is selected to be large enough to ensure that the entire peak is covered. Depending on how the motor positions are aligned the peak may not be exactly in the center of the scan. The first step then is to center the scan. This scan is approximately centered by finding the angular position of the highest intensity point and substrating it from all of the angular positions. Next, using a non-linear least squares fit, the data is fit to the function in Equation D.1. Equation D.1 is the equation for a Gaussian function. The actual estimation of fitting parameters is done by Python code using tools in the Scipy package.

$$y(x) = H + A \exp\left(\frac{-(x - x_0)^2}{2\sigma^2}\right) \tag{D.1}$$

Here, $y(x)$ is the intensity measured at the point x and x is the angular position. H and x_0 respectively serve to allow the entire curve to be translated in y and in x. The parameter A determines the height of the highest point in the curve at $x = x_0$ and σ determines the width of the peak. An appropriate value fitting the raw data is found for each of H, A, x_0, and σ. Then the fitted value for intensity is found at each of the angular positions measured in the rocking curve. Figure D.1 shows plots of the real data nearly indistinguishable from the fitted values.

To quantitatively measure the quality of the fit the coefficient of determination R^2 is calculated for each fit. R^2 is calcated by first calculating the residual sum of squares SS_{res}



(a) (b) (c)

Figure D.1: Rocking Curves (a) double axis rocking curve. (b) triple axis rocking curve. (c) both triple and double axis rocking curves for scale.

and the total sum of squares SS_{tot} given by Equation D.2. For every single fit the R^2 value was greater than 0.997.

$$\begin{aligned} SS_{res} &= \sum_i (y_i - f_i)^2 \\ SS_{tot} &= \sum_i (y_i - \bar{y})^2 \\ R^2 &= 1 - \frac{SS_{res}}{SS_{tot}} \end{aligned} \tag{D.2}$$

From the fitting functions the FWHM of each curve is measured by identifying intersection points between the curve and a horizontal line at half of the maximum value of intensity. Again, this was done using tools in the Scipy package. These horizontal lines, indicating the height of half of the maximum, are drawn in Figure D.1. The FWHM could have theoretically been calculated simply from the fitting parameter σ by $FWHM = 2\sqrt{2\ln 2}\sigma$. Both methods are valid and using the line intersection gives an intuitive graphical approach to measuring the FWHM.