Imperial College London

Analytical Transmission Electron Microscopy of InAs/GaAs quantum dots and GaInNAs/GaAs quantum wells

Shima Kadkhodazadeh

Thesis submitted for the degree of
DOCTOR OF PHILOSOPHY
of Imperial College London

Supervisors: Prof David W McComb
Prof Timothy S Jones

Department of Materials and
London Centre for Nanotechnology
Imperial College London

Abstract

Structure and chemical composition of InAs/GaAs quantum dots and GaInNAs/GaAs quantum wells were investigated using analytical transmission electron microscopy. These material systems are of importance for fabricating optoelectronic devices capable of operating in the 1.3 – 1.55 μm wavelength range. The samples were grown by molecular beam epitaxy. The optical properties of the quantum dots and quantum wells are largely determined by their structure and composition. Properties such as the shape, size and composition of the InAs/GaAs quantum dots and the morphology and composition of the GaInNAs/GaAs quantum wells were studied using high angle annular dark field imaging and electron energy-loss spectroscopy. In the case of the InAs/GaAs quantum dots, in particular, the effects of the GaAs spacer layer thickness in bilayer quantum dots and overgrowing the quantum dots with GaInAs instead of GaAs on the properties of the dots were investigated. The results indicated that the GaAs spacer layer can have a significant influence on the shape, size and nucleation position of the quantum dots in the second layer. In addition to this, it was demonstrated that overgrowing the quantum dots with GaInAs instead of GaAs introduces smaller changes to the structure and composition of the dots. In case of the quantum wells, the influence of the growth temperature and In composition on their morphology and composition were examined. Higher growth temperatures and In contents were shown to lead to the degradation of the 2D morphology and formation of inhomogeneities in the chemical composition of the quantum wells. When correlated with the optical performance of these structures, these results can provide important information about their structure – property relationships. Finally, the possibility of utilising electron energy-loss spectroscopy in monochromated transmission electron microscopy instruments to measure and map the band gap of semiconductor structures was studied and the results discussed.

Acknowledgements

I would like to thank my supervisors Prof David McComb and Prof Tim Jones for their guidance and support throughout my PhD and for always encouraging me to show initiative and direct my research independently.

I am grateful to the physical science research council (EPSRC) for funding and supporting this work.

Many thanks to Dr James Perkins, Dr Catriona McGilvery and Dr Mahmoud Ardakani for training me to operate the TEM instruments used in this work and for their technical support. In particular, I am very grateful to Dr James Perkins for his on-call technical support with the Titan microscope during the evening and weekend sessions. Thanks to Dr Mark Ashwin for the MBE growth of the samples studied. Thanks to Dr Paul Robb for reading and providing invaluable comments on the contents of this thesis.

I would like to mention my colleagues, Sehban Husain, Alex Jasper, Virginie Jantou and the others in office 110 in the Department of Materials for their friendship, great sense of humour and the positive atmosphere of our work place. I have many happy memories of my time in the Department of Materials that will always stay with me. Thanks also to my friends, especially Rebecca Bromely, for her support and for always being willing to hear about my QDs and QWs, despite her artistic background.

Last and not least, I would like to thank my wonderful parents. Madar and pedar, thank you for loving me so unconditionally and always wanting the very best for me. I hope this makes you proud.

Table of Contents

ABSTRACT **2**

ACKNOWLEDGEMENTS 3
LIST OF ACRONYMS 7
LIST OF TABLES 8
LIST OF FIGURES 9

1 INTRODUCTION **16**

1.1 BASICS 16
1.1.1 LIGHT EMITTING SEMICONDUCTOR STRUCTURES 18
1.1.2 EPITAXIAL CRYSTAL GROWTH: MOLECULAR BEAM EPITAXY (MBE) 21
1.1.3 BULK GAAS AND INAS CRYSTALLOGRAPHY 24
1.1.4 MODES OF FILM GROWTH 24
1.2 MOTIVATION 27
1.2.1 INAS /GAAS QUANTUM DOTS 27
1.2.2 GAINNAS / GAAS QUANTUM WELLS 29
1.3 OUTLINE OF THE THESIS 31

2 INSTRUMENTATION AND EXPERIMENTAL DETAILS **33**

2.1 INTRODUCTION 33
2.2 ELECTRON-SPECIMEN INTERACTIONS 34
2.2.1 ELASTIC SCATTERING 35
2.2.2 INELASTIC SCATTERING 37
2.3 COMPONENTS OF A TEM 38
2.3.1 ELECTRON SOURCES 38
2.3.1.1 THERMIONIC GUNS 39
2.3.1.2 FIELD EMISSION SOURCES 41
2.3.2 LENSES, ABERRATIONS AND APERTURES 42
2.3.3 ELECTRON DETECTORS 46
2.3.4 EELS SPECTROMETER 47
2.4 CONVENTIONAL TEM 49
2.5 STEM 50
2.6 FEI TITAN 53
2.6.1 MONOCHROMATOR 55
2.6.2 CONDENSER SYSTEM 57
2.6.3 OBJECTIVE SYSTEM 59
2.6.4 EELS SPECTROMETER 60
2.7 SPECIMEN PREPARATION 61
2.7.1 FOCUSED ION BEAM (FIB) MILLING 61
2.7.2 MECHANICAL POLISHING, DIMPLING AND AR+ ION MILLING 61
2.8 SUMMARY 64

3 ELECTRON ENERGY-LOSS SPECTROSCOPY **66**

3.1 INTRODUCTION 66
3.1.1 BETHE THEORY 69
3.1.2 DIELECTRIC FORMULATION 70
3.2 EXCITATION OF INNER-SHELL ELECTRONS 70
3.3 EXCITATION OF OUTER-SHELL ELECTRONS 73
3.3.1 SINGLE ELECTRON EXCITATIONS 73
3.3.2 PLASMON EXCITATIONS 75
3.3.2.1 VOLUME PLASMONS 75
3.3.2.2 SURFACE AND INTERFACE PLASMONS 77
3.4 QUANTITATIVE ANALYSIS OF EELS 80
3.4.1 REMOVAL OF PLURAL SCATTERING 80
3.4.1.1 FOURIER-LOG DECONVOLUTION 80

3.4.1.2 Fourier-ratio deconvolution 81
3.4.2 Kramers-Kronig analysis 82
3.5 Other considerations 83
3.5.1 Spatial resolution 83
3.5.2 Statistical detection limits 85
3.6 Summary 86

4 STRUCTURAL INVESTIGATIONS OF INAS/GAAS QUANTUM DOTS 88

4.1 Introduction 88
4.1.1 General aspects of growth 89
4.1.2 Shape 90
4.1.3 Composition 95
4.2 Experimental details 96
4.3 Results and Discussion 98
4.3.1 Structural investigations using TEM and STEM 98
4.3.2 Thickness contrast in images 100
4.3.3 Bilayers of QDs 101
4.3.4 Shape 102
4.3.5 Size 108
4.3.6 Composition 109
4.3.7 GaInAs capping 113
4.4 Summary 115

5 STRUCTURAL INVESTIGATIONS OF GAINNAS / GAAS QUANTUM WELLS 117

5.1 Introduction 117
5.1.1 Composition modulation and morphology breakdown of GaInNAs 118
5.2 Experimental details 122
5.3 Results and Discussion 124
5.3.1 Structure and morphology 124
5.3.2 Composition 129
5.3.2.1 Z-contrast imaging 129
5.3.2.2 EELS 135
5.3.3 Discussion 138
5.3.4 Plasmon mapping 142
5.4 Summary 146

6 MEASURING INTERBAND TRANSITIONS AND OPTICAL PROPERTIES OF SEMICONDUCTORS USING EELS 148

6.1 Introduction 148
6.2 Challenges of measuring the band gap and optical properties of semiconductors using VEELS 150
6.2.1 Energy resolution 150
6.2.2 Data acquisition 150
6.2.3 ZLP removal 151
6.2.3.1 Pre-recorded ZLP deconvolution 151
6.2.3.2 Reflected left-hand tail fitting of the ZLP 152
6.2.4 Relativistic radiation losses 152
6.2.4.1 Kröger theory 155
6.3 Modelling the relativistic losses using Kröger's equation 158
6.3.1 Accelerating voltage 158
6.3.2 Spectrometer collection angle 159
6.3.3 Specimen thickness 160
6.4 Removing relativistic losses: off-line removal using Kröger's equation 162
6.5 Band transitions and optical properties of GaAs 164
6.5.1 Experimental conditions 164
6.5.2 Results and discussion 164
6.6 Measuring the band transitions of GaInNAs QWs 170

6.6.1 RESULTS AND DISCUSSION 170
6.7 SUMMARY 173
7 CONCLUSIONS AND FUTURE WORK 175
APPENDIX A 178
APPENDIX B 180
REFERENCES 184

List of Acronyms

AFM	Atomic Force Microscopy
BF	Bright Field
CCD	Charge Coupled Device
DF	Dark Field
DOS	Density Of States
EELS	Electron Energy-Loss Spectroscopy
EDS	Energy Dispersive Spectroscopy
f.c.c.	Face Centred Cubic
FEG	Field Emission Gun
FIB	Focused Ion Beam
FWHM	Full Width Half Maximum
HAADF	High Angle Annular Dark Field
MBE	Molecular Beam Epitaxy
ML	Monolayer
PL	Photoluminescence
RHEED	Reflection High Energy Electron Diffraction
STEM	Scanning Transmission Electron Microscopy
STM	Scanning Tunnelling Microscopy
SI	Spectrum Imaging
TEM	Transmission Electron Microscopy
QD	Quantum Dot
QW	Quantum Well
UHV	Ultra High Vacuum
VEELS	Valence Electron Energy-Loss Spectroscopy
WL	Wetting Layer
ZLP	Zero-Loss Peak

List of Tables

TABLE 3.1: NOMENCLATURE AND QUANTUM NUMBERS ASSOCIATED WITH IONISATION EDGES. 72

TABLE 4.1: THE BASE LENGTH AND HEIGHT OF THE QDS MEASURED IN FIG. 4.8(A) AND (B). 99

TABLE 4.2: THE AVERAGE BASE LENGTH, HEIGHT AND ASPECT RATIO OF THE QDS IN THE SAMPLES CAPPED WITH GAAS AND WITH S = 60 NM, 20 NM AND 10 NM, MEASURED ALONG THE $[110]$ DIRECTION.....109

TABLE 5.1: THE RESULTS OF THE STEM AND STM MEASUREMENTS OF THE MORPHOLOGY OF ..... $GA_{1-x}IN_xN0_{.05}As_{0.95}$ QWS GROWN AT 400 °C AND 450 °C. 127

List of Figures

FIGURE 1.1: SCHEMATIC OF THE ELECTRONIC STATES FOR ATOMIC, DIATOMIC, POLYMERIC AND SOLID STRUCTURES. 17

FIGURE 1.2: A SIMPLIFIED SCHEMATIC OF THE BAND STRUCTURE OF METALS, SEMICONDUCTORS AND INSULATORS AT 0 K TEMPERATURE. 17

FIGURE 1.3: A SCHEMATIC REPRESENTATION OF OPTICAL ABSORPTION, SPONTANEOUS EMISSION AND STIMULATED EMISSION IN SEMICONDUCTORS. 19

FIGURE 1.4: SCHEMATICS SHOWING EXAMPLES OF THE BAND STRUCTURE OF DIRECT AND INDIRECT BAND GAP SEMICONDUCTORS. 19

FIGURE 1.5: A SCHEMATIC SHOWING HOW THE DENSITY OF STATES OF AN ENERGY BAND IN A SEMICONDUCTOR IS ALTERED AS THE SYSTEM'S DIMENSIONS ARE REDUCED....20

FIGURE 1.6: A SCHEMATIC OF THE MBE GROWTH SYSTEM. 23

FIGURE 1.7: A SCHEMATIC OF THE ZINC BLENDE CRYSTAL STRUCTURE OF UNIT CELLS OF GAAS AND INAS. 24

FIGURE 1.8: A SCHEMATIC ILLUSTRATING SOME OF THE PROCESSES INVOLVED IN THIN FILM GROWTH ON A SOLID SUBSTRATE....25

FIGURE 1.9: SCHEMATIC REPRESENTATIONS OF FRANK-VAN DER MERVE (OR LAYER-BY-LAYER), VOLLMER-WEBER AND STRANSKI-KRASTANOV FILM GROWTH MODES. 26

FIGURE 1.10: A SIMPLIFIED DESCRIPTION OF THE SURFACE/INTERFACE TENSIONS ACTING ON A 3D ISLAND GROWN ON A SUBSTRATE. 26

FIGURE 1.11: BAND GAP AS A FUNCTION OF THE LATTICE CONSTANT OF III/V SEMICONDUCTORS. 30

FIGURE 2.1: A SCHEMATIC SHOWING SOME OF THE IMPORTANT ELASTIC AND INELASTIC SCATTERING EVENTS OCCURRING WHEN AN ELECTRON BEAM INTERACTS WITH A THIN SPECIMEN. 34

FIGURE 2.2: A SCHEMATIC OF A THERMIONIC ELECTRON GUN 40

FIGURE 2.3: A SCHEMATIC OF A FEG....41

FIGURE 2.4: A RAY DIAGRAM DISPLAYING THE EFFECT OF SPHERICAL ABERRATION ON THE IMAGE FORMED OF A POINT OBJECT, P....44

FIGURE 2.5: A SCHEMATIC SHOWING THE DEFLECTING, DISPERSING AND FOCUSING EFFECTS OF THE SPECTROMETER ON ELECTRONS. 48

FIGURE 2.6: A SCHEMATIC SHOWING A SPECTRUM IMAGE DATA SET. 49

FIGURE 2.7: SCHEMATICS OF A TEM IN DIFFRACTION AND IMAGING MODES. 50

FIGURE 2.8: A SCHEMATIC SHOWING IMAGE FORMATION IN STEM....51

FIGURE 2.9: A SCHEMATIC SHOWING THE MAIN COLUMNAR COMPONENTS OF THE TITAN. 54

FIGURE 2.10: SIMPLE RAY DIAGRAMS SHOWING EXAMPLES OF THE ELECTRON BEAM TRAJECTORIES IN A WIEN FILTER....55

FIGURE 2.11: A SCHEMATIC SHOWING THE GUN LENS AND MONOCHROMATOR SETTINGS FOR THE NON-EXCITED AND EXCITED MONOCHROMATOR SETTINGS. 56

FIGURE 2.12: A SCHEMATIC SHOWING THE STEM AND TEM OPERATIONAL MODES OF THE TITAN. ...58

FIGURE 2.13: A RAY DIAGRAM OF THE ZOOM-SYSTEM BETWEEN TWO NEIGHBOURING LENSES WHERE THE POSITION OF THE IMAGE IN FRONT OF THE FIRST LENS AND AFTER THE SECOND LENS REMAIN UNCHANGED BY CHANGING THE POSITION OF THE IMAGE IN BETWEEN THE TWO LENSES.....59

FIGURE 2.14: A SCHEMATIC SHOWING THE MAIN COMPONENTS OF GATAN GIF TRIDIEM SPECTROMETER.....60

FIGURE 2.15: DIFFERENT STAGES OF PREPARING A CROSS-SECTIONAL SPECIMEN.....63

FIGURE 3.1: AN EEL SPECTRUM OF GAN, SHOWING ZERO-LOSS PEAK (ZLP), PLASMON AND OUTER AND INNER SHELL EXCITATIONS. 67

FIGURE 3.2: A SCHEMATIC SHOWING THE INELASTIC SCATTERING OF ELECTRONS IN STEM MODE.....68

FIGURE 3.3: SCHEMATICS SHOWING TRANSITIONS ACROSS A DIRECT AND INDIRECT BAND GAP SEMICONDUCTORS. 75

FIGURE 3.4: THE CONTRIBUTIONS FROM JDOS AND MATRIX ELEMENTS FOR DIRECT AND INDIRECT BAND GAP TRANSITIONS.....75

FIGURE 3.5: A SCHEMATIC SHOWING THE ELECTRIC FIELD LINES ASSOCIATED WITH SURFACE PLASMONS IN A THIN TEM SPECIMEN FOR A SINGLE SURFACE MODE AND A COUPLED SURFACE MODE. 78

FIGURE 3.6: A SCHEMATIC SHOWING THE GEOMETRY OF THE BEAM AT DISTANCE X_0 AWAY FROM AN INTERFACE BETWEEN MEDIA A AND B WITH DIELECTRIC RESPONSES E_A AND E_B RESPECTIVELY. ...79

FIGURE 4.1: STM TOPOGRAPHY OF INAS/GAAS QDS GROWN AT 500°C TEMPERATURE AND 0.008 MLs^{-1} GROWTH RATE AND THE CORRESPONDING FACETS OF THE SMALL AND LARGE QDS 92

FIGURE 4.2: PLAN-VIEW ON-ZONE AND OFF-ZONE TEM IMAGES OF UNCAPPED GE QDS ON SI. 92

FIGURE 4.3: A SCHEMATIC SHOWING THE CHANGES IN THE MORPHOLOGY OF INAS QDS AS THE RESULT OF OVERGROWTH WITH GAAS. 93

FIGURE 4.4: CROSS-SECTIONAL BF TEM IMAGES OF INAS/GAAS QDS OVERGROWN AT THREE DIFFERENT TEMPERATURES.....94

FIGURE 4.5: PILLAR SHAPE SPECIMEN PREPARED USING FIB AND Ar^+ ION MILLING AND EMPLOYED TO CARRY OUT HAADF TOMOGRAPHY ON A SINGLE INAS/GAAS QD. 95

FIGURE 4.6: GA AND IN ELEMENTAL MAPS OBTAINED USING THE IN $N_{4,5}$ AND GA $M_{4,5}$ EDGES IN THE ENERGY-LOSS SPECTRUM. 96

FIGURE 4.7: SCHEMATICS OF THE BILAYER QD SAMPLES STUDIED IN THIS WORK. 97

FIGURE 4.8: BF TEM AND HAADF STEM IMAGES OF THE SAME REGION IN THE S = 20 NM SAMPLE. SEE THE TEXT FOR THE EXPERIMENTAL DETAILS. 99

FIGURE 4.9: AN EXAMPLE SHOWING THE ROLLING BALL BACKGROUND SUBTRACTION APPLIED TO AN IMAGE. 100

FIGURE 4.10: HAADF STEM IMAGES OF INAS/GAAS QDS CAPPED WITH GAAS AND WITH S = 60 NM, S = 20 NM AND S = 10 NM SAMPLES. 102

FIGURE 4.11: CROSS SECTIONAL HAADF STEM IMAGE OF A SINGLE QD IN THE SAMPLE WITH S = 60 NM AND CAPPED WITH GAAS. 103

FIGURE 4.12: CROSS-SECTIONAL HAADF STEM IMAGE OF BILAYER INAS/GAAS QDS IN THE S = 10 NM SAMPLE CAPPED WITH GAAS. 104

FIGURE 4.13: A SCHEMATIC REPRESENTING THE METHOD USED TO PREPARE CROSS-SECTIONAL SAMPLES ALONG THE $[110]$ AND $[\bar{1}10]$ DIRECTIONS BY CLEAVING GAAS WAFERS ALONG TWO DIFFERENT <110> DIRECTIONS. 105

FIGURE 4.14: HAADF STEM IMAGES OF THE BILAYER QDS IN THE S = 10 NM SAMPLE CAPPED WITH GAAS. IMAGES WERE TAKEN ALONG THE $[110]$ AND $[\bar{1}10]$ ORIENTATIONS. 106

FIGURE 4.15: HAADF STEM IMAGES OF BILAYER QDS IN THE S = 10 NM SAMPLE CAPPED WITH GAAS. IMAGES WERE ALONG THE $[110]$ AND $[\bar{1}10]$ ORIENTATIONS. 107

FIGURE 4.16: HAADF STEM IMAGES OF THE QDS IN THE SAMPLE CAPPED WITH GAAS AND WITH S = 60 NM VIEWED ALONG THE $[110]$ AND $[\bar{1}10]$ ORIENTATIONS. 108

FIGURE 4.17: HAADF STEM IMAGES OF THE QDS IN THE SAMPLE CAPPED WITH GAAS AND WITH S = 60 NM VIEWED ALONG THE $[110]$ AND $[\bar{1}10]$ ORIENTATIONS. 108

FIGURE 4.18: A HAADF STEM IMAGE OF A QD IN THE SAMPLE WITH S = 60 NM AND CAPPED WITH GAAS ANDTHE INTENSITY PROFILE OF THE SECTION ALONG THE GROWTH DIRECTION OF THE QD MARKED. 110

FIGURE 4.19: (A) A SCHEMATIC SHOWING THE ASSUMPTIONS MADE ABOUT THE SHAPE AND THE GEOMETRY OF THE QD IN FIG. 4.18 WITHIN THE THIN FILM SPECIMEN. (B) THE INTENSITY PROFILE OF THE QD DETERMINED AFTER SUBTRACTING THE CONTRIBUTION FROM THE GAAS MATRIX AND THE CALCULATED INTENSITY PROFILE OF A QD WITH THE SAME GEOMETRY BUT WITH A UNIFORM COMPOSITION. 111

FIGURE 4.20: A HAADF STEM IMAGE OF A QD (IN THE FIRST LAYER) IN THE S = 10 NM SAMPLE CAPPED WITH GAAS AND THE INTENSITY PROFILE OF THE SECTION ALONG THE GROWTH DIRECTION OF THE QD MARKED. 112

FIGURE 4.21: (A) A SCHEMATIC SHOWING THE ASSUMPTIONS MADE ABOUT THE SHAPE AND THE GEOMETRY OF THE QD IN FIG. 4.20 WITHIN THE THIN FILM SPECIMEN. (B) THE INTENSITY PROFILE OF THE QD DETERMINED AFTER SUBTRACTING THE CONTRIBUTION FROM THE GAAS MATRIX AND THE CALCULATED INTENSITY PROFILE OF A QD WITH THE SAME GEOMETRY BUT WITH A UNIFORM COMPOSITION. 113

FIGURE 4.22: HAADF STEM IMAGES OF THE QDS IN THE S = 10 NM SAMPLE CAPPED WITH GAAS IN (A) AND 2 NM OF $GA_{0.9}IN_{0.1}AS$ PRIOR TO FURTHER CAPPING WITH GAAS IN (B). (C) THE INTENSITY PROFILE OF THE REGION SELECTED IN (A). (D) THE INTENSITY PROFILE OF THE REGION SELECTED IN (B). 114

FIGURE 5.1: 002 DF TEM IMAGES OF $GA_{0.62}IN_{0.38}N_{0.023}AS_{0.977}$/GAAS QWS GROWN AT (A) 460 °C AND (B) 400 °C. 119

FIGURE 5.2: A 2D SCHEMATIC SHOWING THE OSCILLATORY STRAIN FIELD DUE TO THE NEIGHBOURING IN-RICH AND N-RICH REGIONS IN A SPINODALLY DECOMPOSED GAINNAS QW LAYER. 121

FIGURE 5.3: A SCHEMATIC SHOWING THE GROWTH DETAILS OF THE QW SAMPLES STUDIED IN THIS CHAPTER. 123

FIGURE 5.4: CROSS-SECTIONAL HAADF STEM IMAGES OF $GA_{1-X}IN_XN_{0.05}AS_{0.95}$ QWS ALONG THE <110> ZONE-AXIS. THE SAMPLES WERE GROWN AT (A) AND (C) T_G = 400 °C AND (B) AND (D) T_G = 450 °C AND WITH (A) AND (B) X = 0.15 AND (C) AND (D) X = 0.3. 125

FIGURE 5.5: (A) CROSS-SECTIONAL HAADF STEM IMAGE OF AN $GA_{0.7}IN_{0.3}N0_{.05}AS_{0.95}$ QW TAKEN ALONG THE <110> ZONE-AXIS. (B) THE INTENSITY PROFILE OF THE TOP 2 NM OF THE QW, INDICATED BY THE BLUE REGION IN 5(A). A SINUSOIDAL FUNCTION IS FITTED THROUGH THE DATA TO DETERMINE THE AVERAGE WAVELENGTH OF THE INTERFACE UNDULATION OF THE QWS. 126

FIGURE 5.6: CROSS-SECTIONAL HAADF STEM IMAGES OF 2 PERIODS OF $GA_{0.8}IN_{0.2}N_{0.01}AS_{0.99}$ QWS (TOP TWO LAYERS) AND ONE $GA_{0.8}IN_{0.2}AS$ QW LAYER (BOTTOM LAYER), GROWN AT (A) 400 °C AND (B) 500 °C. THE SPECIMENS WERE ORIENTED ALONG THE <110> ZONE-AXIS. 127

FIGURE 5.7: (A) CROSS-SECTIONAL HAADF STEM IMAGE OF THE $GA_{0.8}IN_{0.2}N_{0.01}AS_{0.99}$ (TOP) AND $GA_{0.8}IN_{0.2}AS$ (BOTTOM) QWS IN THE SAMPLE IN FIG. 5.6(B). (B) THE WAVELENGTH OF THE UNDULATIONS OF THE TOP $GA_{0.8}IN_{0.2}N_{0.01}AS_{0.99}$ QW IS DETERMINED BY FITTING A SINUSOIDAL FUNCTION THROUGH THE DATA. 128

FIGURE 5.8: HAADF STEM IMAGES OF $GA_{1-X}IN_XN_{0.05}AS_{0.95}$ QWS GROWN AT TEMPERATURE T_G, WITH (A) X = 0.15 AND T_G = 400 °C, (B) X = 0.15 AND T_G = 450 °C, (C) X = 0.3 AND T_G = 400 °C AND (D) X = 0.3 AND T_G = 450 °C. THE INTENSITY PROFILES OF THE MIDDLE 2 NM OF THE QWS IN IMAGES (A) – (D) ARE PLOTTED IN (E) – (H) RESPECTIVELY AND THAT OF THE BACKGROUND IN (I). 130

FIGURE 5.9: (A) HAADF STEM IMAGE OF $GA_{0.7}IN_{0.3}N_{0.05}AS_{0.95}$ QW GROWN AT 450 °C. THE SPECIMEN WAS ORIENTED ALONG THE <110> ZONE-AXIS. (B) THE INTENSITY PROFILES OF THE BOTTOM, MIDDLE AND TOP 2 NM OF THE QW, INDICATIVE OF THE LATERAL IN COMPOSITION MODULATIONS AND THE UNDULATING INTERFACE MORPHOLOGY OF THE QW. (C) THE WAVELENGTH OF THE INTENSITY PROFILE OF THE MIDDLE 2 NM OF THE QW FOUND BY FITTING A SINUSOIDAL FUNCTION THROUGH THE DATA. 132

FIGURE 5.10: (A) AND (B) HAADF STEM IMAGE OF $GA_{0.7}IN_{0.3}N_{0.05}AS_{0.95}$ QW GROWN AT 450 °C. THE SPECIMEN WAS ORIENTED ALONG THE <110> ORIENTATION. (C) AND (D) THE INTENSITY PROFILES OF THE AREAS SELECTED BY THE BLUE AND ORANGE BOXES IN (A) RESPECTIVELY. 133

FIGURE 5.11: HAADF STEM IMAGES OF $GA_{0.8}IN_{0.2}N_{0.01}AS_{0.99}$ QWS (TOP LAYER QWS) AND $GA_{0.8}IN_{0.2}AS$ QWS (BOTTOM LAYER QWS) GROWN AT (A) 400 °C AND (B) AND 500 °C. 134

FIGURE 5.12: (A) AND (B) HAADF STEM IMAGE OF $GA_{0.8}IN_{0.2}N_{0.01}AS_{0.99}$ QW GROWN AT 500 °C. (C) AND (D) THE INTENSITY PROFILES OF THE REGIONS SELECTED BY THE BLUE AND ORANGE BOXES IN (A) RESPECTIVELY. 135

FIGURE 5.13: (A) AN EELS SPECTRUM IMAGE OF THE AREA SHOWN IN THE HAADF IMAGE OF THE $GA_{0.7}IN_{0.3}N_{0.05}AS_{0.95}$ QW WITH T_G = 450 °C. (B) THE BACKGROUND SUBTRACTED IN $M_{4,5}$ EDGES FROM EACH PIXEL OF THE SPECTRUM IMAGE. (C) THE AREA BOUND BY THE IN $M_{4,5}$ SIGNAL AND THE E-AXIS IN (B) FOUND FOR THE SPECTRA IN THE FIRST AND SECOND COLUMNS OF THE SPECTRUM IMAGE PLOTTED AS A FUNCTION OF THE PIXEL POSITION OF THE SPECTRA ALONG THE LENGTH OF THE QW. 137

FIGURE 5.14: (A) A SCHEMATIC REPRESENTING THE COMPOSITION MODULATION OF THE QWS, FURTHER DEVELOPING INTO THE FORMATION OF SURFACE UNDULATIONS. (B) SCHEMATIC REPRESENTATIONS OF THE 1D OSCILLATORY STRAIN FIELD OF THE QWS DURING THEIR 2D GROWTH MODE. 140

FIGURE 5.15: SIMULATED ENERGY-LOSS SPECTRA OF GAAS NEAR A PLANAR INTERFACE WITH INAS FOR DIFFERENT ELECTRON BEAM DISTANCES FROM THE INTERFACE *X*, USING EQN. 3.15. AN ACCELERATING VOLTAGE OF 300 KV AND SPECTROMETER COLLECTION SEMI-ANGLE OF 10 MRAD WERE USED FOR THE SIMULATIONS. 143

FIGURE 5.16: SIMULATED ENERGY-LOSS SPECTRA OF GAAS NEAR A PLANAR INTERFACE WITH $GA_{0.7}IN_{0.3}AS$ FOR DIFFERENT ELECTRON BEAM DISTANCES FROM THE INTERFACE *X*, USING EQN. 3.15. AN ACCELERATING VOLTAGE OF 300 KV AND SPECTROMETER COLLECTION SEMI-ANGLE OF 10 MRAD WERE USED FOR THE SIMULATIONS. 143

FIGURE 5.17: (A) AN EELS LINE SPECTRUM ACROSS THE REGION MARKED IN THE HAADF IMAGE OF A $GA_{0.7}IN_{0.3}N_{0.05}AS_{0.95}$ QW WITH T_G = 400 °C. EEL SPECTRA WERE COLLECTED AT ~ 1.3 NM INTERVALS. (B) THE PLASMON REGION OF THE EEL SPECTRA FROM EACH COLLECTION POINT. (C) THE PLASMON PEAK ENERGY AND THE HAADF IMAGE INTENSITY VALUES OF EACH POINT OF THE LINE SPECTRUM AS A FUNCTION OF THE DISTANCE FROM THE GAAS – QW INTERFACE. 145

FIGURE 6.1: EXAMPLES OF ZLPS OBTAINED WITH A SCHOTTKY FEG, COLD FEG AND A MONOCHROMATED SCHOTTKY FEG. THE FULL-WIDTH AT HALF-MAXIMUM OF THE ZLP RECORDED IN VACUUM IS A MEASURE OF THE ENERGY-RESOLUTION IN TEM. 149

FIGURE 6.2: EXAMPLES OF SPECTRA COMPARING THE FULL-WIDTH AT HALF-MAXIMUM (FWHM) OF THE ZLPS COLLECTED IN VACUUM TO THAT COLLECTED FROM THE SPECIMEN (BOTH RECORDED IN MONOCHROMATED - STEM MODE, 80 KV ACCELERATING VOLTAGE, 9 MRAD PROBE CONVERGENCE SEMI-ANGLE AND 10 MRAD SPECTROMETER COLLECTION SEMI-ANGLE ON THE FEI TITAN). THE ZLP COLLECTED FROM THE SPECIMEN BROADENS DUE TO PHONON SCATTERING AND OTHER ENERGY-LOSS EVENTS BELOW THE INSTRUMENTAL RESOLUTION. 152

FIGURE 6.3: A GRAPH SHOWING THE MAXIMUM REFRACTIVE INDEX FOR WHICH NO ČERENKOV LOSSES OCCUR AS A FUNCTION OF THE ELECTRON BEAM ACCELERATING VOLTAGE. 153

FIGURE 6.4: THE REFRACTIVE INDEX OF GAAS $N(E)$, TAKEN FROM OPTICAL MEASUREMENTS IN [127], AS A FUNCTION OF ENERGY-LOSS. THE CONDITION FOR THE EMISSION OF ČERENKOV RADIATIONS GIVEN BY EQN. 6.1 IS SATISFIED FOR ENERGIES AT WHICH $N(E) > 1.99$ AT 80 KV AND $N(E) > 1.30$ AT 300 KV. 154

FIGURE 6.5: THE ANGULAR DEPENDENCE OF THE SCATTERING CROSS-SECTION FOR A 100 NM THICK GAAS SPECIMEN, 300 KV ACCELERATING VOLTAGE AND 1.4 EV ENERGY-LOSS CALCULATED USING THE NON-RELATIVISTIC AND RELATIVISTIC EXPRESSIONS. 157

FIGURE 6.6: THE ANGULAR DEPENDENCE OF SCATTERING AT DIFFERENT ENERGY-LOSSES FOR A 100 NM THICK GAAS SPECIMEN AND 300 KV ACCELERATING VOLTAGE CALCULATED USING KRÖGER'S EQUATION. 157

FIGURE 6.7: SIMULATED VEEL SPECTRA FOR A 100 NM THICK GAAS SPECIMEN, 8 MRAD COLLECTION SEMI-ANGLE AND DIFFERENT ACCELERATING VOLTAGES. THE PEAK DUE TO ČERENKOV RADIATIONS BECOMES MORE INTENSE WITH INCREASING THE ACCELERATING VOLTAGE. 159

FIGURE 6.8: SIMULATED VEEL SPECTRA FOR A 100 NM GAAS SPECIMEN, 300 KV INCIDENT BEAM ENERGY AND DIFFERENT SPECTROMETER COLLECTION SEMI-ANGLES. THE PEAKS DUE TO RELATIVISTIC LOSSES FLATTEN OUT AT LARGER COLLECTION ANGLES. 160

FIGURE 6.9: THE SIMULATED VEEL SPECTRA FOR DIFFERENT GAAS SPECIMEN THICKNESSES, 300 KV INCIDENT ELECTRON BEAM AND 10 MRAD COLLECTION SEMI-ANGLE. 161

FIGURE 6.10: THE SIMULATED VEEL SPECTRA FOR DIFFERENT GAAS SPECIMEN THICKNESSES, 300 KV INCIDENT ELECTRON BEAM AND 4 MRAD COLLECTION SEMI-ANGLE. 161

FIGURE 6.11: THE RESULT OF THE OFF-LINE CORRECTION METHOD, SHOWING THE RECORDED SPECTRUM, S_{ORIG}, THE OUTPUT OF KRÖGER'S EQUATION IN THE FINAL ITERATIVE STEP, $S_{KRÖGER}$ AND THE RESULT OF THE OFF-LINE CORRECTION PROCESS, CALLED CORRECTED. 163

FIGURE 6.12: VEEL SPECTRUM OF GAAS ACQUIRED USING THE FEI TITAN IN MONOCHROMATED TEM MODE WITH 80 KV INCIDENT ELECTRON BEAM AND 5 MRAD COLLECTION SEMI-ANGLE. 165

FIGURE 6.13: THE ENERGY-LOSS FUNCTION OF GAAS OBTAINED BY APPLYING THE KRAMERS-KRONIG ANALYSIS TO A VEEL SPECTRUM AND THAT OBTAINED FROM OPTICAL MEASUREMENTS. 166

FIGURE 6.14: THE FIRST 8 EV OF THE ENERGY-LOSS FUNCTION OF GAAS OBTAINED BY VEELS AND OPTICAL MEASUREMENTS, SHOWING THE FINE STRUCTURE DUE TO VALENCE TO CONDUCTION ELECTRON BAND TRANSITIONS, INCLUDING THE BAND GAP TRANSITION. 166

FIGURE 6.15: THE THEORETICAL BAND STRUCTURE OF GAAS SHOWING THE MAIN CRITICAL POINTS AND THEIR CORRESPONDING ENERGIES AT ROOM TEMPERATURE. 167

FIGURE 6.16: THE REAL PART OF THE DIELECTRIC FUNCTION OBTAINED FROM VEELS COMPARED TO THAT OBTAINED FROM OPTICAL MEASUREMENTS. 169

FIGURE 6.17: THE IMAGINARY PART OF THE DIELECTRIC FUNCTION OBTAINED FROM VEELS COMPARED TO THAT OBTAINED FROM OPTICAL MEASUREMENTS. THE BAND TRANSITIONS CORRESPONDING TO THE PEAKS IN THE SPECTRUM AND THEIR ENERGIES ARE IDENTIFIED IN THE TABLE. 169

FIGURE 6.18: (A) HAADF IMAGE AND (B) A SERIES OF SSDS OF VEEL SPECTRA OBTAINED ACROSS A $GA_{0.7}IN_{0.3}N_{0.05}AS_{0.95}$/GAAS QW. THE VEEL SPECTRA WERE ACQUIRED AT 6 NM INTERVALS. THE SIGNAL TO NOISE RATIO IS NOT SUFFICIENT TO IDENTIFY THE PEAK DUE TO THE E_0 TRANSITION IN THE SPECTRA. HOWEVER, A SHIFT CAN BE SEEN IN THE ONSET OF THE SPECTRUM COLLECTED FROM THE QW. 171

FIGURE 6.19: (A) THE HAADF IMAGE AND (B) THE VEEL SPECTRA OF FIG. 6.18 (PRIOR TO ZLP REMOVAL), SHOWING THE REDSHIFT IN THE ONSET OF THE E'_1 BAND TRANSITION OF THE SPECTRUM COLLECTED FROM THE QW. 171

1 Introduction

1.1 Basics

Inorganic solids can be categorised into three groups based on their electrical conductivity: metals, semiconductors and insulators [1]. Pure semiconductors behave as insulators at 0 K temperature but exhibit low levels of electrical conductivity (significantly lower than that of metals) at room temperature [2]. The number density of mobile electrons in semiconductors and hence, their electrical conductivity can be altered by conditions such as temperature, presence of impurities and illumination [1]. This explicit control over the conductivity of semiconductors has made them the basis of modern electronic devices.

The physical principle behind the electrical conductivity of inorganic solids is explained in terms of the band structure theory. A complete description of the band structure theory is beyond the scope of this work but can be found, for example in [2]. Continua of energy levels, known as energy bands, are formed by the overlap of individual atomic orbitals in solids [1, 2]. Fig. 1.1 shows how the energy separations between energy states become progressively smaller as more atoms aggregate to form a solid. The highest occupied energy level at 0 K temperature is called the Fermi level [2]. The energy bands below and above the Fermi level are known as the valence and conduction bands respectively. Fig. 1.2 shows a schematic representation of the energy bands of metals, semiconductors and insulators at 0 K temperature. The valence band in metals is partially filled, giving rise to their high electrical

conductivity (having partially filled bands is a necessary condition for conduction) [1]. The valence band in insulators is full and is separated from the conduction band by an energy region in which no electronic states exist, known as the band gap [1]. Semiconductors have a band structure similar to that of insulators but have a smaller band gap (typically less than ~ 3 eV).

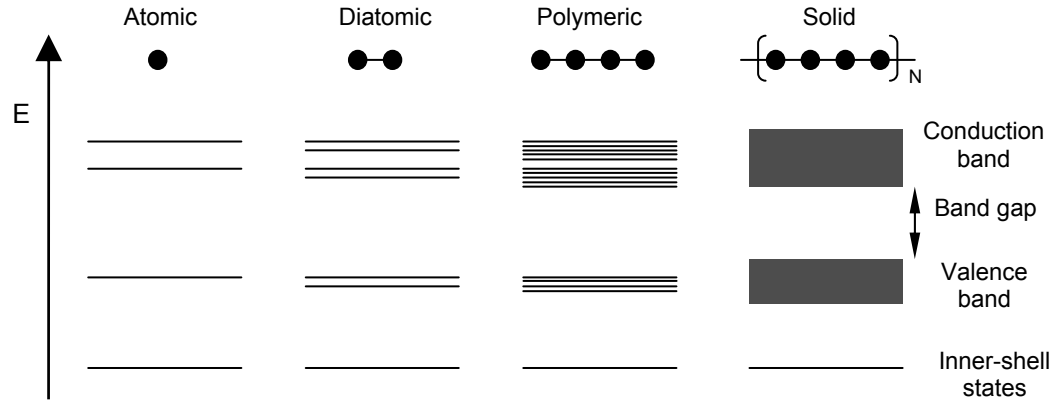


Figure 1.1: Schematic of the electronic states for atomic, diatomic, polymeric and solid structures (E stands for energy). The separation of energy level decreases progressively, until overlapping bands are formed in solids (adapted from [1])

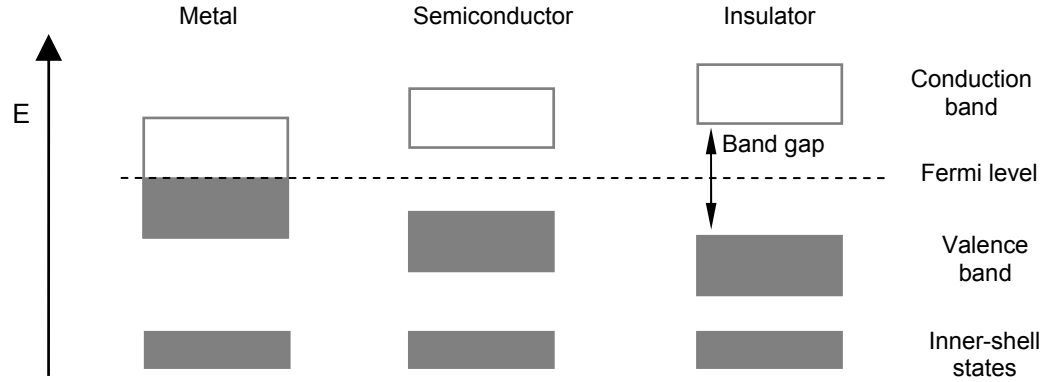


Figure 1.2: A simplified schematic of the band structure of metals, semiconductors and insulators at 0 K temperature (adapted from [2]). The vertical axis represents energy (E stands for energy) and colours grey and white represent occupied and unoccupied bands.

Increasing the thermal energy of electrons in semiconductors can promote them from the valence to the conduction band and thereby results in an increase in their electrical

conductivity [2]. Moreover, the conductivity of semiconductors can be increased by adding small amounts of impurity materials with a different number of valence electrons compared to that of the host semiconductor material [2]. This process is known as doping. Doping materials are typically elements from the groups either to the left or right of the host material in the periodic table. The doping material enhances the conductivity of the host material by introducing deficiency of electrons in the valence band (in p-type semiconductors) or excess of electrons in the conduction band (in n-type semiconductors). Deficiency of valence electrons in semiconductors is generally considered in terms of the presence of positively charged particles, known as holes. Electrons and holes are collectively referred to as carriers. The electrical conductivity of semiconductors may also be increased upon illumination with visible light. This is the principle behind the operation of optoelectronic devices. This concept is explained in the next section.

1.1.1 Light emitting semiconductor structures

Semiconductors are capable of producing an electric current upon interaction with visible light or emitting visible light in response to an electric current. This property of semiconductors is exploited in optoelectronic devices such as light emitting diodes. Fig. 1.3 shows the energy band transitions occurring to the electrons in a semiconductor material as the result of (a) optical absorption and (b) and (c) emission. When illuminated with photons of energy $E = E_c - E_v$, where E_v and E_c are the valence and conduction band energies, valence electrons can make a transition to the conduction band by absorbing energy E (absorption). Conversely, electrons in the conduction band may emit photons of energy $E = E_c - E_v$ upon making a transition to the valence band (spontaneous emission). An additional form of emission exists in which incident photons of energy $E = E_c - E_v$ induce conduction to valence band electron transitions [3]. Since the original incident photons are not absorbed, twice the number of incident photons emerge from the material. This effect, known as stimulated emission, is the basis of light amplification in optoelectronic devices [3].

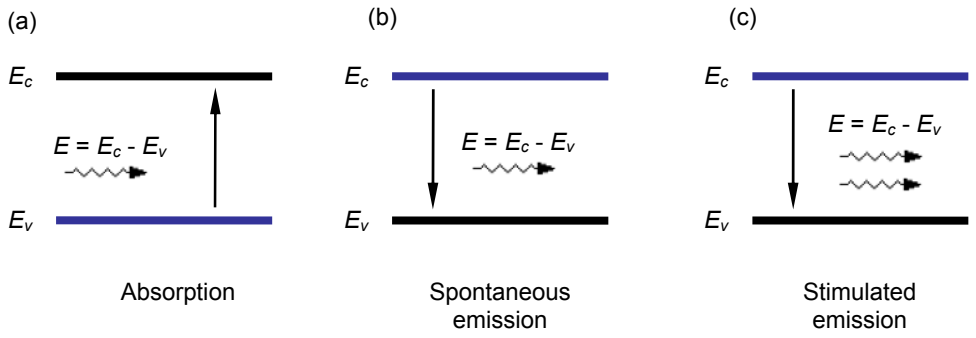


Figure 1.3: A schematic representation of optical absorption, spontaneous emission and stimulated emission in semiconductors (adapted from [3]). The arrows show the transmission undergone by the electrons in the valence and conduction bands in each case.

Efficient photon absorption or emission requires no change in the momentum of the initial and final electronic states [3]. This condition is satisfied in direct band gap semiconductors in which the maxima of the valence band and the minima of the conduction band in the band structure have the same wave vectors associated with them (see Fig. 1.4) [2]. GaAs and InAs are examples of direct band gap semiconductors and have thus been widely used for optoelectronic applications.

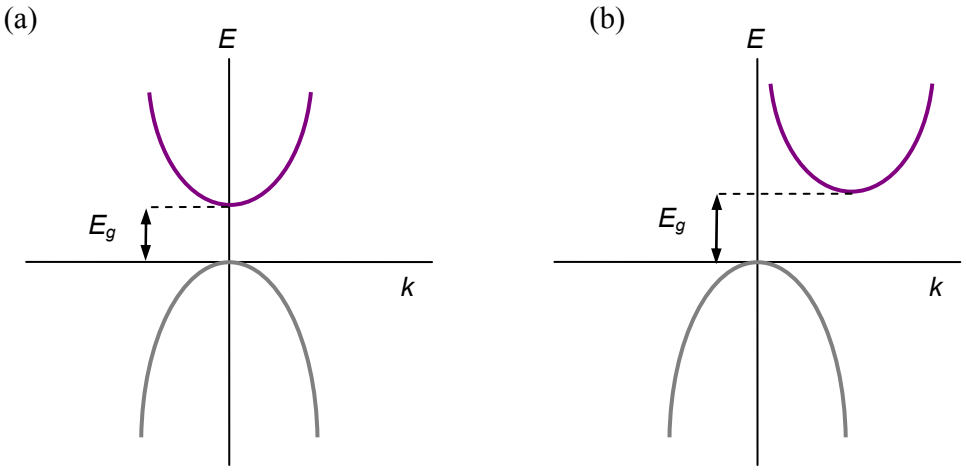


Figure 1.4: Schematics showing examples of the band structure of (a) direct and (b) indirect band gap semiconductors (adapted from [2]). Valence and conduction bands are displayed by colours grey and purple respectively. E, k and E_g represent energy, wave vector and band gap respectively.

In bulk materials it is generally assumed that the electrons are free to move in a three-dimensional space and hence, the energy bands of bulk materials are continua [4].

However, assuming constraints in the movement of electrons in one, two or three dimensions in structures known as quantum wells, quantum wires and quantum dots respectively, modifies the band structure by resulting in quantisation of the energy bands [4]. The effect of this quantisation on the density of states of the energy bands of quantum wells, wires and dots is shown in Fig. 1.5. Quantum effects become significant in semiconductors if their dimensions are in the range ~ 10 – 100 nm [4]. Therefore, structures featuring one dimension of tens of nanometres in length behave as quantum wells, those with two tens of nanometres-length dimensions as quantum wires and those with three tens of nanometres-length dimensions as quantum dots.

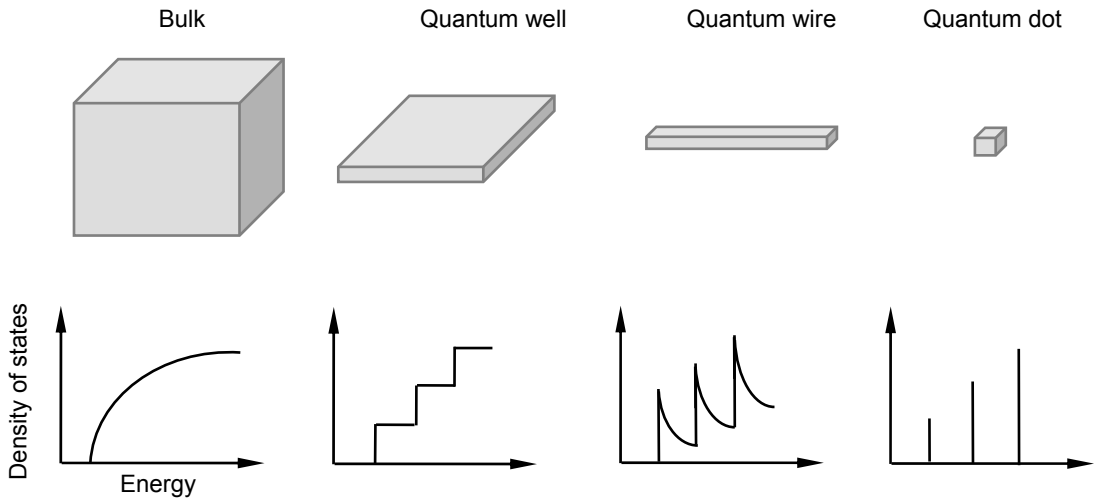


Figure 1.5: A schematic showing how the density of states of an energy band in a semiconductor is altered as the system's dimensions are reduced (adapted from [5]).

The probability of light emission in bulk materials at equilibrium conditions is negligibly low. However, the optical output of semiconductors can be considerably improved by creating non-equilibrium conditions to increase the probability of conduction to valence electron transitions. For example, in light-emitting diodes (in its simplest form is a junction formed between a p-type and a n-type semiconductor) this is achieved by flooding the p-side of the junction with electrons and the n-side of the junction with holes (*i.e.* forward bias of the junction in which the p- and n-sides of the are connected to positive and negative potentials respectively) [3].

Optical efficiency can be further improved by taking advantage of the band structure modifications introduced in quantum structures. Firstly, when buried in a dissimilar bulk material, the difference between the refractive indices of the two materials can be used to provide confinement of the output light (by total internal reflection) and hence, reduce the scattering of the emitted light. Secondly, the confinement of electrons and holes in quantum structures increases the optical output of the system by reducing the probability of their non-radiative energy-loss (for example by phonon scattering (see Ch. 2 section 2.2)). In addition to this, the difference in the band gap of the quantum structure and the surrounding material may be used to provide different potential barriers to the motion of electrons and holes and cause them to be confined in different regions of space. This further improves the probability of light emission [3]. Thirdly, the quantised energy bands of quantum structures can be utilised in devices where specific emission wavelengths are required. Due to these reasons, semiconductor quantum structures are of great interest for optoelectronic applications.

It is important to recognise that semiconductor structures can only operate to their full potential if they are grown with the highest degree of crystallinity and precision. In general, ultra high vacuum (UHV) conditions and slow growth rates have to be adopted in epitaxial crystal growth (monolayer by monolayer crystal growth) in order to achieve accurate control over the composition and dimensions of structures. An additional advantage of the UHV growth condition is the possibility of utilising techniques such as reflection high-energy electron diffraction (RHEED) or scanning tunnelling microscopy (STM) to monitor the growth process. The molecular beam epitaxy (MBE) technique employed for the growth of samples studied in this work, is one such high precision epitaxial growth technique. A brief description of MBE is provided in the next section.

1.1.2 Epitaxial crystal growth: molecular beam epitaxy (MBE)

MBE has proved to be a versatile tool in fabrication of low dimensional structures due to its high precision control over growth parameters such as film thickness and composition. A UHV condition (~ 10^{-11} Torr) is employed in MBE, in which beams of evaporated atoms or molecules of source materials are directed towards a heated

substrate to impinge upon [5]. A necessary condition for MBE growth is that the mean-free path (average distance between collisions) of the particles in the beams is larger than the dimensions of the growth chamber. The UHV condition ensures of this by providing an almost collision-free path for the atomic or molecular beams. Molecular beams are produced by heating the elemental sources contained in effusion cells in the growth chamber. A growth rate of a few Angstroms per second is typically used in MBE [5]. This slow growth rate and the raised temperature of the substrate allow the incident adatoms (adsorbed atoms) to diffuse across the surface of the substrate until they are incorporated into the crystal. It is essential to ensure a uniform composition of the deposited material on the substrate. The compositional uniformity of growth can be improved through a close control over the flux of the molecular beams and also by rotation of the substrate during growth.

The system employed for the growth of the samples studied in this work is essentially an MBE growth chamber combined with a STM unit. Fig. 1.6 is a schematic showing the main components of this system. The STM unit is used for *in-situ* imaging of samples. Samples are entered into and removed from the system via an isolated fast entry chamber (FEC) unit in order to prevent any possible contamination of the MBE and STM units. The MBE, STM and FEC units are independently pumped to the required vacuum level and can be isolated from each other by the gate valves between them. Samples are transferred between different units using magnetically controlled transfer arms.

The MBE chamber (DCA Instrument, Finland) contains three effusion cells (providing materials for growth), a sample stage and a RHEED apparatus (see Fig. 1.6). Group III metals are kept in boron nitride crucibles which are then heated to produce molecular beams. The As source contains solid As which is turned into As_4 upon heating and then cracked (by hot graphite) to produce an As_2 molecular beam. Magnetically controlled shutters are positioned in front of all effusion cells in order to block the beams when required. There is also an additional shutter that can be placed in front of the sample stage. In the case of the growth of the dilute nitride samples, active N species are provided by a radio frequency (RF) plasma source. Ultra-pure $N_{2(gas)}$ (99.9999% purity) is used as the source material. During the ignition and extinguishing of the plasma source, the sample is protected from nitridation by

closing the shutter placed in front of the sample stage. The sample stage is fitted with heater and manipulator systems. The stage manipulator controls the movement of the stage in the x-, y- and z-directions and its rotation around the z-direction (perpendicular to the plane of the stage). The temperature of the sample is measured by an infrared pyrometer (determining the temperature of the sample by measuring its infrared emission), positioned opposite the stage.

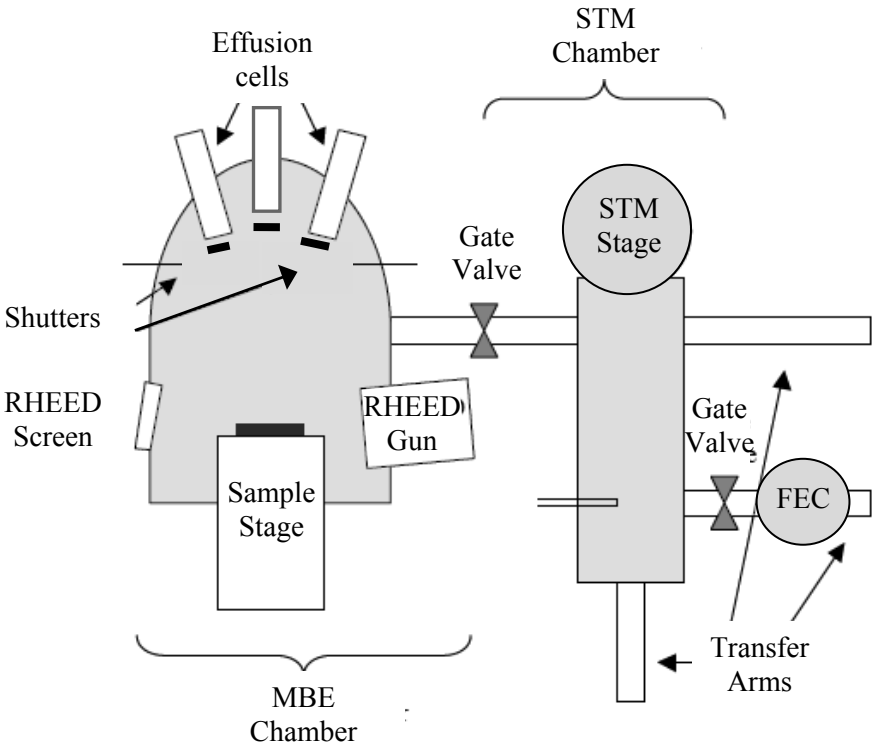


Figure 1.6: A schematic of the MBE growth system, taken from [6].

1.1.3 Bulk GaAs and InAs crystallography

GaAs and InAs are examples of III – V semiconductors (comprised of group III and group V elements). They adopt a zinc blende crystal structure in which each atom is bonded to adjacent atoms of the other group in a tetrahedral arrangement (see Fig. 1.7). Zinc Blende crystal structures can be considered as interpenetrating group III and group V face-centred cubic (fcc) structures, displaced from each other along the <111> direction by a quarter of the lattice constant. The lattice constants of GaAs and InAs are 5.65 Å and 6.06 Å respectively. The thickness of one monolayer along the (001) plane of GaAs and InAs is $a/2$, where a is the lattice constant of the crystal.

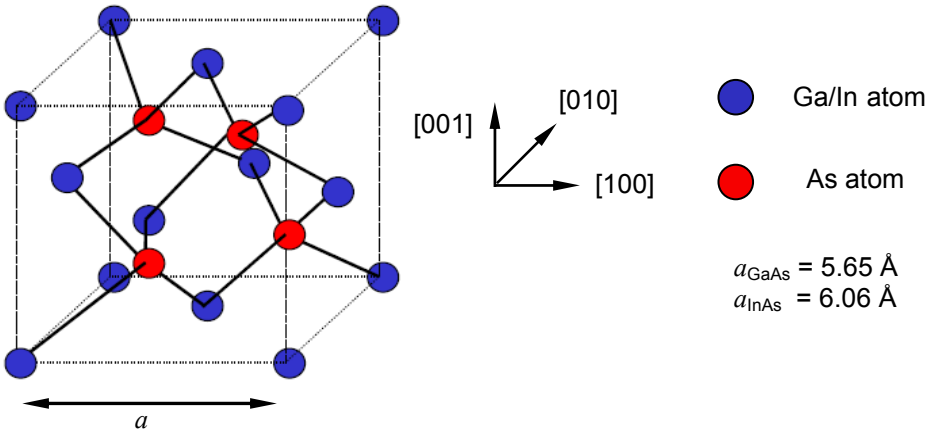


Figure 1.7: A schematic of the zinc blende crystal structure of unit cells of GaAs and InAs (adapted from [5]). The lattice constants of GaAs and InAs are 5.65 Å and 6.06 Å respectively.

1.1.4 Modes of film growth

Fig. 1.8 shows some of the atomic processes that occur during the initial stages of film growth on a solid substrate [7]. During growth, particles can be adsorbed onto, diffuse along or be desorbed by the surface of the substrate. Moreover, more than one adsorbed particles may nucleate and form 3D islands. It is also possible for the particles within an existing island to migrate away from it. Another important process in thin film growth is the exchange of particles between the substrate and the film, known as interdiffusion. It is important to note that film growth is a non-equilibrium

process and the final outcome of the growth is determined by the kinetics of each of the processes involved [7].

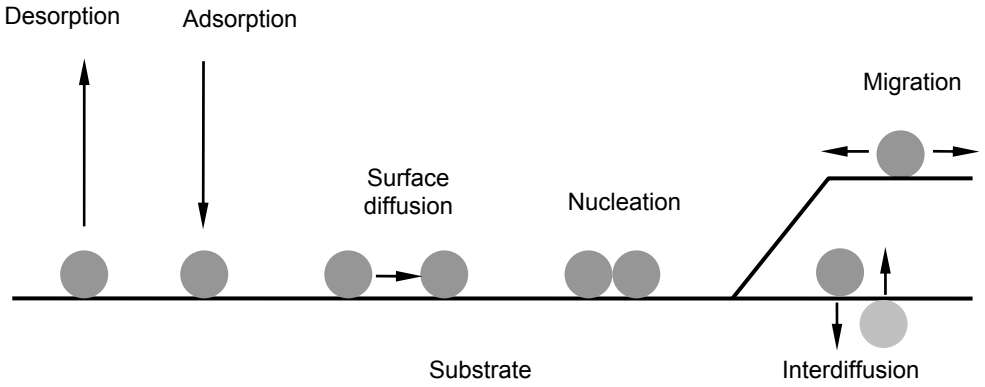


Figure 1.8: A schematic illustrating some of the processes involved in thin film growth on a solid substrate (adapted from [7]).

Viewing the process of film growth phenomenologically, three distinctly different growth modes can be identified. These growth modes, as displayed in Fig. 1.9, are: Frank-van der Merve (or layer-by-layer), Vollmer-Weber (or island) and Stranski-Krastanov (or layer plus island) growth modes [7]. In the Frank-van der Merve growth mode, the interaction between the overlayer particles and the substrate particles is stronger than that between neighbouring overlayer particles and hence, a layer-by-layer growth is adopted. In comparison, in the Vollmer-Weber growth mode a stronger interaction between neighbouring overlayer particles exists, leading to conglomeration of adsorbed particles and island growth. Stranski-Krastanov growth mode is an intermediate case, in which after an initial layer-by-layer growth a transition to island growth occurs. Factors such as a lattice mismatch (dissimilar lattice parameters) between the substrate and the overgrown film may contribute to this growth mode transition.

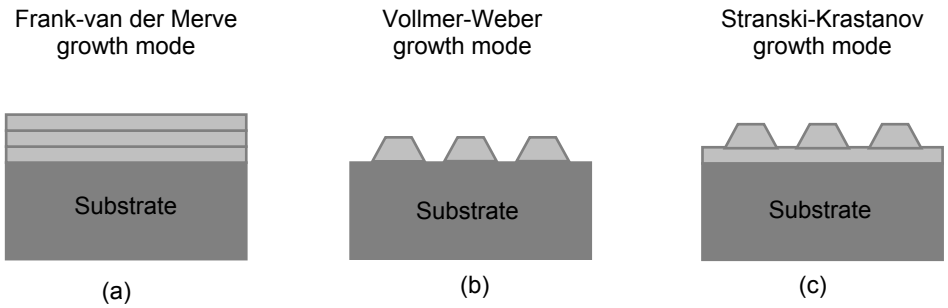


Figure 1.9: Schematic representations of (a) Frank-van der Merve (or layer-by-layer), (b) Vollmer-Weber (or island) and (c) Stranski-Krastanov (or layer plus island) film growth modes (adapted from [7]).

Different film growth modes can be considered in terms of surface (or interface) tension γ of the substrate and the film, defined as the characteristic free energy per unit area required for extending the surface area [7]. Surface tension can also be regarded as a force per unit length of a boundary, in which case the surface tensions acting on a 3D island grown on a substrate can be described by that of the substrate/vacuum interface γ_S, the film/vacuum interface γ_F and the substrate/film interface $\gamma_{S/F}$ (see Fig. 1.10). Force equilibrium requires [7]:

$$\gamma_S = \gamma_{S/F} + \gamma_F \cos\Phi \qquad \textbf{Equation 1.1}$$

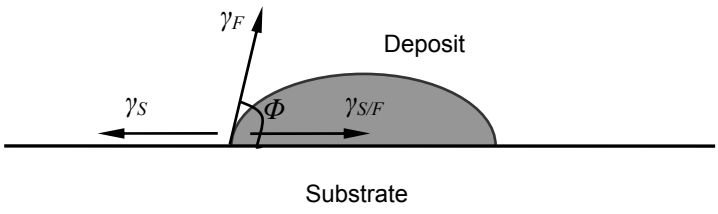


Figure 1.10: A simplified description of the surface/interface tensions acting on a 3D island grown on a substrate (adapted from [7]). γ_S, γ_F and $\gamma_{S/F}$ are the substrate/vacuum, film/vacuum and substrate/vacuum interface tensions.

Therefore, conditions for layer-by-layer and island growth modes are given by [7]:

(I) Frank-van der Merve growth: $\Phi = 0$, $\gamma_S \geq \gamma_F + \gamma_{S/F}$

(II) Vollmer-Weber growth: $\Phi > 0$, $\gamma_S < \gamma_F + \gamma_{S/F}$

The growth mode transition in a layer plus island growth can be explained in terms of the contribution of strain energy to the surface energies of the system. Initially the system favours a layer-by-layer growth. However, the increasing strain energy of the overgrown layer with continual growth (for example due to lattice mismatch between film and substrate) leads to increase in $\gamma_{S/F}$ until at a critical film thickness θ_{crit} the sum $\gamma_F + \gamma_{S/F}$ exceeds γ_S and an island growth mode is adopted instead [7]. The initial 2D layer in Stranski-Krastanov growth is known as the wetting layer. The wetting layer is said to grow pseudomorphically, meaning that the overlayer film adopts the same lattice parameter as that of the substrate.

1.2 Motivation

Optical communication devices such as laser diodes and light detectors, exploit both the electronic and the optical properties of semiconductor materials to convert information into optical signals and vice versa. These optical signals are transmitted through glass fibres. One of the most attractive features of glass fibre transmission is its considerably low propagation loss [3]. Minimum attenuation of light by glass is obtained in the 1.3 μm – 1.55 μm range [3] and hence, in recent years there has been a significant effort in producing semiconductor devices capable of emission in this wavelength range. InAs/GaAs quantum dots and GaInNAs/GaAs quantum wells are two such material systems. They are briefly introduced in the next section.

1.2.1 InAs /GaAs quantum dots

Quantum dot-based light emitting devices offer important advantages over quantum well devices, due to their 3D carrier confinement. For example, they exhibit a weaker temperature sensitivity in their optical performance compared to quantum well devices [8]. Therefore, in recent years there has been a great interest in producing quantum dot devices capable of optical emission in the 1.3 μm – 1.55 μm wavelength range, such as InAs/GaAs quantum dots. The high level of control over composition and dimensions in current epitaxial growth techniques allows the overgrowth of a

crystalline substrate with a different crystalline material. This process is known as heteroepitaxy [7]. Heteroepitaxial growth of InAs on a (001) GaAs substrate under most growth conditions follows a Stranski-Krastanov condition [9]. The dissimilarity between the lattice constants of InAs and GaAs (6.06 Å and 5.65 Å respectively) induces compressive elastic strain in the 2D grown InAs layer and at a critical film thickness θ_{crit} causes a switch to 3D island growth. These 3D islands have typical dimensions of tens of nanometres and hence, exhibit quantum dot-like electronic properties.

It is needless to say that the optical properties of InAs/GaAs quantum dots are largely determined by their structural and compositional properties. Although a substantial volume of work already exists on the morphology and composition of the quantum dots prior to their overgrowth (for example with GaAs), obtaining similar information on overgrown quantum dots has been more challenging. The overgrowth step (*i.e.* capping) of the quantum dots is a necessary stage of any device fabrication and thus, it is of vital importance to have a good understanding of the morphology and composition of capped structures. For instance, it is of interest to investigate the effect of different capping materials on the structure and composition of the quantum dots. Another important aspect of growth of InAs/GaAs quantum dots is the possibility of forming vertically aligned stacks of 3D islands in multilayers, separated by spacer layers of a larger band gap material (for instance GaAs) [*e.g.* 10]. Vertically aligned InAs/GaAs quantum dots in multilayers are of interest for producing structures with higher quantum dot number density and coupled optical and electronic properties. Hence, investigating the morphology and composition of the quantum dots in these structures is necessary for successfully fabricating devices of desired properties. The morphological and compositional properties of single layer and vertically aligned bilayers of InAs/GaAs quantum dots were studied in this work using analytical transmission electron microscopy. The role of the capping material on the morphology and composition of the quantum dots was also examined by comparing structures grown under identical conditions but overgrown with different materials.

1.2.2 GaInNAs / GaAs quantum wells

Addition of small amounts of N to III/V semiconductor materials such as GaAs and GaInAs (*i.e.* dilute nitride GaNAs or GaInNAs) causes significant changes in their electronic properties. Such changes include considerable reduction in band gap [*e.g.* 11, 12, 13] and increase in electron effective mass [*e.g.* 14]. The band gap reduction in dilute nitride III/V semiconductors diverges from Vegard's law, in which the band gap of the resulting alloy varies linearly with the composition of the constituent elements [*e.g.* 15]. Instead, the band gap of dilute nitrides exhibits a second-order relationship with N concentration. This effect is known as band gap bowing [12]. This makes GaInNAs/GaAs quantum wells of great interest for 1.3 μm – 1.55 μm emission wavelength applications.

The unusual modifications to the electronic properties of dilute nitride materials have been ascribed to the large difference in size and electronegativity of As and N [17, 18]. The replacement of As with N atoms in dilute nitride GaNAs or GaInNAs results in the formation of highly localised N-related electronic states, due to the large electronegativity of N [17, 18]. The energy level of these localised states is close to the conduction band of the crystal and as the result of interaction between the two, the conduction band is split into two sub-bands [18]. This effect results in the unusual band gap reduction observed in dilute nitrides.

The band gap as a function of lattice constant of several III/V semiconductors are displayed in Fig. 1.11. In general, a reduction in the band gap of semiconductors is accompanied by an increase in lattice constant. However, in the case of N incorporation in III/V materials, band gap reduction is concomitant with decrease in lattice constant. This implies the possibility of producing lattice matched GaInNAs quantum structures on GaAs substrates. This is of particular importance in device applications since high levels of strain in heterostructures often leads to defects such as dislocations and thereby degradation of the material and optical quality of devices.

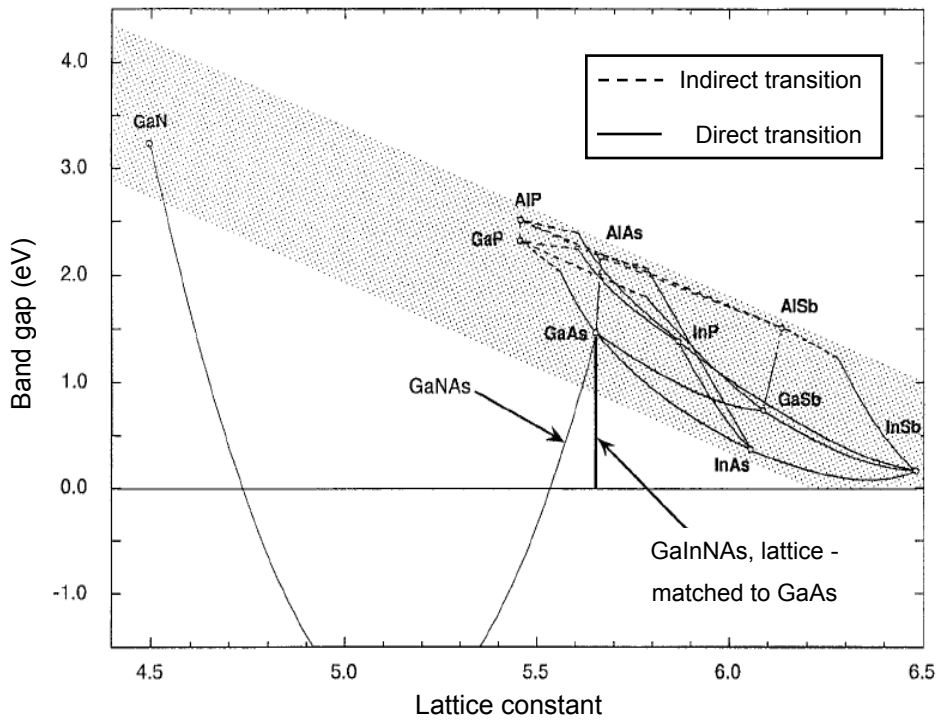


Figure 1.11: Band gap as a function of the lattice constant of III/V semiconductors (taken from [16]).

An important aspect of growth of dilute nitride III/V semiconductors is the highly non-equilibrium conditions required for their growth (due to the metastability of these alloys) [19, 20]. In general, N needs to be provided as highly reactive free radicals (for example by plasma sources in MBE growth) during the growth in order to ensure its incorporation into the deposited film [19]. Successful application of GaInNAs/GaAs quantum wells for 1.3 μm emission wavelength devices has already been demonstrated [21]. Increasing the emission wavelength of these quantum well structures to 1.55 μm requires incorporation of higher levels of N and In (typically larger than 3% for N and 30% for In). However, there are still several issues concerning the morphology and composition of higher In and N content GaInNAs/GaAs quantum wells that need to be addressed before their successful application as light-emitting devices. The morphology and composition of higher N and In content GaInNAs/GaAs quantum wells were investigated in this work using analytical transmission electron microscopy.

1.3 Outline of the thesis

The principles of light emission in semiconductor materials were introduced in this Chapter, followed by an account of the motivation behind the research carried out in this work. The second Chapter in this thesis provides a description of transmission electron microscopy (TEM), the instruments employed and the related techniques used in this work. For instance, the important features of the FEI Titan analytical transmission electron microscope, scanning transmission electron microscopy (STEM) and electron energy-loss spectroscopy (EELS) are described. Furthermore, this Chapter includes an outline of the (S)TEM specimen preparation procedure employed. The theoretical aspects of EELS, used for compositional and electronic structure analyses carried out in this work, are presented in Chapter 3.

The results of the analytical TEM investigations of InAs/GaAs quantum dots and GaInNAs/GaAs quantum wells are presented in the following three Chapters. Chapter 4 includes the results of the STEM analysis of the structure and composition of the InAs/GaAs quantum dot samples. The effects of the GaAs spacer layer thickness and the capping material are investigated using high angle annular dark field (HAADF) STEM imaging. The advantages of utilising HAADF STEM for imaging highly strained structures such as quantum dots are also discussed in this Chapter. Chapter 5 contains the results of the morphological and compositional studies of GaInNAs/GaAs quantum wells. The effects of growth parameters such as growth temperature, In content and N content on the morphology of the quantum wells are investigated by STEM imaging. This is followed by the results of STEM and EELS analyses of possible In composition inhomogeneities in these structures. Finally, the validity of EELS plasmon mapping, used in several reports in the literature, for examining the N composition of the quantum wells is verified in this Chapter. In Chapter 6 a discussion on the advantages and challenges of employing monochromated EELS for measuring band gap and other interband transitions is provided. This is followed by the results of monochromated EELS experiments carried out on GaInNAs/GaAs quantum wells. The results of Chapters 4, 5 and 6 are summarised and concluded in Chapter 7, accompanied by an outline of possible future works.

2 Instrumentation and Experimental Details

2.1 Introduction

Less than a century ago, the concept of electron microscopy was proposed to overcome the limited image resolution of optical microscopy. Since then electron microscopy has been shown to be a powerful technique in characterising the micro- and nano-structure of materials with the possibility of obtaining important information on crystallography, local composition and electronic structure. This chapter introduces the basic theory of transmission electron microscopy (TEM), followed by a description of the key components of a TEM instrument and the principles of different TEM imaging techniques. The majority of the results presented were obtained using the FEI Titan microscope. Therefore, a more detailed description of the important components of this instrument is also provided. Producing artefact-free, electron-transparent specimens is a crucial step in obtaining unambiguous results in TEM. The details of the specimen preparation technique used in this work are outlined in the final section before summarising the contents of this chapter.

2.2 Electron-specimen interactions

In TEM an electron beam interacts with a thin specimen and after undergoing several scattering processes the majority of the electrons are transmitted through the specimen. The scattering processes undergone by the electron beam traversing a specimen can be of two distinct types: elastic and inelastic. In elastic scattering the energy of the electrons before and after the scattering event remains unchanged while in inelastic scattering electrons experience an energy-loss. The nature of these elastic and inelastic scattering processes provide the basis of the imaging and analytical techniques available in TEM. Fig. 2.1 is a schematic showing some of the scattering events of an electron beam interacting with a thin crystalline specimen. These scattering events are discussed in section 2.1.

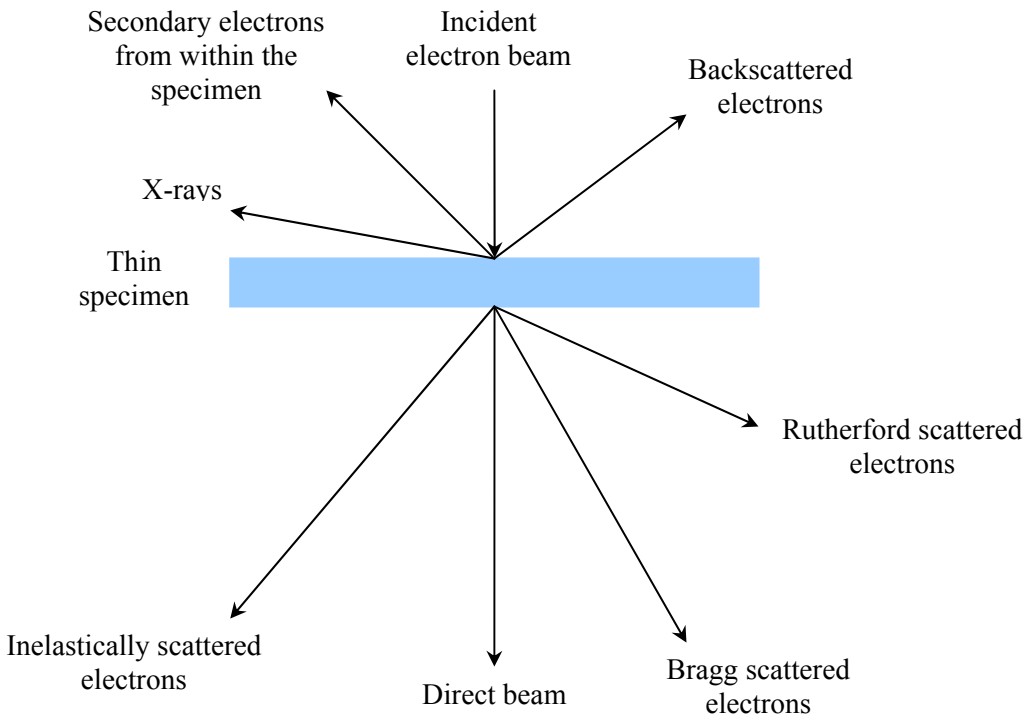


Figure 2.1: A schematic showing some of the important elastic and inelastic scattering events occurring when an electron beam interacts with a thin specimen (adapted from [22])

2.2.1 Elastic scattering

Elastic scattering of the electron beam gives rise to the contrast observed in TEM images. In scattering from an isolated atom, the electron beam can interact either with the atomic nucleus or the electron cloud and is scattered to different angles depending on the type of interaction [22]. The relationship between the angle at which electrons are scattered and the amplitude of scattering, known as the atomic scattering factor *f(θ)* is given by [23]:

$$f(\theta) = \frac{m_0 e^2}{2h^2}\left(\frac{\lambda}{\sin\theta}\right)^2 \left(Z - f_x(\theta)\right) \qquad \textbf{Equation 2.1}$$

Here m_0 is the rest mass of the electron, e the electron charge, h Plank's constant, λ the wavelength of the electron beam, θ the scattering angle, Z the atomic number and $f_x(\theta)$ the atomic scattering factor for x-rays. It is clear in this equation that there are two scattering phenomena associated with scattering from an atom. The part involving the atomic number, Z, signifies scattering by the nucleus, known as Rutherford scattering and the part involving f_x signifies scattering by the electron cloud. Electron-electron interactions have small scattering angles associated with them while Rutherford scattering results in higher scattering angles.

The description given so far has been based on the particle nature of electrons. However, an electron beam can also be viewed as a wave and hence, it can be diffracted by a periodic crystal. When interacting with a crystal, coherent (in phase) diffracted beams are produced in specific directions defined by Bragg's law [22]:

$$n\lambda = 2d\sin\theta_B \qquad \textbf{Equation 2.2}$$

where n is an integer, λ the electron wavelength, d the interplanar spacing and θ_B the angle of the beam with the diffracting crystal planes, known as Bragg's angle. Bragg diffraction is the main cause of contrast observed in TEM images of crystalline materials. In line with the atomic structure factor defined above, the amplitude of scattering from the unit cell of a crystal, known as the structure factor *F(θ)*, is defined as a sum of the atomic scattering factors of all the atoms in the unit cell [22]:

$$F(\theta) = \sum_i f_i e^{\frac{2\pi i}{\lambda}(h x_i + k y_i + l z_i)} \qquad \textbf{Equation 2.3}$$

Here θ is the scattering angle, f_i the atomic structure factor and (x_i, y_i, z_i) the atomic coordinates for each unit cell and (*hkl*) the Miller indices of specific diffracting atomic planes.

It is important to note that diffraction is seldom a single scattering event (assumed in the kinematical theory of diffraction) and there is almost always interaction between the diffracted beams as they pass through the specimen [24]. This concept, known as the dynamical theory of diffraction, means, for example, that a diffracted beam can return to its original direction after undergoing diffraction a second or third time. The probability of a dynamical diffraction increases with increasing specimen thickness.

Bragg diffraction becomes less significant at higher scattering angles since larger θ_B implies smaller interplanar spacings, d (Eqn. 2.2). Thus the smallest thermal vibration of the lattice planes or defects can cause irregularity in d and disrupt the coherent diffraction from the diffracting planes. Scattering at large angles is mainly due to Rutherford scattering. The intensity of Rutherford scattering is strongly dependent on the atomic number Z. Rutherford scattering can be shown to have a probability of scattering to angles greater than θ (scattering cross-section) $\sigma_{ADF}(\theta)$ of [22]:

$$\sigma_{ADF}(\theta) = 1.62x10^{-24}\left(\frac{Z}{E_0}\right)^2 \cot^2\frac{\theta}{2} \qquad \textbf{Equation 2.4}$$

where E_0 is the energy of the electron beam. For a TEM specimen of thickness t this equation can be re-written as [23]:

$$\sigma_{ADF}(\theta) = 1.62x10^{-24}\left(\frac{N\rho t}{A}\right)\left(\frac{Z}{E_0}\right)^2 \cot^2\frac{\theta}{2} \qquad \textbf{Equation 2.5}$$

where N is Avogadro's number, Z the atomic number, A the atomic weight and ρ the density of the specimen. Rutherford scattering is an incoherent process as the scattered beams are not in phase. It is clear that at large angles it is possible to solely collect the Rutherford scattered electrons and hence to obtain images of crystals such that the intensity is only a function of Z and not the orientation of the specimen (Z-contrast imaging). This is the principle behind the high angle annular dark field imaging in scanning transmission electron microscopy (STEM) as discussed in section 5 in this chapter.

2.2.2 Inelastic scattering

Several interaction processes can cause the incident electron beam to lose energy and the energy to be transferred to the electrons or atoms in the specimen. Studying the energy-loss experienced by the electrons constitutes analytical electron microscopy (AEM) in which a wealth of important information about the chemistry and electronic structure of specimens can be obtained. There are four types of scattering mechanisms by which electrons can lose energy. These are: phonon scattering; plasmon scattering; single electron scattering; and direct radiation losses [23]. A brief description of each of these events is given here. Plasmon and single electron scattering will be considered in detail in Ch. 3.

Phonons are defined as quantised atomic vibrations within a solid [24]. In solids the atoms oscillate about their mean atomic positions and the magnitude of this oscillation varies with the thermal energy of the material. Upon interaction with the specimen the electron beam can lose energy by exciting phonons and in effect increasing the thermal energy of the specimen. Phonon scattering occurs at energies of the order of $k_B T$, where k_B is the Boltzmann constant and T is the temperature (~ 0.025 eV at room temperature) [23] and hence the energy-loss associated with them is small but they have a large scattering angle (up to ~ 175 mrad) [24]. The intensity removed from the diffracted beams by the high angle phonon scattered electrons forms the diffuse background intensity observed in diffraction patterns.

The Coulomb interaction between the traversing electron beam and the valence band electrons causes a temporary charge displacement and collective oscillation of the valence electrons [25]. This phenomena known as plasmon scattering can cause the primary electrons to suffer an energy-loss of ~ 5 – 30 eV. Plasmon scattering has a relatively small scattering angle associated with it (typically ~ 0.1 mrad with a 100 kV electron beam) [25]. Although the plasmon energy-loss is not strictly characteristic of the scattering element, it can still prove useful in characterising materials. Plasmon excitations will be discussed further in the following chapter.

It is also possible for the fast moving electron beam to lose energy by exciting single electrons in the valence or inner-shell states. Electron energy-loss spectroscopy

(EELS) (see Ch. 3) is the analysis of the energy-loss spectrum of the electron beam. Loosely held valence electrons can be ejected upon transfer of relatively small amounts of energy from the beam. Such secondary electrons are used for imaging in scanning electron microscopy (SEM). A larger amount of energy is required to eject inner-shell electrons due to their higher binding energy (*e.g.* 283 eV to remove carbon K electrons [24]) and hence the cross-sections (*i.e.* probability) associated with such excitation events are small (for example compared with plasmon excitations). The secondary events (*e.g.* x-ray or Auger electron emission) occurring when the excited atoms relax can also be exploited for analysis (for example in energy dispersive X-ray spectroscopy (EDS)) [22].

The primary electrons can lose energy and emit radiation directly if they are decelerated by the specimen. Besides interacting with the electron cloud, the beam may interact with the Coulomb field of the atomic nuclei and lose energy by emitting X-rays known as Bremsstrahlung X-rays. The relativistic effects can be another cause of direct radiation losses as such effects become significant for an electron beam of typically 100 – 300 kV energy travelling through materials with high refractive indices. The effect of relativistic radiation losses on EELS will be explored in Ch.6.

2.3 Components of a TEM

2.3.1 Electron sources

Electron sources are required to produce an electron beam which is then directed towards the illumination system of the microscope by a gun assembly (gun lens) positioned below the source. Electron sources are either required to produce a large total current in the beam (in conventional TEM) or a high intensity electron probe of small diameter (in STEM). There are two main mechanisms by which electron sources operate: thermionic emission and field emission [22]. In thermionic sources the electrons are given sufficient thermal energy to overcome the potential barrier (work function) and escape the source. The source material in thermionic sources either has a high melting temperature (tungsten, W) or a low work function

(lanthanum hexaboride, LaB_6). In field emission sources a large electric field is applied on the source to allow the electrons to tunnel through the potential barrier and emit. Before describing each of these electron source types, some of the characteristics of sources are considered.

The term source brightness, β_S, is defined as the current density per unit solid angle of the source [22]. It is important to note that the source brightness is different from the source intensity which is a measure of the total electron flux emitted from the source. The source brightness is an important parameter in STEM because it determines the total current that can be focused into a probe and onto the specimen. Ignoring the lens aberrations, the maximum achievable current, I_{max}, in a probe of diameter d is given by [23]:

$$I_{\max} = \frac{\beta_S \left(\pi^2 \alpha^2 d^2\right)}{4} \qquad \textbf{Equation 2.6}$$

where α is the convergence angle of the probe.

Another important property of electron sources is their energy spread, ΔE_S. The energy spread of the source is of particular importance in EELS where the energy resolution is directly related to ΔE_S. A large energy spread in the beam can also render image degradation in TEM and an increase in the probe size in STEM [23]. The energy spread of the beam can be improved by utilising a monochromator as discussed in section 6. Field emission guns (FEG) in general have larger β_S values, smaller ΔE_S values and longer lifetimes than thermionic guns [22].

2.3.1.1 Thermionic guns

The source in thermionic guns is either a W wire bent into a hairpin or a LaB_6 crystal. LaB_6 sources are grown in <110> orientation to enhance emission [22]. Fig. 2.2 shows a schematic diagram of a thermionic gun. Both W and LaB_6 sources form the cathode in the gun assembly and are connected to the high voltage power supply. The source is heated to emission resistively by a current passed through it. The anode is kept at earth potential. The considerable potential difference between the source and

the anode causes the electrons to accelerate towards the anode. Moreover, the emitted electron beam is brought into focus by a small negative bias applied to the Wehnelt cylinder. The Wehnelt cylinder in effect acts as an electrostatic lens, directing the electron beam through the hole in the anode. The electronic circuitry driving the Wehnelt bias and heating the source are connected such that increasing the emission current also increases the bias on the Wehnelt cylinder. This arrangement is known as self-biasing [22]. In self-biasing guns the emission current reaches a saturation point beyond which increasing the heating current through the source does not increase the emission current. Thermionic sources are operated at or just below this saturation point.

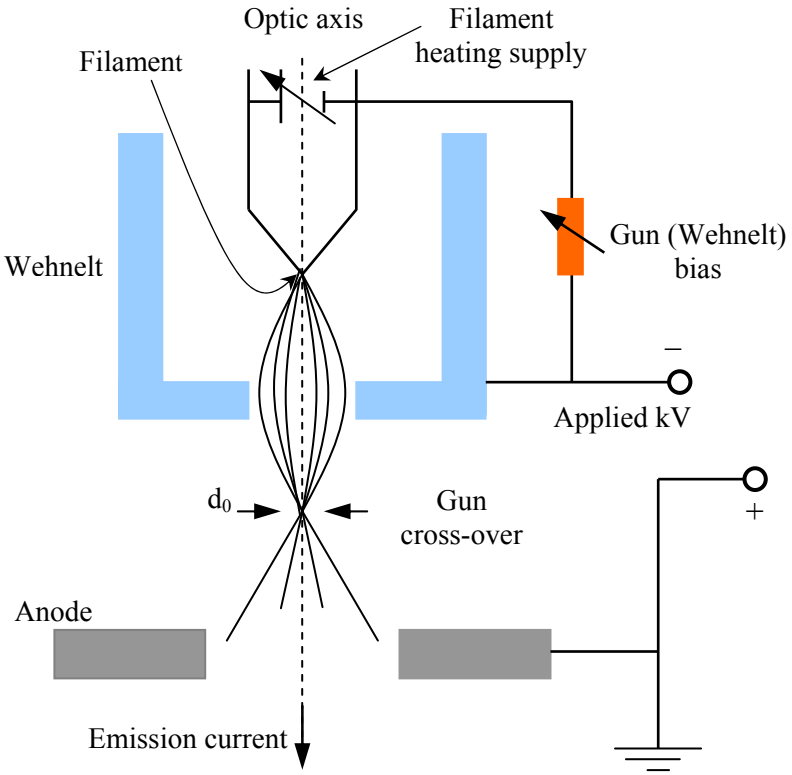


Figure 2.2: A schematic of a thermionic electron gun (adapted from [22])

LaB_6 sources require a more stringent vacuum (~ 10^{-4} Pa) than tungsten sources (~ 10^{-2} Pa) [22]. This is because LaB_6 is a highly reactive material. Another advantage of Lab_6 sources is their lower energy spread compared to tungsten sources. The low work function of LaB_6 implies a lower operating temperature and hence, a lower energy spread in the emitted electron beam. LaB_6 sources also have higher brightness values than tungsten sources [22].

2.3.1.2 Field emission sources

In field emission an electron beam is produced when a voltage is applied to a sharp needle tip, creating a large electric field, E, according to [22]:

$$E = \frac{V}{r} \qquad \textbf{Equation 2.7}$$

where V is the applied voltage (extraction voltage) and r is the radius of the needle tip. The significantly large electric field placed at the sharp tip lowers the potential barrier between the tip and vacuum sufficiently for some electrons to tunnel through the barrier and out of the tip. The FEG emitter is a tungsten wire with a fine tip. Tungsten wire is chosen because it can be readily shaped into a fine tip of radius < 0.1 μm [22]. Fig. 2.3 is a schematic of a FEG unit.

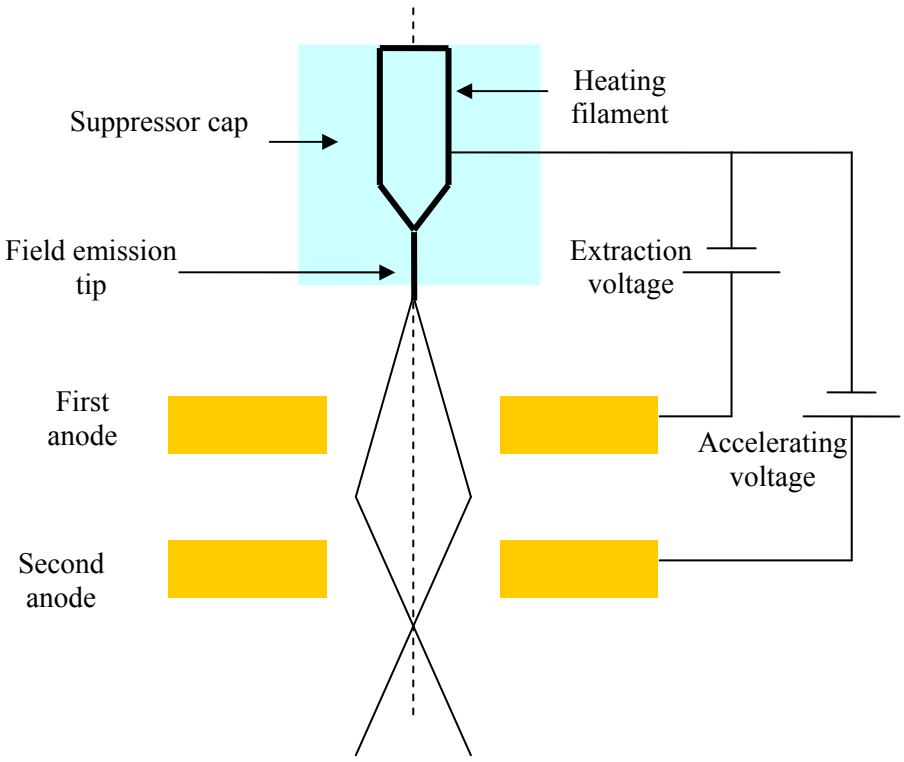


Figure 2.3: A schematic of a FEG (adapted from [22]).

The source in FEGs is made a cathode followed by two anodes. The first anode has a positive potential of several kV, known as the extraction voltage, with respect to the tip and provides the electric field required for field emission and extraction of electrons. The second anode in turn accelerates the beam to typically 100 – 300 kV. The tip of the emitter is covered with a suppressor cap kept at a constant negative

voltage, in order to block the electrons emitted from anywhere else apart from the protruding tip of the emitter.

Field emission requires the surface of the emitter to be completely free from contamination and oxides. In cold field emission this is achieved by operating at room temperature but under UHV conditions (< 10^{-11} Torr) [22]. In thermal field emission however, the surface is kept free from contamination by heating the tip to ~ 1800 K and without the need to operate under such high vacuum conditions [23]. In thermal FEGs the surface of the emitter is often covered with a thin layer of ZrO_2 in order to lower its work function (Schottky FEG) [23].

The FEI Titan microscope used in this work is fitted with a Schottky FEG. The FEG is typically operated with an extraction voltage of 4 – 4.5 kV. The emitter area is pumped through a separate pumping aperture by an ion getter pump positioned directly above the emitter to ~ 10^{-9} – 10^{-8} Torr to ensure a high vacuum [26]. The Titan is also fitted with a monochromator which sits below the gun lens assembly. For reasons explained in section 2.6.1, the monochromator software usually controls the strength of the gun lens.

2.3.2 Lenses, aberrations and apertures

In TEM lenses and apertures are used to direct the trajectory of the electron beam. Lenses are largely magnetic lenses of fixed positions and variable strengths. The effect of the magnetic field of the lens on an electron beam is described by [23]:

$$\mathbf{F} = -e(\mathbf{v} \times \mathbf{B}) \qquad \textbf{Equation 2.8}$$

where **F** is the force exerted on the beam, **v** the electron velocity and **B** the magnetic field of the lens. The exerted force is along a direction perpendicular to both **v** and **B**. For a magnetic field perpendicular to the electron beam it can be shown that electrons travelling through the lens along the optic axis and precisely centrally continue undeflected while electrons travelling along the optic axis but not centrally are deflected by the magnetic field of the lens to spiral towards the centre, through the

centre and then away from the centre before spiralling back [23]. This leads to the focusing action of the lens on the beam.

Magnetic lenses in TEM comprise of a cylindrically symmetrical core of soft magnetic material (*e.g.* soft iron) with a hole drilled through it, known as the pole piece and coils of copper wire surrounding the pole piece [22]. When a current is passed through the coils, a magnetic field is generated in the hole through the pole piece. This magnetic field is homogeneous about the optic axis but inhomogeneous along the length of the pole piece. A water circulating system is required to cool down the lens due to the resistive heating of the coils. Magnetic lenses suffer from many imperfections or aberrations compared to optical lenses. Spherical aberration, chromatic aberration and astigmatism are the most limiting forms of aberrations in TEM.

Spherical aberration arises because of the inhomogeneous effect of the lens on the off-axis beams such that beams further off-axis are deflected more strongly towards the optic axis by the magnetic field of the lens [22]. This effect is shown in Fig. 2.4.

The net result of spherical aberration is the degradation of image resolution as it blurs the image of a point object into a disc of finite size. This distorted image is known as the disc of least confusion and has a radius of [22]:

$$r_{\min} = \tfrac{1}{4} C_S \beta^3 \qquad \textbf{Equation 2.9}$$

where C_S is a constant known as the spherical aberration coefficient and β is the collection semi-angle of the imaging lens.

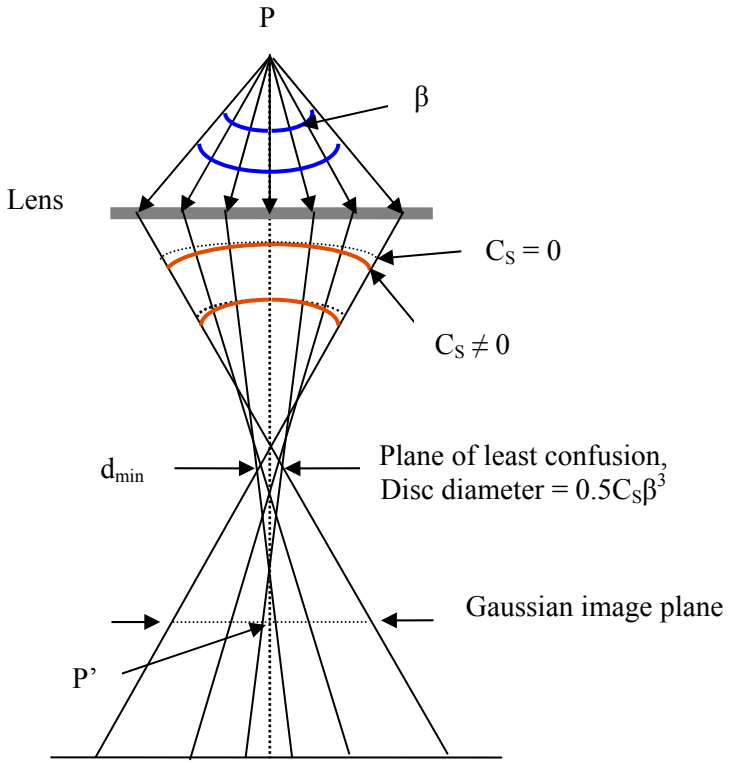


Figure 2.4: A ray diagram displaying the effect of spherical aberration on the image formed of a point object, P (adapted from [22]).

As discussed in the previous section, the electron beam produced by the source is not monochromatic. An energy spread of ΔE_S in the beam causes the electrons with different energies come to focus by the lenses at different points. In the absence of a specimen ΔE_S is negligible compared with beam energy E_0 and can be safely ignored. However, when the beam passes through the specimen several inelastic events can occur, resulting in a much larger energy spread in the beam and thus the image becomes distorted. The radius of the disc image caused by chromatic aberration is given by [22]:

$$r_{chr} = C_C \frac{\Delta E}{E_0} \beta \qquad \textbf{Equation 2.10}$$

where ΔE is the energy-loss of the electron and C_C is known as the chromatic aberration coefficient of the imaging lens. Chromatic aberration is more of a limiting factor in EELS than in imaging and several lenses in EELS spectrometers are designed to correct it. It is also important to note that chromatic aberration increases

with specimen thickness as ΔE increases when the beam interacts with thicker specimens.

Astigmatism is caused by the non-symmetry of the magnetic field of the lens. Defects present in the pole piece structure, the presence of the apertures or contamination can disturb the magnetic field of the lens and cause it not to be perfectly symmetrical [24]. The inhomogeneities in the magnetic fields of lenses and hence the astigmatism is corrected by applying compensatory magnetic fields provided by stigmators.

In the absence of aberrations the theoretical resolution of a lens, r_{th}, is expressed in terms of the Rayleigh criterion [22]:

$$r_{th} = 0.61\frac{\lambda}{\beta} \qquad \textbf{Equation 2.11}$$

where λ is the wavelength of the beam and β the collection semi-angle of the lens. However, the practical resolutions of magnetic lenses differ largely from this expression due to aberrations. In deriving the practical resolution of the microscope astigmatism and chromatic aberration are ignored (they can become negligible by using stigmators and thin specimens respectively) and the practical resolution is described in terms of the Rayleigh criterion and spherical aberration [22]:

$$r = \sqrt{r_{th}^2 + r_{sph}^2} \qquad \textbf{Equation 2.12}$$

This expression is then minimised in terms of an optimal β value to give:

$$r_{\min} \approx 0.91\left(C_S\lambda^3\right)^{1/4} \qquad \textbf{Equation 2.13}$$

The resolution in TEM is expressed as the r_{min} of the objective lens as this lens is the most important image forming in TEM mode (see section 2.4). TEM instruments typically have r_{min} values of ~ 0.2 nm [22].

Apertures limit the angular range of the beam. This action is of particular importance when selecting certain diffracted beam for imaging in TEM, controlling the convergence angle of the probe in STEM and, as explained above, controlling the resolution of the microscope.

2.3.3 Electron detectors

Any analysis in TEM requires the intensity variation of the electron beam resulting from different scattering events to be detected and converted into a viewable signal. Electron detection is carried out differently in different parts of the microscope depending on the specific requirements. The viewing screen is the most basic electron detector in TEM. It is coated with a material such as ZnS and visible light is emitted when it is illuminated by electrons [22]. The coating material is doped with impurities so that it emits a green shade of light as the colour green is considered to be more relaxing for eyes. Users are protected from the X-rays generated from the screen alongside the visible light by lead coated glass.

Two other types of detectors widely used in TEM are semiconductor detectors and scintillator-photomultiplier systems. Semiconductor detectors are Si wafers doped to form a p-n junction. A reverse bias is applied to the p-n junction and the incoming electron beam creates electron-hole pairs, resulting in a current detected by an external circuit [22]. A drawback of semiconductor detectors is the large dark current (the readout current in the absence of an incident beam due to thermally generated electron-hole pairs) associated with them. Scintillators operate on a similar principle to the viewing screen as visible light is emitted when they are struck by an electron beam. However, a material with a higher light emission than ZnS is used in scintillators (typically Ce-doped yttrium aluminium garnet (YAG)) to allow more rapid readouts [22]. The generated light signal is then amplified by a photomultiplier system. Although more expensive, scintillator-photomultiplier detectors suffer from lower noise levels compared to semiconductor detectors and are the preferred choice for electron detection in STEM. The high angle annular dark field detector on the Titan is a scintillator-photomultiplier-type detector.

Charge coupled devices (CCDs) are widely utilised as viewing devices in TEM. Each pixel on a CCD is a metal-oxide-semiconductor (MOS) capacitor [22]. The pixels store a charge proportional to the intensity of the electron beam incident on them. An image is formed either by transferring and amplifying the charge of each pixel serially or, in case of 2D array CCDs, by shifting an array of pixels into a storage array and leaving the main array available for acquisition. The analogue output of the pixels

(charge current) is digitised using an analogue-to-digital converter to form a grey scale. In 2D array CCDs it is possible to bin (sum) a column array into a single pixel to improve the signal to noise ratio. Similar to semiconductor detectors there is a dark current associated with each pixel on the CCD. All dark current contributions should be measured and removed from the output signal of the CCD.

2.3.4 EELS spectrometer

As mentioned previously, EELS concerns analysing the energy-loss suffered by the inelastically scattered electrons. This requires separating (dispersing) electrons of different energies after the beam has passed through the specimen. In a magnetic prism spectrometer electrons are deflected through ~ 90° by the magnetic field of the spectrometer. It can be shown that electrons of velocity **v** travelling in a circular trajectory through a magnetic field **B** perpendicular to the direction of the beam exhibit a radius of curvature **R** given by [27]:

$$\mathbf{R} = {}^{m_0\mathbf{v}}\!\!\diagup\!\!_{e\mathbf{B}} \qquad \textbf{Equation 2.14}$$

where m_0 is the rest mass of the electron and e the electron charge. It is clear from Eqn. 2.14 that the radius of curvature of electrons through the magnetic prism is determined by their velocity, resulting in the dispersion of the beam. Beams of equal energy are then brought into focus by the magnetic prism further along the spectrometer in the dispersion plane. The deflecting, dispersing and focusing effects of the magnetic prism are shown in Fig. 2.5. The design of the magnetic prism has to ensure that beams of equal energy but different paths (in both x-z and x-y planes) are brought into focus at the same point in the dispersion plane. This focusing effect of the prism is optimised when the entering and exit faces of the magnet are not normal to the optic axis [27]. The magnetic prism, like all magnetic lenses, suffers from aberrations. Second order aberrations (varying with β^2) are reduced by devising curved faces of appropriate radii for the magnetic prism [27].

The energy resolution in EELS is limited by the focusing power of the spectrometer and the energy spread of the beam. Therefore, ways of improving the energy

resolution include correcting various spectrometer aberrations to optimise its focusing power and utilising a monochromator to reduce the energy spread of the source. The Titan is fitted with a monochromator (see section 2.6.1) and a Gatan GIF Tridiem 865 spectrometer (see section 2.6.4) by which up to 4th order aberrations are corrected.

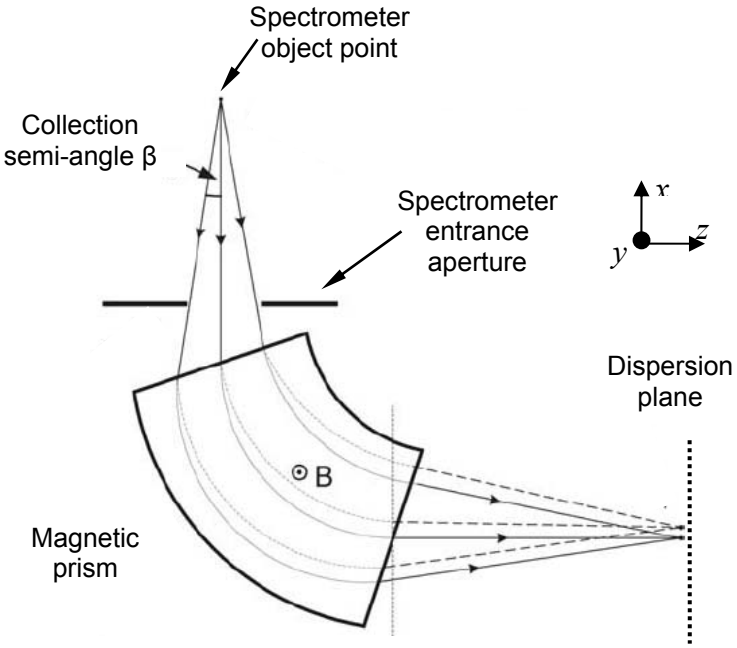


Figure 2.5: A schematic showing the deflecting, dispersing and focusing effects of the spectrometer on electrons (taken from [28]).

Ideal EELS characterisation would involve obtaining energy-loss spectra as a function of position within the specimen (x, y, z coordinates). Current TEM instruments are not capable of data acquisition through the thickness of the specimen (z-direction). However, it is possible to acquire EELS data as a function of the x- and y-positions on the specimen in spectrum imaging (SI). In SI an EELS spectrum is acquired as the STEM probe is scanned over each pixel over the specimen (see Fig. 2.6), resulting in a 3D data set where the x-position, y-position and the energy-loss spectrum form the axes.

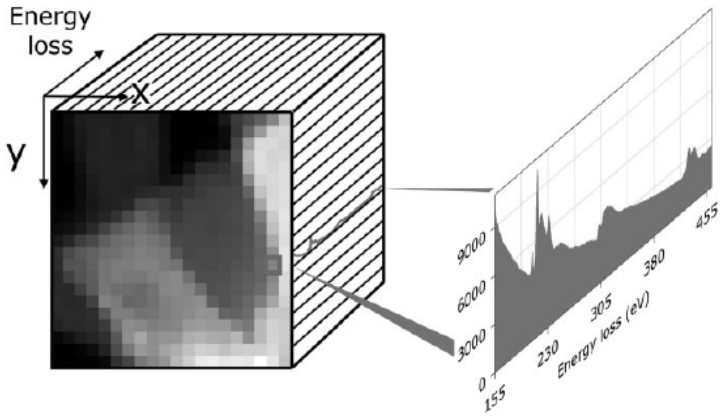


Figure 2.6: A schematic showing a spectrum image data set (taken from [29]).

2.4 Conventional TEM

In conventional TEM an area of the specimen is uniformly illuminated by the electron beam. As described in section 2.3.1 the electron beam generated by the gun is accelerated by a stable potential difference of typically 80 – 300kV. The demagnification of the image of the source and the diameter and the convergence angle of the beam are adjusted by two or three (see section 2.6.2) lenses and apertures in the condenser system, allowing the user to control the intensity and size of the illuminated area on the specimen [24]. The specimen chamber is situated below the condenser system. The objective lens follows the specimen and forms the first image of the specimen [24]. This image is then magnified by the following intermediate and projector lenses and displayed on the screen. Fig. 2.7 shows the formation of (a) a diffraction pattern and (b) an image with the specimen illuminated with a parallel electron beam. In diffraction mode, the intermediate lens is focused on the diffraction pattern formed in the back focal plane of the objective lens and this diffraction pattern is then enlarged and displayed on the screen.

In imaging mode, the objective aperture situated in the back focal plane of the objective lens selects one diffracted beam to contribute to the image. A bright field (BF) image is formed when the directly transmitted beam is selected. Conversely, a dark field (DF) image is formed when any of the diffracted beams are selected [23]. If, by the use of a larger objective aperture, two or more diffracted beams are selected to form the image, the resulting interference between these beams gives a phase-

contrast or high resolution image (HRTEM) [22]. In HRTEM images contrast due to the periodicity of crystalline specimens in the form of lattice fringes can be observed. Interpretation of HRTEM images can be rather complicated and often image simulation studies are carried out to assist the interpretation of images.

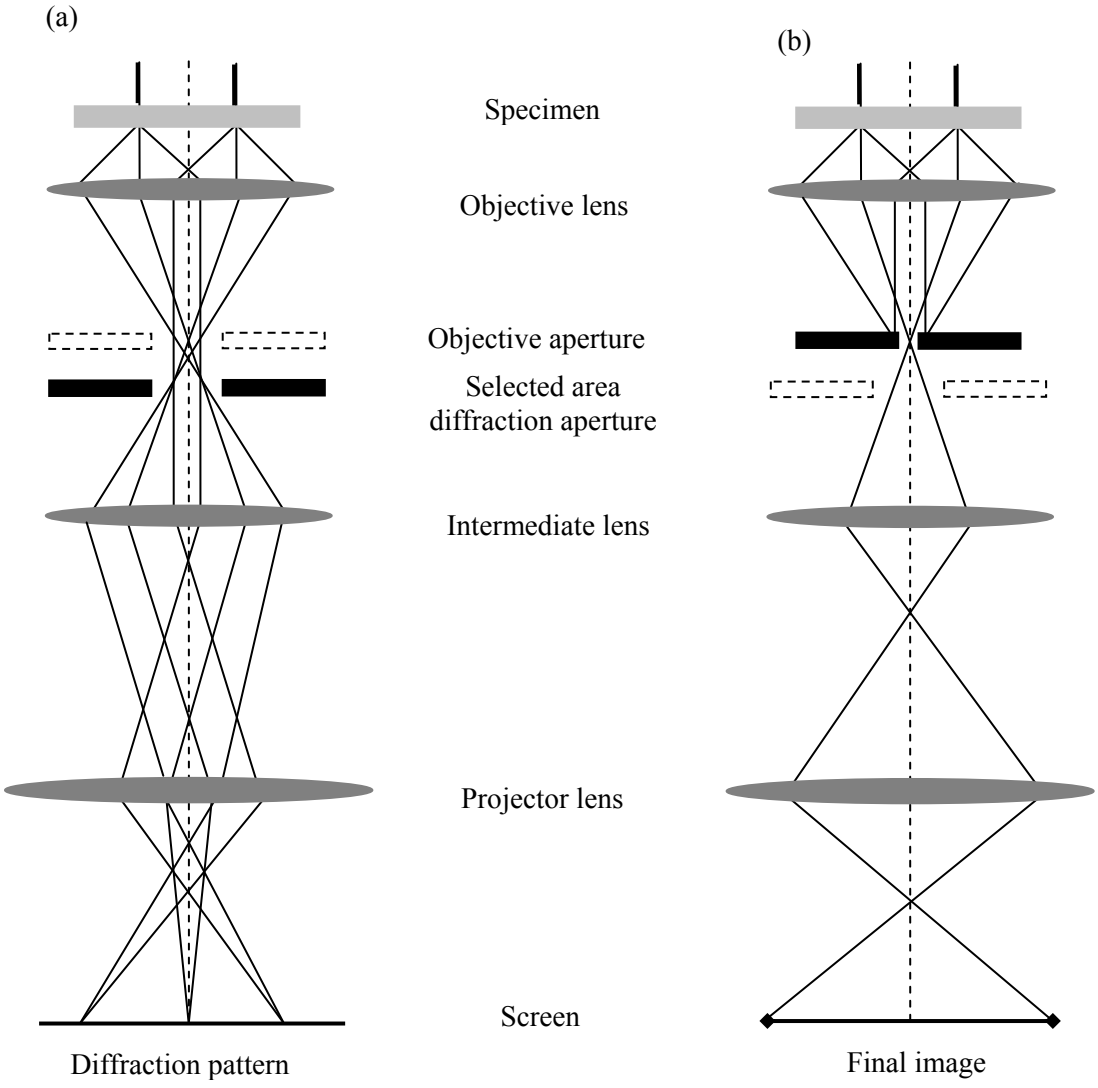


Figure 2.7: Schematics of a TEM in (a) diffraction and (b) imaging modes adapted from [22]. The specimen is illuminated with a parallel electron beam.

2.5 STEM

In STEM the electron beam is focused into a small probe by the condenser and objective lenses and is scanned point by point across the sample. Scan coils are used

to move the electron probe across the sample by tilting it twice such that the electron beam path stays parallel to the optic axis of the microscope. The transmitted beams are detected by on-axis BF or annular DF detectors to form a STEM image. Fig. 2.8 is a schematic representing image formation in STEM.

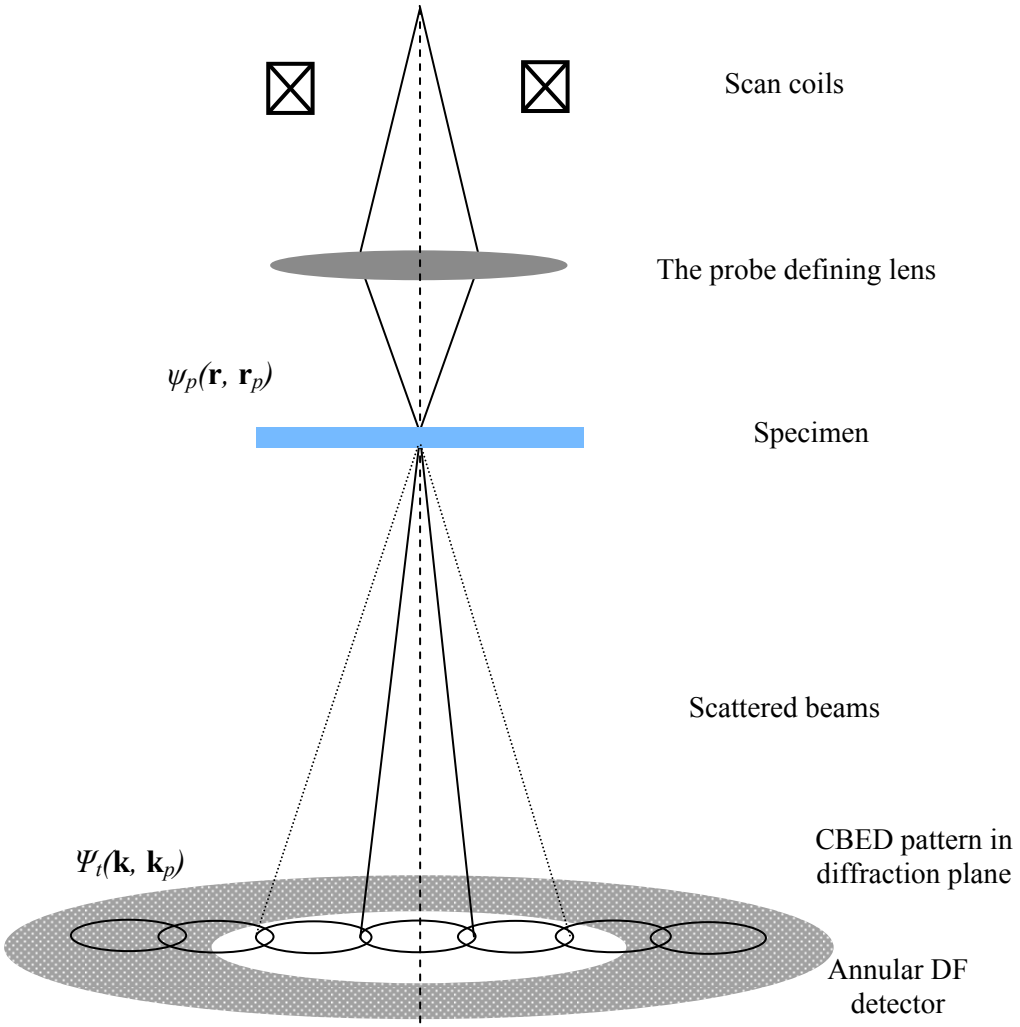


Figure 2.8: A schematic showing image formation in STEM (modified from [30]).

Assuming that the electron source is a fully coherent point source, the wave function of an electron probe $\psi_p(\mathbf{r}, \mathbf{r}_p)$ in the specimen plane can be expressed by [31]:

$$\psi_p(\mathbf{r},\mathbf{r}_p) = A \int_{Aperture} \exp\left[-iW(\mathbf{k}) - 2\pi i\,\mathbf{k}.(\mathbf{r}-\mathbf{r}_p)\right] d\mathbf{k} \qquad \textbf{Equation 2.15}$$

where $W(\mathbf{k})$ is the aberration function, $\mathbf{r} = (r_x, r_y)$ is the position vector in the image plane with corresponding Fourier spatial frequency vector $\mathbf{k} = (k_x, k_y)$, $\mathbf{r}_p$ is the probe

position and A is a normalisation constant found such that $\int \left|\psi_p(\mathbf{r}, \mathbf{r}_p)\right|^2 d\mathbf{r} = 1$. Eqn. 2.15 is integrated over the spatial frequency range $|\mathbf{k}| \leq k_{\max}$ (k_{max} is determined by the aperture limiting the angular range of the probe). During transmission through the specimen, the probe function is modulated by the specimen. This modulation is represented by the specimen transmission function, $t(\mathbf{r})$, (in general a complex valued function) and results in a transmitted wave function, $\psi_t(\mathbf{r}, \mathbf{r}_p)$, approximately given by $\psi_t(\mathbf{r}, \mathbf{r}_p) = t(\mathbf{r})\psi_p(\mathbf{r}, \mathbf{r}_p)$. In case of a crystalline specimen this wave function is diffracted onto the detectors and a coherent convergent beam diffraction pattern (CBED) is formed in the detector plane (see Fig. 2.8). The diffracted wave function is represented by the Fourier transform of the transmitted wave function, $\Psi_t(\mathbf{k}, \mathbf{k}_p)$. The intensity of a single STEM image pixel is obtained by integrating $\Psi_t(\mathbf{k}, \mathbf{k}_p)$ over the angular range of the detector. Similar to HRTEM, the interference between the CBED beams can result in high resolution crystallographic information about the specimen.

As stated in section 2.1, Bragg diffraction has a small scattering angle associated with it and the majority of electrons scattered to high angles have undergone Rutherford-type scattering. Further to this, it can be argued that when collecting electrons at sufficiently large angles in DF imaging in STEM, coherent correlation between the beams can average to zero [32]. This can be understood by considering several beams of intensity f_i and phase φ_i contributing to the resulting image intensity I [32]:

$$I = \sum f_i^2 + \sum_{i \neq j} f_i f_j \exp\left(i(\varphi_i - \varphi_j)\right) \qquad \textbf{Equation 2.16}$$

The interference between the beams can be constructive or destructive. In the case where a large number of beams contribute to I, the positive and negative phase correlations included in the second part of Eqn. 2.16 average to zero, eliminating the coherent contributions to I. Therefore, in high angle annular DF imaging (HAADF) in STEM, it is possible to achieve incoherent imaging where the coherence relationship between the beams emitted from different points on the specimen becomes negligible. It has been shown that for a probe size smaller than the distance between the atomic columns of the specimen and large STEM detector inner and outer collection angles, the image intensity can be approximated by [*e.g.* 31, 33]:

$$I(\mathbf{r}) = \left|\Psi_p(\mathbf{r})\right|^2 \otimes \left|V(\mathbf{r})\right|^2 \qquad \textbf{Equation 2.17}$$

where $V(\mathbf{r})$ is the specimen object function and is related to the projected atomic potentials within the specimen. This expression confirms the strong Z-contrast nature of HAADF imaging in STEM. The exact relationship between the image intensity and Z in HAADF depends on the accelerating voltage of the electron beam, specimen thickness and inner and outer collection angles of the detector [34], but it is often approximated by $I \propto Z^n$ where $1.5 < n < 2$. Typically inner and outer collection semi-angles larger than 25 mrad and 75 mrad respectively are used in HAADF.

2.6 FEI Titan

The majority of the microscopy work carried out in this thesis was done using the FEI Titan 80 – 300 kV microscope. The Titan is a state-of-the-art analytical (S)TEM, fitted with a FEG (thermal FEG) and a monochromator. The Titan can be operated with the monochromator excited or unexcited. An energy resolution < 0.2 eV is attainable with the monochromator on. In STEM mode with the monochromator unexcited a probe size < 0.14 nm can be attained (0.3 nm with the monochromator excited). Some of the important components of the Titan microscope are described in the next section.

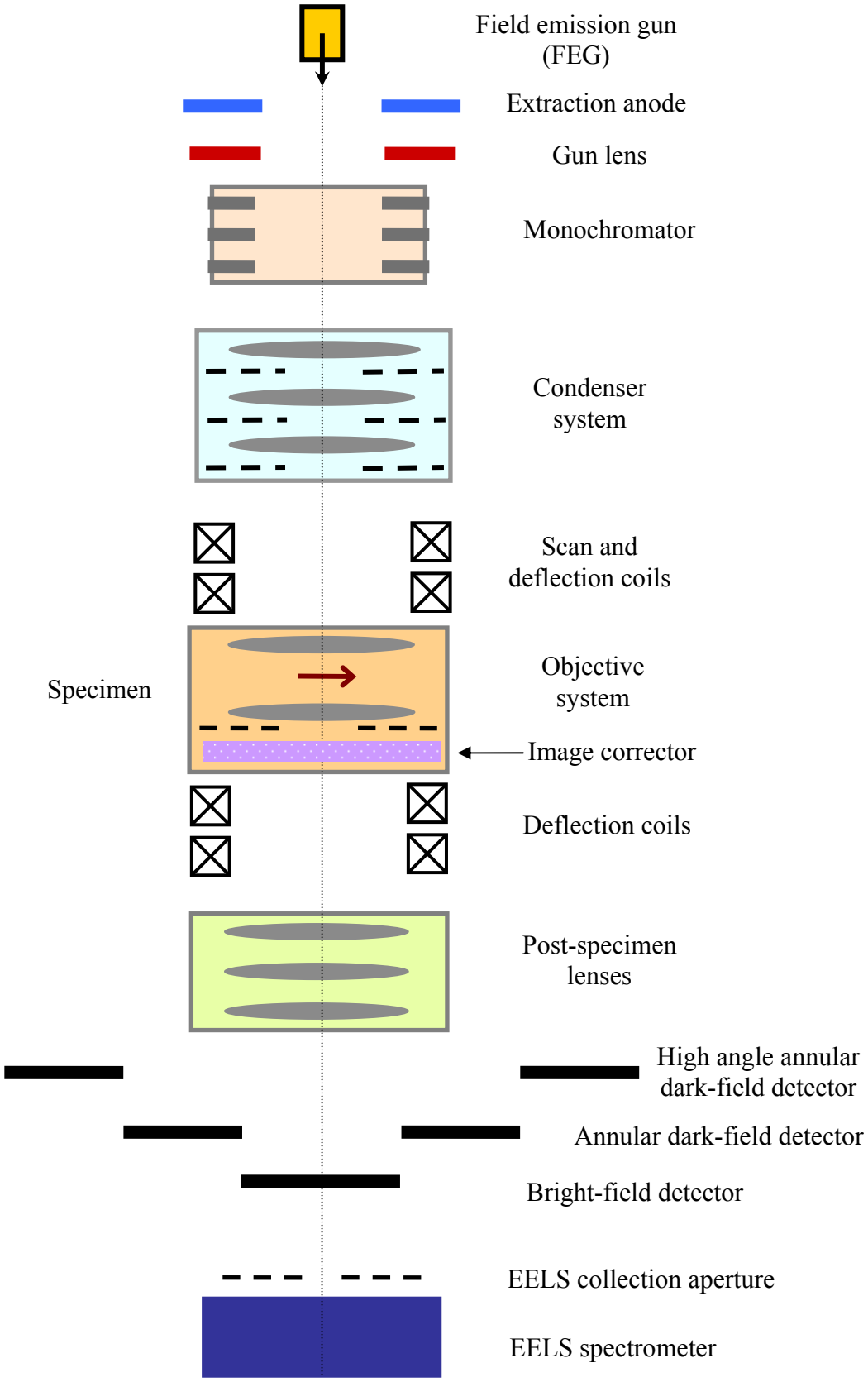


Figure 2.9: A schematic showing the main columnar components of the Titan.

2.6.1 Monochromator

A monochromator selects a narrow energy range of the electron beam generated by the source and thereby improves the energy resolution of the microscope. This is of particular importance for EELS analysis where sometimes a high energy resolution is required to resolve certain excitations. The Titan features a Wien filter type monochromator placed in the gun lens system. A Wien filter simply provides a homogeneous magnetic field $\mathbf{B}_1$ and electric field $\mathbf{E}_1$ perpendicular to each other and to the electron beam, over a length L of the filter [27]. The magnetic and electric fields exert deflection forces of magnitudes $-e\mathbf{B}_1\mathbf{v}$ and $-e\mathbf{E}_1$ respectively on electrons of velocity $\mathbf{v}$ passing through the filter. Electrons with velocity $\mathbf{v}_0 = \mathbf{E}/\mathbf{B}$ pass through the filter undeflected while electrons with velocity $(\mathbf{v}_0 - dv)$ are first drawn towards the positive electrode by the dominant electric field until the increase in their kinetic energy increases the magnetic force exerted on them and they are forced back to the optic axis (Fig. 2.10(a)). Similarly, when an electron beam enters the filter off-axis and closer to the positive electrode, it is first deflected towards and accelerated by the positive electrode until the increasing magnetic force pulls the beam along the trajectory shown in Fig. 2.10(b). Thus it can be said that a Wien filter acts as a strong lens, dispersing electrons according to their energies and focusing beams with the same energy to a single cross-over point. The strengths of the deflecting magnetic and electric fields in a Wien filter are determined by the excitation of the filter and the degree of dispersion increases with the excitation as shown in Fig. 2.11.

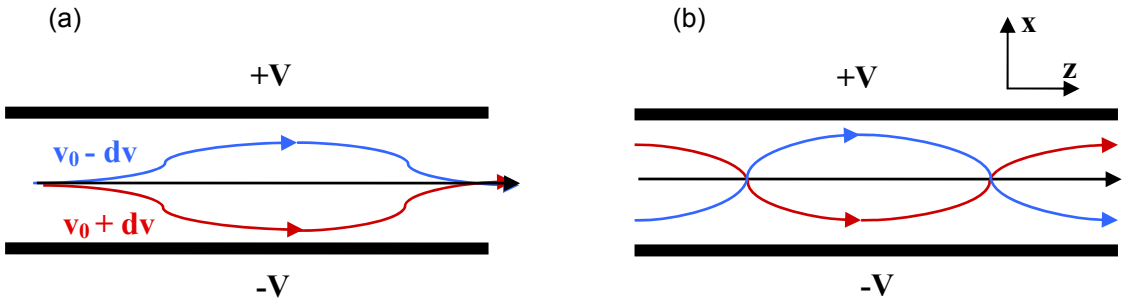


Figure 2.10: Simple ray diagrams showing examples of the electron beam trajectories, showing the dispersing and focussing effects of a Wien filter (modified from [27]).

The beam is required to be brought into focus in the xz- and yz-planes at the monochromising slit. In order to bring the monochromated beam to a point focus, an electric quadruple field $\mathbf{E}_2$ is applied over the length of the monochromator, focusing the beam in the yz-plane such that the focusing in the xz- and yz-planes become equal.

Unless specified otherwise by the user, the monochromator software controls the strength of the gun lens such that when the monochromator is excited, the beam is brought into focus at the monochromising slit. When the monochromator is not excited, the focusing of the monochromator is set such that the C_2 aperture becomes the beam limiting aperture and not the monochromising slit. This condition can be satisfied with the gun lens over-focused (positive monochromator focus value) or under-focused (negative monochromator focus value). The resulting beam has a larger beam current with the gun lens under-focused but has increased brightness with the gun lens over-focused. Therefore, the monochromator focus is set to a negative value in TEM mode and a positive value in STEM mode.

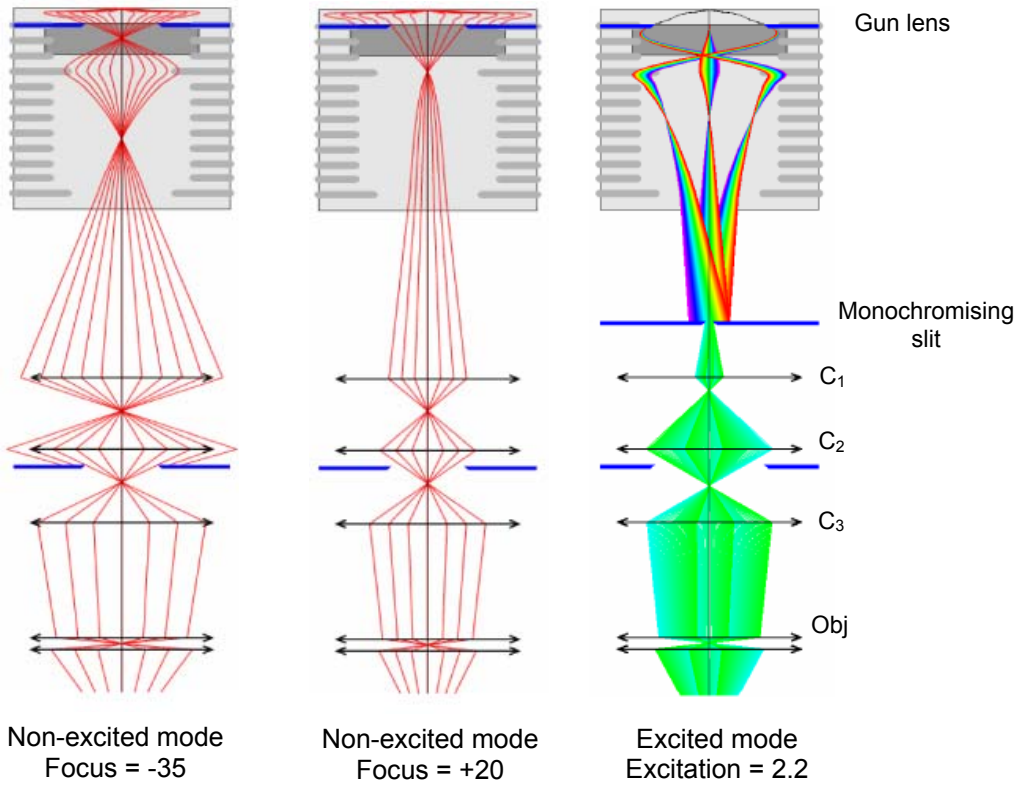


Figure 2.11: A schematic showing the gun lens and monochromator settings for the non-excited and excited monochromator settings (taken from [35]).

With the monochromator excited, the dispersion is focused at the monochromising slit and hence, the object distance of the C_1 lens is reduced (see Fig. 2.11). Using some of the previous C_1 settings, this lens will not be strong enough to bring the beam into focus before C_2. Therefore, some spot sizes become unavailable in monochromated mode but extra spot sizes, corresponding to stronger C_1 settings, are provided to compensate for this effect.

2.6.2 Condenser system

The condenser system of the Titan consists of three lenses (Fig. 2.12). The additional lens allows better control over experimental parameters such as the beam current, the convergence angle and the probe size. For example, the probe size in STEM can be changed whilst keeping the convergence angle constant or without needing to change the condenser apertures. The Titan can be operated with the third condenser lens (C_3) switched off, in which case it operates as a conventional two-condenser lens system.

The specimen is positioned in the middle of the magnetic field of the objective lens and therefore, the objective lens has an effect on the beam both below and above the specimen. This is shown in Fig. 2.12 by representing the objective lens as having an upper and a lower component. The upper objective lens plays an essential role in STEM mode but makes it difficult to get a good parallel illumination on very large or very small areas of the specimen in TEM mode. Therefore, a mini-condenser lens is added above the upper objective lens to compensate for it when required. In TEM mode the beam is brought into focus on the front focal plane of the upper objective lens in order to get parallel illumination on the specimen by fine-tuning the C_3 lens. In STEM mode the probe is focused on the specimen by fine tuning the C_3 lens.

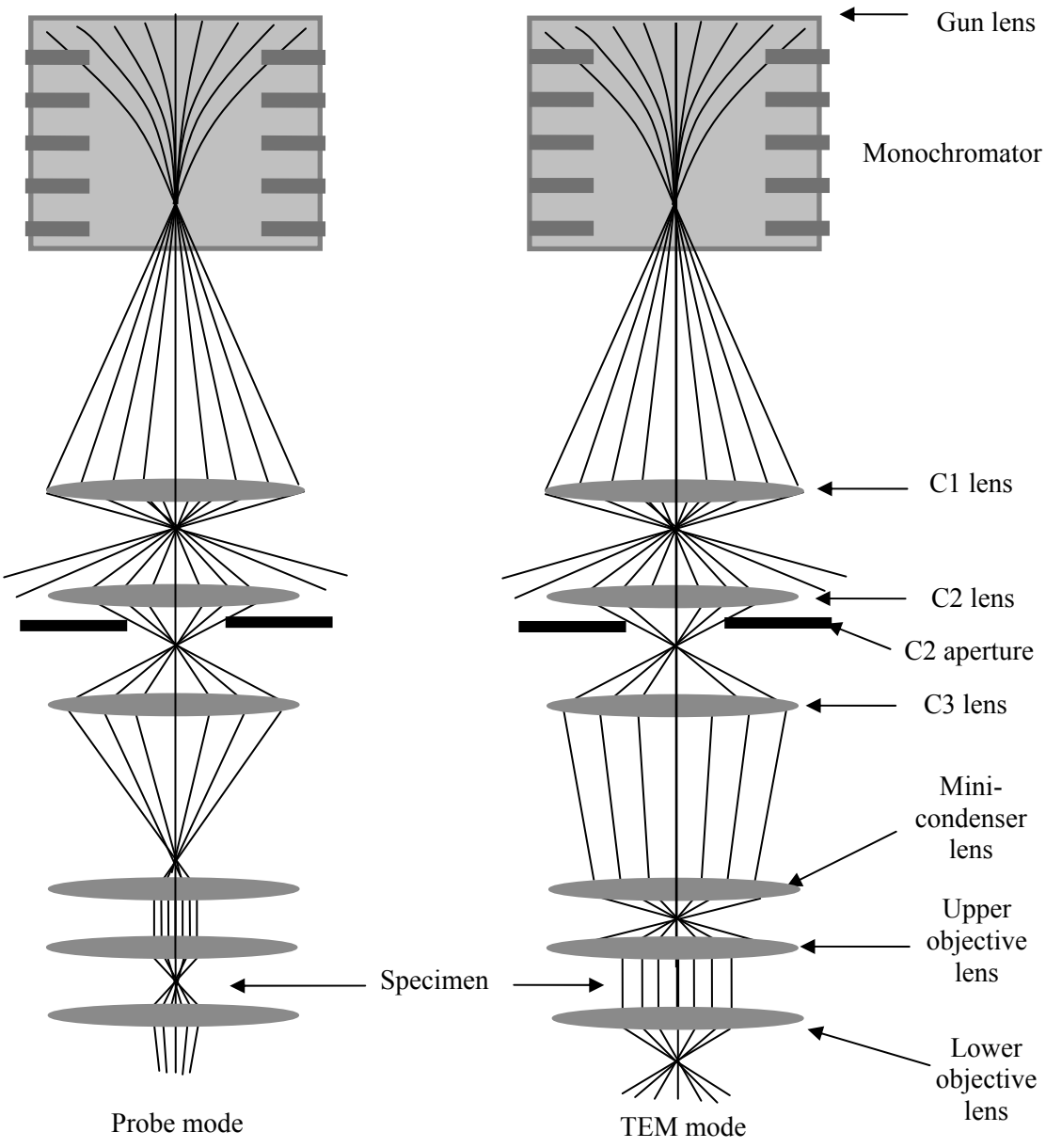


Figure 2.12: A schematic showing the STEM and TEM operational modes of the Titan (adapted from [35]).

When operated in the three-condenser lens mode, the condenser system operates as a double zoom-system. The term zoom-system is used when in two neighbouring lenses, the position of the image after the second lens or the image in front of the first lens is kept unchanged by changing the position of the image in between them (Fig. 2.13). In a three-lens condenser system the C_1-C_2 zoom-system determines the spot-size of the beam and the C_2-C_3 zoom system changes the beam diameter on the screen in TEM mode and the convergence angle in STEM mode. This becomes of particular

importance in monochromated mode when the position of the cross-over prior to the condenser system is moved closer to C_1, resulting in a larger probe-size formed by a conventional two-lens condenser system. However, because in a three-lens condenser system the probe-size is determined by C_1 and C_2 a smaller probe-size can be achieved despite the change in the object distance of C_1.

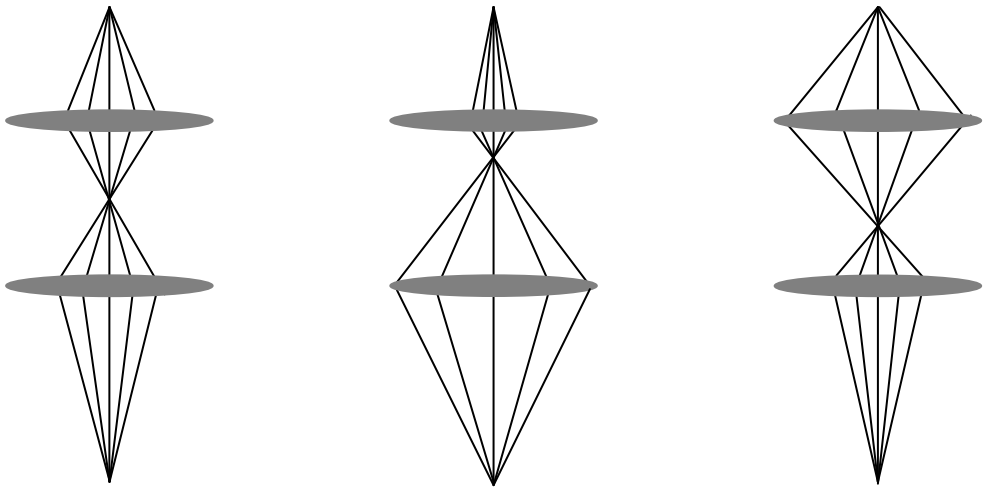


Figure 2.13: A ray diagram of the zoom-system between two neighbouring lenses where the position of the image in front of the first lens and after the second lens remain unchanged by changing the position of the image in between the two lenses (adapted from [35]).

2.6.3 Objective system

The objective lens system in the Titan is similar to that of a conventional TEM. In May 2008 an aberration corrector unit was added to the Titan to correct the spherical aberration of the lower part of the objective lens (image corrector). After the installation, the image corrector's software controls the objective stigmation settings and hence objective astigmatism was corrected using the corrector's software. Since this work does not concern the use of phase contrast imaging for characterisation, this unit is not described in detail here.

2.6.4 EELS spectrometer

The EELS data in this work was acquired in STEM mode using a Gatan GIF Tridiem 865 spectrometer. This spectrometer features additional pre-prism multipole lenses to correct the 3rd order and partially correcting the 4th order spectral aberrations [36]. Fig. 2.14 shows the main features of this spectrometer. An entrance aperture of 5 mm, 2.5 mm or 1 mm diameter can be selected to limit the angular range of the electrons entering the spectrometer. Only the 2.5 mm and 1 mm entrance apertures were used for data acquisition due to the larger aberrations present with the 5 mm aperture (larger aberrations are associated with larger spectrometer collection angles). Two sets of quadrupole, sextupole, octupole and decapole lenses situated above the magnetic prism correct the 1st, 2nd, 3rd and 4th order aberrations respectively. The quadruple lenses after the magnetic prism further change the dispersion of the beam and focus it into a line at the detector. The detector consists of an yttrium aluminium garnet (YAG) scintillator coupled with a 2k x 2k CCD. The energy dispersion per pixel on the CCD is controlled by the focusing strength of the quadrupole lenses.

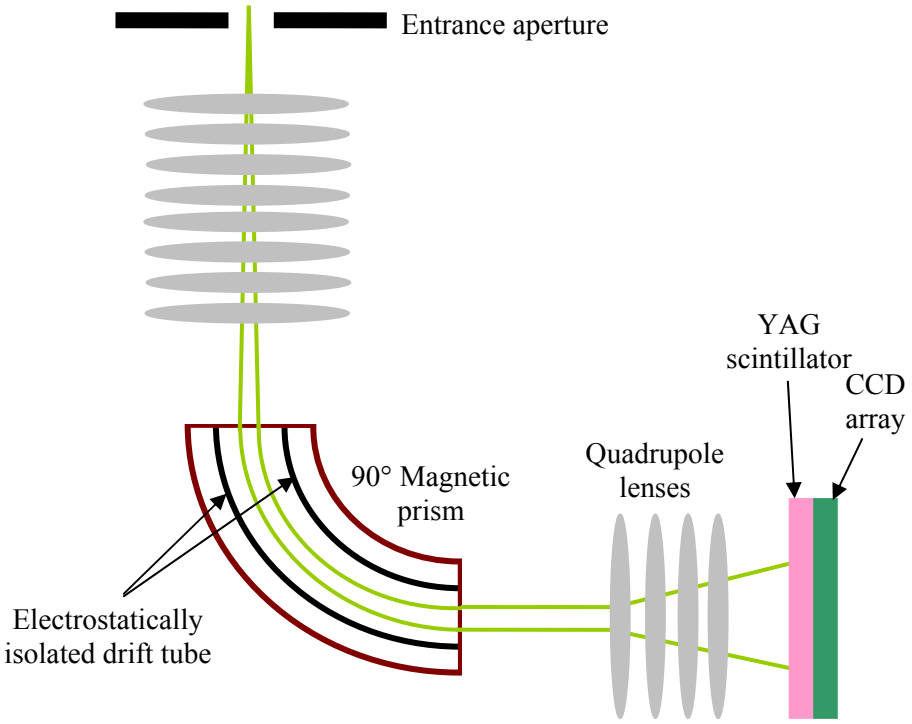


Figure 2.14: A schematic showing the main components of the Gatan GIF Tridiem spectrometer.

2.7 Specimen preparation

2.7.1 Focused ion beam (FIB) milling

The use of focused ion beam (FIB) for preparing TEM specimens has increased rapidly in recent years. The advantages of FIB include the ability to cut the specimen at any user-defined direction in a relatively short time and viewing the specimen while the operation is taking place. A FIB instrument operates in a similar way to a scanning electron microscope (SEM), where a beam of focused ions (rather than electrons) is rastered across the sample, creating secondary ions that can be used to image the sample. Ga is often used to form the ion beam in FIB. This is because Ga has a melting point close to room temperature and can be focused to form a very fine probe.

FIB was initially used to prepare cross-sectional samples in this work but was abandoned after a few attempts because of the high level of damage introduced to the specimens by it. The high energy beam used to remove material in FIB can also cause severe damage to the specimen. This damage includes gallium implantation, formation of amorphous layers (in semiconductors), formation of dislocations and redeposition of material [37]. The amount of damage caused can be reduced by lowering the beam current during the final stages of thinning but can not be completely eliminated. Therefore, all the specimens studied here were prepared using the mechanical polishing, dimple grinding and ion milling method described next.

2.7.2 Mechanical polishing, dimpling and Ar+ ion milling

The procedure used in this work is based on the method introduced by Newcomb *et al.* [38]. In this method a better mechanical stability is provided for the specimens by inserting two strips of the sample into a slotted molybdenum rod which is then able to slide into a brass tube. Molybdenum is chosen here because of its low sputtering rate in an ion-beam miller. Cross-sectional TEM specimens were prepared by first cleaving wafers of ~ 900 μm along two perpendicular <110> directions with a diamond scribe and ~ 2 mm x 10 mm sections are cut of the material. Two thin strips

were turned through 90° and the top surfaces were glued together with an epoxy (*e.g.* EPO-TEK 353ND) (see Fig. 2.15). After the epoxy has been applied to the surfaces of interest, the strips are pressed together to remove excess epoxy and ensure a minimum epoxy thickness between the two strips once set. The piece was then left on a hot plate to cure.

Sections of length ~ 2 cm were cut from a molybdenum rod of length 1 m and diameter 2 mm and 1 cm long slots are made in each section. The width of the slots made in the molybdenum rod sections using a 400 μm diamond blade was ~ 450 μm and hence, after the epoxy between the two strips of material has cured, the top and bottom surface of the piece had to be polished to fit the slot. The specimen was ground down manually to 450 μm thickness parallel to the plane of interface using 1200 grit SiC paper. The two sides of the piece were ground down equally so that the interface remains in the middle of the structure.

The piece was then covered in epoxy and was fitted into the slotted molybdenum rod. The molybdenum rod was slid into a brass tube of a 2 mm and 3 mm inner and outer diameter respectively. The brass tube was wetted with a few drops of epoxy prior to the insertion of the molybdenum. The assembly was then left to cure. This process is illustrated in Fig. 2.15.

The encapsulated sample was sliced into disks of ~ 0.3 mm thickness using an IsoMet low speed diamond saw. A thin blade (200 μm) was used to achieve a more precise cut. The samples were mechanically ground down using 1200 grit size SiC paper using a Gatan disk grinder until the discs are ~ 120 μm thick. The Gatan disk grinder is fitted with a micrometer dial, giving control over the specimen thickness to ~ 5 μm accuracy. The surface damage caused by SiC grinding was removed by successive polishing with 9 μm – 1 μm particle size diamond paste.

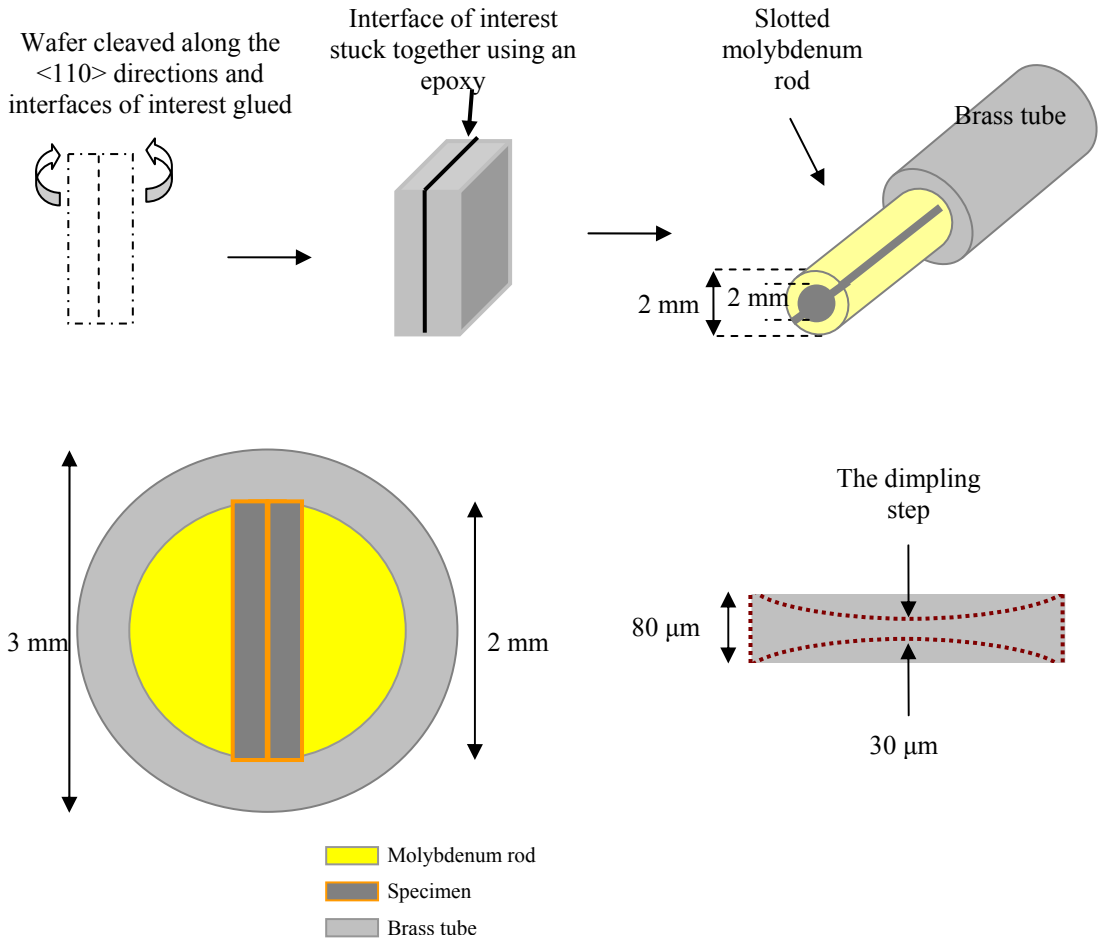


Figure 2.15: Different stages of preparing a cross-sectional specimen.

The disks by this stage were foils of ~ 80 µm thickness and appeared smooth and shiny under an optical microscope. The polished disks were cleaned using warm isopropanol for ~ 10 min to ensure that the resulting specimens were free from any contamination. After cleaning the specimen, the foil was dimpled using a dimple grinder machine (*e.g.* Gatan model 656). A 15 mm wheel was used to polish the central region of the sample (forming a dimple in the sample) successively with 1 µm and 0.5 µm diamond paste. An advantage of dimple grinding is the mechanical stability of the sample provided by the thicker rim around the sample. The dimple grinder is fitted with a micrometer for *in-situ* monitoring of the thickness of the sample. Both sides of the specimen foil were ground down equally using the dimple grinder, producing a concave shaped specimen with the central region ~ 30 µm thick.

The specimen was again cleaned thoroughly with warm isopropanol after the dimpling stage. The final stage of specimen preparation is ion milling. A Gatan model 651 precision ion polishing system (Gatan 651 PIPS) was used for this purpose. Ion milling involves bombarding a pre-thinned sample with high energy ions (e.g. Ar^{+}) in order to sputter material and achieve the desired thickness. PIPS 651 allows single or double sector milling with two guns. Hence, the sample can be milled from both back and front and a more homogeneously thinned sample can be produced. The guns are switched on and off sequentially when the double sector milling is selected. This greatly reduces specimen heating and eliminates the need for specimen cooling. The low power optical microscope on PIPS permits *in-situ* inspection of the specimen. Specimens were initially ion polished with double beams at 4 keV energy, 15 μA current and 4° incident angle, milling the top and bottom sides of the specimen. After a perforation appeared, specimens were further ion milled at 2 kV energy and 3° incident angle for typically 2 hr.

2.8 Summary

An outline of the principles behind electron microscopy was provided in this chapter followed by a description of the main components of a TEM instrument. The majority of the microscopy work was carried out using the FEI Titan and therefore a particular emphasis was placed on describing the main features of this instrument. Finally, the methods of specimen preparation used in this work were briefly introduced.

3 Electron Energy-Loss Spectroscopy

3.1 Introduction

The concept of elastic and inelastic scattering by a thin specimen was introduced in Ch. 2. Electron energy-loss spectroscopy (EELS) is the analysis of the energy distribution of the electron beam after interacting with and traversing a thin (typically < 100 nm) specimen. An EEL spectrum represents the scattering intensity as a function of the energy-loss of the transmitted electron beam.

Fig. 3.1 is an example of an EEL spectrum [29]. The most intense feature in an EEL spectrum is the zero-loss peak (ZLP), arising from the elastically scattered electrons. Although elastic scattering was defined previously as interactions by which no exchange of energy between the electron beam and the specimen occurs, in practice interactions by which the beam suffers an energy-loss less than the energy resolution of the TEM instrument are also classified as elastic events [27]. This results in events such as phonon excitations (see section 2.2.2) being included in the zero-loss region since the energy resolution of TEM instruments is seldom < 0.2 eV.

The fast electron beam can lose energy by exciting electrons in the outer or inner atomic shells of atoms to higher unoccupied energy states. Outer-shell single electron excitations can be seen as peaks at energies up to ~ 50 eV in the energy-loss spectrum. At higher energies the scattering intensity decreases according to a high power of the energy-loss and inner-shell electron excitations are seen as edges

superimposed on this smoothly decreasing background [27] (marked as ionisation edges in Fig. 3.1).

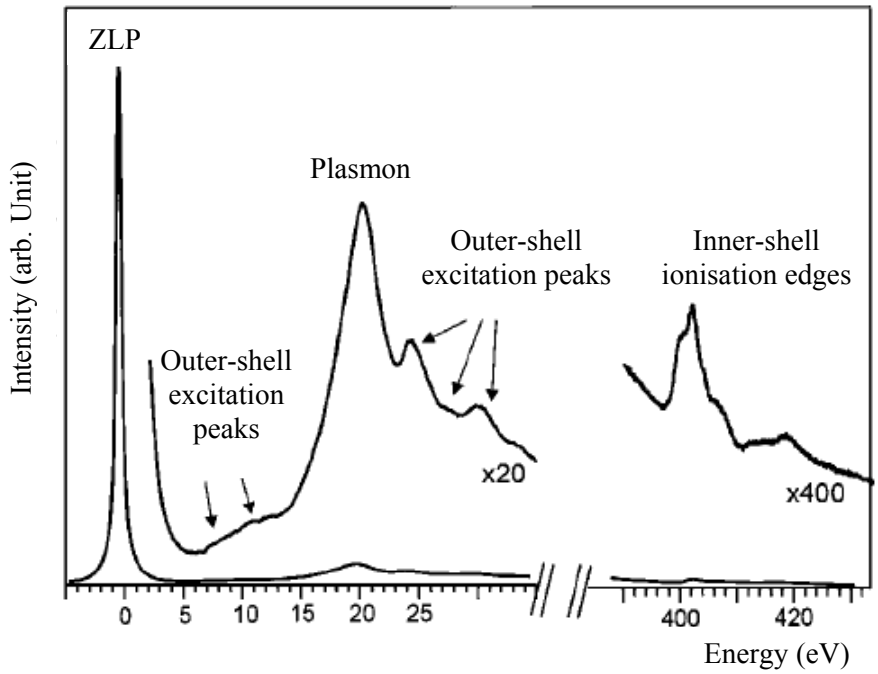


Figure 3.1: An EEL spectrum of GaN, showing zero-loss peak (ZLP), plasmon and outer and inner shell excitations [29].

In addition, the energy-loss spectrum can contain peaks corresponding to plasmon excitations (see section 2.2.2). Plasmons can be excited within the specimen, *i.e.* a bulk plasmon, or at an interface, *i.e.* an interface plasmon. Plasmons generally occur in the range 5 – 30 eV, at energies higher than the band gap of the material [27]. The first 50 eV of an energy-loss spectrum is usually referred to as the low-loss region. Higher energy-loss regions containing inner-shell ionisation edges are usually referred to as the core-loss region.

Before considering each of these scattering events in detail, it is important to understand the geometry of inelastic scattering in a TEM instrument. Fig. 3.2 shows a schematic of the scattering process in STEM mode (EELS data presented in this thesis were collected in STEM mode). The probe illuminating the specimen is formed of coherent plane-waves with a range of wave vectors $\mathbf{k_0}$ defined by the size of the probe-limiting aperture. After interacting inelastically with the specimen, a momentum $\hbar\mathbf{q} = \hbar(\mathbf{k_0} - \mathbf{k_1})$ is transferred to the specimen and the beam is scattered to an angle θ. The resulting scattered beam has a wave vector $\mathbf{k_1}$ (see Fig. 3.2). Given

that a relatively small spectrometer collection angle is used (typically collection semi-angles < 20 mrad for accelerating voltage 100 – 300 kV), **q** becomes negligible and is usually assumed to be zero.

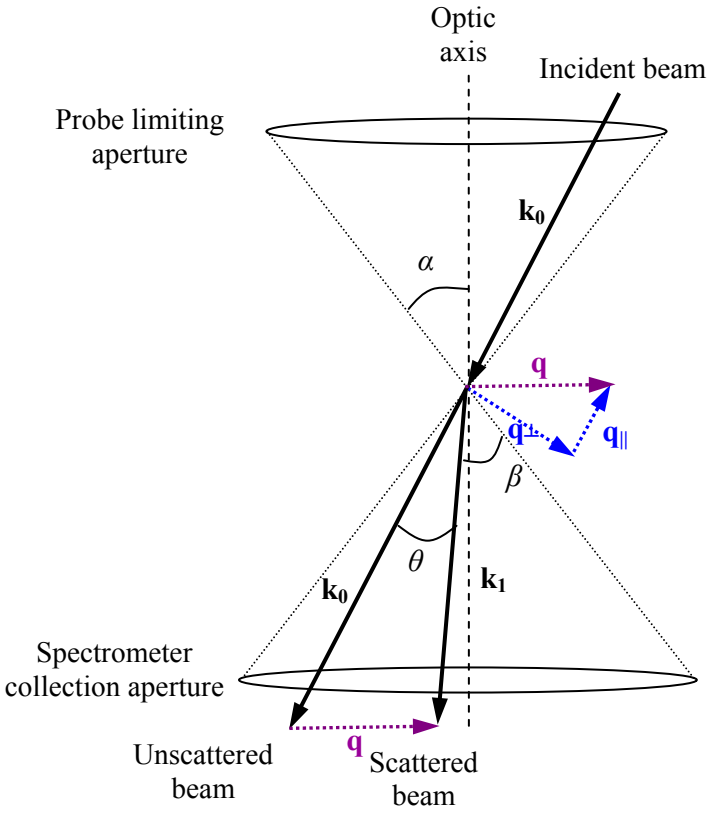


Figure 3.2: A schematic showing the inelastic scattering of electrons in STEM mode. α, β and θ are the convergence semi-, spectrometer collection semi- and scattering angles. The energy and momentum of the beam are altered after being scattered inelastically by the specimen. The momentum transferred to the specimen, $\hbar\mathbf{q}$, can be resolved into components parallel and perpendicular to the direction of the incident beam. The relative magnitude of **q** is exaggerated here (modified from [39]).

What is measured in EELS is the fraction of electrons $d\sigma$ with a particular energy range dE scattered into a specific solid angle Ω limited by the collection angle of the spectrometer 2β and is often described in terms of the double differential cross-section, $\frac{d^2\sigma}{dEd\Omega}$ (*i.e.* the measured intensity $\propto \int_{-\beta}^{\beta} \frac{d^2\sigma}{dEd\Omega} d\Omega$). Two main theories are often used to understand and predict this cross-section. The first theory, referred to as Bethe theory, is a quantum mechanical model and considers single electron excitations [27, 39]. The second theory, known as the dielectric formulation,

considers both collective and single electron excitations by modelling the specimen as a homogeneous dielectric, through which the fast electron beam passes [27, 41]. These theories are introduced briefly in the next section.

3.1.1 Bethe theory

Bethe theory considers the response of atomic electrons to the beam in terms of transitions from an initial state with wave function $\psi_i(\mathbf{r})$ to a final state (previously unoccupied) with wave function $\psi_f(\mathbf{r})$. Wave functions describing the electronic states within a material are obtained by solving the Schrödinger equation. Bethe theory can be manipulated to show that the double differential cross-section of inelastic scattering has the form [29, 40]:

$$\frac{d^2\sigma}{d\Omega\, dE} \propto \frac{\gamma^2}{\mathbf{q}^4} \times \left| \sum_{j=1}^{N} \left\langle \psi_i \left| \exp\left(i\mathbf{q}.\mathbf{r_j}\right) \right| \psi_f^* \right\rangle \right|^2 \rho(E) \qquad \textbf{Equation 3.1}$$

where γ is the relativistic factor given by $\gamma = (1 - \frac{v^2}{c^2})^{-1/2}$, $\mathbf{r_j}$ the coordinate of the jth atomic electron, N the number of electrons in the target atom and $\rho(E)$, known as the density of states (DOS), the number of electron states available within a small energy increment. In order to describe the angular distribution of inelastic scattering, the relationship between the scattering angle, θ, and scattering vector, $\mathbf{q}$, must be determined. Bethe theory shows that for $\theta << 1$ and energy-loss, E, notably smaller than the energy of the incident beam, E_0, this relationship can be approximated to a Lorentzian relationship by [27]:

$$\mathbf{q}^2 = \mathbf{k_0^2}\left(\theta^2 + \theta_E^2\right) \qquad \textbf{Equation 3.2}$$

where θ_E is known as the characteristic scattering angle and for an electron beam of velocity v is defined by:

$$\theta_E = \frac{E}{\left(E_0 + m c^2\right)\left(v/c\right)^2} \qquad \textbf{Equation 3.3}$$

Here m is the electron mass and c the speed of light. The criterion of small scattering angle and relatively low energy-losses is known as the dipole approximation [27].

3.1.2 Dielectric formulation

An alternative approach to the quantum mechanical formulation discussed above is the classical method of considering the interaction of the fast electron beam with the entire solid in terms of the dielectric response function $\varepsilon(\mathbf{q},\omega)$. $\varepsilon(\mathbf{q},\omega)$ is a complex function with real and imaginary parts $\varepsilon_1(\mathbf{q},\omega)$ and $\varepsilon_2(\mathbf{q},\omega)$ respectively. This formulation, derived by Ritchie [41] views the fast electron beam of velocity *v* and position **r** as a point charge with charge distribution $-e\delta(\mathbf{r}-\mathbf{v}t)$. The beam generates an electrostatic potential $\phi(\mathbf{r},t)$ according to the Poisson equation $\varepsilon(\mathbf{q},\omega)\nabla^2\phi(\mathbf{r},t)=e\delta(r,t)$. The electrostatic potential $\phi(\mathbf{r},t)$ acts as a perturbation on the solid and can cause the atomic electrons to make transitions from occupied states to unoccupied states after gaining energy *E*. The probability of such transitions per unit path length is found to be [41]:

$$\frac{d^2\sigma}{d\Omega\, dE}=\frac{e^2}{\pi^2\hbar v^2}\mathrm{Im}\left(\frac{-1}{\varepsilon(\mathbf{q},E)}\right)\frac{1}{\mathbf{q}^2} \qquad \textbf{Equation 3.4}$$

$\mathrm{Im}\left(\frac{-1}{\varepsilon(\mathbf{q},E)}\right)$, known as the energy-loss function, provides a complete description of the response of the solid to the electron beam. Within the dipole approximation, the same method can be applied to describe the interaction of photons with a solid, allowing a direct comparison between EELS and optical measurements [27]. This will be further discussed in section 3.3.

3.2 Excitation of inner-shell electrons

Interaction with the electron beam can result in atomic ionisation if an energy larger than or equal to the critical ionisation energy E_c is transferred to inner-shell electrons to cause them to escape the nucleus. E_c, characteristic of the atomic element and the electron shell, is the minimum energy required to excite an atomic electron from its initial state to the nearest empty final state and indicates the onset energy of the ionisation edge in the energy-loss spectrum. Thus the appearance of certain ionisation edges in the energy-loss spectrum can be used to detect the presence of certain

elements in the specimen. The intensity of such ionisation edges reflects the amount of the element present and hence can be used for quantitative analysis [42].

The excited inner-shell electrons can be given energies exceeding E_c (hence the ionisation signal extending above the onset energy to form an edge). The range of possible energies imparted from the beam to the excited electrons is controlled by the energy distribution of the atomic unoccupied states which itself is determined by the atom and its bonding with the neighbouring atoms [42]. Thus the fine structure of an ionisation edge above the onset, known as energy-loss near-edge structure (ELNES), is a function of the energy distribution of unoccupied electron states and can be interpreted in terms of the local bonding of the ionised atom.

The excited electron leaves behind an empty inner-shell state (*i.e.* core-hole) which can affect the energy distribution of the electron states. However, the effect of the core-hole on the atomic electrons is ignored here (one-electron approximation). Returning to Bethe theory of scattering cross-section (section 3.1.1), $\frac{d^2\sigma}{d\Omega\, dE} \propto \frac{\gamma^2}{\mathbf{q}^4} \times \left| \sum_{j=1}^{N} \langle \psi_i | \exp(i\,\mathbf{q}.\mathbf{r_j}) | \psi_f^* \rangle \right|^2 \rho(E)$, the exponential part can be expanded in a Taylor series as:

$$\exp(i\,\mathbf{q}.\mathbf{r}) = 1 + i\,\mathbf{q}.\mathbf{r} - \frac{1}{2!}(\mathbf{q}.\mathbf{r})^2 + \cdots \qquad \textbf{Equation 3.5}$$

The first part of the expansion disappears as the initial and final states are orthogonal to one another and within the dipole approximation the $(\mathbf{q}.\mathbf{r})^2$ part and higher order parts can be ignored. The scattering cross-section can then be simplified to [42]:

$$\frac{d^2\sigma}{d\Omega\, dE} \propto \frac{\gamma^2}{\mathbf{q}^4} \times \left| \sum_{j=1}^{N} \langle \psi_i | (i\,\mathbf{q}.\mathbf{r_j}) | \psi_f^* \rangle \right|^2 \rho(E) \qquad \textbf{Equation 3.6}$$

There are two main parts in Eq. 3.6: the DOS and the summed transition matrix part (approximately the radial overlap of the initial and final wave functions). The DOS term is strongly energy dependent and generates the ELNES. The transition matrix term is in general a slowly varying function of energy and determines the overall shape of an edge. The transition matrix is only non-zero for transitions with $\Delta l = \pm 1$ (known as the dipole selection rule) where Δl is the change in the angular momentum

quantum number associated with the transition. Further to this, the product of the initial and final states only has a significant magnitude when there is some degree of overlap between the initial and final states. These statements indicate that the DOS measured in ELNES reflects a subset of the total DOS, said to be a site- and symmetry-projected DOS [42].

Table 1 displays the nomenclature and quantum numbers for the initial and final states of some ionisation edges [27]. For example, according to the dipole selection rule atomic electrons in 2p-states ($L_{2,3}$ shells, $l = 1$) can make a transition to either s- ($l = 0$) or d-states ($l = 2$). In this case, the degree of overlap between the initial and final states determines which of the possible transitions is more likely to occur, *e.g.* the greater overlap between 2p- and 4s-states in Si means that the $L_{2,3}$ edge probes the s-DOS [42].

		Initial state quantum number			
Edge	Initial state	*n*	*l*	*j*	Final state symmetry
K	$1s^{1/2}$	1	0	1/2	p
L_1	$2s^{1/2}$	2	0	1/2	p
L_2	$2p^{1/2}$	2	1	1/2	s or d
L_3	$2p^{3/2}$	2	1	3/2	s or d
M_1	$3s^{1/2}$	3	0	1/2	p
M_2	$3p^{1/2}$	3	1	1/2	s or d
M_3	$3p^{3/2}$	3	1	3/2	s or d
M_4	$3d^{3/2}$	3	2	3/2	p or f
M_5	$3d^{5/2}$	3	2	5/2	p or f
N_1	$4s^{1/2}$	4	0	1/2	p
N_2	$4p^{1/2}$	4	1	1/2	s or d
N_3	$4p^{3/2}$	4	1	3/2	s or d
N_4	$4d^{3/2}$	4	2	3/2	p or f
N_5	$4d^{5/2}$	4	2	5/2	p or f
N_6	$4f^{5/2}$	4	3	5/2	d
N_7	$4f^{7/2}$	4	3	7/2	d

Table 3.1: Nomenclature and quantum numbers associated with ionisation edges adapted from [42]. '*n*' is the principle quantum number, '*l*' the orbital angular momentum and '*j*' the total angular momentum quantum number, defined as the sum of the orbital angular momentum and the spin quantum numbers.

Although the above description is for that of an isolated atom and not a solid (formed of an aggregate of atoms), it still provides an accurate indication of inner-shell ionisation edges since there is little difference between the wave function of inner-shell electrons in isolated atoms and solids [27].

3.3 Excitation of outer-shell electrons

3.3.1 Single electron excitations

Interaction of the electron beam with the specimen can result in the excitation of outer-shell electrons. Since this process involves single-electron transitions, it is best described by the Bethe theory. However, unlike inner-shell electrons, usually there is a considerable difference between the wave function of outer-shell electrons in single atoms and in solids and therefore, the electronic states are considered in terms of the band structure of solids. The scattering cross-section for valence band (outer-shell) single electron transitions can be written as:

$$\frac{d^2\sigma}{d\Omega\, dE} \propto \frac{\gamma^2}{\mathbf{q}^4} \times \left| \sum_{j=1}^{N} \left\langle \psi_i \left| \exp\left(i\,\mathbf{q}.\mathbf{r_j}\right) \right| \psi_f^* \right\rangle \right|^2 \rho_{joint}(E) \qquad \textbf{Equation 3.7}$$

where $\rho_{joint}\,(E)$, known as the joint density of states (JDOS), is a convolution of the valence and conduction-band density of states. JDOS arises because with typically low binding energies of valence electrons to the nucleus, there can be a considerable degree of overlap between the valence and conduction states. Valence to conduction band transitions (interband transitions) occur at critical points of $\rho_{joint}\,(E)$ (sharp peaks occurring in $\rho_{joint}\,(E)$) [43]. In addition to this, JDOS can be related to the imaginary part of the dielectric function of the solid, $\varepsilon_2(E)$, as $\varepsilon_2 \propto \left| M_{vc} \right|^2 \rho_{joint}$ [43], where M_{vc} is the transition matrix for the valence-conduction transition. Because M_{vc} is a slowly varying function of energy, the critical points of JDOS, corresponding to allowed interband transitions, also appear in $\varepsilon_2(E)$.

The smallest energy separation between the valence and conduction bands of semiconductors and insulators is known as the band gap E_g and hence, the first

intensity appearing in the energy-loss spectrum of these materials is that of the band gap transition. Fig. 3.3 is a schematic showing band gap transitions for direct and indirect band gap semiconductors. Since the band gap of most semiconductors is at most only a few eV, the perpendicular components of the scattering wave vector $\mathbf{q}_\perp$ typically has a larger magnitude than the parallel component $\mathbf{q}_\parallel$. Given that the spectrometer collection angle is sufficiently large, the $\mathbf{q}_\perp$ transferred to the valence electrons can lead to indirect interband transitions (*i.e.* a change in **k** between the valence and conduction states). An extension of Bethe theory for valence electron transitions was used in [39] to demonstrate that the contribution of the transition matrix to the scattering cross-section differs for direct and indirect transitions. The intensities of direct and indirect transitions were shown to vary with energy-loss according to:

$$I_{direct} \propto \left(E - E_g\right)^{1/2}, \qquad I_{indirect} \propto \left(E - E_g\right)^{3/2} \qquad \textbf{Equation 3.8}$$

This leads to the direct and indirect transitions displaying different shapes in the energy-loss spectrum, with direct transitions appearing to have an abrupt onset but indirect transitions a smooth onset (see Fig. 3.4). Although only purely direct or indirect transitions are discussed above, using typical spectrometer collection angles, the $\mathbf{q}_\perp$ contribution produces both direct and indirect transitions across the transitional points in the band structure. This smoothes out the interband transition peaks in the energy-loss spectrum compared with optical measurements where the momentum transferred to the electron beam by photons is negligible.

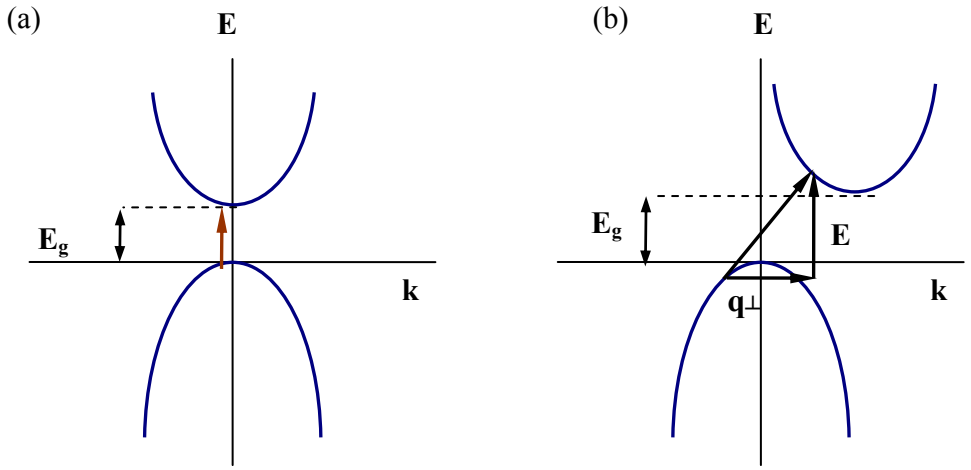


Figure 3.3: Schematics showing transitions across a (a) direct and (b) indirect band gap semiconductors (adapted from [39]).

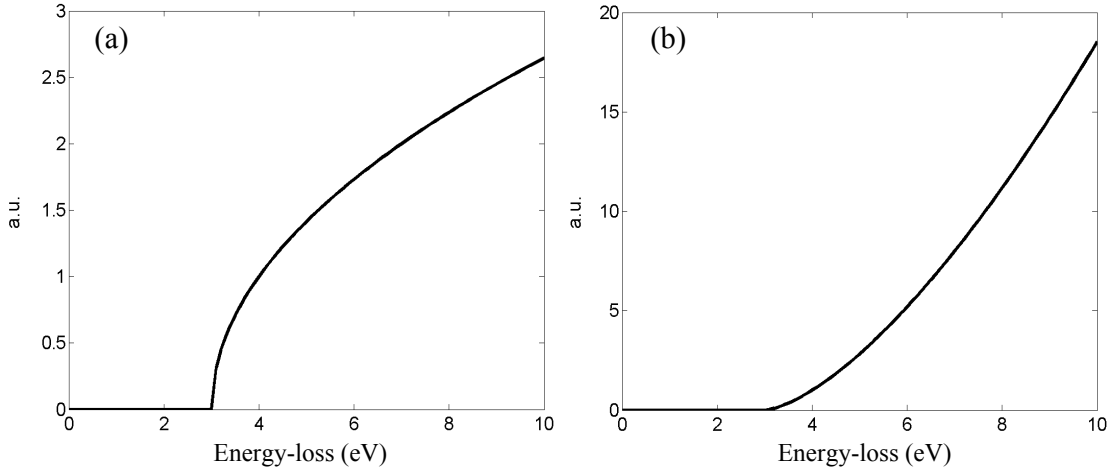


Figure 3.4: The contributions from JDOS and matrix elements for (a) direct and (b) indirect band gap transitions (adapted from [39]).

3.3.2 Plasmon excitations

3.3.2.1 Volume plasmons

Valence electrons in a solid can be viewed as an ensemble that can interact electrostatically with the electron beam. One simple approach to model this collective response is the jellium model where the valence electrons are assumed to behave as free particles and form a free electron gas. Upon interaction with the external

oscillatory electric-field of the electron beam, the valence electrons in the free electron gas begin to oscillate with a characteristic angular frequency ω_p and damping constant Γ. Considering the equation of motion $m\ddot{\mathbf{x}} + m\Gamma\dot{\mathbf{x}} = -e\mathbf{E}$ for a free electron gas of mass *m* experiencing a displacement **x** in response to the electric field **E** of the form $\mathbf{E} = \mathbf{E}\exp(-i\omega t)$ of the electron beam, the following solution for **x** is found [27]:

$$\mathbf{x} = \left(e\mathbf{E}/m\right)\left(\omega^2 + i\Gamma\omega\right)^{-1} \qquad \textbf{Equation 3.9}$$

This charge density displacement results in a polarisation $\mathbf{P} = -n\,e\,\mathbf{x} = \varepsilon_0\chi\,\mathbf{E}$ where *n* is the number of valence electrons per unit volume and χ is the dielectric susceptibility, defined as $\chi(\omega) = \varepsilon(\omega) - 1$. One can then find an expression for the complex dielectric function as:

$$\varepsilon(\omega) = \varepsilon_1(\omega) + \varepsilon_2(\omega) = 1 - \frac{\omega_p^2}{\omega^2 + \Gamma^2} + \frac{i\Gamma\omega_p^2}{\omega\left(\omega^2 + \Gamma^2\right)} \qquad \textbf{Equation 3.10}$$

where ω ($\omega = E/\hbar$) is the angular frequency associated with the energy-loss and ω_p is the natural resonant frequency of the oscillation or the plasmon frequency, defined as $\omega_p = \left(\frac{ne^2}{\varepsilon_0 m}\right)^{1/2}$. Using Eq. 3.10, the energy-loss function is found to be:

$$\mathrm{Im}\left(\frac{-1}{\varepsilon(\omega)}\right) = \frac{\varepsilon_2}{\varepsilon_1^2 + \varepsilon_2^2} = \frac{\omega\Gamma\omega_p^2}{\left(\omega^2 - \omega_p^2\right) + \left(\omega\Gamma\right)^2} \qquad \textbf{Equation 3.11}$$

The jellium model provides a rather simplistic description of plasmon oscillations by assuming a free electron gas model for valence electrons. A more realistic description of the energy-loss function can be derived by modifying the jellium model to include the effect of the bound electrons so that $\varepsilon(\omega) = 1 + \chi_f(\omega) + \chi_b(\omega)$ where, χ_f and χ_b are the dielectric susceptibilities of the free and bound electrons respectively with χ_b defined as [44]:

$$\chi_b = \frac{n_b\,e^2}{m\,\varepsilon_0}\,\frac{1}{\omega_b^2 - \omega^2 + i\omega\Gamma_b} \qquad \textbf{Equation 3.12}$$

The bound electrons have a characteristic angular frequency, ω_b, number density, n_b and damping constant, Γ_b. It can be shown that for insulating materials where the binding energy of valence electrons, E_g, is small compared with the plasmon energy $E_p = \hbar\omega_p$ the plasmon energy is shifted to a higher value according to [44]:

$$\widetilde{E}_p = \left(E_p^2 + E_g^2\right)^{1/2} \qquad \textbf{Equation 3.13}$$

This result is apparent in insulators and semiconductors as the valence electrons in these materials are bound at least by E_g before they can undergo a collective oscillation.

3.3.2.2 Surface and interface plasmons

Surface plasmons can be excited on each surface of the specimen as shown in Fig. 3.5. Unless the specimen is very thin, the surface plasmons of each surface are excited independently of each other (Fig 3.5(a)). For very thin specimens the electrostatic fields of the top and bottom surfaces interact and couple, resulting in a different surface plasmon mode (Fig. 3.5(b)). The resonant frequencies associated with the single and coupled surface modes are $\omega_a \approx \frac{\omega_p}{\sqrt{2}}$ and $\omega_b \approx \frac{\omega_p\sqrt{\mathbf{q}_\perp t}}{\sqrt{2}}$ respectively [44]. It was shown in [41] that with the beam travelling normal to the surface of a thin film specimen of thickness t, the coupled surface mode is the dominant mode when $\mathbf{q}_\perp t$ << 1, but otherwise the single surface mode is the main mode of surface plasmon excitation.

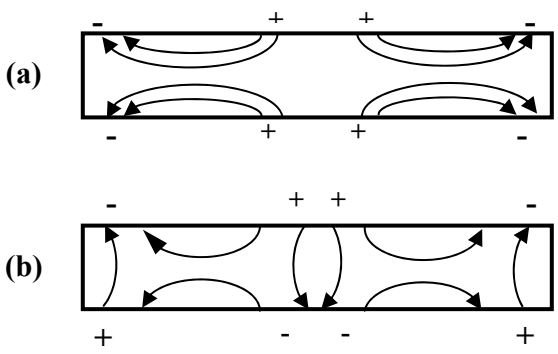


Figure 3.5: A schematic showing the electric field lines associated with surface plasmons in a thin TEM specimen for (a) a single surface mode and (b) a coupled surface mode (adapted from [27]).

Similar to surface plasmons, charge density can oscillate along an interface between two different materials, generating interface plasmons. An expression for the probability of energy-loss at a single planar interface is derived in [45] by solving Maxwell's dynamical equations in terms of the Hertz vector. The interface is assumed to be infinitely long and formed by two media with dielectric response functions ε_A and ε_B, and the beam travelling in medium *B* at a constant speed *v* and a distance x_0 from the interface as displayed in Fig. 3.6. The differential scattering cross-section is given by [45]:

$$\frac{d^2\sigma}{dz\,d\omega} = \frac{e^2}{2\pi^2\varepsilon_0\hbar v^2}\int_0^{k_y^m} \mathrm{Im}\left[G - \frac{1-\varepsilon_A\left(\frac{v}{c}\right)^2}{\alpha_A\varepsilon_A}\right] dk_y \qquad \textbf{Equation 3.14}$$

With

$$G = \left[\frac{2\alpha_A^2\left(\varepsilon_B-\varepsilon_A\right)}{\varepsilon_A\alpha_B+\varepsilon_B\alpha_A} + \left(\alpha_B-\alpha_A\right)\left(1-\varepsilon_A\left(\frac{v}{c}\right)^2\right)\right]\frac{\exp\left(-2\alpha\left|x_0\right|\right)}{\varepsilon_A\alpha_A\left(\alpha_A+\alpha_B\right)}$$

where $\mathbf{k} = \left(\mathbf{k_x}, \mathbf{k_y}, 0\right)$ is the wave vector of the scattered beam and k_y^m is the maximum value of $\mathbf{k_y}$, defined by the collection angle of the spectrometer. α_A and α_B are defined as:

$$\alpha_A = \left[k_y^2 + k_x^2 \left(1 - \varepsilon_A \left(\frac{v}{c} \right)^2 \right) \right]^{1/2}$$

$$\alpha_B = \left[k_y^2 + k_x^2 \left(1 - \varepsilon_B \left(\frac{v}{c} \right)^2 \right) \right]^{1/2}$$

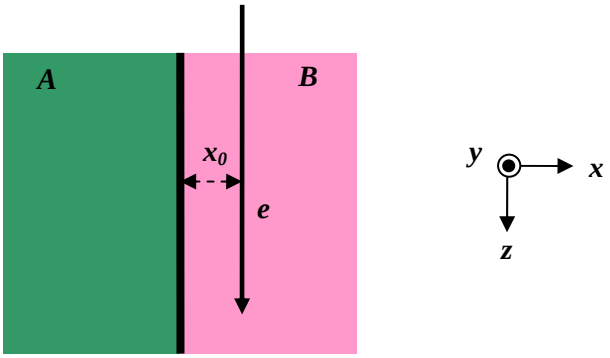


Figure 3.6: A schematic showing the geometry of the beam at distance x_0 away from an interface between media A and B with dielectric responses ε_A and ε_B respectively.

The integral in Eqn. 3.14 can only be solved numerically. This expression takes into account the losses due to relativistic effects (see Ch. 6) as well as bulk and interface losses. A non-relativistic expression can be derived by allowing $\frac{v}{c} \to 0$:

$$\frac{d^2\sigma}{dz\,d\omega} = \frac{e^2}{2\pi^2 \varepsilon_0 \hbar v^2} \left\{ \mathrm{Im}\left(\frac{-1}{\varepsilon_B} \right) \left[\ln \frac{k_y^m v}{\omega} - K_0\left(\frac{2\omega x_0}{v} \right) \right] + \mathrm{Im}\left(\frac{-2}{\varepsilon_A + \varepsilon_B} \right) K_0\left(\frac{2\omega x_0}{v} \right) \right\}$$

Equation 3.15

where $K_0(y)$, a modified Bessel function of second kind (discussed in Appendix A), replaces the integral term in Eqn. 3.14. There are two main parts in Eqn. 3.15: the bulk plasmon term $\mathrm{Im}\left(\frac{-1}{\varepsilon_B} \right)$ and the interface plasmon term $\mathrm{Im}\left(\frac{-2}{\varepsilon_A + \varepsilon_B} \right)$. The bulk plasmon contribution to the scattering cross-section increases with increasing distance from the interface x_0 whilst the interface contribution decreases, indicating that beyond a certain limit of x_0 the interface contribution becomes negligible. Moreover, the bulk plasmon contribution is expected to become negligible at the interface

(x_0 = 0), causing $\ln\left(\frac{k_y^m v}{\omega}\right) = K_0\left(\frac{2\omega x_0}{v}\right)$. This condition is satisfied for $k_y^m = \frac{1}{2x_0}$. However, the finite value of k_y^m means that this can not be true for $x_0 \rightarrow 0$, making the Bessel function approximation in Eqn. 3.15 invalid for vanishingly small values of x_0.

3.4 Quantitative analysis of EELS

3.4.1 Removal of plural scattering

So far in describing the scattering cross-section of inelastic events, it has been assumed that the specimen is sufficiently thin and only one scattering event occurs to each scattered beam (single scattering). However, this is often not the case and the resulting plural scattering complicates the interpretation of the energy-loss spectrum by introducing multiple plasmon peaks in the low-loss region and reducing the jump ratio (the ratio of the initial rise of an edge and the preceding background) of ionisation edges in the core-loss region [27]. It is, therefore, desirable to remove the effects of plural scattering from energy-loss spectra. Two methods widely used for the removal of plural scattering are described here.

3.4.1.1 Fourier-log deconvolution

Viewing inelastic events as independent collisions, the probability of a transmitted beam suffering n collisions, P_n, is given by the Poisson distribution [27]:

$$P_n = I_n / I_t = \left(\frac{1}{n!}\right)\left(\frac{t}{\lambda}\right)^n \exp\left(\frac{-t}{\lambda}\right) \qquad \textbf{Equation 3.16}$$

where I_n and I_t are the energy-integrated intensity of n-fold scattering and the total integrated intensity respectively, t the specimen thickness and λ the average distance between the collisions, known as the mean-free path for inelastic scattering. Since scattering events are assumed to be independent, the measured spectrum can be expressed as a sum of single and plural scatterings:

$J(E) = Z(E) + S(E) \otimes R(E) + D(E) \otimes R(E) + \cdots$, where *Z(E)* is the zero-loss contribution, *R(E)* the instrumental resolution, *S(E)* the single scattering contribution, *D(E)* the double scattering contribution, *etc*. Using Eqn. 3.16, the zero-loss peak, single and double scattering distributions are given by $Z(E) = I_0 R(E)$, $\int S(E)\, dE = I_1 = I_0 \left(\frac{t}{\lambda} \right)$ and $\int D(E)\, dE = I_2 = \frac{I_0}{2!} \left(\frac{t}{\lambda} \right)^2$ respectively [27]. Given that the double scattering distribution has an energy dependence of the form $S(E) \otimes S(E)$, the measured spectrum can be written as:

$$J(E) = R(E) \otimes \left[I_0 + S(E) + D(E) + \cdots \right] = Z(E) \otimes \left[\delta(E) + S(E)/I_0 + S(E) \otimes S(E)/2! I_0^2 + \cdots \right] \qquad \textbf{Equation 3.17}$$

where *δ(E)* is the unit-area delta function. Taking Fourier transforms of both sides of Eqn. 3.17, the convolution operations transform to products, simplifying Eqn. 3.17 to an exponential $j(v) = z(v) \left[1 + s(v)/I_0 + s^2(E)/2! I_0^2 + \cdots \right] = z(v) \exp\left(\frac{s(v)}{I_0} \right)$, in which the lower-case terms denote the Fourier transform of equivalent terms in Eqn. 3.17 and *v* is the frequency term with units eV^{-1}. An expression for the single scattering distribution is found by taking logarithm of both sides [46, 27]:

$$s(v) = I_0 \ln \left[j(v)/z(v) \right] \qquad \textbf{Equation 3.18}$$

This method of deconvolution requires a complete spectrum, from the zero-loss peak to the ionisation edges of interest, and therefore can not be applied to core-loss spectra. In order to prevent introducing truncation errors in Fourier transformation of the measured spectra, spectra are either collected up to energies well beyond the ionisation edge of interest or artificially extrapolated to zero at a high energy-loss value.

3.4.1.2 Fourier-ratio deconvolution

This method of deconvolution requires acquiring the low-loss and core-loss regions of the energy-loss spectrum under the same experimental conditions and dividing the

Fourier transform of the core-loss region by that of the low-loss to remove the effects of plural scattering. Since scattering events are viewed to be independent of each other, the total intensity within the core-loss spectrum *K(E)* can be defined as a sum of the intensity purely due to the core-loss ionisation $K^1(E)$ and the plural scattering contributions from the low-loss and core-loss regions *M(E)*: $K(E) = K^1(E) + M(E)$ [27]. The core-loss intensity $J_k(E)$ measured with an instrument of resolution *R(E)* is then given by [27]:

$$J_k(E) = K(E) \otimes R(E) = K^1(E) \otimes {J_l(E)}/{I_0} \qquad \textbf{Equation 3.19}$$

Where $J_l(E)$ is a low-loss spectrum recorded under the same experimental conditions as $J_k(E)$ (*e.g.* spectrometer collection angle, probe convergence angle and energy dispersion). Taking Fourier transforms of both sides of Eqn. 3.19, the deconvoluted intensity of the core-loss spectrum is given by:

$$k^1(v) = I_0 \, {j_k(v)}/{j_l(v)} \qquad \textbf{Equation 3.20}$$

3.4.2 Kramers-Kronig analysis

The dielectric formulation, described in section 3.1.2, allows the single scattering distribution to be related to the complex dielectric function of the specimen by (not including surface losses and broadening due to instrumental resolution) [27]:

$$S(E) = \int_0^\beta \frac{d^2\sigma}{d\Omega\, dE} d\Omega = \frac{2\pi e^2 I_0 t}{\pi^2 \hbar^2 v^2} \; \mathrm{Im}\left[\frac{-1}{\varepsilon(E)}\right] \ln\left[1 + \left(\frac{\beta}{\theta_E}\right)^2\right] \qquad \textbf{Equation 3.21}$$

where I_0 is the zero-loss intensity, t specimen thickness, v the speed of the electron beam, β the collection semi-angle of the spectrometer and θ_E the characteristic scattering angle defined by Eqn. 3.2. Since $\varepsilon(E)$ is a causal function, $\mathrm{Re}\left[\frac{1}{\varepsilon(E)}\right]$ can be calculated from $\mathrm{Im}\left[\frac{-1}{\varepsilon(E)}\right]$ using the Kramers-Kronig transformation [47, 27]:

$$\mathrm{Re}\left[\frac{1}{\varepsilon(E)}\right] = 1 - \frac{2}{\pi} P \int_0^\infty \mathrm{Im}\left[\frac{-1}{\varepsilon(E)}\right] \frac{E'\, dE'}{E'^2 - E^2} \qquad \textbf{Equation 3.22}$$

where P is the Cauchy principle part of the integral, preventing a pole at $E = E'$. The recovered real and imaginary parts of $\varepsilon(E)$ given by $\varepsilon(E) = \varepsilon_1(E) + i\varepsilon_2(E) = \frac{\mathrm{Re}\left[1/\varepsilon(E)\right] + i\,\mathrm{Im}\left[-1/\varepsilon(E)\right]}{\left(\mathrm{Re}\left[1/\varepsilon(E)\right]\right)^2 + \left(\mathrm{Im}\left[-1/\varepsilon(E)\right]\right)^2}$ can then be compared with those obtained from optical measurements (since the refractive index, n, and the extinction coefficient, k, are related to the dielectric function by $\varepsilon_1 = n^2 - k^2$ and $\varepsilon_2 = 2nk$ [43]). Although EELS has a poorer energy resolution compared to optical techniques (~ 0.2 eV compared to ~ 0.002 eV respectively), its superior spatial resolution enables studying the local optical properties of materials. When drawing comparison between EELS and optical data, it is important to note that $\varepsilon(E)$ obtained from EELS is a longitudinal mode dielectric response to the incident electron beam while optically measured $\varepsilon(E)$ is a transverse mode dielectric response to the incident photons. This is because the polarising electric field in EELS is in the direction of the incident beam but is perpendicular to the direction of photons in optical experiments. However, the longitudinal and transverse dielectric responses of isotropic materials are believed to be equivalent within the dipole approximation [48].

3.5 Other considerations

3.5.1 Spatial resolution

The attainable spatial resolution in EELS is determined by several factors, including those of instrumental conditions such as the probe convergence angle and diameter in STEM-EELS and the aberrations of lenses, or fundamental scattering phenomena such as beam broadening in the specimen and the localisation diameter of inelastic scattering [49]. The effect of instrumental conditions on the spatial resolution of EELS is clear as, for example, irradiating a minimal volume of the specimen in STEM requires an incident probe of minimal size, limited by the spherical and chromatic aberrations of the lenses.

The volume of the specimen irradiated by the electron beam is also limited by the lateral spreading of the beam (mainly caused by elastic scattering for which the average scattering angle is larger) as it enters the specimen. In EELS the collection aperture of the spectrometer mitigates this effect by rejecting electrons scattered to angles larger than β. In amorphous specimens ~ 90% of the electrons collected by the spectrometer are estimated to emanate from a volume of diameter ~ βt within the specimen, where t is the specimen thickness [27]. In case of crystalline specimens the spectrometer collection angle can be chosen to block Bragg diffracted beams. Further to this, the channelling effect in STEM can reduce the radial spreading of the beam by confining the beam to atomic columns and in effect collimating the inelastic scattering [49].

Even in the absence of instrumental aberrations or beam broadening, inelastic scattering is not perfectly localised at the centre of each scattering atom, but has a localisation diameter, defined as the spatial distribution of the scattering probability [27]. This imposes a physical limit to the achievable spatial resolution in EELS. The localisation diameter varies with energy-loss. For core-loss energy-loss excitation events atomic resolution EELS data has been demonstrated [*e.g.* 50] but valence electron excitations are believed to have larger localisation diameters [27].

An early suggestion for the localisation diameter of inelastic scattering, Δx, used Heisenberg's uncertainty principle and the change in momentum as $\Delta p = \sqrt{2} k_0 \theta_E$ to give $\Delta x \sim \frac{h}{\sqrt{2} k_0 \theta_E}$ [51], equivalent to the Rayleigh criterion of spatial resolution $\Delta x \sim \frac{0.61\lambda}{\theta_E}$. These expressions overestimate the localisation diameter of inelastic scattering because in practice, only a small fraction of electrons are scattered through to θ_E. A better estimate of Δx is achieved by using the median angle of inelastic scattering given by $\tilde{\theta} \sim \left(\sqrt{\frac{E}{E_0}} \theta_E \right)^{1/2}$, to determine the localisation diameter containing 50% of the scattered electrons d_{50} [52]:

$$d_{50} \sim \frac{0.6\lambda}{\tilde{\theta}} \sim 0.8\lambda \left(\frac{E_0}{E} \right)^{3/4} \qquad \textbf{Equation 3.23}$$

where E and E_0 are the energy-loss and the beam energy respectively. Eqn. 3.23 gives a median localisation diameter of ~ 2 nm for E = 20 eV and E_0 = 300 kV. Based on the classical model of scattering, in which the incident electrons are viewed as point charges interacting with the atomic Coulomb field, the delocalisation of inelastic scattering is caused by the long-range nature of Coulomb interactions [53]. In the quantum mechanical view of inelastic scattering, this delocalisation reflects the radial distribution of electronic wave functions, which describes the probability of finding an electron as a function of the radial distance from a given point [54].

3.5.2 Statistical detection limits

Since the inner-shell ionisation edges have small cross-sections, only a small number of inelastically scattered electrons may contribute to their signal. This means that in practice, detection of inner-shell ionisation edges is strongly affected by statistical constraints. Considering N number of atoms per unit area of an element in a matrix of total N_t atoms per unit area, the minimum detectable atomic fraction of an analysed element is limited by factors such as the jump-ratio of the ionisation edge of interest, the detector quantum efficiency (*i.e.* the noise performance of the detector) and statistical errors associated with the background subtraction (expressions for the minimum detectable atomic fractions and minimum detectable number of atoms can be found in [27]). In general, a small probe diameter, large probe current and long acquisition times are desired for improved detection but radiation damage and specimen drift can prove limiting to these conditions.

3.6 Summary

The theoretical aspects of EELS were introduced in this chapter, including Bethe theory and the dielectric formulation, descriptions of some of the important inner and outer-shell electron excitation events and important quantitative analyses applied to EELS data. Finally, the spatial resolution and statistical limitations of EELS and the constraints they impose on data acquisition were considered.

4 Structural Investigations of InAs/GaAs Quantum Dots

4.1 Introduction

The optical properties of InAs/GaAs quantum dots (QDs) are largely determined by their structural properties such as shape, size and composition. Successful application of this material system in fabrication of optoelectronic devices requires a precise control over the number density, nucleation site and uniformity of size and composition of the QDs. Several analytical techniques such as scanning tunnelling microscopy (STM), reflection high-energy electron diffraction (RHEED), atomic force microscopy (AFM) and TEM have been employed in the past to characterise the structural properties of the QDs [*e.g.* 55-60]. The majority of these studies have involved the characterisation of the QDs prior to their overgrowth (*i.e.* capping) [*e.g.* 55-60, 8, 63-66]. The capping of the QDs with a material of larger band gap is a necessary step in any device fabrication. Given the strain induced on the QDs by capping (due to the lattice mismatch between the QDs and the capping material), this process can introduce significant changes in the morphology and composition and thereby, modifying the optical properties of the QDs. Therefore, it is of vital importance to understand the effects of capping on the structural properties of the QDs.

An interesting feature of InAs/GaAs QD structures is the possibility of producing vertically aligned QDs in multilayers [10, 61, 62]. Vertically aligned multilayer QD structures offer important advantages such as increased number density of the QDs

(and thereby enhancement of their optical efficiency) and control over the nucleation site of the QDs. Despite the existence of a large volume of works on the structural properties of single-layer InAs/GaAs QDs, little work has been carried out comparing the structural properties of single layer and vertically aligned multilayer QDs.

A brief summary of the results of structural investigations of InAs/GaAs quantum dots published in the literature is provided in the next section. A particular emphasis is placed on the results of TEM investigations of the QDs. This is followed by the experimental details and results of analytical TEM studies carried out in this work and finally a summary of the contents of this chapter.

4.1.1 General aspects of growth

As introduced in section 1.1.4, heteroepitaxial growth of InAs on (001) GaAs in general follows a Stranski-Krastanov growth mode, in which 3D InAs islands are formed after the growth of an initial 2D InAs layer of thickness θ_{crit}, [*e.g.* 8]. Under typical experimental conditions θ_{crit} was measured to be in the range 1.5 – 1.8 monolayer (ML) [8, 63]. Further InAs deposition results in an increase in the dimensions and number density of the 3D islands, until the number density begins to saturate beyond a total InAs coverage of ~ 3 ML [8, 64, 65]. The number density saturation of the islands is caused by possible coalescence of neighbouring islands [65]. The size distribution of the islands has been shown to undergo two distinct stages [57]. Immediately after the 2D to 3D growth transition, the islands exhibit a wide size distribution. However, with further growth and the resulting reduction in the distance between neighbouring islands, the size distribution range of the islands becomes increasingly uniform [57, 65]. This has been explained in terms of different migration rates experienced by In adatoms in regions between neighbouring 3D islands of different size, causing In adatoms to favour migrating towards smaller islands [57]. The net result of this effect is a tendency for size equalisation of the islands. The size distribution of the 3D islands widens again with further InAs deposition beyond their number density saturation point, due to the formation of larger islands by coalescence [65].

The resulting properties of the 3D islands, such as number density, dimensions and size distribution, are the consequence of the kinetics of at least four important processes involved during their growth: I) the surface diffusion and migration rates of the adatoms II) the attachment and detachment rates of the adatoms by the islands III) any difference between the adsorption rates of the adatoms by the islands and the regions between them and IV) the possible interdiffusion rate between the wetting layer and the underlying GaAs substrate [66]. Therefore, it is clear that growth parameters, such as the growth temperature and rate, play a significant role in the subsequent properties of the 3D islands. In general, lower growth temperatures result in higher number density and smaller dimensions of the 3D islands. For example, reducing the growth temperature from 500 ºC to 420 ºC was shown to cause an increase in number density from 140 μm^{-2} to approximately 900 μm^{-2}, accompanied by considerable decrease in the size of the islands [66]. Furthermore, smaller number density and larger dimensions of the 3D islands have been demonstrated for lower InAs growth rates [56].

An additional important aspect of the growth of InAs/GaAs quantum dots is the possibility of forming vertically aligned stacks of 3D islands in multilayers, separated by spacer layers of a dissimilar material (typically a material with a larger band gap than the 3D islands, for instance GaAs) [10, 61, 62]. When overgrown with a spacer layer, the lattice mismatch between the 3D islands and the overgrown material induces strain in the spacer layer. The magnitude of this strain is maximal directly above the 3D islands and in cases in which the spacer layer is sufficiently thin (typically of thickness less than ~ 20 nm), these regions provide preferential nucleation sites for 3D islands in the second and subsequent InAs layers [10]. Vertically-aligned InAs/GaAs quantum dots are of interest for producing structures with higher quantum dot number density and coupled optical and electronic properties.

4.1.2 Shape

Accurate determination of the 3D shape of the QDs using (S)TEM imaging techniques has proved to be a challenging task. This is due to the fact that (S)TEM

plan-view and cross-sectional images are 2D projections of 3D structures and therefore, interpretation of such images is often ambiguous. Moreover, QD structures are in general highly strained, due to the lattice mismatch between the substrate and the QD material. This strain causes distortion of the InAs and GaAs lattice and gives rise to a strong diffraction contrast effect in bright field (BF) TEM/STEM images (see section 2.2.1), further complicating the interpretation of these images. Hence, image simulations must often be employed in order to understand the contrast present in images.

The 3D shapes reported for InAs/GaAs QDs mainly fall into two categories: pyramids and domes [*e.g.* 55, 57, 58, 60]. Theoretical study of the shape of heteroepitaxially grown 3D islands under thermodynamic equilibrium condition indicated the coexistence of these two shapes for the 3D islands [67]. It was predicted that 3D islands in thermodynamic equilibrium undergo a pyramid to dome shape transition at a critical volume (by forming steeper facets), in order to minimise their total strain energy [67]. STM investigation of InAs/GaAs and Ge/Si QDs grown under conditions close to thermodynamic equilibrium (In the case of InAs/GaAs QDs, 500 ºC growth temperature and 0.008 MLs^{-1} deposition rate) confirmed the coexistence of small pyramid shape QDs formed of shallow facets and larger dome shape QDs of steeper facets in both material systems [68]. Fig. 4.1 shows a STM topography map and the corresponding facets obtained for InAs/GaAs QDs in [68]. Although typical growth conditions of InAs/GaAs QDs are far from thermodynamic equilibrium, similar {137}, {101} and {111} facets have been reported by other STM and TEM studies [55, 79].

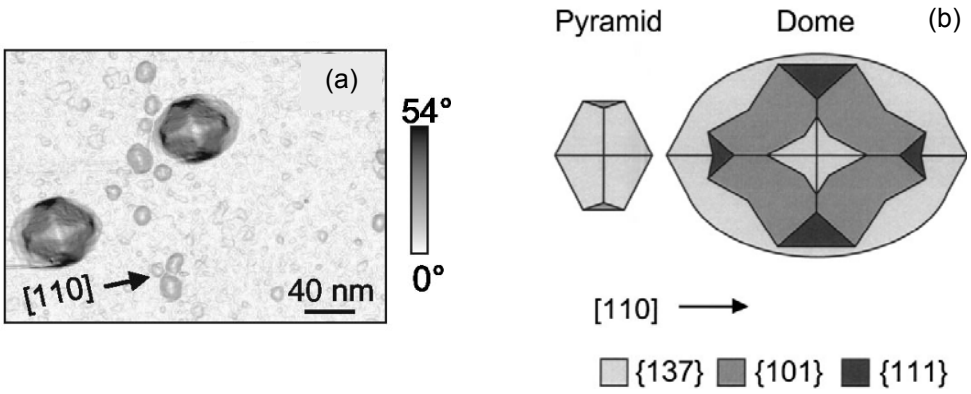


Figure 4.1: (a) STM topography of InAs/GaAs QDs grown at 500ºC temperature and 0.008 MLs^{-1} growth rate and (b) the corresponding facets of the small and large QDs in (a) (taken from [68]).

Fig 4.2 shows on-zone and off-zone plan-view BF TEM images of uncapped Ge QDs on Si [59]. The contrast arising from diffraction is reduced in Fig. 4.2 (b) compared to (a) and the coexistence of smaller pyramid shape QDs and larger dome shape QDs can be clearly seen in this image. This is in agreement with the results of STM studies in [68].

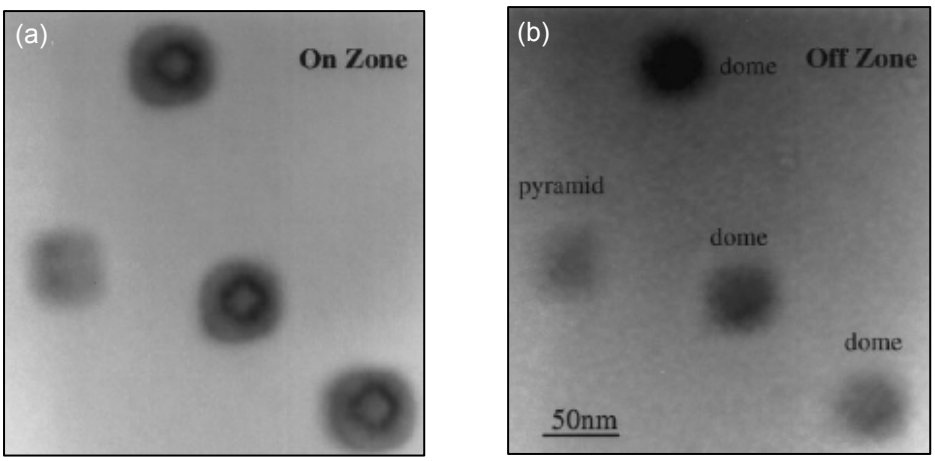


Figure 4.2: Plan-view (a) on-zone and (b) off-zone TEM images of uncapped Ge QDs on Si. The shape of the QDs can be distinguished more clearly in the off-zone image because of the suppressed diffraction effects compared with on-zone images [59].

The capping process has been demonstrated to have a significant effect on the morphology of the 3D islands [69-72]. *In-situ* STM study of the overgrowth of InAs QDs with GaAs indicated two distinct stages of capping [69]. The first stage includes

a rapid decrease in the height of the 3D islands, accumulation of the deposited material around their base and elongation of the base length of the islands. The second stage is characterised by a complete coverage and overgrowth of the 3D islands. Prior to capping the islands were bound by steep {101} and {111} facets near the base and shallow {137} facets near the apex (see Fig. 4.1) [68, 69]. During the first stage of overgrowth, the {137} facets near the apex of the 3D islands were found to become larger while the {101} and {111} facets were found to reduce in size (see Fig. 4.3) [69]. The net result of this effect is a rapid decrease in height and the shape evolution of the islands from multi-faceted domes to truncated pyramids [68, 69]. The changes occurring to the morphology of the islands during this stage of capping are caused by the intermixing of In and Ga near the base of the islands and migration of In from the top to the base of the islands, in order to relieve strain [69, 72]. The deposited Ga adatoms favour the base of the islands over the apex, since the lattice parameter of the 3D islands is maximal near the top. The Ga-rich regions around the base of the islands impose compressive strain on the overgrown regions and cause intermixing of In and Ga in order to reduce the strain energy of the structure (see Fig. 4.3). Given that bulk diffusion of materials is prohibited under typical growth conditions, this process includes migration of In adatoms away from the apex and towards the base of the islands, leading to the observed flattening and elongation of them [69, 72].

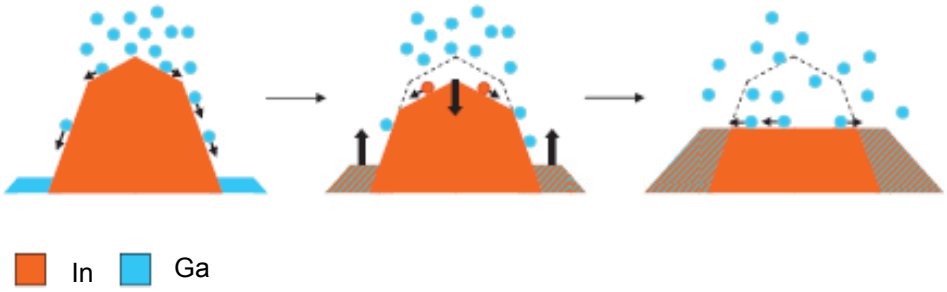


Figure 4.3: A schematic showing the changes in the morphology of InAs QDs as the result of overgrowth with GaAs (taken from [69]).

The elongation of the base length of the 3D islands during capping is an anisotropic process, favouring the $[\bar{1}10]$ direction [69, 70]. This is due to the preferential diffusion of the In and Ga adatoms along the $[\bar{1}10]$ direction [73]. Given that the changes occurring to the morphology of the QDs are driven by the surface diffusion or

migration of In and Ga adatoms, increasing the growth temperature of capping is expected to result in higher degrees of flattening and elongation of the QDs. Fig. 4.4 shows cross-sectional BF TEM images of InAs/GaAs QDs capped at different temperatures [71]. The results of the analysis carried out in [71] indicated that QDs overgrown at higher temperatures on average exhibited smaller heights and larger base lengths.

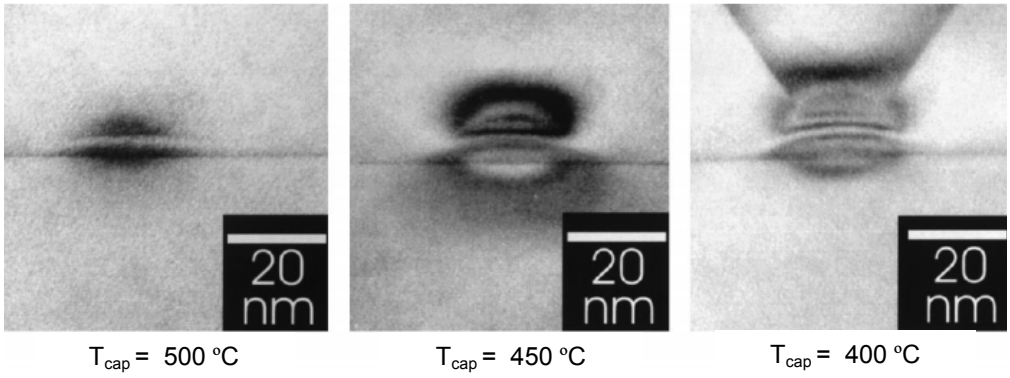


Figure 4.4: Cross-sectional BF TEM images of InAs/GaAs QDs overgrown at three different temperatures (taken from [71]).

The anisotropic shape of capped InAs/GaAs QDs has been demonstrated by high angle annular dark field (HAADF) STEM (see section 2.5) tomography in [74], in which a 3D image is reconstructed from a series of 2D images of an object tilted to different angles (for details of electron tomography see *e.g.* [75]). A pillar-shaped specimen of diameter ~ 70 nm (with a single QD lying in the centre of the pillar) was prepared by a combination of FIB and Ar^+ ion milling [74]. A series of HAADF STEM images were obtained by rotating the specimen along the [001] axis. The reconstructed 3D image of the QD exhibited an elongated base and a flat top along the $[\bar{1}10]$ direction [74].

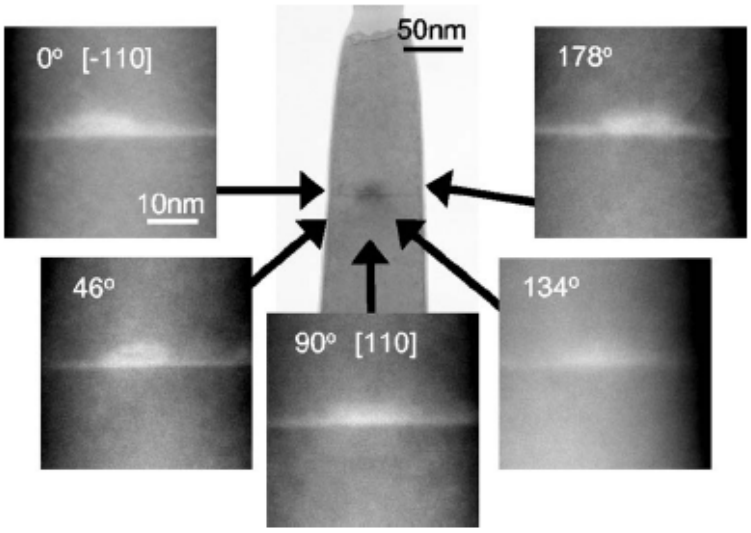


Figure 4.5: Pillar shape specimen prepared using FIB and Ar^+ ion milling and employed to carry out HAADF tomography on a single InAs/GaAs QD (taken from [74]).

4.1.3 Composition

(S)TEM offers an important advantage over other composition characterisation techniques: it can be combined with analytical techniques such as EDS and EELS to determine chemical composition with a high spatial resolution. However, the application of (S)TEM – EELS in composition characterisation of InAs/GaAs QDs has proved to be challenging. Ga, As and In core-loss edges ($L_{2,3}$ edges of Ga and As and $M_{4,5}$ edge of In) are broad in shape and feature a delayed maximum (up to ~ 100 eV above the edge onset). This gives rise to the low jump ratio associated with these edges and makes their detection constrained by the level of noise present. Therefore, these edges have scarcely been used to obtain information on the composition profile of InAs/GaAs QDs. A number of studies have used the low-energy occurring Ga $M_{4,5}$ (onset at ~ 20 eV) and In $N_{4,5}$ (onset at ~ 18 eV) edges or variations in the plasmon peak energy (~ 16 eV for GaAs and ~ 14 eV for InAs) to map compositional variations in QDs [76-78]. Besides the poor spatial resolution associated with these excitation events (approximately 2-3 nm (see section 3.5.1)), these analyses are complicated by the overlap of the Ga $M_{4,5}$ edge, In $N_{4,5}$ edge and the plasmon peak in the energy-loss spectrum. One further complication that is inherent in most TEM studies of capped QDs is the geometry of the buried QDs in thin film (S)TEM specimens. Since TEM-based composition analyses are generally carried out on cross-

sectional samples, the capping material can also contribute to the recorded data, and unless the exact shape and position of QDs within thin film specimens are known, errors are introduced in the results.

The EDS composition analysis of capped and uncapped QDs carried out in [79] indicated a significant reduction in the In content of QDs as the result of capping, with an average In concentration of ~ 65-67% found for QDs capped with GaAs. The In concentration in the wetting layer was reported to be as low as ~ 12% [79]. This is in agreement with the strain driven In-Ga intermixing mechanism proposed in [69] to explain the morphological evolution of the QDs with capping. The EELS signal from the In $M_{4,5}$-edge was analysed in [80] to examine the variation in the In content of QDs along the growth direction. The analysis revealed a similar In deficiency in the wetting layer but In enrichment of up to ~ 100% at the top of the QDs [80]. This is consistent with the results of atomistic modelling of truncated pyramid shape InGaAs/GaAs QDs carried out in [81]. The quantitative EELS analysis of Ga $M_{4,5}$ and In $N_{4,5}$ edges from InGaAsSb/GaAs QDs in [76] demonstrated a similar In deficiency in the wetting layer but In enrichment in the core of the QDs (see Fig. 4.6).

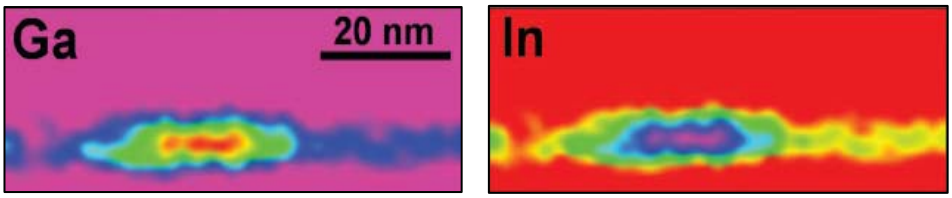


Figure 4.6: Ga and In elemental maps obtained in [76] using the In $N_{4,5}$ and Ga $M_{4,5}$ edges in the energy-loss spectrum. The colour scheme used follows the rainbow colour scheme with colours red and purple representing minimal and maximal concentrations respectively.

4.2 Experimental details

Samples were grown on epi-ready (001) GaAs substrates using the MBE system introduced in section 1.1.2. The oxide layer of the substrate was thermally desorbed at 620 ºC. This was followed by deposition of 0.5 μm thick GaAs buffer layer at 580 ºC. Two sets of samples were studied in this work. Fig. 4.7 shows schematics of the growth details of these samples. The first set consists of three samples, each

comprised of 2 periods of InAs/GaAs QDs grown at 500 ºC. 2.5 ML of InAs was deposited on GaAs in each period. The growth occurs in accordance to a Stranski-Krastanov mode and 3D InAs islands are formed after an initial 2D wetting layer. The two InAs layers are separated from each other by a GaAs spacer layer of thickness *S*. Samples were grown with *S* = 60 nm, 20 nm and 10 nm. The second InAs layer was capped with 100 nm of GaAs. The second set consists of three samples, identically grown to the samples in the first set but with the second layer QDs overgrown with 2 nm of $Ga_{0.9}In_{0.1}As$ prior to a final capping with 100 nm of GaAs.

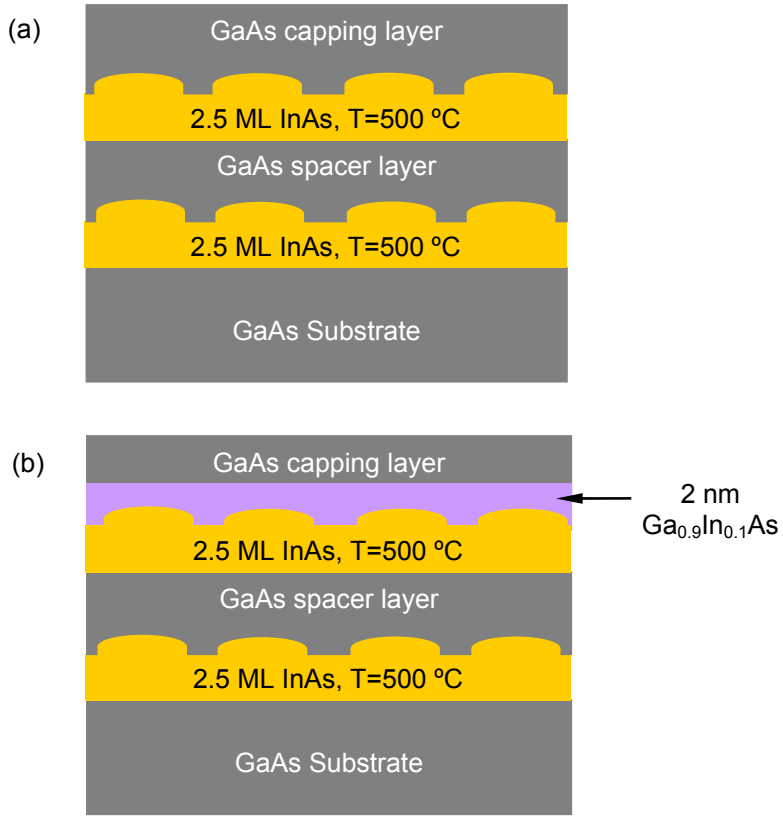


Figure 4.7: (a) and (b) Schematics of the bilayer QD samples studied in this work. Samples were grown with GaAs spacer layer thickness *S* = 60 nm, 20 nm and 10 nm in each set. See the text for details.

Cross-section (S)TEM samples were prepared by mechanical grinding and polishing followed by dimple-grinding and argon ion milling as described in section 2.7.2. STEM imaging was carried out using the FEI Titan microscope (see section 2.6 for details) operated at accelerating voltage of 300 kV and with a FEG extraction voltage of 4 kV. Other experimental details are stated for individual experiments in the results section.

4.3 Results and Discussion

4.3.1 Structural investigations using TEM and STEM

In conventional TEM the acceptance angle of the objective aperture is typically limited to a few mrad and hence, the contrast arising in conventional BF or DF TEM images of crystalline structures is largely due to diffraction effects (see section 2.2.1). This makes the interpretation of conventional TEM images of strained structures complicated, since the contrast in such images is affected by the lattice distortion induced by strain (see Eqn. 2.2). However, as demonstrated in section 2.5, *Z*-contrast incoherent imaging can be achieved in HAADF STEM. The exact *Z*-dependence of the contrast in HAADF is determined by factors such as the accelerating voltage of the electron beam, specimen thickness and the inner and outer collection angles of the detector. A criterion for establishing incoherent imaging in HAADF STEM was estimated in [83] as:

$$\beta = \frac{1.22\lambda}{\Delta R} \qquad \textbf{Equation 4.1}$$

where β is the minimum inner collection semi-angle, λ the wavelength of the incident electron beam and ΔR the distance between the two objects wished to be imaged incoherently. This gives a minimum inner collection semi-angle of ~ 20 mrad for atomic resolution images of GaAs ($\Delta R \approx 1.4$ Å) with a 300 kV electron beam accelerating voltage. However, this criterion does not take into account strain effects and in practice larger collection angles were required for achieving incoherent imaging of the QD structures under the stated conditions. Inner and outer collection semi-angles of approximately 40 mrad and 160 mrad were employed in HAADF STEM imaging of the QDs in this work.

Fig 4.8 shows BF TEM and HAADF STEM images of the same region in a bilayer InAs/GaAs QD sample with S = 20 nm. The BF TEM image was recorded using a JEOL instrument fitted with a LaB_6 filament and operated at 200 kV. The HAADF STEM image was recorded using a FEI Titan instrument operated at 300 kV. Electron probe convergence angle of ~ 10 mrad, probe diameter of ~ 2 Å and inner and outer HAADF collection semi-angles of 40 mrad and 160 mrad were used respectively.

Samples were oriented along the <110> zone-axis. The In-rich QDs appear as regions of high intensity in the HAADF image due to the higher atomic number of In compared with that of Ga. The numbers on the images represent the same QDs in each case. Comparing the dimensions of the QDs measured from the BF TEM and HAADF STEM images, the measurements taken from the BF TEM image suggest larger dimensions for the QDs than the HAADF STEM image. This difference is likely to be due to the considerable diffraction contrast present in BF TEM images, leading to the overestimation of the dimensions of strained QDs. This demonstrates that *Z*-contrast imaging in HAADF STEM provides a more accurate measure of the shape and size of the QDs than conventional TEM imaging.

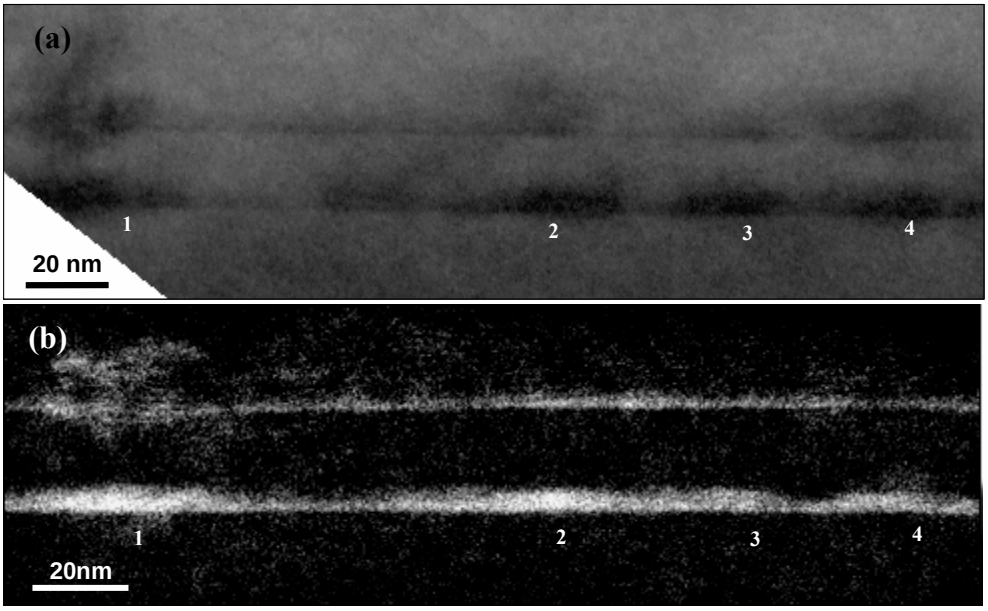


Figure 4.8: (a) BF TEM and (b) HAADF STEM images of the same region in the $S = 20$ nm sample. See the text for the experimental details. The numbers 1 – 4 correspond to the same features in images.

	BF TEM		**HAADF STEM**	
	Base length	Height	Base length	Height
QD 2	32 ± 1 nm	8 ± 1 nm	30 ± 1 nm	4.5 ± 1 nm
QD 3	22 ± 1 nm	7 ± 1 nm	20 ± 1 nm	5.0 ± 1 nm
QD 4	24 ± 1 nm	7 ± 1 nm	22 ± 1 nm	4.0 ± 1 nm

Table 4.1: The base length and height of the QDs measured in Fig. 4.8(a) and (b).

4.3.2 Thickness contrast in images

The thin film (S)TEM specimens, prepared according to the procedure described in section 2.7.2, typically have a wedge shape and therefore, the thickness of the area viewed in images is not constant. The thickness variation in wedge shape specimens leads to a smoothly increasing background intensity in *Z*-contrast STEM images (see Fig. 4.9). This background was removed from images using the rolling ball algorithm described in [88]. The principle behind background normalization in this method is understood in terms of a sphere rolling on the surface of a 3D image, with the pixel positions on the image and the intensity at each point forming the (x, y, z) coordinates of the 3D image. The ball follows the smooth background contours of the surface but does not detect any sharp peaks due to the real features in the image and thereby determines the change in the background intensity.

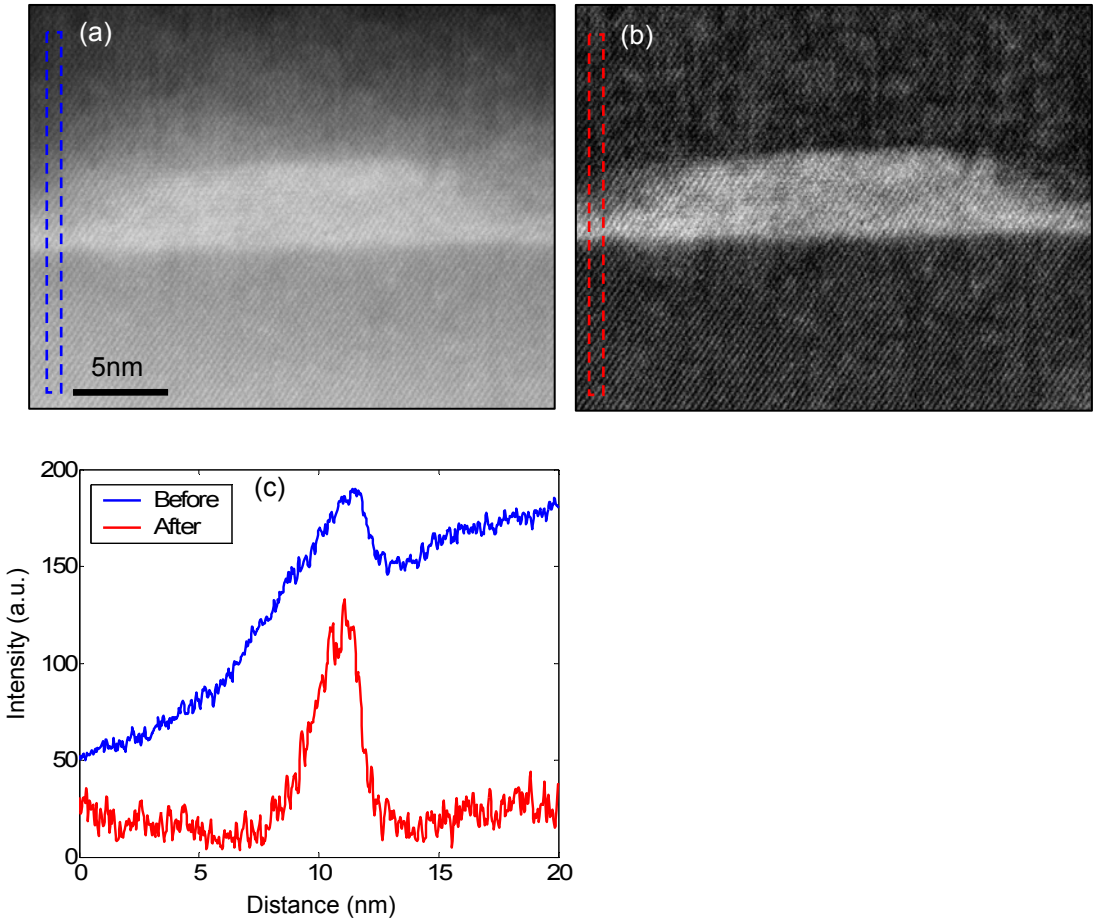


Figure 4.9: An example showing the rolling ball background subtraction applied to an image. (a) An as recorded HAADF STEM image. (b) The image in (a) after removing the background. (c) Intensity profiles of the same regions appearing in the image before and after the background subtraction.

4.3.3 Bilayers of QDs

Fig 4.10 shows HAADF STEM images of bilayer InAs/GaAs QD samples capped with GaAs and with (a) S = 60 nm, (b) S = 20 nm and (c) S = 10 nm. Images were taken along the <110> zone-axis. The electron probe convergence semi-angle and diameter were approximately 10 mrad and 2 Å respectively. The inner and outer HAADF collection semi-angles were 40 mrad and 160 mrad respectively. Examining the vertical position of the QDs in the first and second layers, no evident correlation is observed between the position of the QDs in the first and second layer in the S = 60 nm sample. This implies an independent nucleation of the QDs in the first and second layers in this sample. In comparison, a full correlation is seen between the position of the QDs in the first and second layers in the S = 10 nm sample, with every QD in the second layer formed directly above a QD in the first layer. Such vertical correlation between the position of the QDs in the first and second layers is not observed everywhere in the S = 20 nm sample. For instance, a QD can be seen directly above the QD marked 1 in Fig. 4.10 (b) but not above the QD marked 2.

The vertical stacking of the QDs in multilayers has been explained in terms of the strain interaction between the subsequent QD layers through the GaAs space layer [10, 85, 86]. When S is sufficiently small, regions in the GaAs spacer layer directly above the buried QDs are under more tensile strain and have a lattice parameter closer to that of InAs. This forms preferential sites for In atom migration and hence, nucleation of the QDs in the second layer. Given that the origin of vertical correlation is the strain field induced in the GaAs spacer layer by the underlying QDs, the partial correlation observed in the S = 20 nm sample suggests that all the dots in the first layer do not have the same magnitude of strain field around them. This non-uniformity of the strain fields around the first layer dots could be caused by non-uniformity in either their size or their composition, or both.

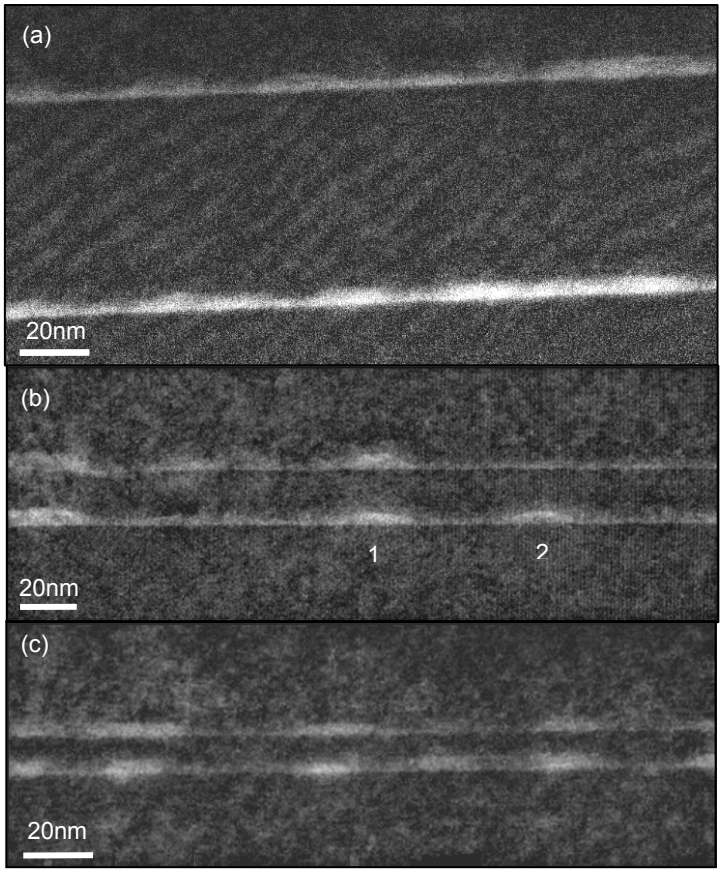


Figure 4.10: HAADF STEM images of InAs/GaAs QDs capped with GaAs and with (a) S = 60 nm, (b) S = 20 nm and (c) S = 10 nm samples. Images were taken along the <110> zone-axis. See the text for experimental details.

4.3.4 Shape

As stated previously, cross-sectional (S)TEM images of QDs are 2D projections of 3D structures along the oriented zone-axis and therefore, it is not possible to determine the 3D shape of the QDs directly from (S)TEM images. The projected shape of the QDs in the S = 60 nm and S = 10 nm samples capped with GaAs are considered in this section. These samples provide examples of QDs in multilayer and single-layer structures. In the S = 60 nm sample the QDs in the bilayers are grown independently of each other and the effect of the strain field from the bilayer on the shape of the QDs can be ignored. In the S = 10 nm sample, however, the strain field from the bilayer structure causes the vertical self-alignment of the QDs and is likely to influence the shape of the QDs.

Fig. 4.11 shows a HAADF STEM image of a QD in the second layer of the S = 60 nm sample. The specimen was oriented along the <110> zone-axis. The probe convergence semi-angle, probe diameter and inner and outer HAADF collection semi-angles used are identical to those stated in section 4.3.3. Some surface damage, caused by the ion milling process, is present in the image. The contrast due to specimen thickness variation is removed according to the procedure described in section 4.3.2. The shape of the contour around the QD is traced in Fig. 4.11(b). The QD appears to have a flat-top shape and has a height and a base length of 5 ± 1 nm and 24 ± 1 nm respectively.

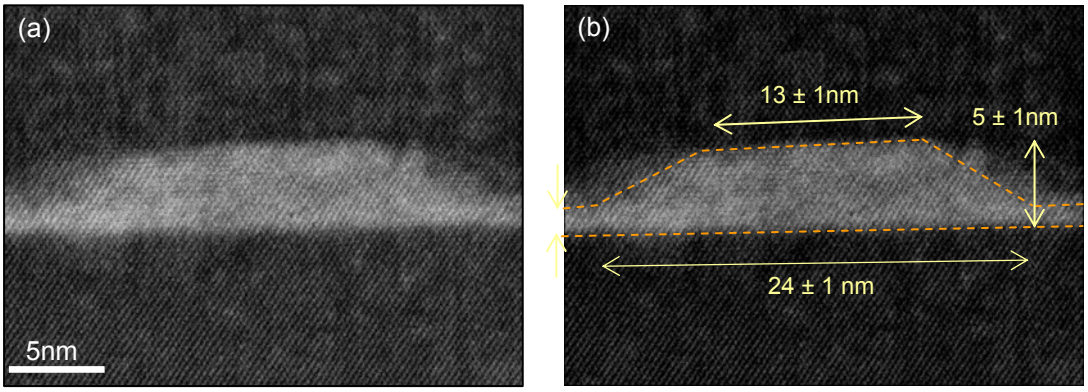


Figure 4.11: (a) Cross sectional HAADF STEM image of a single QD in the sample with S = 60 nm and capped with GaAs. The image was taken along the [110] zone-axis. See the text for other experimental details. (b) The shape of the contour around the QD is traced and its dimensions measured.

Fig. 4.12 is a HAADF STEM image of InAs/GaAs QDs in the first and second layers of the S = 10 nm sample. The image was taken along the [110] zone-axis. The contours around the QDs in the first and second layers and the dimension of the QDs are marked in Fig. 4.12(b). The first layer QD in Fig. 4.12 appears to have a similar flat-top shape to the QD in Fig. 4.11 and the two QDs exhibit similar base lengths and heights. However, the facets bounding the first-layer QD in Fig. 4.12 make smaller angles with the plane of the substrate compared to the QD in Fig. 4.11 and the length of the flat top of the first-layer QD in Fig. 4.12 is smaller than that of the QD in Fig. 4.11. In comparison, the second-layer QD in Fig. 4.12 appears to have a multi-faceted dome-like contour but with a base length and height of similar dimensions to those of the first-layer QD.

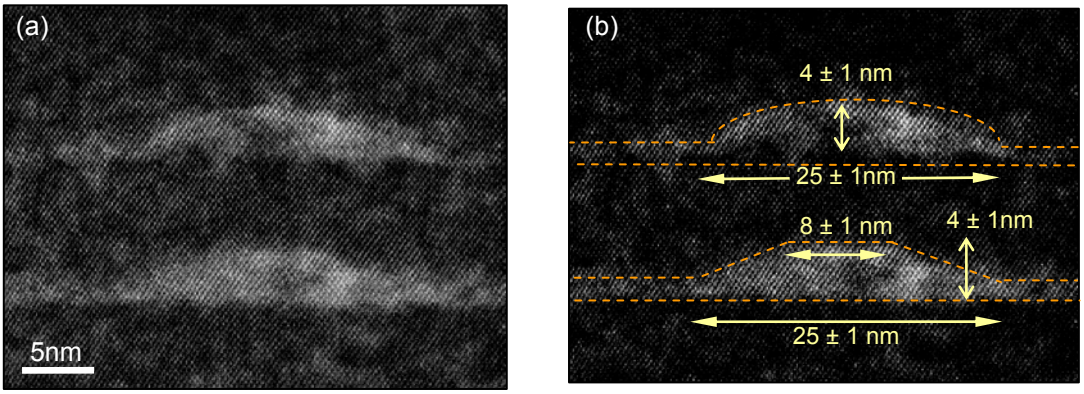


Figure 4.12: (a) Cross-sectional HAADF STEM image of bilayer InAs/GaAs QDs in the S = 10 nm sample capped with GaAs. The image was taken along the [110] zone-axis. See the text for other experimental details. (b) The contours around the first and second layer QDs and their dimensions are marked.

The observed flat-top shape of the QD in Fig. 4.11 and the first layer QD in Fig. 4.12 is consistent with the results of STM studies of single layer QDs discussed in section 4.1.2. In case of the QDs in the S = 60 nm sample, the QDs in the first and second layers were shown to grow independently of each other. This indicates that the strain interaction between the first and second layer QDs in this sample is negligible and does not have a significant effect on the properties of the QDs. Thus the shape of the QDs in this sample is expected to be similar to those grown in a single layer. In comparison, the correlation between the vertical position of the QDs in the first and second layers in the S = 10 nm sample is indicative of a strong strain interaction between the two. In addition to the nucleation position of the QDs in the second layer, this strain interaction can have an impact on the shape of the second-layer QDs in this sample.

Vertical self-alignment of QDs in multilayers has been explained in terms of preferential diffusion of In adatoms towards the regions directly above the QDs in the underlying layers [85, 86]. In bilayer QDs, the lattice parameter of regions in the spacer layer directly above the 3D islands is closest to that of InAs, leading to the favoured diffusion of In adatoms and nucleation of 3D islands in these regions. In addition to vertical alignment of the QDs, this effect has been shown to result in reduction in the critical thickness and an increase in the dimensions of the second layer QDs [86, 87]. Although it is difficult to determine changes in the critical

thickness and total volume of the QDs in the first and second layers from the image in Fig. 4.12, the observed difference in the shape of the first and second layer QDs may be concurrent with difference in the total volumes of the two.

In order to investigate anisotropy in the shape of the QDs along the $[\bar{1}10]$ and $[110]$ directions, TEM specimens formed of two strips of material, one cleaved along the $[\bar{1}10]$ and the other along the $[110]$ direction, were prepared (see section 2.7.2). Fig 4.13 illustrates this procedure. This allows imaging the QDs along the $[110]$ and $[\bar{1}10]$ orientations, otherwise impossible to achieve by tilting the thin film specimens. Nevertheless, this method involves imaging two different regions of the sample, rather than imaging the same QDs along the desired orientations and comparisons should only be made about the average properties of the QDs in the two different regions viewed.

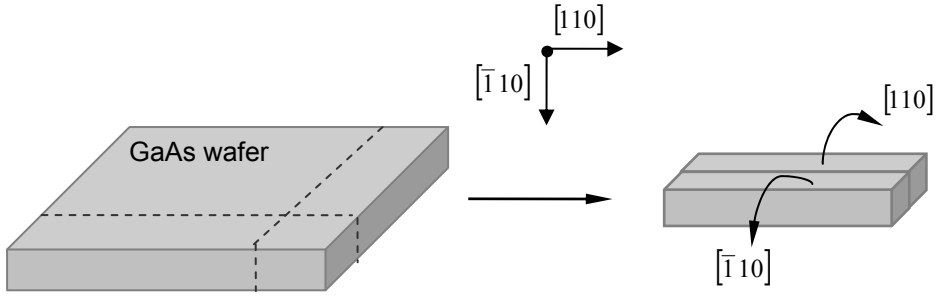


Figure 4.13: A schematic representing the method used to prepare cross-sectional samples along the $[110]$ and $[\bar{1}10]$ directions by cleaving GaAs wafers along two different <110> directions. For details of thin film (S)TEM specimen preparation see section 2.7.2

Fig. 4.14 shows low magnification HAADF STEM images of bilayer QDs in the S = 10 nm sample, taken along the (a)$[110]$ zone-axis and (b) $[\bar{1}10]$ zone-axis. High resolution HAADF STEM images of vertically aligned QDs in this sample oriented along the$[110]$ and$[\bar{1}10]$ zone-axes are displayed in Fig. 4.15. The probe convergence semi-angle, diameter and the inner and outer HAADF collection semi-angles were identical to those stated for Fig. 4.11. Comparing the shape of the QDs projected along the$[110]$ and$[\bar{1}10]$ directions, the bounding facets of the QDs imaged along the $[\bar{1}10]$ direction appear to be less distinct than those imaged along the $[110]$ direction.

This difference is unlikely to be solely a specimen preparation artefact, since the two regions viewed were polished and ion-milled together as parts of the same thin film specimen. The halo-like contrast present around the QDs in Fig. 4.14(b) and 4.15(b) can be evidence of a greater extent of In/Ga intermixing of the QDs along the $[\bar{1}10]$ direction during capping. Further to this, the QDs viewed along the $[\bar{1}10]$ direction on average appear to have a smaller aspect ratio (height to base length ratio) than those viewed along the $[110]$ direction. The average aspect ratios of the QDs appearing in images taken along the $[110]$ and $[\bar{1}10]$orientations were measured to be 0.17 ± 0.1 and 0.14 ± 0.1 respectively. This is in agreement with the *in-situ* STM analysis of the shape and size of the QDs along $[110]$ and $[\bar{1}10]$directions in [55, 69, 70].

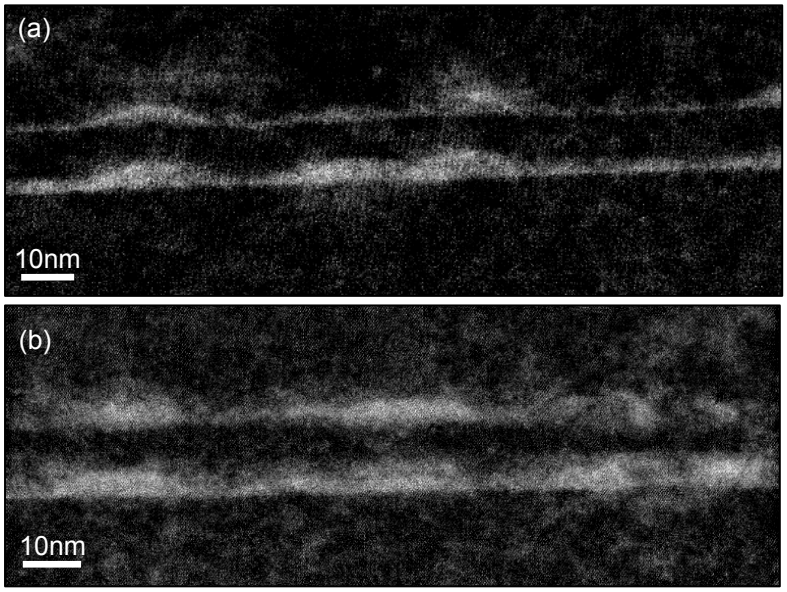


Figure 4.14: HAADF STEM images of the bilayer QDs in the S = 10 nm sample capped with GaAs. Images were taken along the (a) $[110]$ and (b) $[\bar{1}10]$ orientations. See the text for other experimental details.

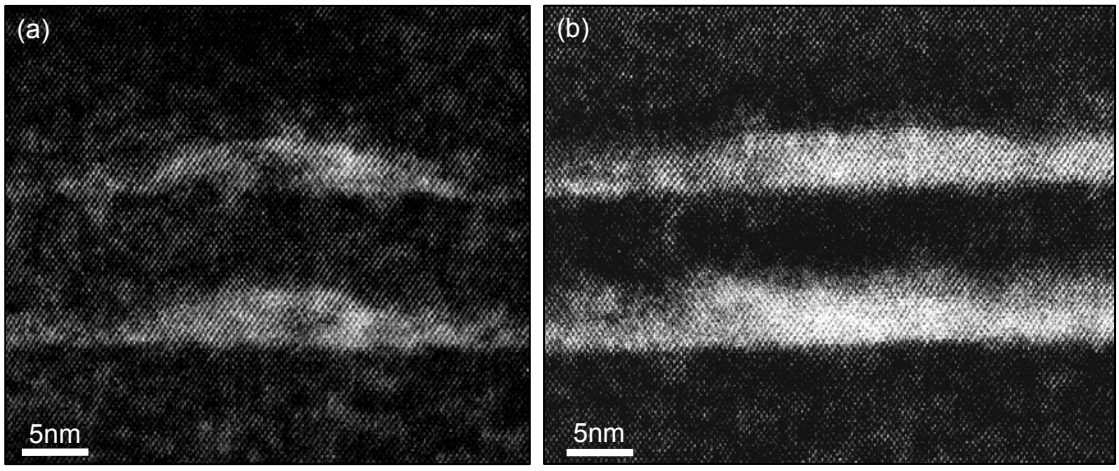


Figure 4.15: HAADF STEM images of bilayer QDs in the S = 10 nm sample capped with GaAs. Images were along the (a) $[110]$ and (b) $[\bar{1}10]$ orientations. See the text for other experimental details.

Images of the QDs in the S = 60 nm sample along the $[110]$ and $[\bar{1}10]$ orientations are displayed in Fig 4.16 and 4.17. Little difference is observed in the shape of the QDs along these two orientations, with the QDs in both orientations appearing as flat-top pyramids. The average aspect ratio of the QDs along the $[110]$ and $[\bar{1}10]$ directions were found to be 0.20 ± 0.1 and 0.19 ± 0.1 respectively. This indicates a lower degree of anisotropy in the shape of the QDs in the S = 60 nm sample than the S = 10 nm sample. However, it is possible that this difference in the shape of the QDs in the S = 60 nm and S = 10 nm samples is due to statistical errors. As mentioned previously, ideal characterisation of the 3D morphology of QDs requires imaging the same QDs along different orientations, (*i.e.* tomography). However, the information presented in this section represents the average shape and dimensions of a limited number of QDs in thin film specimens and therefore, may not be fully representative of the 3D properties of the QDs. Assuming that the results obtained are representative of the true 3D structure of the QDs, the larger degree of anisotropy observed in the shape of the QDs in the S = 10 nm sample compared to that in the S = 60 nm sample can be indicative of stronger chemical and morphological changes occurring to the QDs in this sample during capping. Thus it can be concluded that a larger magnitude of strain energy acts on the QDs in the S = 10 nm sample, since the chemical and morphological evolution of QDs during capping occur as a consequence of the strain energy exerted on them.

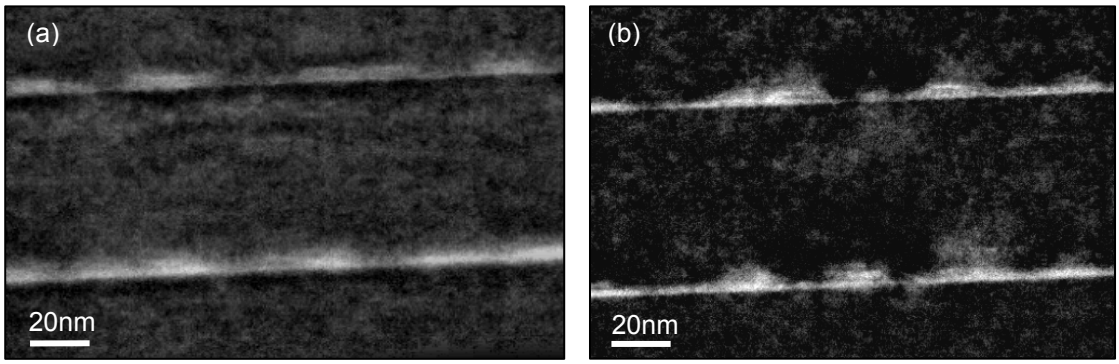


Figure 4.16: HAADF STEM images of the QDs in the sample capped with GaAs and with S = 60 nm viewed along the (a) $[1\,1\,0]$ and (b) $[\bar{1}\,1\,0]$ orientations. See the text for other experimental details.

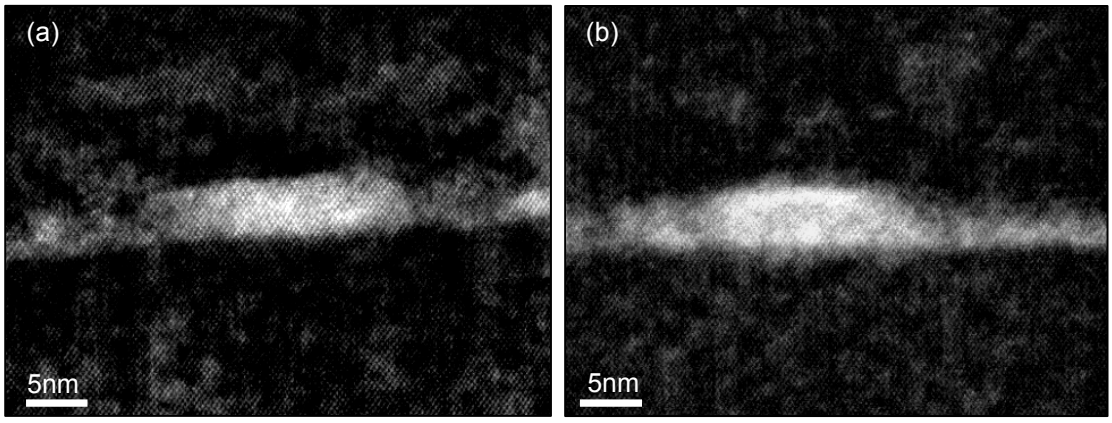


Figure 4.17: HAADF STEM images of the QDs in the sample capped with GaAs and with S = 60 nm viewed along the (a) $[1\,1\,0]$ and (b) $[\bar{1}\,1\,0]$ orientations. See the text for other experimental details.

4.3.5 Size

In general only small areas of thin-film specimens that are of the desired electron-transparent thickness are viewed in (S)TEM. In cross-sectional (S)TEM imaging a further problem arises due to the unknown position of the QDs within the thin film specimen. Therefore, (S)TEM imaging is not considered to be an accurate technique for acquiring statistical information on the average size of the QDs. The accuracy of the measurements can be improved by examining and averaging the dimensions of a large number of QDs. The results presented here are obtained by averaging the measurements carried out on 15-20 QDs in each sample.

A summary of the measurements of the average base length, height and aspect ratio (height to base length) of the second layer QDs in the samples capped with GaAs and with S = 60 nm, 20 nm and 10 nm, imaged along the $[1\bar{1}0]$ direction is provided in Table 4.2. These results indicate that the aspect ratio of the QDs decreases with decreasing the GaAs spacer layer thickness between the bilayer QDs. This can be due to a higher degree of material intermixing and morphological evolution of the QDs in samples with smaller *S* values during the capping step, which leads to more significant elongation and flattening of the QDs in these samples. This further implies that the magnitude of the strain energy acting on the QDs increases with decreasing *S* in bilayers. This is however, contrary to the results of other studies in which the second layer QDs in vertically aligned samples were shown to have larger dimensions due to the reduction in the critical thickness of the QDs in the second layer [86, 87].

Sample	Base length	Height	Aspect ratio
S = 60 nm	24.3 ± 2.5 nm	5.0 ± 1.0 nm	0.20 ± 0.1
S = 20 nm	24.1 ± 2.5 nm	4.3 ± 1.5 nm	0.18 ± 0.1
S = 10 nm	23.9 ± 3.0 nm	4.1 ± 1.5 nm	0.17 ± 0.1

Table 4.2: The average base length, height and aspect ratio of the QDs in the samples capped with GaAs and with *S* = 60 nm, 20 nm and 10 nm, measured along the $[1\bar{1}0]$ direction.

4.3.6 Composition

As discussed in section 4.3.1, the experimental conditions employed here are expected to result in *Z*-contrast imaging in HAADF STEM. Therefore, HAADF STEM images can be used directly to analyse compositional variations within the QDs. If the chemical composition within the QDs was uniform, the resulting intensity of the QDs in *Z*-contrast images was expected to decrease monotonically along the growth direction. This is because the cross-sectional area of the 3D QDs decreases along the growth direction, resulting in a smaller number of In atoms contributing to the intensity of the recorded image. Fig. 4.18 shows a HAADF STEM image of a QD in the *S* = 60 nm sample capped with GaAs along with the intensity profile of the

selected region. Examining the intensity profile of the QD along the growth direction reveals that the intensity decreases initially, reaches a minimum in the middle and increases again towards the top of the QD (see Fig. 4.18(c)). This suggests a non-uniform chemical composition of the QD along the growth direction, with significantly higher In concentrations near the top. This is in agreement with the results of other studies of the local composition of the QDs in [70, 80, 81].

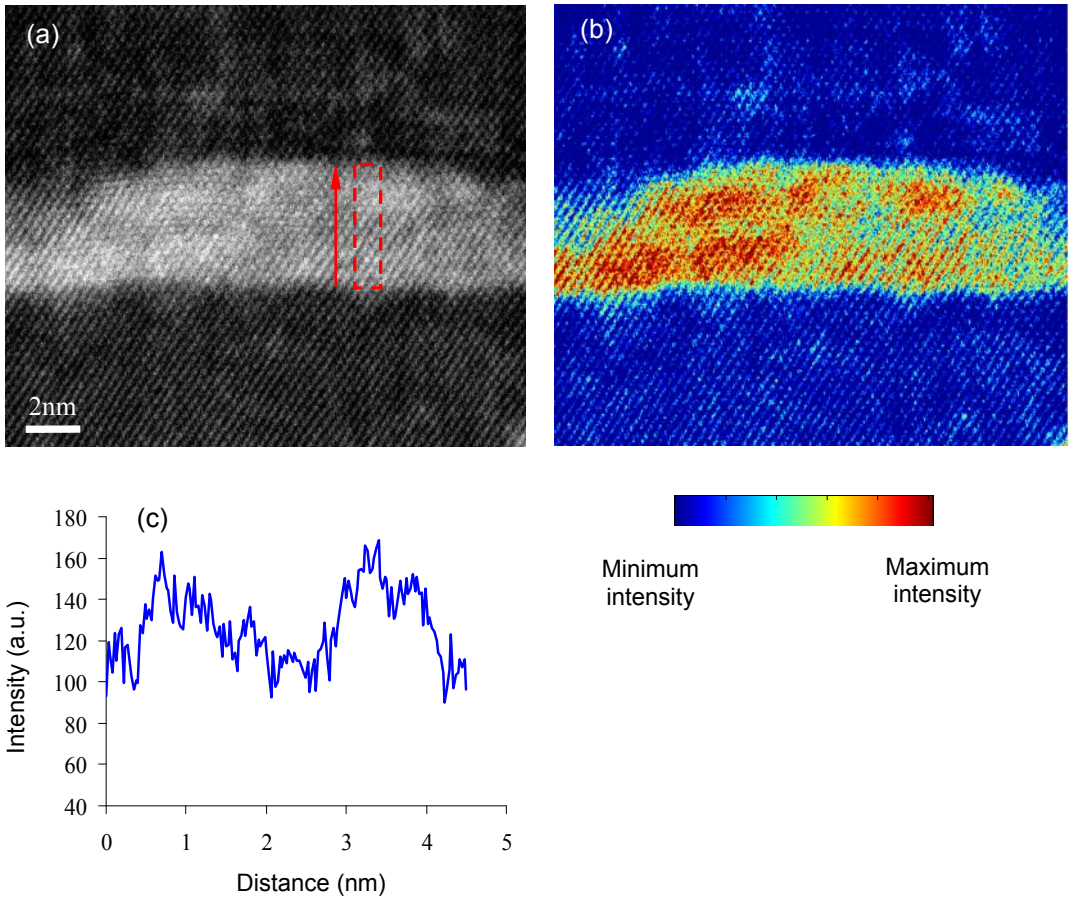


Figure 4.18: (a) A HAADF STEM image of a QD in the sample with S = 60 nm and capped with GaAs. (b) The colour scheme of the image in (a) is changed to that represented by the colour map in order to emphasise intensity variations. (c) The intensity profile of the section along the growth direction of the QD marked in (a).

The intensity profile of the QD in Fig. 4.18(c) contains contributions from the QD and the surrounding GaAs matrix. The contribution from the GaAs matrix can be removed if the 3D shape and the geometry of the QD within the thin film specimen are known. In this example, the QD is assumed to have a square-base pyramid shape of the dimensions displayed in Fig. 4.19(a) and to be fully contained within the thin film

specimen. Since the contrast arising from the change in the thickness of the wedge-shape specimen has been removed, the specimen is assumed to have a constant thickness of 25 nm. Based on the assumptions made, the total intensity at any point of the QD in the image can be defined as the sum of the contributions from the QD material (InAs) and the surrounding matrix (GaAs):

$$I_{total} = I_{GaAs} + I_{InAs} \qquad \textbf{Equation 4.2}$$

The contribution of the surrounding GaAs is estimated by using the average intensity value of GaAs obtained from a region in the image not containing the QD and weighting it according to the estimated thickness of the GaAs matrix at each point along the QD. This is then subtracted from the total image intensity values. The intensity profile of the QD in Fig. 4.18(a) after the subtraction of the contribution from the GaAs matrix is plotted in Fig. 4.19(b). Fig. 4.19(b) also displays the intensity profile of a QD of the same geometry but with a uniform composition. Comparing the two intensity profiles in Fig. 4.19(b) a clear deviation from a homogeneous composition is observed, with significant In deficiency within the first ~ 1 nm of the QD (approximately the wetting layer) and In enrichment near the top of the QD.

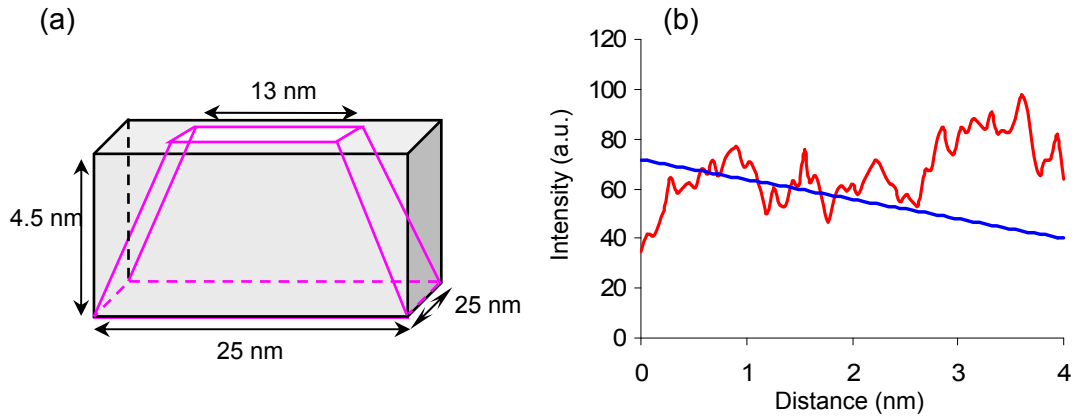


Figure 4.19: (a) A schematic showing the assumptions made about the shape and the geometry of the QD in Fig. 4.18 within the thin-film specimen. (b) The intensity profile of the QD determined after subtracting the contribution from the GaAs matrix and the calculated intensity profile of a QD with the same geometry but with a uniform composition.

Fig. 4.20 shows a HAADF image of a QD in the S = 10 nm sample. The intensity profile of the region along the growth direction of the QD marked in Fig. 4.20(a) is displayed in Fig. 4.20(b). The intensity variation of the QD in this example exhibits a

similar trend to that of the QD in the S = 60 nm sample (see Fig. 4.18). The same procedure as that described for Fig. 4.19 was applied here to remove the contribution from the surrounding GaAs matrix to the image intensity. The 3D shape assumed for the QD is displayed in Fig.4.21. The intensity profile of the selected region along the QD after the subtraction of GaAs contribution and of that calculated for a uniform composition QD of identical dimensions are shown in Fig. 4.21(b). The analysis indicates a similar In distribution to that of the QD in the S = 60 nm sample, with In deficiency in the wetting layer and In enrichment near the top of the QD.

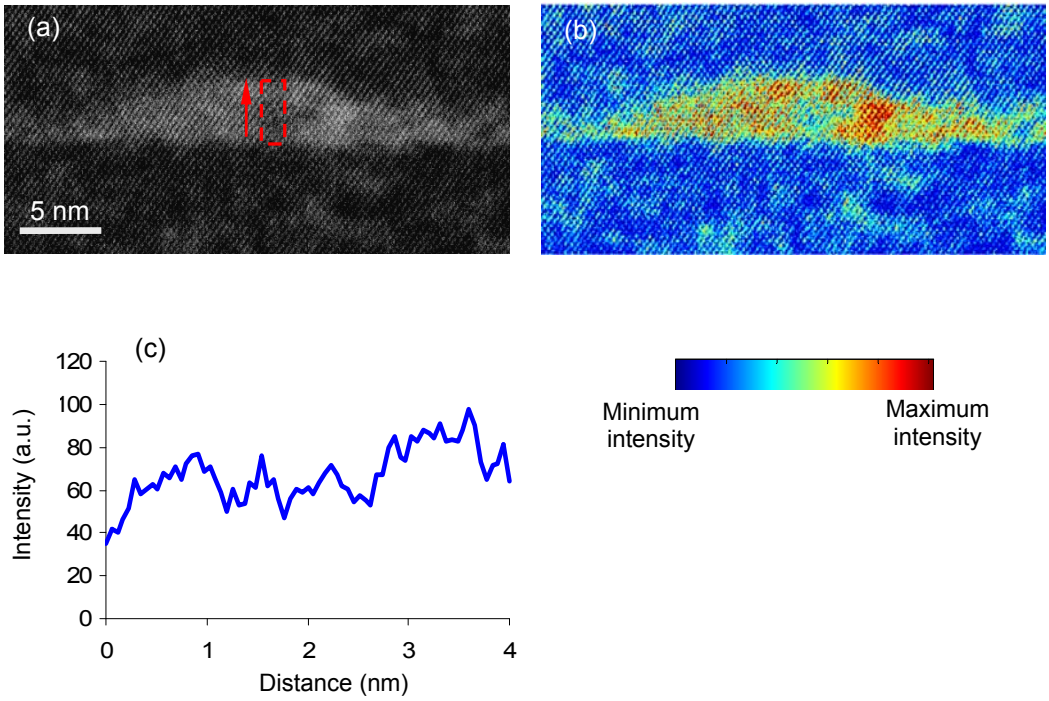


Figure 4.20: (a) A HAADF STEM image of a QD (in the first layer) in the S = 10 nm sample capped with GaAs. (b) The colour scheme of the image in (a) is changed to that represented by the colour map in order to emphasise intensity variations. (c) The intensity profile of the section along the growth direction of the QD marked in (a).

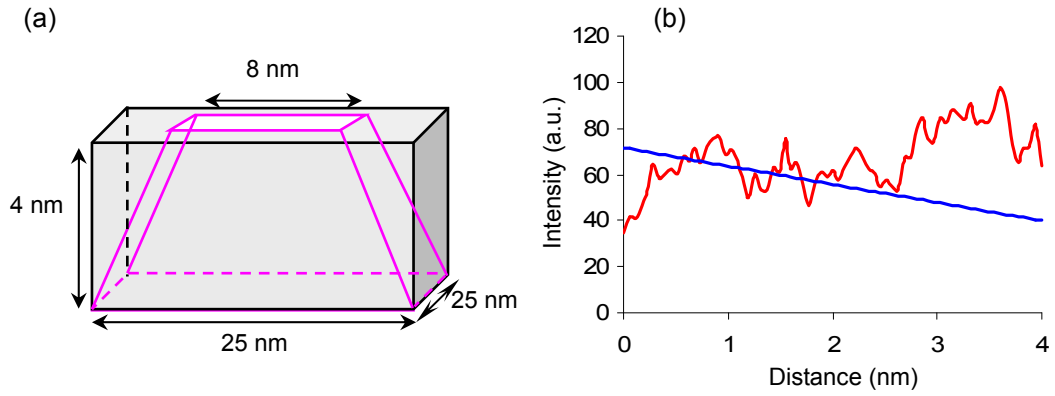


Figure 4.21: (a) A schematic showing the assumptions made about the shape and the geometry of the QD in Fig. 4.20 within the thin-film specimen. (b) The intensity profile of the QD determined after subtracting the contribution from the GaAs matrix and the calculated intensity profile of a QD with the same geometry but with a uniform composition.

4.3.7 GaInAs capping

Given that the changes occurring to the QDs during capping are driven by strain, the choice of the capping material is expected to have a significant effect on the morphology and composition of capped QDs. Fig 4.22 shows HAADF STEM images of bilayer QDs in the S = 10 nm sample with the second layer QD capped with GaAs in (a) and 2 nm of $Ga_{0.9}In_{0.1}As$ prior to further capping with GaAs in (b). The corresponding intensity profiles of the selected regions within the QDs are displayed in Fig. 4.22(c) and (d). The second layer QD in Fig. 4.22(b) appears to have larger dimensions than the second layer QD in Fig. 4.22(b). The base length and height of the QD in Fig. 4.22(a) are measured to be 25 ± 1 nm and 4 ± 1 nm respectively and those of the QD in (b) are measured to be 27 ± 1 nm and 6 ± 1 nm respectively. The average aspect ratio of the second layer QDs in the sample capped with GaInAs was found to be 0.21 ± 0.1 compared to the 0.17 ± 0.1 found for the GaAs capped QDs. In general, a larger aspect ratio was found for the QDs in samples capped with GaInAs than those capped with GaAs.

Comparing the intensity profiles of the QDs in Fig. 4.22, the GaInAs capped QD exhibits a more uniform intensity profile along the growth direction, implying a more uniform chemical composition in the QD capped with GaInAs compared to the QD

capped with GaAs. Although a similar In deficiency in the bottom ~ 1 nm (approximately the wetting layer region) of the QD in Fig. 4.22(b) is observed, the In enrichment near the top of the 3D islands, demonstrated for the QDs capped with GaAs in section 4.3.6, is not seen here. A similar trend was seen in the composition profile of the GaInAs capped QDs in the S = 60 nm and S = 20 nm samples.

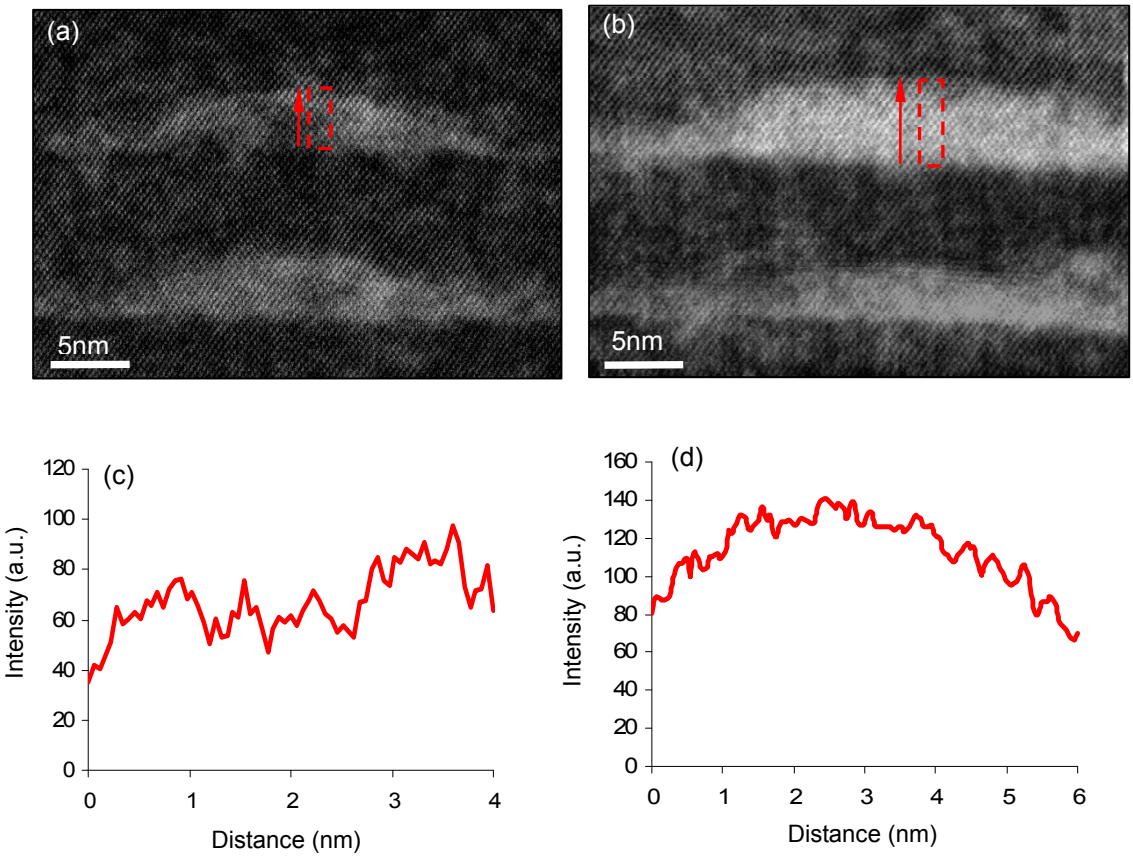


Figure 4.22: HAADF STEM images of the QDs in the S = 10 nm sample capped with GaAs in (a) and 2 nm of $Ga_{0.9}In_{0.1}As$ prior to further capping with GaAs in (b). (c) The intensity profile of the region selected in (a). (d) The intensity profile of the region selected in (b).

The larger aspect ratio and more homogeneous composition of the QDs in samples capped with GaInAs can be evidence of a lower degree of material intermixing and morphological evolution of the QDs during the capping step. Since the lattice mismatch between InAs and GaInAs is smaller than that between InAs and GaAs, a smaller magnitude of strain energy is induced on the QDs by capping and thus, the capping step results in smaller changes to the morphology and composition of the QDs.

4.4 Summary

The results of morphological and compositional investigations of bilayer InAs/GaAs QDs with different GaAs spacer layer thicknesses *S* were presented in this chapter. The nucleation site of the QDs in the second layer was shown to be largely influenced by the thickness of the GaAs spacer layer, with the first- and second-layer QDs in samples with $S < 20$ nm exhibiting a vertical self-alignment. The thickness of the spacer layer was also shown to have an effect on the shape and average size of the second layer QDs. The shape of the QDs was found to be anisotropic along the $[110]$ and $[\bar{1}10]$ directions, as the result of larger extent of In/Ga intermixing occurring along the $[\bar{1}10]$ direction during the capping of the QDs. The local In concentration of the QDs was assessed using *Z*-contrast images in STEM. It was shown that the QDs capped with GaAs exhibit In deficiency near the base (approximately the wetting layer region) and In enrichment near the top. Finally, the effect of capping the QDs with GaInAs instead of GaAs was investigated. It was shown that GaInAs capped QDs exhibit larger aspect ratios and have more uniform compositions compared to GaAs capped QDs. This can be suggestive of a smaller extent of changes occurring to the morphology and composition of QDs during capping with GaInAs.

5 Structural Investigations of GaInNAs / GaAs Quantum Wells

5.1 Introduction

The unusual electronic properties of dilute nitride III-V semiconductor alloys introduced in section 1.2.2, makes them of considerable interest for applications as the active layer of a number of long wavelength (1.3 μm – 1.55 μm) optoelectronic devices. Although significant progress has been made in the development of 1.3 μm emission devices using GaInNAs [*e.g.* 89, 90], the successful application of 1.55 μm emission devices has been hindered by problems arising from increasing the N and In content of the alloys. It has been shown that increasing the N and In concentrations to those required for 1.55 μm emission (typically > 3% for N and > 30% for In) leads to significant lateral material segregation and poor morphology in the quantum wells (QWs), which then results in a substantial decrease in the emission intensity [91-94]. The segregation effects can be attenuated by reducing the growth temperature, but this can further decrease the emission intensity by causing the formation of point defects such as N interstitials [95, 96].

Given the direct effect of the structure, morphology and local composition of the dilute nitride GaInNAs/GaAs QWs on their optical properties, it is of utmost importance to acquire a good understanding of the morphological and compositional changes that occur in QWs of higher In and N content. A number of existing scanning

tunnelling microscopy (STM), reflection high-energy electron diffraction (RHEED) and TEM imaging studies have attempted to characterise the structural properties of GaInNAs/GaAs QWs [*e.g.* 92, 98, 95]. Although several works have explained the morphological changes of the dilute nitride QWs in terms of their composition, little work has so far been carried out to directly examine the composition of these QWs. A summary of the results of the structural and compositional studies of the QWs reported in the literature are provided in the next section. This is followed by the results of (S)TEM and EELS analysis of the morphology and composition of GaInNAs/GaAs QWs carried out in this work and a discussion on their growth mechanism. Finally, the validity of using plasmon mapping to determine changes in the local composition of the QWs is investigated and discussed.

5.1.1 Composition modulation and morphology breakdown of GaInNAs

Several experimental studies have reported a deterioration in the morphology and optical properties of GaInNAs/GaAs QWs with increasing N and In content [91-93, 97, 98]. Clearly-defined pits were observed in *in-situ* plan-view STM images of $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$ (450 ºC growth temperature) after ~ 2 nm material deposition, which developed into a 3D growth mode in the form of surface undulations with further growth [92]. In contrast, GaInAs QWs grown under identical experimental conditions exhibited a 2D growth mode throughout. Similar interface undulations of wavelength of approximately 20 nm were observed in TEM investigations of $Ga_{0.62}In_{0.38}N_{0.023}As_{0.977}$/GaAs QWs (see Fig. 5.1) [99]. Reducing the growth temperature to 400 ºC was shown to decrease the degree of interface roughening of the dilute nitride QWs and maintain a 2D growth mode but also lead to the deterioration of the optical properties of the QWs by creating defects such as N interstitials [99].

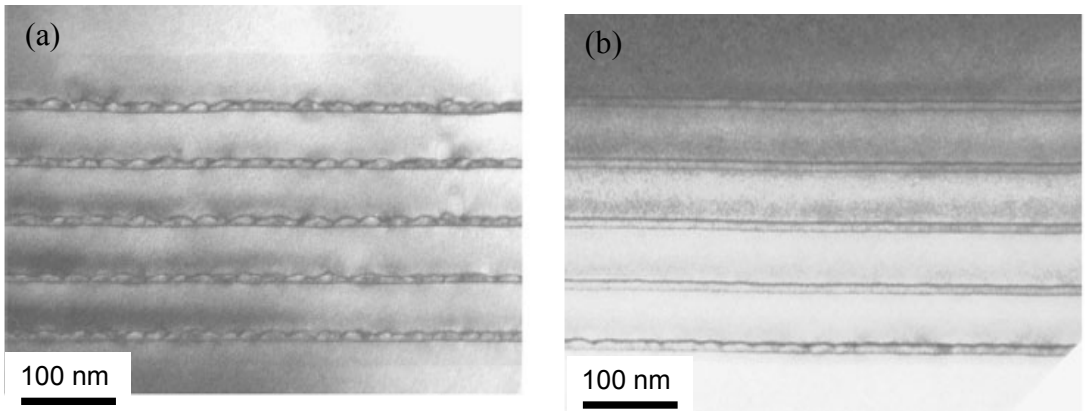


Figure 5.1: 002 DF TEM images of $Ga_{0.62}In_{0.38}N_{0.023}As_{0.977}$/GaAs QWs grown at (a) 460 ºC and (b) 400 ºC, taken from [99].

Theoretical studies of dilute nitride III-V alloys have indicated low miscibility of N in GaAs and GaInAs [100, 101], suggesting the instability of homogeneous GaInNAs alloys at certain temperatures and compositions. In line with the phase separation occurring in metal alloys, the composition modulation in ternary or quaternary semiconductors has often been ascribed to a spinodal decomposition process in which an unstable alloy gradually decomposes into different binary or ternary phases [102, 103]. Several theoretical works have attempted to characterise the thermodynamical stability of GaInNAs in terms of a spinodal decomposition model [*e.g.* 93, 101, 104]. Despite the relative success of such models, it is important to point out that there are clear differences between the spinodal decomposition occurring in bulk metal alloys over a long period of time and the surface diffusion driven, phase separation occurring during the epitaxial growth of semiconductor alloys. Therefore, spinodal decomposition does not seem the ideal term to describe the phase separation process in semiconductors and is not used to explain the phase separation of GaInNAs in this work.

At higher growth temperatures or higher N and In compositions, the metastable GaInNAs alloy is believed to form mutually exclusive In-rich and N-rich regions (occurring within the first few monolayers of growth) [92, 93, 105]. In addition to this lateral phase segregation, the results of reflection high-energy electron diffraction (RHEED) and photoluminescence (PL) studies of GaInNAs QWs in [98] inferred an enhancement in In surface segregation with N incorporation. Surface segregation of In in the ternary GaInAs alloy stems from the difference in the elastic and binding

energies of Ga-As and In-As in forming GaInAs [106]. At the initial stage of growth of GaInAs on GaAs, the In atoms tend to stay at the surface (*i.e.* floating layer) and seek lower energy configurations rather than being buried by the incoming flux. In the case of GaInNAs, the incorporation of N further increases the difference in binding energies for In and Ga forming bonds with nitrides (the difference in binding energy between Ga-N and In-N is 0.31 eV while the difference in binding energy between Ga-As and In-As is 0.08 eV) [98]. This enhances the tendency of In atoms towards surface segregation in GaInNAs compared to GaInAs. The same principle of the large difference between the binding energies of Ga-N and In-N can cause the lateral phase separation of the QWs by promoting the formation of exclusive GaN-rich and InAs-rich regions.

Lateral phase separation of the QWs results in lattice-mismatched GaInAs (compressively strained with respect to the GaAs substrate) and GaNAs (tensile strained with respect to the GaAs substrate) neighbouring regions within the QW. With further growth, the lateral strain energy between the neighbouring regions increases and at a critical thickness a 3D growth mode is adopted in order to lower the total strain energy of the system (similar to the Stranski-Krastanov growth of quantum dots, described in section 1.1.4) [92]. This laterally varying strain energy was represented in [92] by an oscillatory strain field (see Fig. 5.2). Returning to the thermodynamical model of thin film growth described by Eqn. 1.1, at a critical film thickness the additional contribution of the oscillatory strain field to γ_F facilitates a 3D growth mode transition and leads to the formation of surface undulations in the QWs.

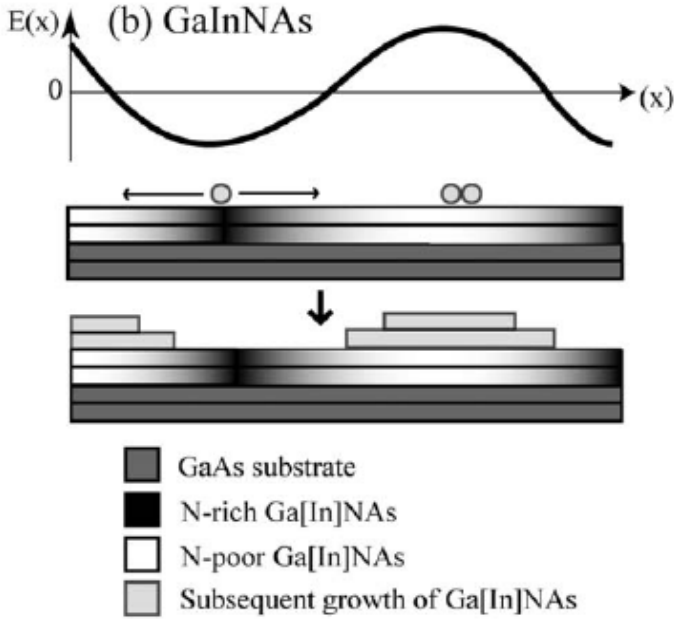


Figure 5.2: A 2D schematic showing the oscillatory strain field due to the neighbouring In-rich and N-rich regions in a spinodally decomposed GaInNAs QW layer (E(x) represents the strain energy of the thin film), taken from [92]. This oscillatory strain field facilitates a 2D to 3D growth mode transition since overgrowth of the tensile strained N-rich regions is inhibited.

Despite the proposed models describing the composition modulation and morphological breakdown of GaInNAs/GaAs QWs, little direct evidence of such composition modulations can be found in the literature and the accuracy of some of the existing results is questionable. The EELS analysis of the compositional inhomogeneities in [107, 108], utilised the In $N_{4,5}$ and Ga $M_{4,5}$ edges and the shift in the plasmon energy to show lateral fluctuations in the In and N composition of the QWs. Given that the In $N_{4,5}$ and Ga $M_{4,5}$ edges (occurring at 18 eV and 20 eV respectively) are close to each other and superimposed on the tail of the plasmon peak (occurring at ~ 16 eV), measuring the relative changes in the intensities of the In $N_{4,5}$ and Ga $M_{4,5}$ edges can be ambiguous. The relative intensities of the In $N_{4,5}$ and Ga $M_{4,5}$ edges were examined in the imaginary part of the dielectric function $\varepsilon_2(E)$, in which the plasmon excitations do not appear. However, as demonstrated in Ch. 6, the relativistic losses present in the energy-loss spectra of semiconductors can introduce artefacts in the dielectric function obtained from EELS. Since the relativistic contributions to the energy-loss spectra were not considered, such artefacts would have been present in the quantitative analysis carried out. It was argued in [108] that

due to the large increase in the effective mass of electrons in dilute nitride GaInNAs QWs (see section 1.2.2), the shift in the plasmon energy primarily reflects the local N composition ($E_p = \hbar\left(\frac{ne^2}{\varepsilon_0 m^*}\right)^{1/2}$, where n is the valence electron number density, e the electron charge, ε_0 the permittivity of free space and m^* the effective electron mass in the material, as described in section 3.2.1). Besides the poor spatial resolution of plasmon mapping, the effect of the QWs interface plasmons was ignored. The experimental and theoretical results presented in section 5.3.4 show that interface plasmon contributions can also cause a shift in the peak plasmon energy and hence, plasmon mapping does not provide an accurate measure of the N composition.

5.2 Experimental details

Samples were grown using a molecular beam epitaxy (MBE) (DCA Instrument) system on epi-ready GaAs (001) substrates. The oxide layer of the substrate was thermally desorbed at 620 ºC, followed by the growth of a 0.5 μm thick GaAs buffer layer at 590 ºC. Active N species were provided by an EPI Unibulb rf plasma source.

Two sets of samples were studied here (see Fig. 5.3). The first set consists of four samples, each sample containing five periods of $Ga_{1-x}In_xN_{0.05}As_{0.95}$ QWs of 8 nm nominal thickness and grown at temperature T_g, where x = 0.15 and 0.3 and T_g = 400 ˚C and 450 ˚C (see Fig. 5.3(a)). The QW layers are separated by GaAs spacer layers of 50 nm nominal thickness. The final QW layers were further overgrown with 200 nm of GaAs. During the growth of the GaAs spacer layer the N plasma source was isolated from the growth chamber by a shutter. However, the shutter is not 100% effective and GaAs spacer layers are expected to contain up to ~ 0.05% N.

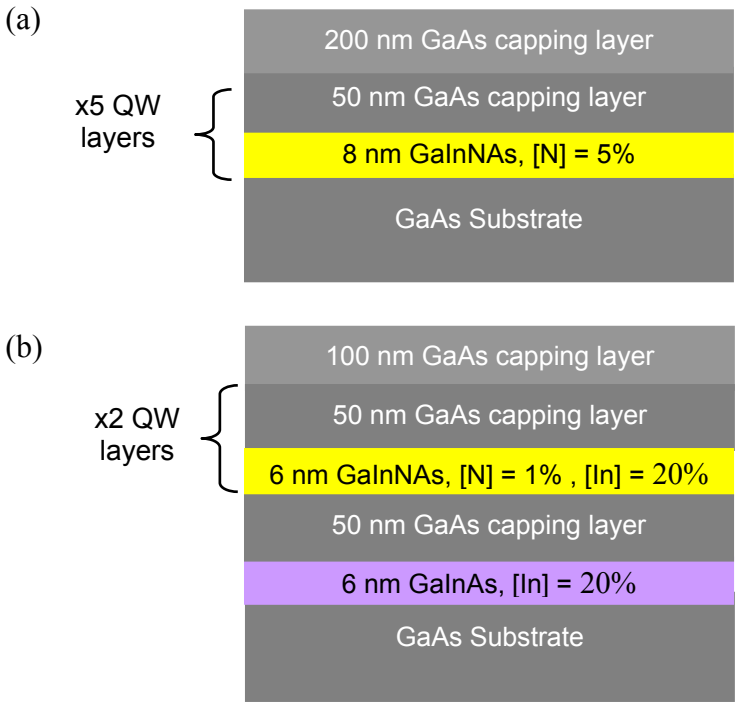


Figure 5.3: A schematic showing the growth details of the QW samples studied in this chapter.

The second set consists of two samples, each sample containing two $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ and one $Ga_{0.8}In_{0.2}As$ QW layers of 6 nm nominal thickness and grown at temperatures T_g = 400 ˚C and 500 ˚C (see Fig. 5.3(b)). The GaInAs layer was grown as a reference to examine the effect of N on the morphology and composition of the QWs. The QW layers are separated by GaAs spacer layers of 50 nm nominal thickness. The final QW layers were further overgrown with 100 nm of GaAs.

Cross-section (S)TEM samples were prepared by mechanical grinding and polishing followed by dimple-grinding and argon ion milling as described in section 2.7.2. STEM imaging and EELS were carried out using the FEI Titan microscope operated at an accelerating voltage of 300 kV and with a FEG extraction voltage of 4 kV. Other experimental details such as the beam convergence angle and spectrometer collection angles are stated for each experiment in the results section.

5.3 Results and Discussion

5.3.1 Structure and morphology

The results of the morphological investigations of the QW samples are presented in this section. These results will be further discussed in section 5.3.3. Fig. 5.4 shows HAADF STEM images of $Ga_{1-x}In_xN_{0.05}As_{0.95}$ QWs grown at temperatures T_g = 400 ºC and 450 ºC and with x = 0.15 and 0.3. Images were acquired with a STEM probe convergence semi-angle of ~ 10 mrad and a probe diameter of ~ 2 Å. The inner and outer HAADF collection semi-angles were ~ 40 mrad and ~ 160 mrad respectively. The QWs in the images appear to have a higher intensity than the surrounding GaAs. This is due to the higher atomic number of In compared to that of Ga. As discussed in section 4.3.1, the experimental conditions employed here result in strong *Z*-contrast HAADF STEM images and hence, these images can be used directly to examine the morphology of the QWs. Samples in 4(a), 4(b) and 4(c) do not exhibit significant interface roughness and the QWs appear to have a good 2D morphology. However, a clear 3D morphology in the form of interface undulations is observed in 4(d) (*i.e.* the sample with the highest T_g and In composition). Fig. 5.5 shows a HAADF STEM image of a QW in this sample along with the intensity profile of the top 2 nm of the QW. The intensity profile reflects the periodicity of the interface undulations of the QWs in this sample. An average wavelength of ~ 31.3 ± 3.1 nm is found for the undulations in this sample by fitting a sinusoidal function through the data in Fig. 5.5(b). The upper interface of the QW appears to oscillate with average amplitude of ~ 3.0 ± 1.0 nm within each period.

It is important to note that cross-sectional STEM images do not provide a statistically accurate measure of the interface roughness of the QWs since only small areas that are of the desired electron transparent thickness in each specimen are imaged. Moreover, cross-sectional (S)TEM images are 2D projections of 3D structures and hence, it is difficult to unambiguously determine the depth of any pits in the QWs from them. The results of *in-situ* surface topography STM studies carried out on these samples [109] confirmed a 2D morphology for the $Ga_{0.85}In_{0.15}N_{0.05}As_{0.95}$ and $Ga_{0.70}In_{0.30}N_{0.05}As_{0.95}$ QWs grown at 400 ºC with root-mean-square (r.m.s) surface

roughness values of 0.177 nm and 0.225 nm (relative to the nominal QW thickness) respectively. Sporadic pits of up to ~ 2 nm deep and r.m.s surface roughness of 0.341 nm were observed in the STM images of $Ga_{0.85}In_{0.15}N_{0.05}As_{0.95}$ QWs grown at 450 ºC, while $Ga_{0.70}In_{0.30}N_{0.05}As_{0.95}$ QWs grown at 450 ºC exhibited a clear 3D morphology, containing pits of up to 6 nm deep and r.m.s surface roughness of 1.167 nm [109]. These results are summarised in Table 5.1.

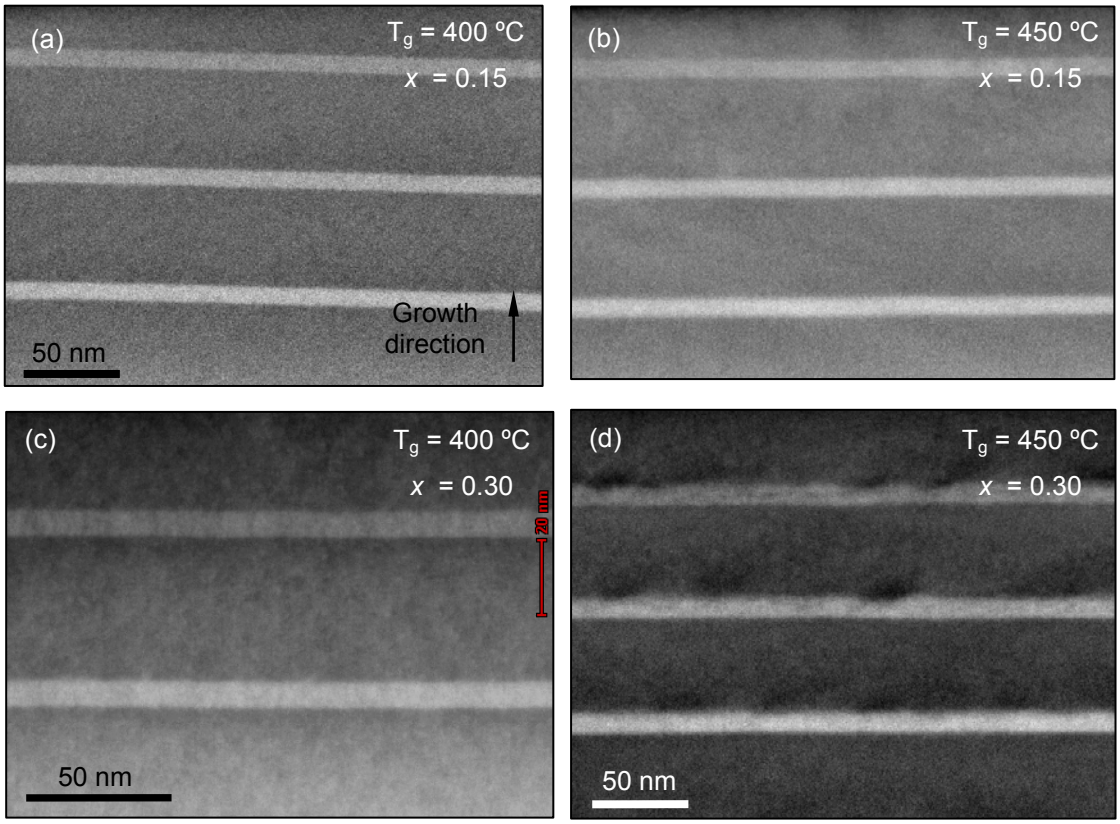


Figure 5.4: Cross-sectional HAADF STEM images of $Ga_{1-x}In_xN_{0.05}As_{0.95}$ QWs along the <110> zone-axis. The samples were grown at (a) and (c) T_g = 400 ºC and (b) and (d) T_g = 450 ºC and with (a) and (b) x = 0.15 and (c) and (d) x = 0.3. See the text for experimental conditions.

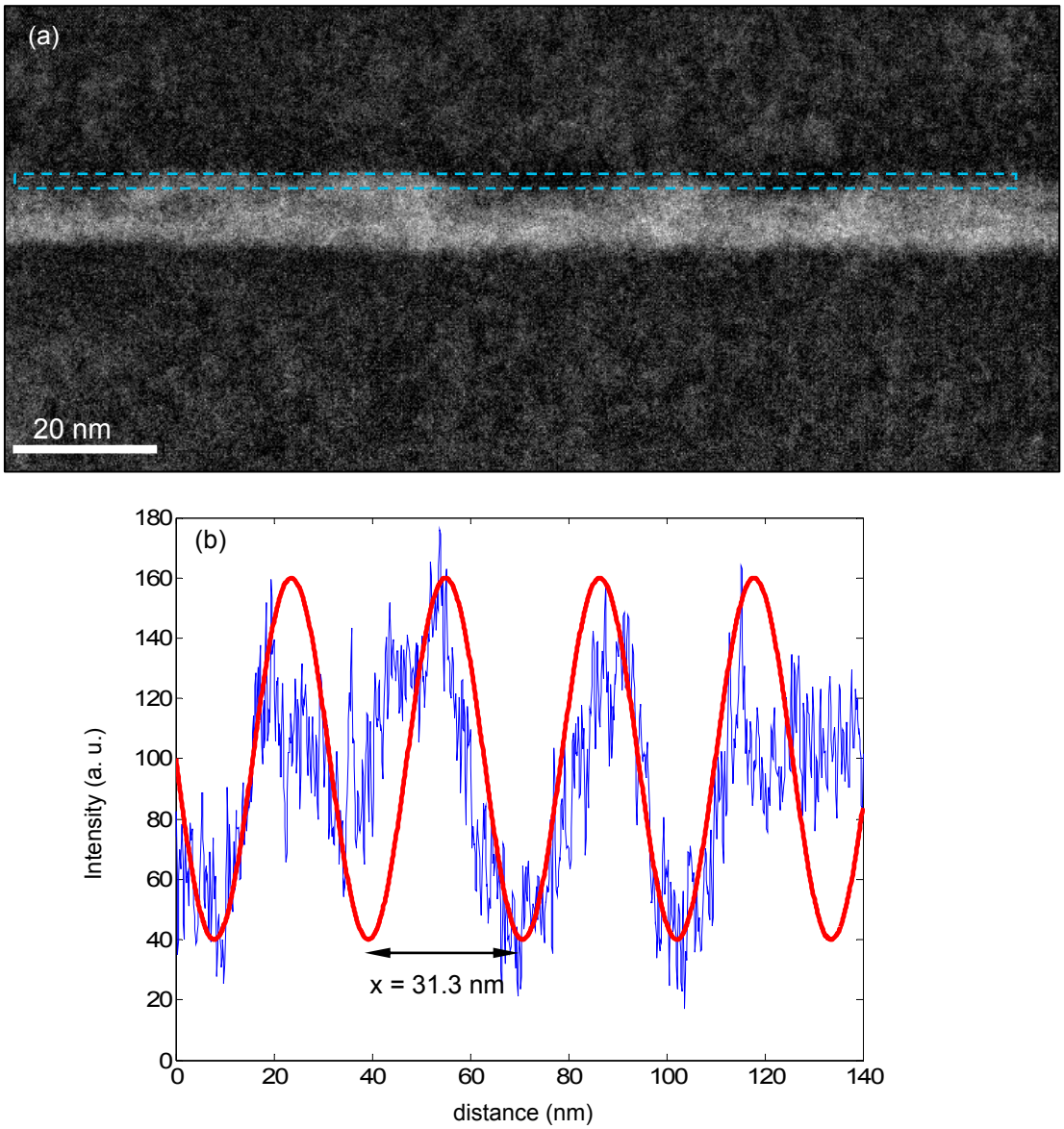


Figure 5.5: (a) Cross-sectional HAADF STEM image of an $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$ QW taken along the <110> zone-axis. See the text for experimental conditions. (b) The intensity profile of the top 2 nm of the QW, indicated by the blue region in 5(a). A sinusoidal function is fitted through the data to determine the average wavelength of the interface undulation of the QWs.

Specimen	Tg (ºC)	Roughness r.m.s. (nm)	Morphology
$Ga_{0.85}In_{0.15}N_{0.05}As_{0.99}$	400	0.177	2D
$Ga_{0.85}In_{0.15}N_{0.05}As_{0.99}$	450	0.341	2D but sporadic pits of up to 2 nm deep observed.
$Ga_{0.70}In_{0.30}N_{0.05}As_{0.99}$	400	0.225	2D
$Ga_{0.70}In_{0.30}N_{0.05}As_{0.99}$	450	1.167	3D with pits of up to 6 nm deep. The interface undulates with a wavelength of ~ 31.3 nm.

Table 5.1: The results of the STEM and STM measurements of the morphology of $Ga_{1-x}In_xN_{0.05}As_{0.95}$ QWs grown at 400 ºC and 450 ºC. The STM results can be found in [109].

Fig. 5.6 shows HAADF STEM images of $Ga_{0.8}In_{0.2}$(N)As QWs grown at (a) T_g = 400 ºC and (b) 500 ºC. The top two QW layers contain 1% N and the bottom QW layer is pure GaInAs. The upper interfaces of the GaInAs QW layers in both samples appear to be of similar smoothness as their lower interfaces. The $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QWs in Fig. 5.6(a) exhibit the same 2D morphology as the GaInAs QW layer in the same sample. However, the $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QWs in Fig. 5.6(b) exhibit a clear 3D morphology with an undulating interface similar to that seen in the sample in Fig. 5.4(d). The wavelength of the interface undulations in Fig. 5.7 is found to be ~ 101.0 ± 9.4 nm.

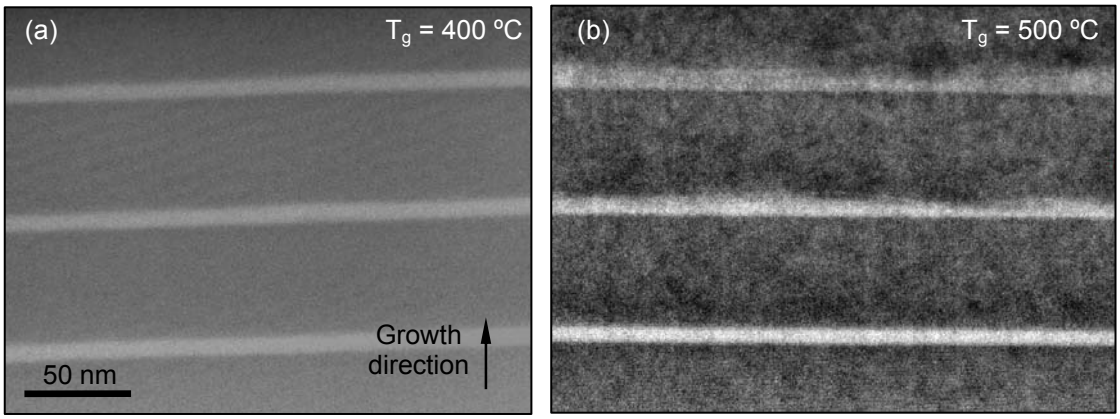


Figure 5.6: Cross-sectional HAADF STEM images of 2 periods of $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QWs (top two layers) and one $Ga_{0.8}In_{0.2}As$ QW layer (bottom layer), grown at (a) 400 ºC and (b) 500 ºC. The specimens were oriented along the <110> zone-axis. See the text for other experimental conditions.

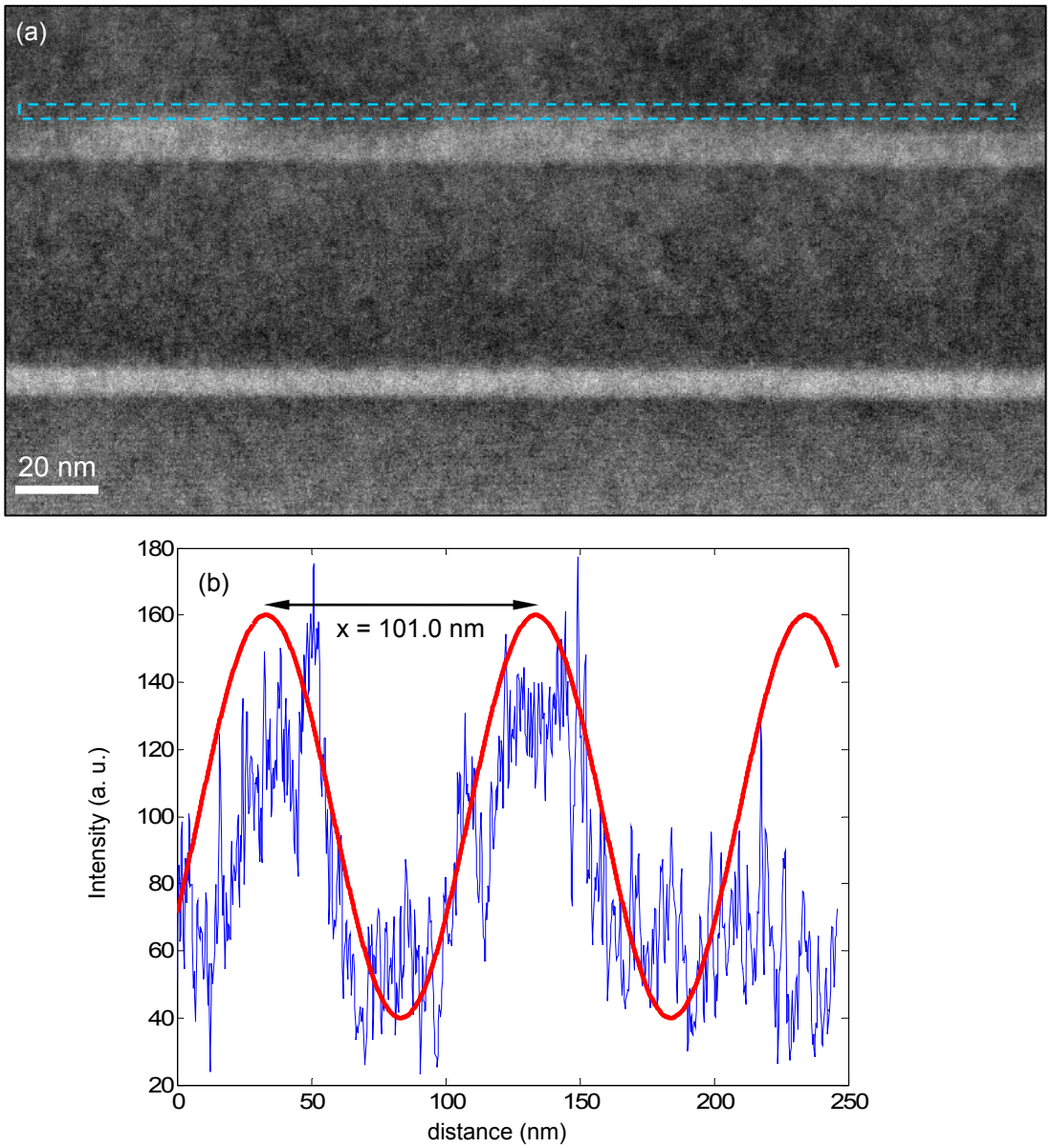


Figure 5.7: (a) Cross-sectional HAADF STEM image of the $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ (top) and $Ga_{0.8}In_{0.2}As$ (bottom) QWs in the sample in Fig. 5.6(b). See the text for experimental conditions. (b) The wavelength of the undulations of the top $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QW is determined by fitting a sinusoidal function through the data.

5.3.2 Composition

5.3.2.1 Z-contrast imaging

Given the strong *Z*-contrast characteristic of HAADF STEM imaging, it is possible to directly interpret changes in image intensity in terms of changes in composition. However, such interpretation becomes more complicated when dealing with alloys composed of multiple elements, with the possibility of composition fluctuations of all constituent elements, since a number of different compositions can result in the same image intensity. In case of the quaternary $Ga_{1-x}In_xN_yAs_{1-y}$ QWs, although the geometry of the structure is simpler compared to that of the QDs discussed in Ch. 4, the presence of four elements in the alloy complicates quantitative compositional analysis of the QWs from HAADF images. Nonetheless, regions of highest intensity in images of $Ga_{1-x}In_xN_yAs_{1-y}$ QWs can be associated with high In concentrations and those of lowest intensity to poor In concentrations.

Images of $Ga_{1-x}In_xN_{0.05}As_{0.95}$ QWs with x = 0.15 and 0.3 grown at 400 ºC and 450 ºC are shown in Fig. 5.8. The colour setting of the images were changed from grey scale to that displayed by the colour map in order to emphasise intensity variations within the QWs. The intensity profiles of the middle 2 nm of the QWs along the lateral direction are shown in Fig.5.8(e) – (h). A wider change of intensity between the QWs and the surrounding GaAs is observed in Fig. 5.8(c) and 5.8(d), due to the higher In content of these samples compared to those in Fig. 5.8(a) and (b). A background intensity variation due to specimen surface damage is present in all images (see Fig. 5.8(i)). The intensity variations in Fig. 5.8(e) and (f) are of similar order to that of the background intensity and hence, it is difficult to resolve any lateral compositional modulation in these samples. Nonetheless, the intensity profile in Fig. 5.8(e) shows less modulation than that in Fig. 5.8(f), indicating a more uniform composition profile in this sample. A more evident periodic lateral intensity modulation is seen in Fig. 5.8(g) and 8(h). This implies a periodic lateral In segregation in these samples. The amplitude of the intensity oscillations in Fig. 5.8(h) is larger than that in Fig. 5.8(g), implying a more considerable lateral In segregation in this sample compared to the sample in Fig. 5.8(c). Average wavelengths of ~ 12.28 ± 1.8 nm and ~ 19.39 ± 3.6 nm were found for the periodic intensity modulations in Fig. 5.8(g) and 8(h) respectively.

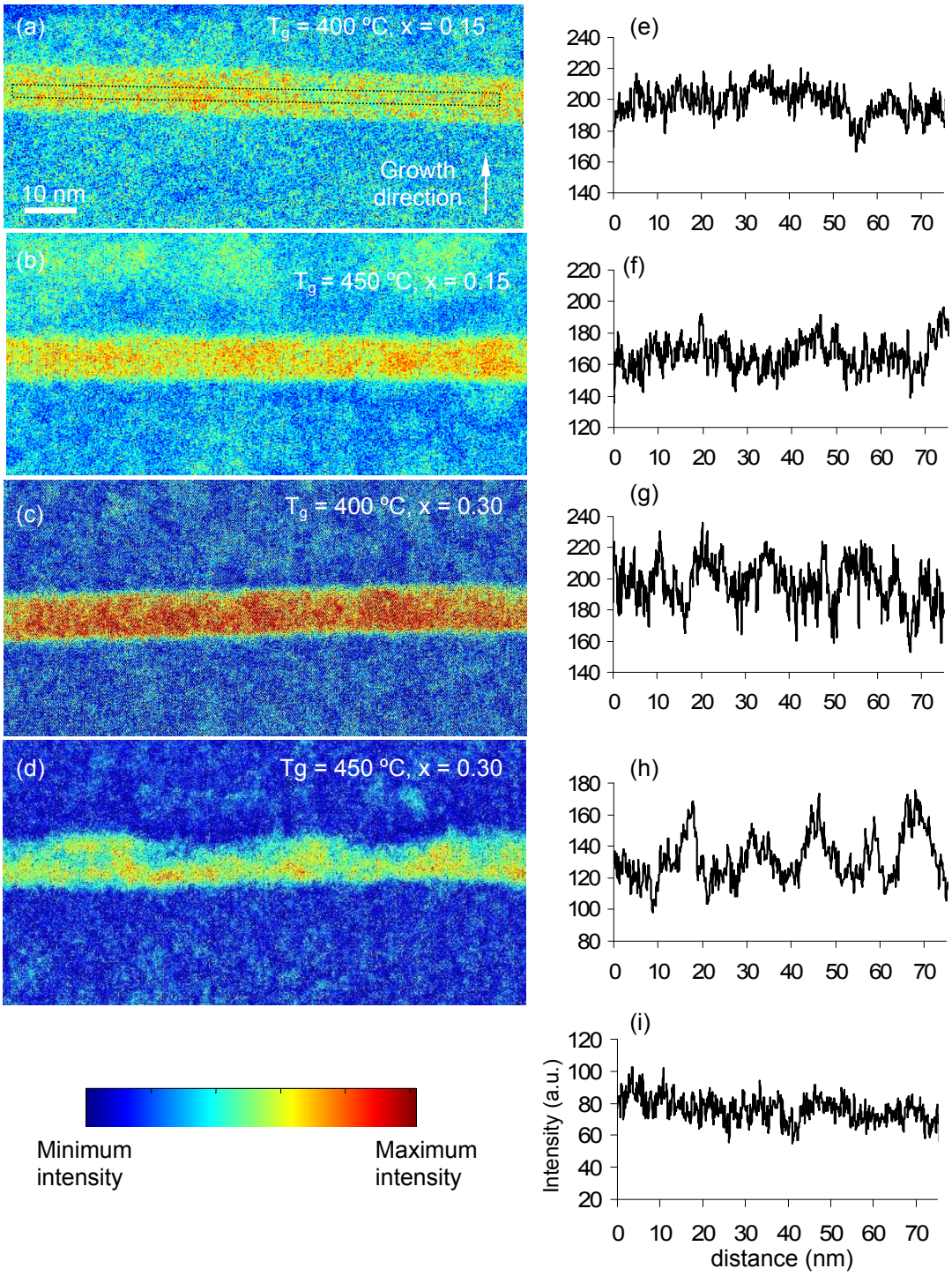


Figure 5.8: HAADF STEM images of $Ga_{1-x}In_xN_{0.05}As_{0.95}$ QWs grown at temperature T_g, with (a) x = 0.15 and T_g = 400 ºC, (b) x = 0.15 and T_g = 450 ºC, (c) x = 0.3 and T_g = 400 ºC and (d) x = 0.3 and T_g = 450 ºC. The colour setting was changed from grey scale to that shown by the colour map. The intensity profiles of the middle 2 nm of the QWs (see the box drawn in (a)) in images (a) – (d) are plotted in (e) – (h) respectively and that of the background in (i). See the text for experimental conditions.

The intensity profiles of the top, middle and bottom 2 nm of the QW in Fig. 5.8(d) are displayed in Fig. 5.9(b). The intensity profiles of the bottom and middle sections exhibit the same periodic intensity variation with similar wavelengths but the intensity profile of the middle section has a larger amplitude. This demonstrates that lateral phase separation of the QWs is present throughout their growth but becomes more significant with further growth of the QWs. An average wavelength of ~ 19.39 ± 3.6 nm is found for the intensity oscillations of the middle section by fitting a sinusoidal function through the data in Fig. 5.9(c). Comparing the intensity profiles of the bottom and middle sections of the QW with that of the top 2 nm section (see Fig. 5.9(b)), shows the wavelength of the intensity oscillations of the top profile (see Fig. 5.5) is almost twice the wavelength of the middle and bottom profiles, with the maxima and minima of the top intensity profile coinciding with maxima of the middle and bottom profiles. Given that the top 2 nm intensity profile of the QW reflects its undulating interface morphology, this result implies that the interface undulations of the QWs in this sample have a wavelength approximately twice that of the lateral In composition modulations, with the crests and troughs of the undulations occurring at points of maximum In composition.

A high resolution HAADF image of one period of the interface undulations of the sample in Fig. 5.9 is displayed in Fig. 5.10. Neighbouring regions of high and low intensity are clearly seen in the middle section of the QW (see Fig. 5.10(d)). In addition to this, the dark regions within the QW exhibit a non-uniform intensity profile along the growth direction (see Fig. 5.10(c)), revealing a minimum In composition in the middle region of the QW. While an intensifying phase separation along the growth direction of the QW was demonstrated in Fig. 5.9(b), this result also indicates an In interface segregation in this sample.

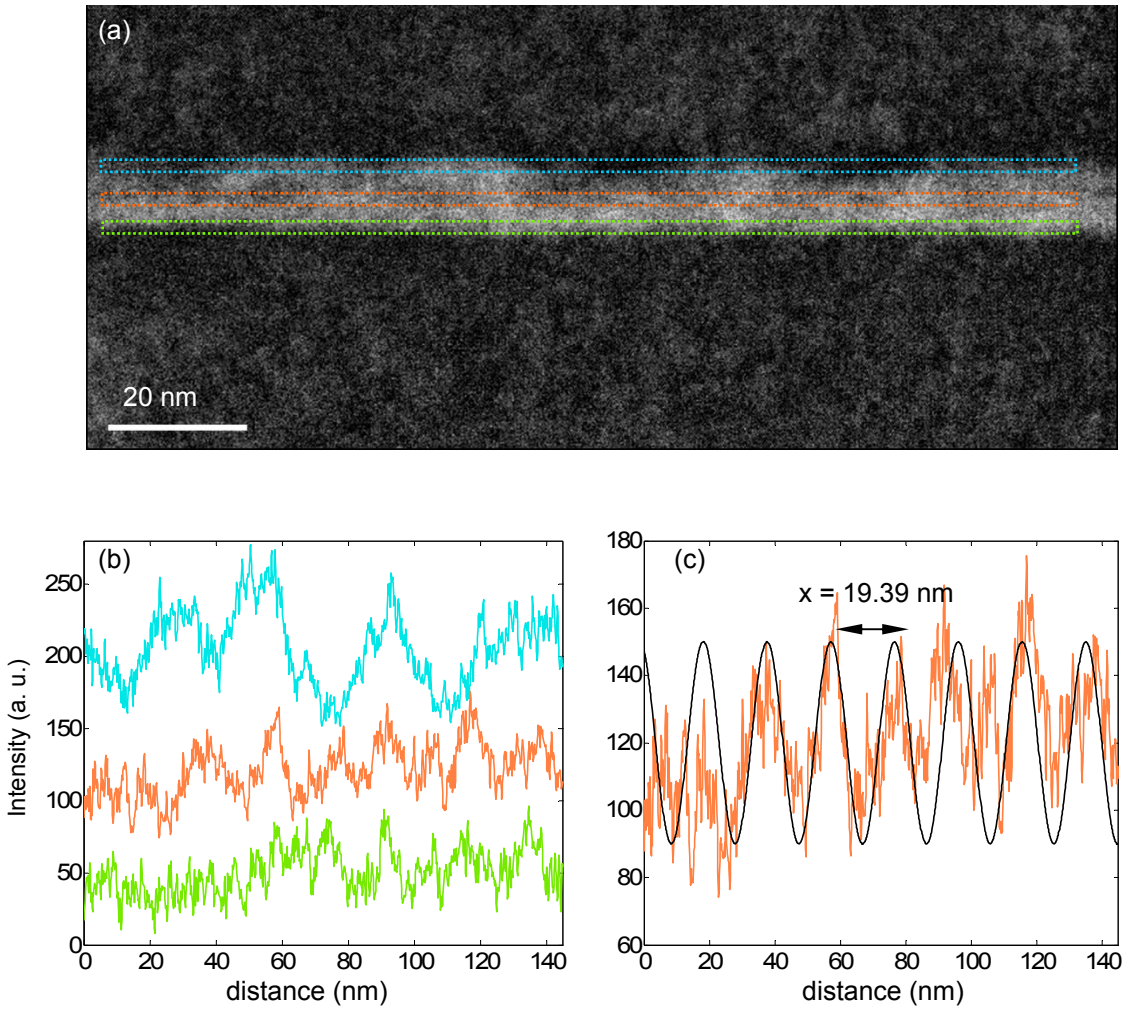


Figure 5.9: (a) HAADF STEM image of $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$ QW grown at 450 ºC. The specimen was oriented along the <110> zone-axis. See the text for experimental conditions. (b) The intensity profiles of the bottom, middle and top 2 nm of the QW (regions selected by the green, orange and turquoise boxes in (a) respectively), indicative of the lateral In composition modulations and the undulating interface morphology of the QW. (c) The wavelength of the intensity profile of the middle 2 nm of the QW found by fitting a sinusoidal function through the data.

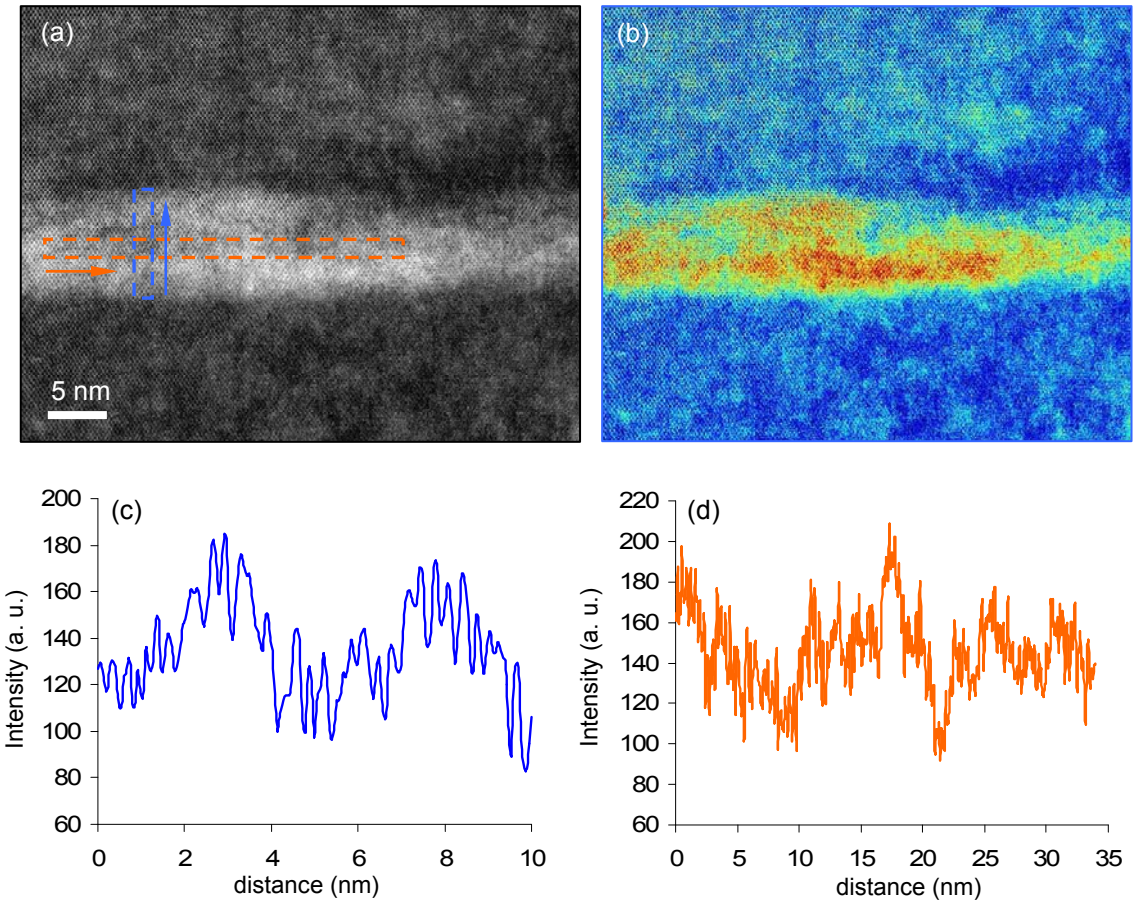


Figure 5.10: (a) HAADF STEM image of $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$ QW grown at 450 ºC. The specimen was oriented along the <110> orientation. See the text for experimental conditions. (b) The grey scale colour setting of the image in (a) changed to that defined by the colour map in Fig. 5.8. (c) and (d) The intensity profiles of the areas selected by the blue and orange boxes in (a) respectively.

HAADF STEM images of $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QWs grown at 400 ºC and 500 ºC are shown in Fig. 5.11(a) and (b) respectively (top QW layers). The bottom QW layers in these images are $Ga_{0.8}In_{0.2}As$. No evident lateral intensity modulation is observed in the $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ or the $Ga_{0.8}In_{0.2}As$ QWs in sample in Fig. 5.11(a). In comparison, a considerable variation is observed in the intensity profile of the $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QW in Fig. 5.11(b), indicating a lateral segregation of In in the QW. This In composition modulation is caused by the presence of N in the QW since the $Ga_{0.8}In_{0.2}As$ QW in this sample does not exhibit a similar intensity variation. A closer examination of the intensity profile of the $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QW in this sample reveals similar neighbouring regions of high and low intensity along the lateral direction of the QW (see Fig. 5.12) to that seen in Fig. 5.10(c). The intensity profile of this sample along the growth direction also demonstrates a similar phase separation and In interface segregation to that in Fig. 5.10(b).

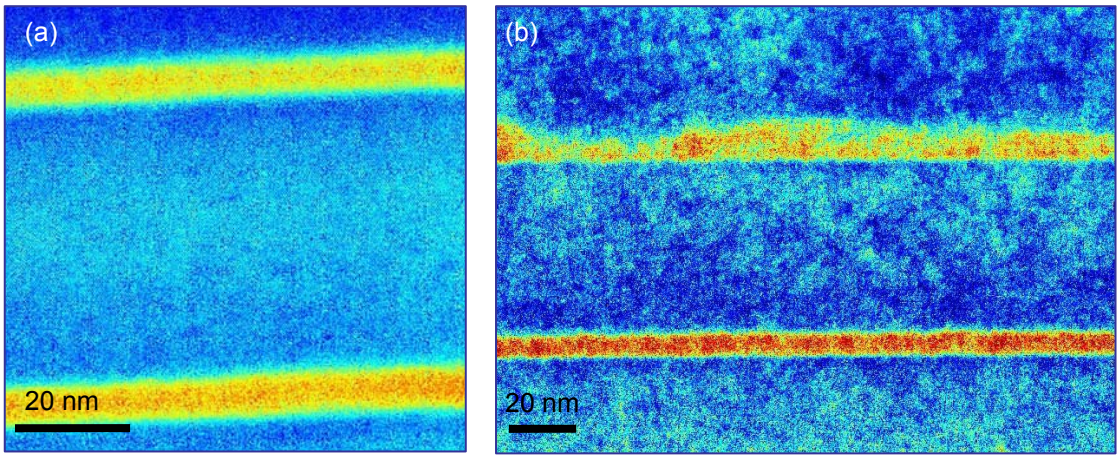


Figure 5.11: HAADF STEM images of $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QWs (top layer QWs) and $Ga_{0.8}In_{0.2}As$ QWs (bottom layer QWs) grown at (a) 400 ºC and (b) and 500 ºC. The specimens were oriented along the <110> orientation. The grey scale colour scheme of the images was changed to that defined by the colour map in Fig. 5.8. See the text for other experimental conditions.

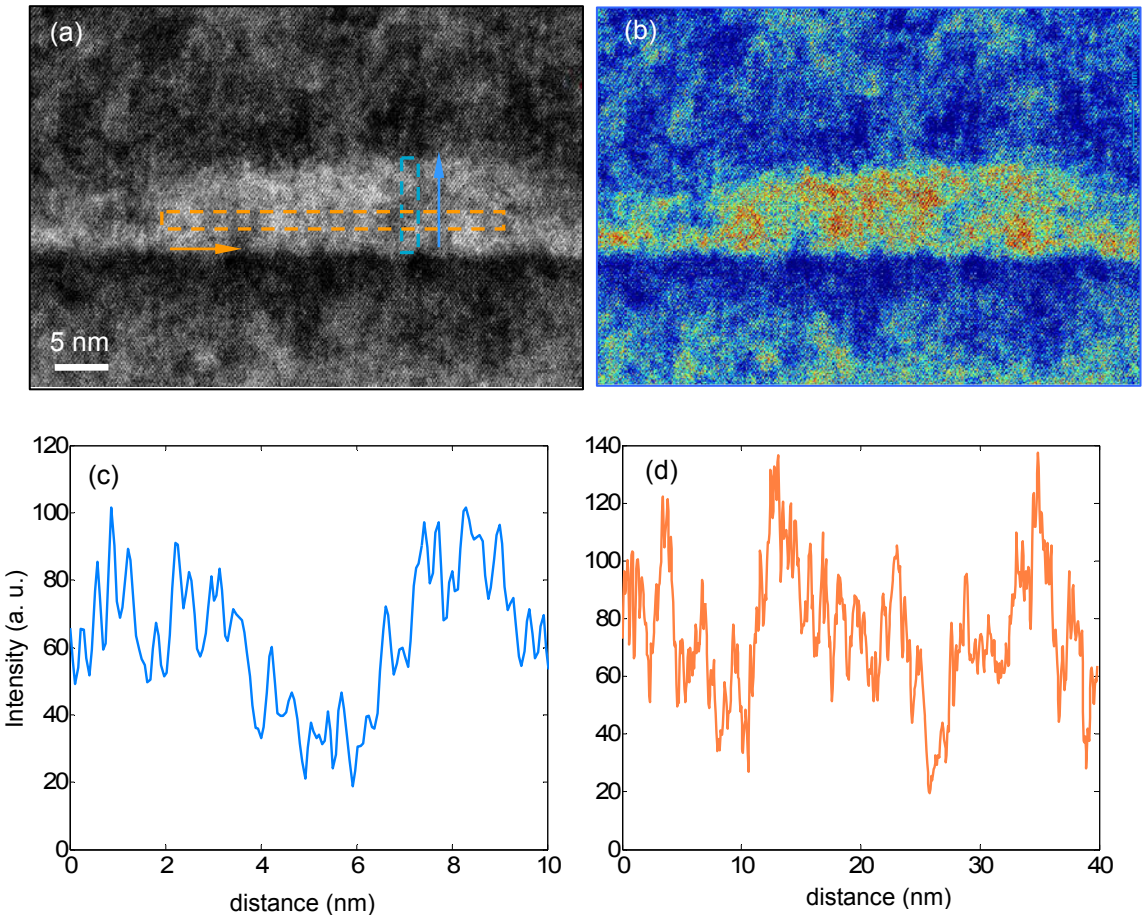


Figure 5.12: (a) and (b) HAADF STEM image of $Ga_{0.8}In_{0.2}N_{0.01}As_{0.99}$ QW grown at 500 ºC. The specimens were oriented along the <110> orientation. See the text for other experimental conditions (c) and (d) The intensity profiles of the regions selected by the blue and orange boxes in (a) respectively.

5.3.2.2 EELS

The results of EELS quantitative analysis based on the In $M_{4,5}$ edge signal from the QWs are presented in this section. Ideal compositional characterisation of the QWs entails obtaining maps of all the constituent elements by a high spatial resolution analytical technique such as EELS. However, in practice such an analysis was constrained by statistical limitations such as the low jump-ratio of Ga, As and In core-loss edges and the low concentration of N in the QWs (see section 3.5.2). Ga and As $L_{2,3}$ and In $M_{4,5}$ edges (occurring at 1115 eV, 1323 eV and 443 eV, respectively) feature a broad shape, with the edge maximum occurring at an energy as high as ~ 100 eV above the onset. This results in the relatively low jump-ratio of these edges

(the ratio of the initial rise of the edge and the preceding background), which then makes the detection of these edges challenging. Longer acquisition times and large probe currents had to be adopted in order to improve the detection of these edges. In case of the Ga and As $L_{2,3}$ edges, these experimental conditions lead to significant specimen drift and beam damage, limiting the accuracy of the acquired data. In contrast, the N K edge (occurring at ~ 401 eV) has a saw-tooth shape and a considerably larger jump-ratio than those of Ga, In and As edges. However, the small fraction of N atoms per unit area of the QWs proved to be below the minimum detectable atomic fraction of the system (see section 3.5.2) and despite attempts, the N K edge was not observed in the recorded spectra.

Fig. 5.13 shows a 6 x 2 spectrum image (each pixel is ~ 4.5 nm x 4.5 nm) of the area selected in Fig 5.13(a) of the $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$ QW sample grown at 450 ºC. The probe convergence semi-angle and diameter were ~ 9.5 mrad and ~ 2 nm respectively, the HAADF detector's inner and outer collection semi-angles ~ 65 mrad and ~ 260 mrad respectively, the spectrometer collection semi-angle ~ 10 mrad and the energy-dispersion of the acquisitions 0.1 eV per channel. The corresponding In $M_{4,5}$ edge from each collection point in the image are displayed in Fig. 5.13(b). The signal due to the In $M_{4,5}$ edge in each spectrum was found by fitting a power law function through the background intensity and then subtracting it from the spectrum. Assuming that the thickness of the specimen stays unchanged over the spectrum image area, the relative intensity of the extracted In $M_{4,5}$ edges can be used to measure variations in the In concentration. The area bound by the In $M_{4,5}$ signal and the energy axis for the spectra in Fig. 5.13(b) are plotted in Fig. 5.13(c). The results indicate neighbouring In-rich and In-deficient regions along the lateral and the growth directions of the QWs. This is in agreement with the results of *Z*-contrast imaging of this sample presented in section 3.2.1. Further to this, the changes in the In composition of regions in the first and second columns of the spectrum image (top-half and bottom-half of the QW along the growth direction, respectively) do not appear to be in phase with each other at every point. For example, an In-deficient region in pixel position (3, 1) is met by an In-rich region in pixel position (3, 2), while pixel positions (5, 1) and (5, 2) both appear to be In-rich compared to their neighbouring regions. Although the length over which this data was acquired is insufficient to comment on the periodic length of In

composition modulations in the top-half and bottom-half of the QW, this result can be due to the different periodicities of the top and middle sections of the QWs in this sample obtained from *Z*-contrast imaging in section 3.2.1.

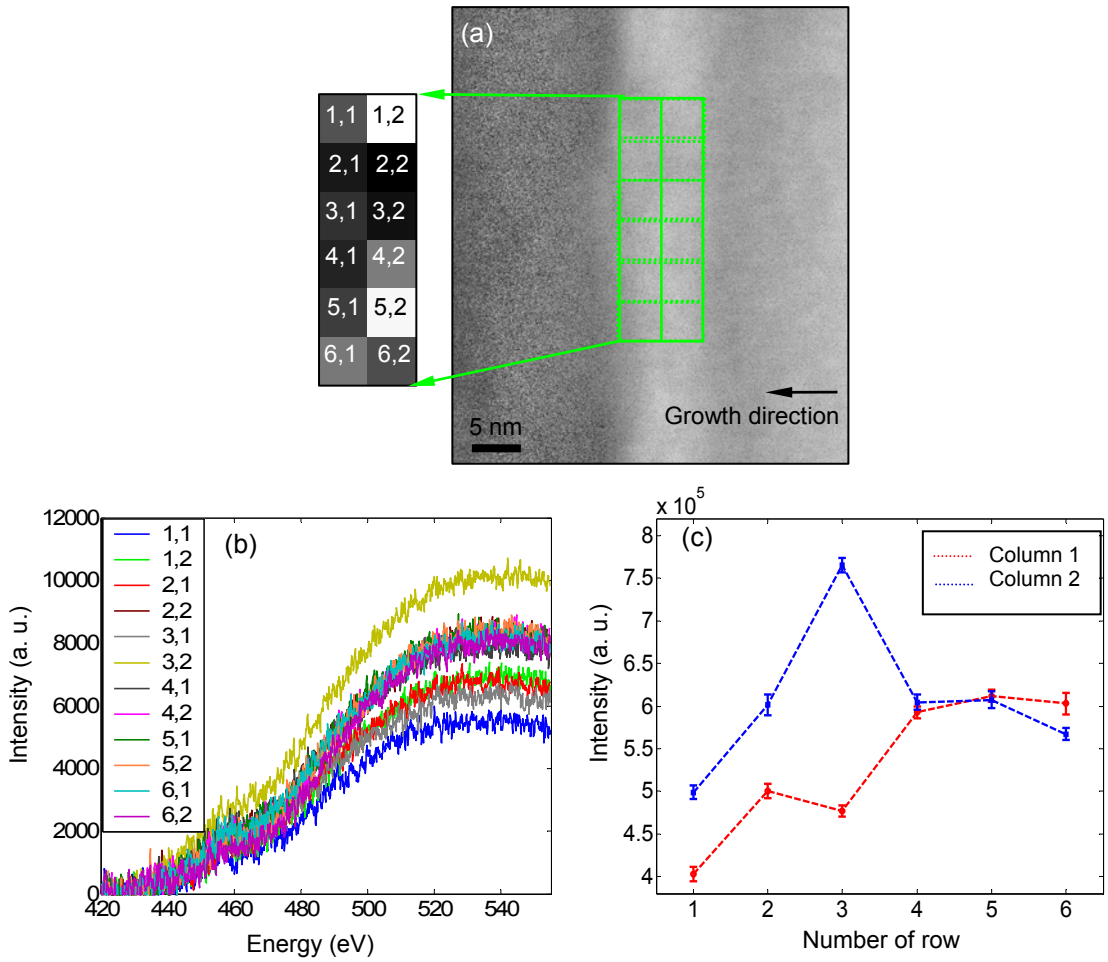


Figure 5.13: (a) An EELS spectrum image of the area shown in the HAADF image of the $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$ QW with T_g = 450 ºC. See the text for experimental conditions of the acquisition. (b) The background subtracted In $M_{4,5}$ edges from each pixel of the spectrum image. (c) The area bound by the In $M_{4,5}$ signal and the E-axis in (b) found for the spectra in the first and second columns of the spectrum image plotted as a function of the pixel position of the spectra along the length of the QW.

5.3.3 Discussion

The results of the morphological and compositional investigations of the dilute nitride QWs in sections 5.1 and 5.2 can be summarised in the following points:

I. Increasing the growth temperature T_g, or the In content of GaInNAs QWs leads to a breakdown of their 2D morphology and formation of periodic interface undulations. Increasing T_g increases the wavelength of these interface undulations.

II. Comparison with identically grown nitrogen-free GaInAs QWs confirms these morphological changes to be due to N incorporation. The intensity of these morphological changes is expected to increase with increasing the N content.

III. GaInNAs QWs with 3D interface morphology exhibit composition modulations in the form of periodic lateral In-poor and In-rich regions. A similar lateral composition modulation is detected in a QW sample of high In content but low growth temperature which has a 2D morphology. The composition modulation in this sample has a smaller amplitude and wavelength than that in the sample with 3D morphology.

IV. In QWs with composition modulations, such modulations exist throughout their thickness (along the growth direction) but intensify along the growth direction of the QWs. In QWs with a 3D morphology a degree of In segregation at the top interface is also found to exist.

V. Unlike composition modulations, interface undulations appear to occur beyond a critical thickness, below which the QWs have a 2D growth.

VI. The interface undulations and composition modulations in samples that exhibit both do not have the same wavelength. The interface undulations in samples studied were shown to have a wavelength approximately twice as long as that of the composition modulations, with the crests of the undulations consistently occurring at In-rich regions.

The concepts of instability of GaInNAs alloys and the phase separation into neighbouring In-rich and N-rich regions were introduced in section 5.1. Although it was not possible to obtain information on the N composition changes of the QWs by EELS, the periodic In composition profile demonstrated by *Z*-contrast imaging and EELS is evidence of such phase separation. Within the first few monolayers of growth, mutually exclusive In-rich and N-rich regions are formed in the alloy, whilst maintaining a 2D growth mode. This phase separation intensifies at higher growth temperatures or In compositions, as predicted by the theoretical studies of the thermodynamical stability of GaNAs and GaInNAs [100, 105]. Moreover, the enhanced migration length of adatoms at higher temperatures increases the lateral distance between the neighbouring In-rich and N-rich regions. Since incorporation of N and In into GaAs results in a decrease and increase in lattice constant respectively, the formation of tensile strained N-rich regions and compressively strained In-rich regions induces a modulating strain energy across the epitaxially grown alloy. This was represented in [92] by an oscillatory strain field. The amplitude and wavelength of this oscillatory strain field, determined by the amount of N and In present in the neighbouring regions and the lateral distance between them respectively, increases with increasing In composition and growth temperature of the QWs, as displayed in Fig. 5.14.

The amplitude of this oscillatory strain field continues to increase with further material deposition and in effect, increases the tendency of the newly deposited In and N adatoms to segregate away from each other. This explains the intensified In lateral segregation in the middle 2 nm of the QW compared to that in the bottom 2 nm of the QW measured in Fig. 5.9. However, similar to the Stranski-Krastanov growth mechanism introduced in section 1.1.4, beyond a certain thickness, the strain energy induced in the film can no longer be accommodated by a 2D growth mode and at some stage during the growth a 3D growth mode is adopted. In the case of the samples exhibiting interface undulations, this critical thickness is below the nominal thickness of the QWs and therefore, a strong lateral phase separation combined with a 3D morphology is observed in these samples. In the case of the samples with a 2D morphology, the magnitude of the strain energy of the film throughout the growth is not sufficient to cause a 2D to 3D growth mode transition. Nevertheless, such a

transition is expected to take place in these samples at film thicknesses beyond those studied here.

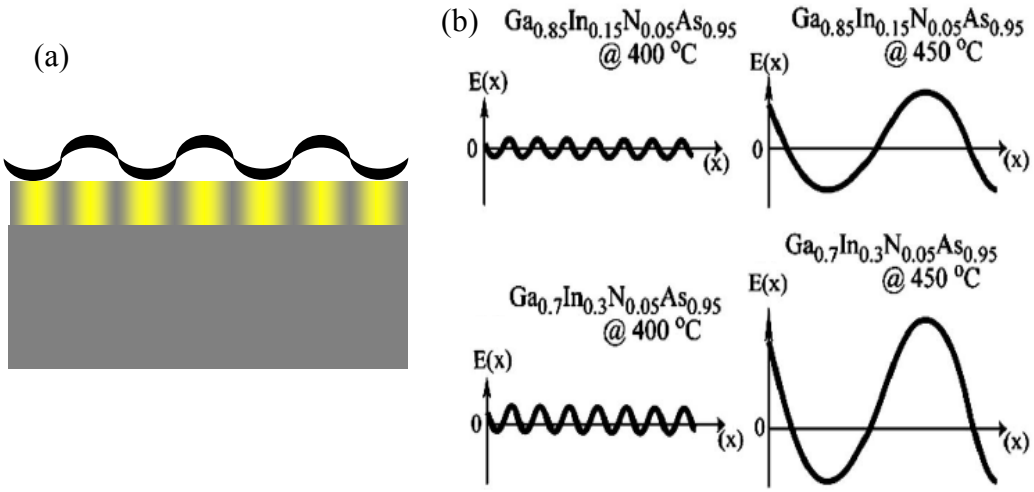


Figure 5.14: (a) A schematic representing the composition modulation of the QWs, further developing into the formation of surface undulations. (b) Schematic representations of the 1D oscillatory strain field of the QWs during their 2D growth mode. The amplitude and wavelength of the oscillations are determined by the In content and growth temperature of the QWs.

While the composition modulation of the QWs is caused by phase separation of the GaInNAs alloy, interface undulations are caused by the oscillatory strain field described above. The two phenomena have previously been assumed to have the same wavelengths and be in phase [104, 109]. However, the results of the measurements carried out in this work showed that they have different wavelengths, with the undulations having a wavelength almost twice as long as that of the composition modulations. Given that the wavelengths of the composition modulations and interface undulations indicate the characteristic lengths of the phase separation and the strain field present in the QWs respectively, the measurements here reveals that the strain field induced in the QWs due to phase separation extend to almost twice as far as the length of the phase separation. Further theoretical studies are required in order to verify and fully explain this observation.

Although a mechanism based on the oscillatory strain field of phase separated QWs has been used to explain the morphological changes occurring to QWs of higher In

content and growth temperature here, no theoretical studies of this model has been carried out to predict the relationship between the 3D growth of QWs and the oscillatory strain field. Therefore, further investigations are required in order to verify and fully explain the results obtained. In addition to this, a better understanding of the 3D phase separation and surface undulations of the QWs is required in order to interpret images of the QWs more accurately. (S)TEM images are 2D projections of 3D structures and accurate interpretation of them often requires knowledge of the 3D structures they represent. No distinction has been made in the results presented in this chapter between the composition modulation of QWs along different orientations. However, given the orientation-dependent surface diffusion length of In adatom on a (001) GaAs surface discussed in Ch. 4, an anisotropy in the phase separation of the QWs seems also possible and should be investigated. An additional factor that needs to be considered is the possible morphological changes introduced as the result of overgrowing the QWs with a material of smaller lattice parameter such as GaAs.

5.3.4 Plasmon mapping

In recent years several works have used the shift in the plasmon energy of the QWs or QDs to map changes in the local composition or strain of these structures [*e.g.* 78, 107, 108]. However, interface plasmon excitations, neglected in these analyses, can also produce a shift in plasmon energy and interfere with the results of plasmon mapping of QW and QD structures. The results of simulations and experiments investigating the effect of interface plasmon excitations on the energy-loss spectrum of QWs are provided in this section. Simulations based on the expression for the probability of energy-loss near planar interfaces introduced in section 3.3.2.2, were carried out in Matlab using the code provided in Appendix A. Empirical dielectric functions of GaAs, InAs and GaInAs, derived from EELS, were used as inputs for the simulations.

Fig. 5.15 shows simulated energy-loss spectra of GaAs with the electron beam travelling at a distance x from an interface with InAs. The spectra obtained for x = 1000 nm and 10 nm are identical to the energy-loss spectrum of bulk GaAs. However, decreasing the distance of the beam from the interface causes a redshift in the plasmon energy. This is the result of the contribution of the interface plasmon term in Eqn. 3.15 to the energy-loss spectrum. A redshift of ~ 0.05 eV is observed in the plasmon peak energy of the x ~ 4 nm spectrum (0.05 eV is the energy dispersion per channel typically used in low-loss EELS data acquisition in this work and hence, is the smallest experimentally measurable shift in plasmon energy). The redshift of the plasmon peak increases with decreasing x until the plasmon energy reaches that of InAs when the beam travels in the InAs medium and sufficiently away from the interface. A similar trend is observed in the simulated energy-loss spectra of a GaAs – $Ga_{0.7}In_{0.3}As$ interface with the beam travelling in GaAs at distance x away from the interface (see Fig. 5.16). Decreasing x produces a smaller redshift in the plasmon energy here compared to that seen in Fig. 5.15 due to the lower In content of the material forming the interface with GaAs in this example. A redshift of ~ 0.05 eV of the plasmon energy is observed with x ~ 1.5 nm.

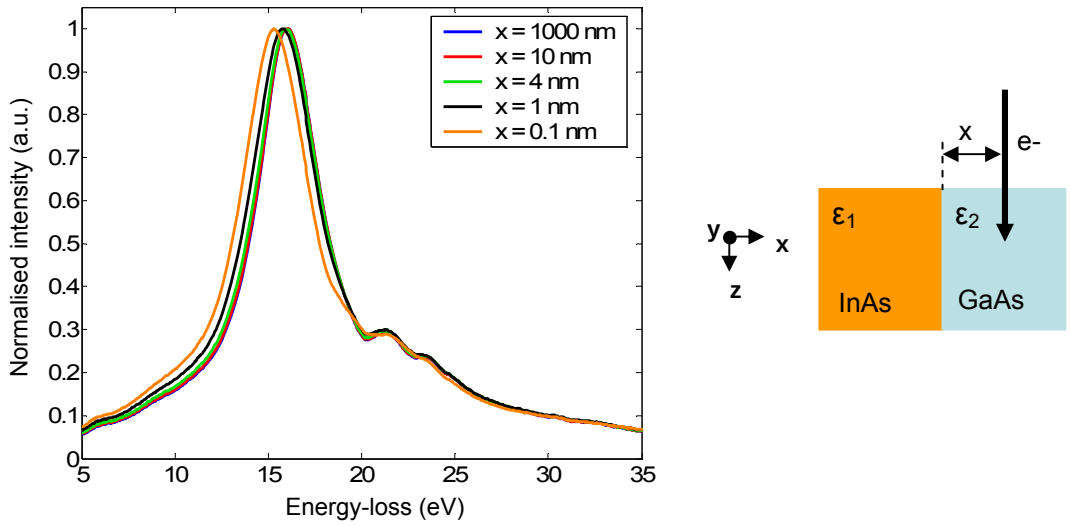


Figure 5.15: Simulated energy-loss spectra of GaAs near a planar interface with InAs for different electron beam distances from the interface *x*, using Eqn. 3.15. An accelerating voltage of 300 kV and spectrometer collection semi-angle of 10 mrad were used for the simulations.

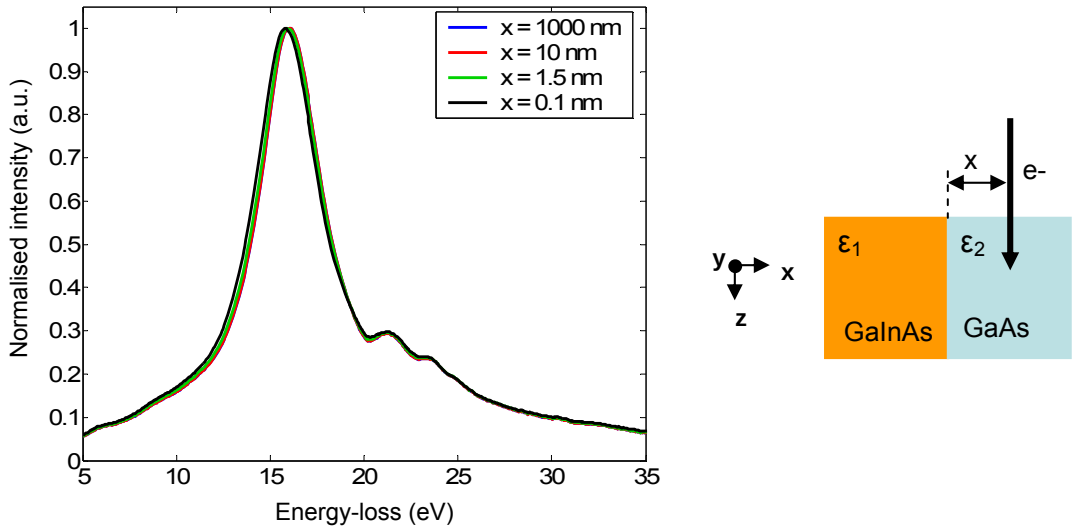


Figure 5.16: Simulated energy-loss spectra of GaAs near a planar interface with $Ga_{0.7}In_{0.3}As$ for different electron beam distances from the interface *x*, using Eqn. 3.15. An accelerating voltage of 300 kV and spectrometer collection semi-angle of 10 mrad were used for the simulations.

In order to experimentally investigate the changes to the energy-loss spectrum of GaAs near a planar interface with a QW, an EELS line-spectrum was acquired perpendicular to the interface (see Fig. 5.17). The experiment was carried out on the $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$ QW sample grown at 400 ºC. The electron probe convergence semi-angle and diameter were ~ 9.5 mrad and ~ 2 Å respectively, the HAADF detector's inner and outer collection semi-angles ~ 65 mrad and 260 mrad respectively, the spectrometer collection semi-angle ~ 10 mrad and the energy-dispersion of the acquisitions 0.05 eV per channel. The plasmon regions of the energy-loss spectra acquired from each point on the line-spectrum are displayed in Fig. 5.17(b). The plasmon peak energies E_p and the HAADF STEM intensity values of the points on the line-spectrum are plotted as a function of distance from the GaAs – QW interface x in Fig. 5.17(c). The position of the interface was determined as the median of the upward slope of the intensity profile of the interface region along the direction of the line-spectrum. The points on the left-hand side of the interface marked in Fig. 5.17(a) are denoted with a negative x-value and those on the right-hand side of the interface with a positive x-value.

The energy-loss spectra recorded from the first five points of the line-spectrum (-9.3 nm < x < -4.0 nm) exhibit a constant E_p of 16.25 ± 0.05 eV. A clear red-shift of the plasmon peak is observed at x = -2.7 nm, with further reduction in E_p with increasing x in the following collection points on the line-spectrum. This gives a localisation distance of ~ 3.35 ± 0.65 nm for the interface plasmon excitation in this sample. One source of error in the experimental results can be the GaAs – QW interface roughness. In addition to introducing errors in determining the position of the interface, interface roughness can also alter the probability of energy-loss by interface plasmon excitation since this phenomenon is affected by the shape of the interface [*e.g.* 44].

A further contributing factor to the decrease in the plasmon peak energy near the GaAs – QW interface in the experimental data is the tensile strain present in GaAs due to the underlying QW. This tensile strain causes a decrease in the number density of valence electrons and thereby produces a redshift in E_p (see section 3.3.2.1). Similarly, a change in the band gap of the material can lead to a shift in plasmon peak energy. These effects have been previously used to comment on the local strain or

chemistry of the QWs and QDs [*e.g.* 78, 107, 108]. The results of the simulations and experiments carried out in this work show that interface plasmon excitation can have a localisation distance of similar length to the dimensions of the QWs and QDs. Therefore, variation in plasmon peak energy cannot be viewed to be characteristic of changes in the local strain or chemistry of these structures and interface plasmon losses should also be considered when analysing E_p variation within nanometre-scale structures.

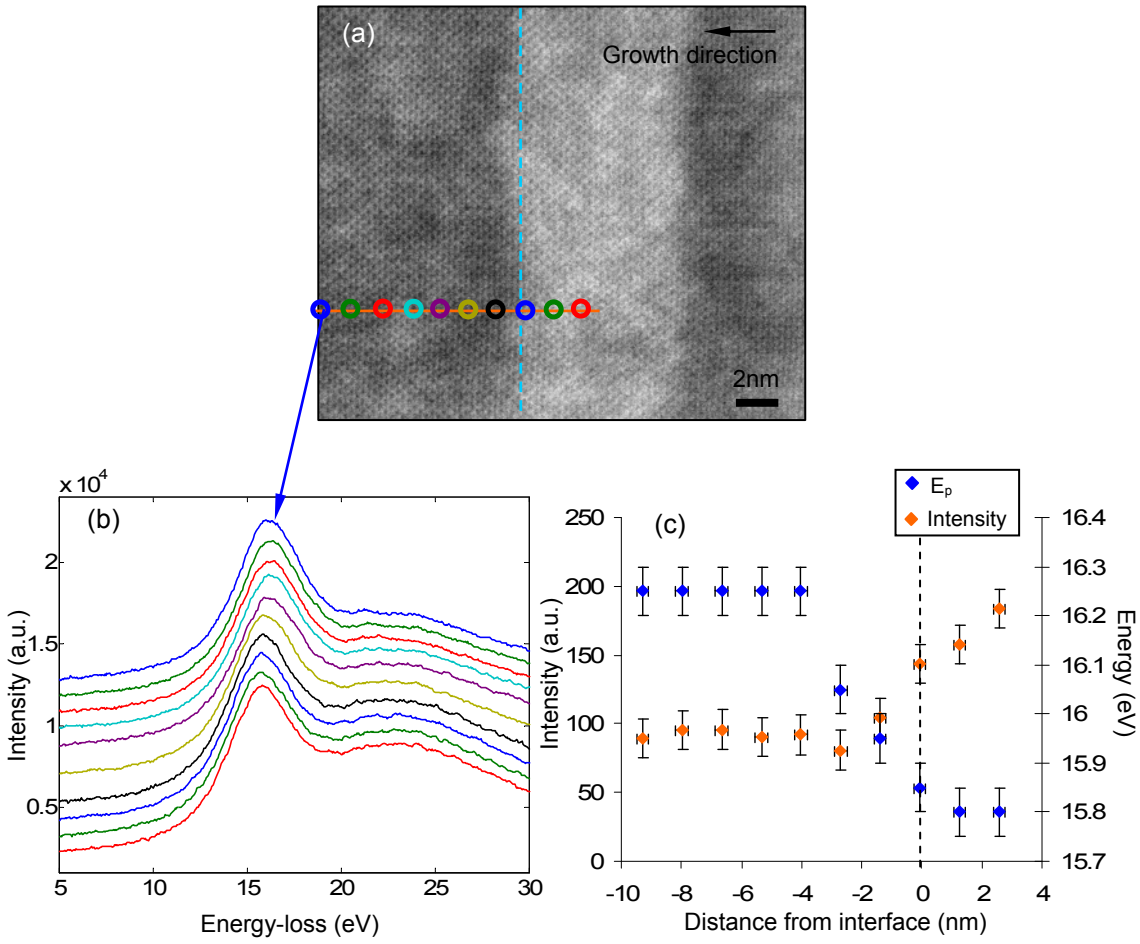


Figure 5.17: (a) An EELS line-spectrum across the region marked in the HAADF image of a $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$ QW with T_g = 400 ºC. EEL spectra were collected at ~ 1.3 nm intervals. See the text for experimental conditions of acquisition. (b) The plasmon region of the EEL spectra from each collection point. (c) The plasmon peak energy and the HAADF image intensity values of each point of the line spectrum as a function of the distance from the GaAs – QW interface.

5.4 Summary

The results of morphological and compositional investigation of dilute nitride GaInNAs QWs by HAADF STEM imaging and EELS were presented and discussed in this chapter. QWs grown at higher temperature were found to exhibit a 3D morphology in the form of periodic interface undulations. The results of STEM *Z*-contrast imaging and EELS indicated a periodic In composition variation, evidence of a lateral phase separation in QWs. A higher degree of In composition modulation was shown to exist in higher growth temperature and higher In content samples. This phase separation is believed to induce an oscillatory strain field in the QWs during their growth and facilitate a 2D to 3D growth mode transition at a certain film thickness. The interface undulations of the QWs and the composition modulations were found to have different wavelengths, with that of the interface undulations almost twice as long as the composition modulations. Understanding this effect requires further theoretical and experimental studies. Finally, the contribution of interface plasmons on the energy-loss spectrum of QWs was investigated. Interface plasmons were shown to cause a shift in the peak plasmon energy in the energy-loss spectra from regions near the interface. It was then concluded that measuring shifts in the plasmon energy of the QWs cannot be considered to be solely due to compositional variations and be used to map composition inhomogeneities within the QWs.

6 Measuring interband transitions and optical properties of semiconductors using EELS

6.1 Introduction

Valence electron energy-loss spectroscopy (VEELS) is a powerful technique for studying the band structure of materials. Although the theory of the relationship between VEELS, the band structure and the dielectric response function (see section 3.3) was developed several decades ago [41, 110, 111], the application of this technique to probe the band structure of materials was hindered by the poor energy and spatial resolution of the existing electron microscopes. Traditionally, optical techniques have been used to measure the interband transitions in semiconductors, with the most important interband transition being the band gap. However, despite the high energy resolution of these techniques (~ 2 meV), their poor spatial resolution (~ 200 nm) means that they are not ideal for measuring the band gap of modern nanostructure devices.

In recent years, with the advent of monochromated analytical TEMs, there has been much interest in employing VEELS to measure the interband transitions and optical properties of semiconductors with high spatial resolution [112-115]. The improved energy resolution of monochromated instruments allows the fine structure in the energy-loss spectrum to be resolved more clearly. More importantly for band gap measurement, the improved energy resolution reduces the width of the zero-loss peak (ZLP), making it simpler to separate the energy-loss signal from the tail of the ZLP.

An additional advantage of the monochromated instruments is the more symmetrical shape of the ZLP, which simplifies its removal from the energy-loss spectrum. The energy resolution is often measured as the full-width at half-maximum of the ZLP recorded in vacuum [27]. Fig. 6.1 shows how in general the shape and width of the ZLP varies for different electron sources [113].

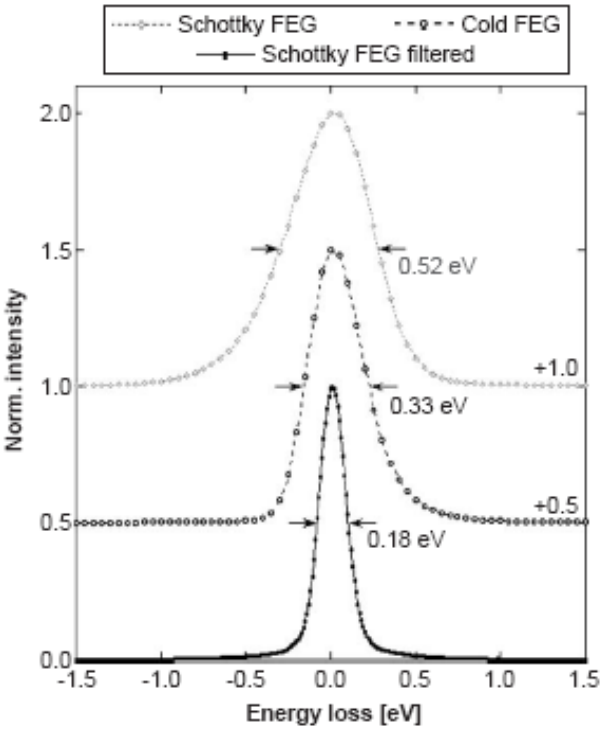


Figure 6.1: Examples of ZLPs obtained with a Schottky FEG, cold FEG and a monochromated Schottky FEG (taken from [113]). The full-width at half-maximum of the ZLP recorded in vacuum is a measure of the energy-resolution in TEM. The ZLPs obtained with non-monochromated Schottky and cold FEGs have asymmetric shapes with broad tails in the negative and positive energy-loss regions respectively. The ZLP obtained with a monochromated Schottky FEG exhibits a more symmetrical shape.

Despite improvements in the attainable energy and spatial resolution in TEM, challenges remain in applying VEELS to measure the band gap of semiconductors. In this chapter some of these challenges and possible solutions to them are discussed. This is followed by the theoretical and experimental VEELS results of the interband transitions and optical properties of GaAs and GaInNAs/GaAs QWs obtained in this work.

6.2 Challenges of measuring the band gap and optical properties of semiconductors using VEELS

6.2.1 Energy resolution

Despite the improved resolution of monochromated (S)TEM instruments, the energy resolution of EELS is still two orders of magnitude poorer than that of optical techniques. This leads to the loss of information at low energies because of the high intensity of the ZLP compared to other energy-loss features. For instance, it is generally impractical to determine band gap energies if the tail of the ZLP is more than an order of magnitude intense than the band gap signal. This means that even with monochromated TEM instruments the measurement of band gap energies less than ~ 0.8 eV remains unfeasible.

6.2.2 Data acquisition

Even with a monochromator the tail of the ZLP extends to ~ 3-4 eV and if there are interband transitions below this range, it is necessary to record the ZLP in the VEEL spectrum in order to accurately determine the onset energy of the interband transitions. However, given that for a thin specimen the ZLP is typically two or three orders of magnitude more intense than the rest of the VEEL spectrum, the spectral acquisition time becomes limited by the dynamic range of the spectrometer detector. With the dynamic range of most CCDs (commonly 14 or 16 bits [113, 116]) it is not possible to record a single VEEL spectrum with a sufficiently high signal to noise ratio, without saturating the detector. One way of overcoming this limitation is by acquiring cumulatively in order to improve the signal to noise ratio whilst using the full dynamic range of the detector (the total acquisition time is then constrained by spectral drift and beam damage). In addition, a small energy dispersion per channel must be used (0.05 eV/channel was used in this work) in order to achieve the optimal energy resolution of the spectrometer [117].

6.2.3 ZLP removal

It is necessary to remove the ZLP from the recorded VEEL spectrum in order to determine the onset of the band gap and other interband transition features. Several subtraction or deconvolution methods have been suggested for this purpose [114, 118, 119]. Deconvolution has been shown to be the most accurate method of ZLP removal, as subtraction fails to completely eliminate the effect of the ZLP from the spectrum [118]. More crucially, an accurate removal of the ZLP requires the exact shape of the ZLP to be known. This has been achieved either by pre-recording the ZLP in vacuum or by modelling its shape [118, 119], for example, using a reflected left-hand tail model in the case of symmetrically shaped ZLPs. A brief description of each method is given in the next section.

6.2.3.1 Pre-recorded ZLP deconvolution

The principle behind this method is similar to the Fourier-ratio deconvolution method of removing plural scattering, as described in section 3.4.1. The recorded spectrum is assumed to be a convolution of the ZL spectrum and the rest of the VEEL spectrum and the deconvoluted VEEL spectrum is found as the ratio of the Fourier transforms of the recorded VEEL and ZL spectra [119]. Without the use of monochromators this method is the most effective way of removing the ZLP [119]. However, due to phonon scattering and other energy-losses below the instrumental energy resolution, the shape of a ZLP recorded from the specimen differs from that recorded in vacuum (see Fig. 6.2). The broadening of the ZLP by phonon scattering can be minimised by selecting a small spectrometer collection angle (due to the large scattering angle of phonon scattering (see section 2.2.2)) but this also has the effect of degrading the signal to noise ratio and, in case of most semiconductors, intensifying the relativistic effects present in the VEEL spectrum (see section 6.3.2).

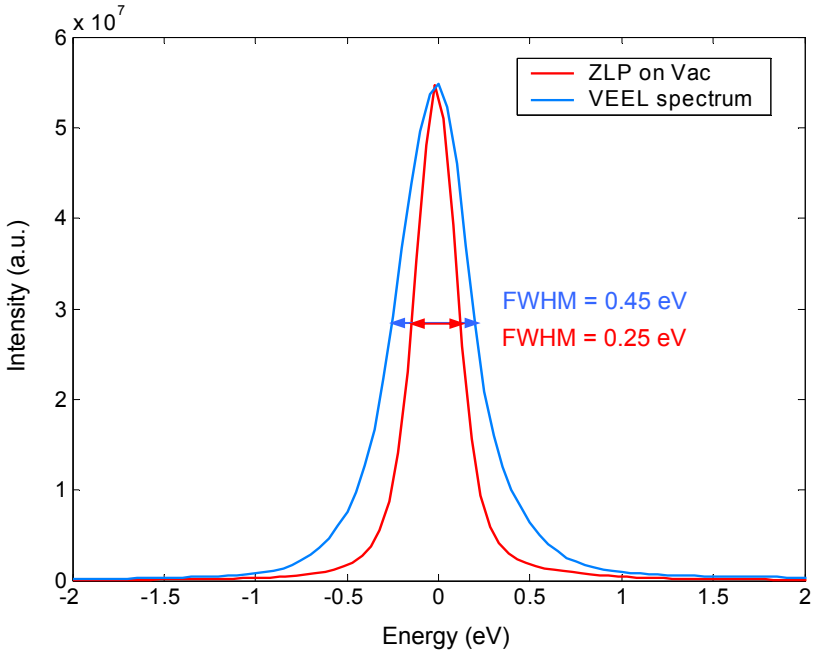


Figure 6.2: Examples of spectra comparing the full-width at half-maximum (FWHM) of the ZLPs collected in vacuum to that collected from the specimen (both recorded in monochromated - STEM mode, 80 kV accelerating voltage, 9 mrad probe convergence semi-angle and 10 mrad spectrometer collection semi-angle on the FEI Titan). The ZLP collected from the specimen broadens due to phonon scattering and other energy-loss events below the instrumental resolution, as described in the text.

6.2.3.2 Reflected left-hand tail fitting of the ZLP

This method can only be applied to data collected with a monochromated instrument in which the ZLP has a symmetrical shape. The right-hand side tail of the ZLP is found as the mirror reflection of the left-hand tail and then subtracted or deconvoluted from the spectrum. This method can introduce additional noise from the left-hand tail to the resulting subtracted or deconvoluted spectrum, and therefore, a recorded spectrum with a high signal to noise ratio is desired.

6.2.4 Relativistic radiation losses

At the accelerating voltages commonly used in (S)TEM, the speed of electrons in most semiconductor specimens is greater than the speed of light in them. This causes the electron beam to lose energy through the emission of electromagnetic radiation,

known as Čerenkov radiation. The criterion for the emission of Čerenkov radiation can be written as [27]:

$$v > \frac{c}{n(E)} \qquad \textbf{Equation 6.1}$$

where v is the speed of the electrons in vacuum, c the speed of light in vacuum and $n(E)$ the refractive index of the specimen material ($\varepsilon_1 = n^2 - k^2$ and $\varepsilon_2 = 2nk$, where ε_1 and ε_2 are the real and imaginary parts of the dielectric function and n and k are the refractive index and the extinction coefficient respectively). The emitted photons follow a hollow cone of semi-angle $\phi = \cos^{-1}\left(\frac{c}{v\,n}\right)$, known as the Čerenkov cone [27]. In thin specimens the electron beam can further interact with Čerenkov radiation and lose energy by resonating guided light modes [123]. The graph in Fig. 6.3 shows the maximum material refractive index for which no Čerenkov radiation is emitted, as a function of the electron beam accelerating voltage. It reveals that for GaAs specimens ($n \sim 3.4$) unless the accelerating voltage is below ~ 24 kV, Čerenkov losses are suffered.

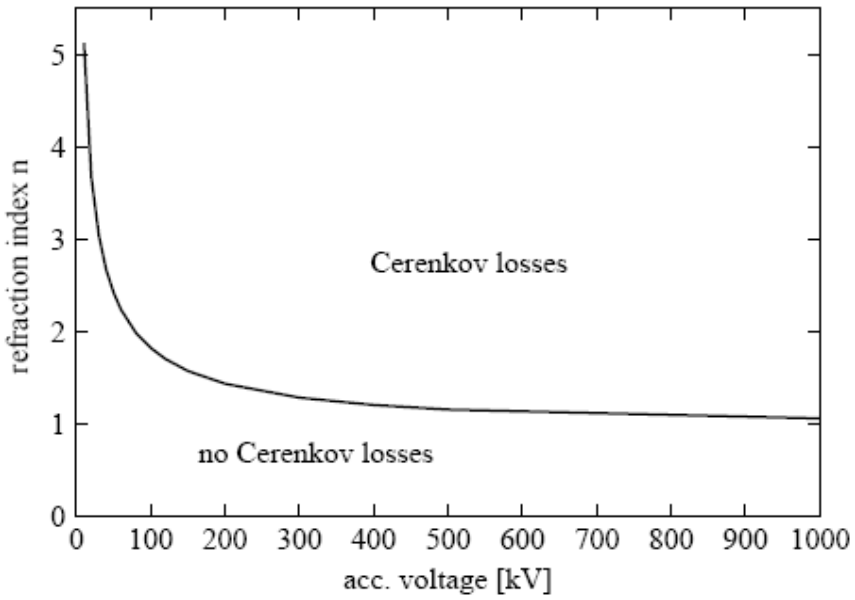


Figure 6.3: A graph showing the maximum refractive index for which no Čerenkov losses occur as a function of the electron beam accelerating voltage (taken from [125]).

Fig. 6.4 shows the energy-loss range for which Čerenkov radiation is emitted in GaAs at 300kV and 80kV accelerating voltages. Although the condition for Čerenkov

emission can be fulfilled at higher energy-losses at which ε_1 tends to unity, Fig 6.4 shows that this is not the case for GaAs.

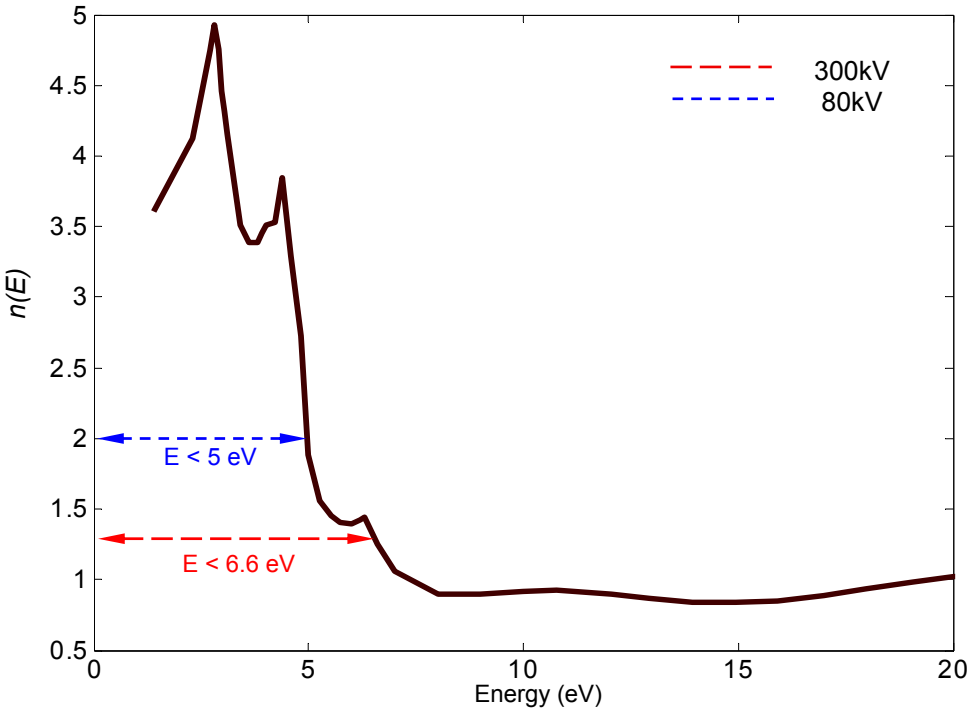


Figure 6.4: The refractive index of GaAs *n(E)*, taken from optical measurements in [127], as a function of energy-loss. The condition for the emission of Čerenkov radiations given by Eqn. 6.1 is satisfied for energies at which $n(E) > 1.99$ at 80 kV and $n(E) > 1.30$ at 300 kV.

Radiation losses, also known as retardation effects, alter the scattering distribution in the energy-loss – scattering angle plane ((E, θ) – plane) from that stated in section 3.1.1. Since the probability of relativistic losses is directly related to *n(E)*, it is clear that they can be emitted even at energies below the band gap and hence, complicate band gap measurements in semiconductors. This makes it necessary to have a full understanding of the scattering cross-section of relativistic losses.

6.2.4.1 Kröger theory

This semi-classical formulation of the inelastic cross-section of a fast electron beam by a thin specimen takes into account the bulk, surface, Čerenkov and guided light mode losses. The electron beam is assumed to be a point charge with charge distribution $-e\delta(\mathbf{r}-\mathbf{v}t)$, at position **r** and time *t*, passing at a normal incident angle and constant speed **v**, through a thin specimen of thickness *D*, and with a dielectric function *ε(ω)*. The following expression is found for the probability of energy-loss by solving Maxwell's equations and revising the solution to include relativistic losses [119, 121]:

$$\frac{\partial^3 P(\omega,\mathbf{k}_\perp)}{\partial\omega\,\partial^2\mathbf{k}_\perp} = \frac{e^2}{\pi^2\hbar v^2}\,\mathrm{Im}\left[\left\{\frac{\mu^2}{\varepsilon\varphi^2}\cdot D - \frac{2\mathbf{k}_\perp^2(\varepsilon-\varepsilon_0)^2}{\varphi_0^4\varphi^4}\cdot\left(\frac{\varphi_{01}^4}{\varepsilon\varepsilon_0}\left(\frac{\sin^2\left(\frac{\omega D}{2v}\right)}{L^+}+\frac{\cos^2\left(\frac{\omega D}{2v}\right)}{L^-}\right)+\beta^2\frac{\lambda_0}{\varepsilon_0}\cdot\frac{\omega}{v}\cdot\varphi_{01}^2\cdot\left(\frac{1}{L^+}-\frac{1}{L^-}\right).\sin\left(\frac{\omega D}{v}\right)\right)\right\}\right.$$
$$\left.-\beta^4\frac{\omega^2}{v^2}\cdot\lambda_0\lambda\left(\frac{\cos^2\left(\frac{\omega D}{2v}\right)\tanh\left(\frac{\lambda D}{2}\right)}{L^+}+\frac{\sin^2\left(\frac{\omega D}{2v}\right)\coth\left(\frac{\lambda D}{2}\right)}{L^-}\right)\right]$$

Equation 6.2

where $\mathbf{k}_\perp$ is the wave vector of the perpendicular component of the momentum transferred to the electron beam ($\mathbf{q}_\perp=\hbar\mathbf{k}_\perp$), ε_0 the dielectric constant of vacuum and:

$$\beta=\frac{v}{c},\quad \mu^2=1-\varepsilon\beta^2,\quad \mu_0^2=1-\varepsilon_0\beta^2,\quad \lambda=\sqrt{\mathbf{k}_\perp^2-\frac{\varepsilon\omega^2}{c^2}},\quad \lambda_0=\sqrt{\mathbf{k}_\perp^2-\frac{\varepsilon_0\omega^2}{c^2}}$$

$$L^+=\lambda_0\varepsilon+\lambda\varepsilon_0\tanh\left(\frac{\lambda D}{2}\right),\quad L^-=\lambda_0\varepsilon+\lambda\varepsilon_0\coth\left(\frac{\lambda D}{2}\right)$$

$$\varphi^2=\lambda^2+\frac{\omega^2}{v^2},\quad \varphi_0^2=\lambda_0^2+\frac{\omega^2}{v^2},\quad \varphi_{01}^2=\mathbf{k}_\perp^2+\frac{\omega^2}{v^2}-(\varepsilon+\varepsilon_0)\frac{\omega^2}{c^2}$$

The first term in Eqn. 6.2 is the bulk energy-loss and has a magnitude proportional to the thickness of the specimen. The remaining terms in Eqn. 6.2 are the surface, Čerenkov and guided light mode losses, with the zeros of the denominators giving the resonance frequencies for each (for example, at the zeros of φ^2 the condition for Čerenkov radiation is satisfied) [119, 120]. The probability of energy-loss is found by integrating Eqn.6.2 over the solid collection angle of the spectrometer.

A simpler expression for the scattering probability is found by ignoring the surface losses in Eqn. 6.2, giving [124]:

$$\frac{d^2P}{d\Omega\, dE} = D\,\frac{e^2}{\pi^2\hbar^2v^2}\,\mathrm{Im}\left(\frac{-1}{\varepsilon(E)}\right)\,\frac{\theta^2+\theta_E^2\left[\left(1-\beta^2\varepsilon_1(E)\right)^2+\left(\beta^2\varepsilon_2(E)\right)^2\right]}{\left[\theta^2+\theta_E^2\left(1-\beta^2\varepsilon_1(E)\right)^2+\left(\theta_E^2\beta^2\varepsilon_2(E)\right)^2\right]}$$

Equation 6.3

where θ_E is the characteristic scattering angle defined by Eqn. 3.3. Comparing Eqn. 6.3 with Ritchie's dielectric formulation introduced in section 3.1.2 indicates that in the presence of relativistic losses, the Lorentzian angular distribution of scattering no longer holds and is replaced by a more complicated distribution. Fig. 6.5 shows the angular distributions for a ~ 100 nm thick GaAs specimen, calculated using the relativistic expression in Eqn. 6.3 and Ritchie's expressions. The relativistic angular distribution largely differs from a Lorentzian shape at small angles due to retardation effects. Čerenkov losses typically peak around an angle of 10^{-2} mrad, which is significantly smaller than the scattering angles associated with other energy-loss events in VEELS. The calculated angular distributions for GaAs at several different energy-losses are displayed in Fig. 6.6. As expected, the scattering intensity due to Čerenkov losses decreases with increasing energy-loss.

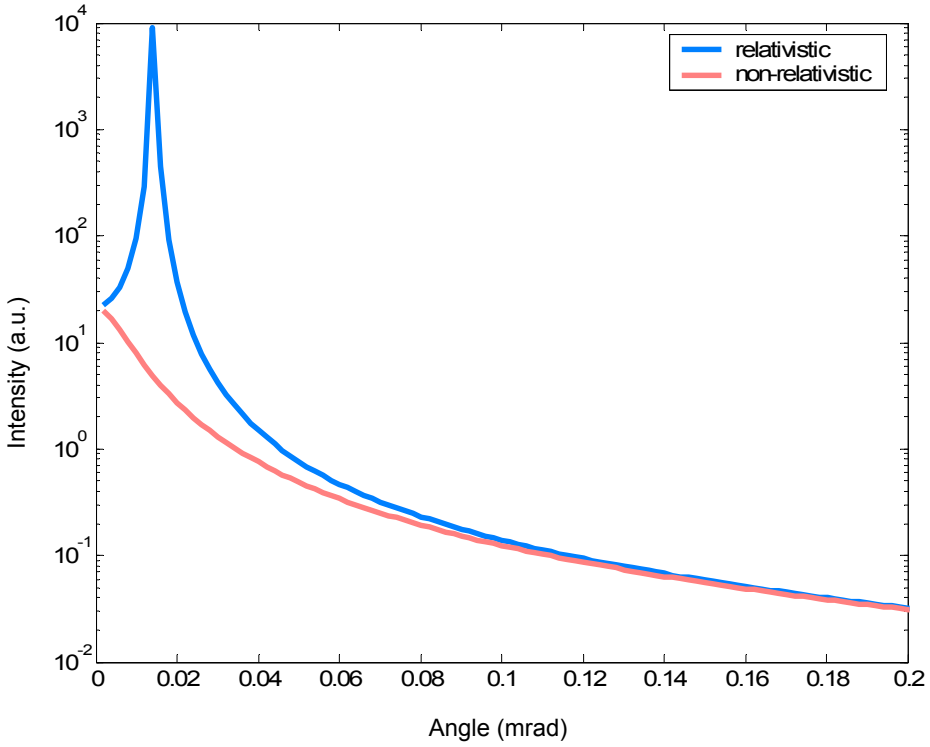


Figure 6.5: The angular dependence of the scattering cross-section for a 100 nm thick GaAs specimen, 300 kV accelerating voltage and 1.4 eV energy-loss calculated using the non-relativistic (Ritchie's dielectric formulation, Eqn. 3.4) and relativistic expressions (Kröger's equation, Eqn. 6.2).

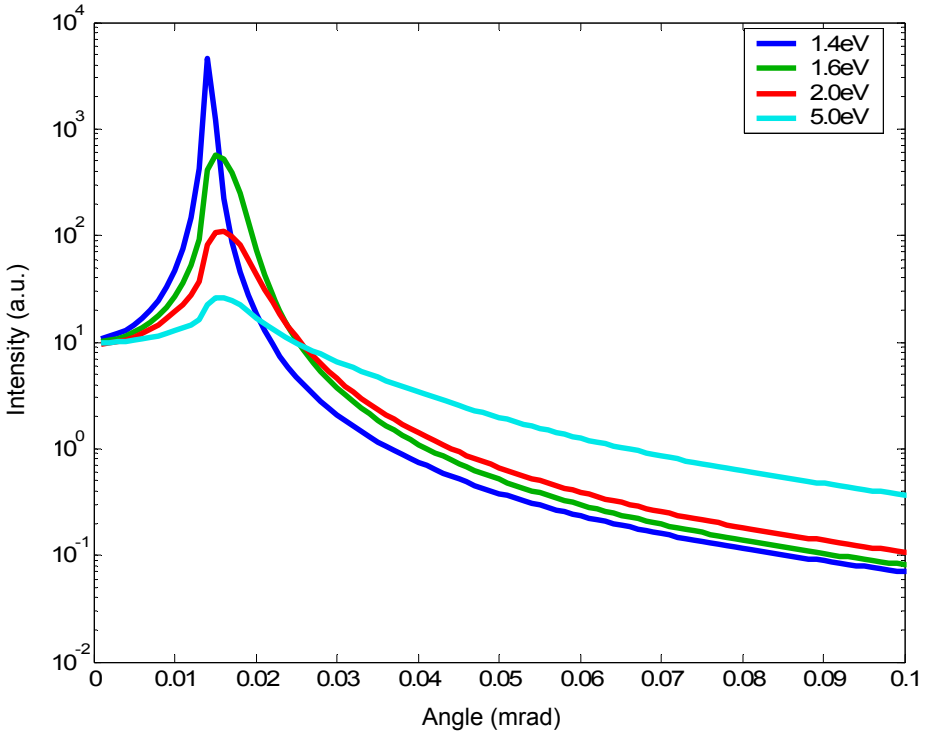


Figure 6.6: The angular dependence of scattering at different energy-losses for a 100 nm thick GaAs specimen and 300 kV accelerating voltage calculated using Kröger's equation (Eqn. 6.2).

6.3 Modelling the relativistic losses using Kröger's equation

The results of the simulations, modelling the effects of experimental conditions on relativistic losses, using Kröger's equation are presented in this section. Simulations were carried out in MATLAB using the code provided in Appendix B. An empirical dielectric function of GaAs was used as the input for the simulations. This was derived from the single scattering distribution of GaAs, modified by fitting a Lorentzian function to the lower energy tail of the plasmon peak. The onset of the Lorentzian tail was chosen as 1.4 eV, approximately the band gap of GaAs [128, 129]. This was done in order to ensure that the input data is free from relativistic losses. Although this modification does not produce an exact representation of the actual dielectric function of GaAs (since it ignores interband transitions), it can still be used to investigate the effects of experimental conditions on the scattering probability of relativistic losses.

6.3.1 Accelerating voltage

Altering the accelerating voltage of the electron beam is expected to change the scattering probability of relativistic losses by changing the velocity v of the incident electrons. Fig. 6.7 compares the calculated VEEL spectra for a GaAs specimen using 20 kV, 100 kV, 200 kV and 300 kV incident beam accelerating voltages. The calculated spectrum for the 20 kV electron beam does not exhibit any features due to relativistic losses, since the condition for Čerenkov radiation is not satisfied at 20 kV. A peak can be seen at ~ 5 eV in the 80 kV, 200 kV and 300 kV spectra. In these spectra the radiation peak appears to shift to higher energies and becomes more intense with increasing accelerating voltage. The shift in the energy of the radiation peak occurs because at higher accelerating voltages the condition for Čerenkov radiation is satisfied for a wider range of energies (see Fig. 6.4). In general, a lower accelerating voltage is desired in order to reduce the probability of relativistic losses. However, reducing the operating accelerating voltage requires using increasingly thinner specimens and in most cases, producing high quality specimens for operating at accelerating voltages < ~ 80 kV is a major challenge.

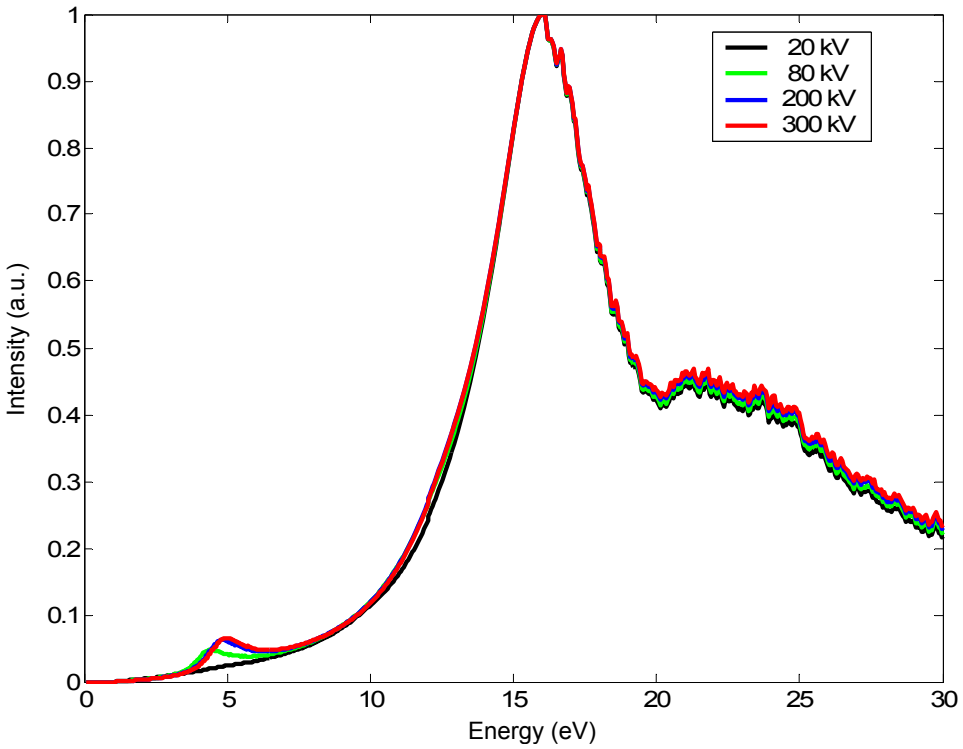


Figure 6.7: Simulated VEEL spectra for a 100 nm thick GaAs specimen, 8 mrad collection semi-angle and different accelerating voltages (the intensity of the spectrum is normalised). The peak due to Čerenkov radiations becomes more intense with increasing the accelerating voltage.

6.3.2 Spectrometer collection angle

Fig. 6.8 shows the simulated VEEL spectra for various spectrometer collection angles using a 100 nm thick GaAs specimen and 300 kV accelerating voltage. The intensity of the relativistic features (seen in the first ~ 5 eV of the spectrum) appears to increase with decreasing collection angle. In addition, at collection semi-angles below 1 mrad, the spectrum onset appears to shift to an energy below 1.4 eV (band gap energy). Given the small scattering angle of relativistic losses discussed in section 6.2.4.1, at very small collection angles the relativistic losses become the main contributor to the spectrum intensity and alter the energy-loss spectrum accordingly. Therefore, a large spectrometer collection angle is desired in order to minimise the contribution of relativistic effects to the VEEL spectrum and to measure the correct band gap.

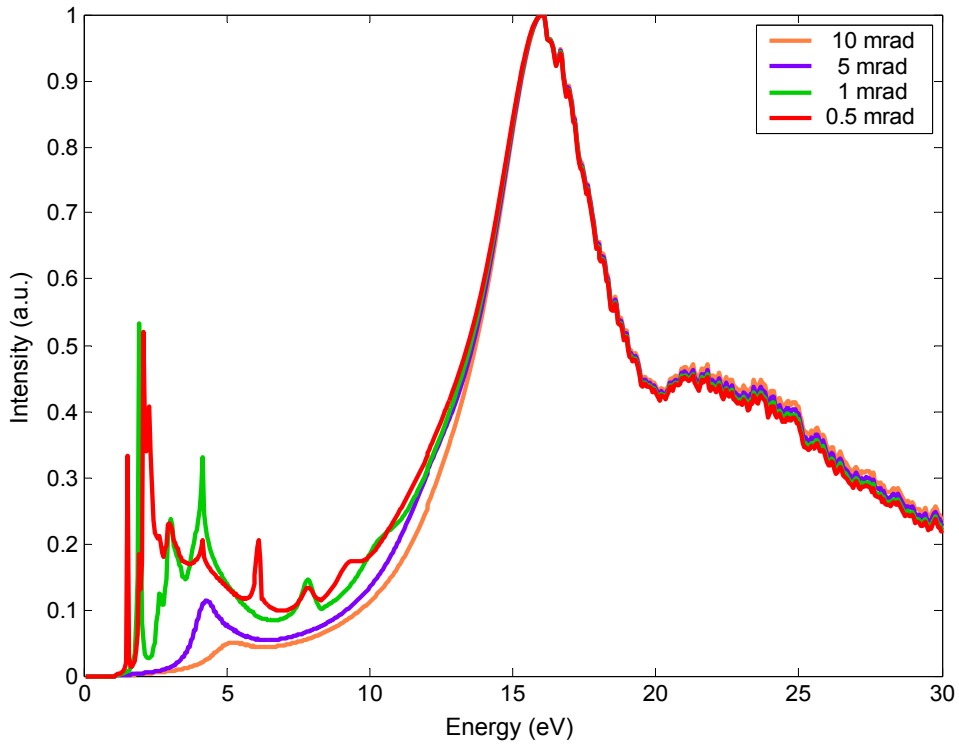


Figure 6.8: Simulated VEEL spectra for a 100 nm GaAs specimen, 300 kV incident beam energy and different spectrometer collection semi-angles (the intensity of the spectrum is normalised). The peaks due to relativistic losses flatten out at larger collection angles.

6.3.3 Specimen thickness

Fig. 6.9 shows the simulated VEEL spectra for GaAs specimens of three different thicknesses, using a 300 kV accelerating voltage and 10 mrad spectrometer collection semi-angle. Not much change seems to occur to the energy-loss spectrum with changing specimen thickness in these calculations. However, this is due to the flattening of the relativistic features at larger collection angles, as demonstrated in section 6.3.2 The effect of specimen thickness on spectra calculated for smaller collection angles is more evident. Fig. 6.10 shows the calculated VEEL spectra for GaAs specimens of different thicknesses, using a 300 kV incident electron beam and 4 mrad spectrometer collection semi-angle. The peak seen at ~ 11 eV in the 50 nm spectrum is the surface plasmon peak. The higher intensity of the surface plasmon in the 50 nm specimen compared to those in the 100 nm and 200 nm specimens is due to the increasing probability of coupling between the top and bottom surface plasmon excitations with decreasing specimen thickness (see section 3.3.2.2).

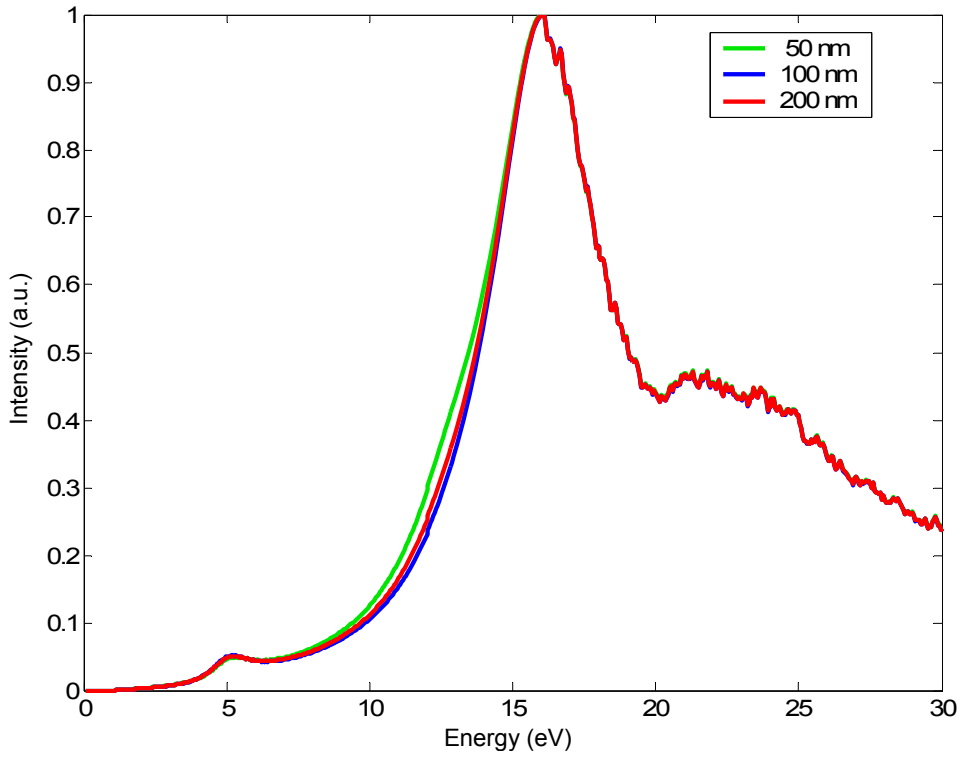


Figure 6.9: The simulated VEEL spectra for different GaAs specimen thicknesses, 300 kV incident electron beam and 10 mrad collection semi-angle (the intensity of the spectrum is normalised).

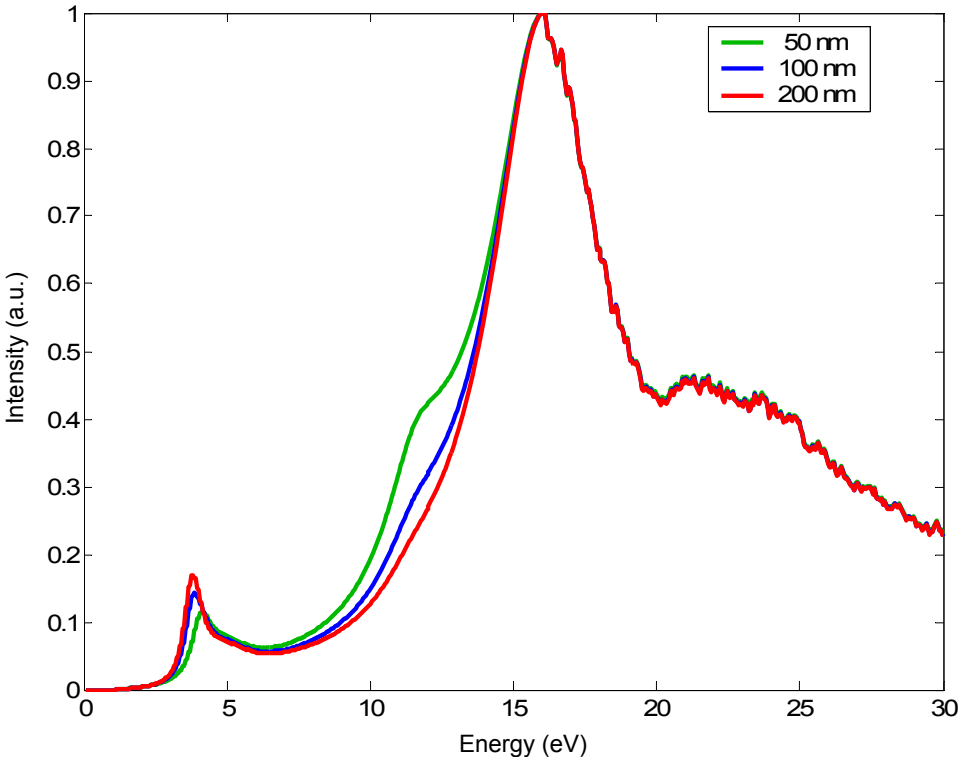


Figure 6.10: The simulated VEEL spectra for different GaAs specimen thicknesses, 300 kV incident electron beam and 4 mrad collection semi-angle (the intensity of the spectrum is normalised).

The relativistic peaks at ~ 4 eV appear to become more intense and shift to a lower energy with increasing specimen thickness. The shift in the relativistic peak can be understood by considering the contribution from the guided light modes. The results of the calculations carried out in [120, 123] indicated that relativistic radiations emitted along the Čerenkov cone can only be excited in thicker specimens. In thinner specimens, guided modes are excited at discrete energies (larger than the energy-loss of Čerenkov radiation [120]) and momentum transfers instead. The net effect is the contribution of two energy-loss events of different origins to the radiation peak in the simulated spectra. In the energy-loss spectrum of the 50 nm specimen guided modes are the dominant mode of radiation loss, and therefore, the peak occurs at a higher energy-loss value. The probability of Čerenkov radiation increases with increasing specimen thickness and the radiation peak gradually shifts to a smaller energy. Further calculations showed that the radiation peak energy is little changed beyond specimen thickness ~ 300 nm, at which Čerenkov radiation becomes the main mode of radiation loss.

6.4 Removing relativistic losses: off-line removal using Kröger's equation

Given the strong contribution of relativistic losses to the VEEL spectrum of semiconductors, it is crucial to suppress or remove their effects from the energy-loss spectrum. As demonstrated above, reducing the operating accelerating voltage, producing thinner specimens and using a larger spectrometer collection angle can reduce the intensity of relativistic features in the VEEL spectrum. Furthermore, off-axis collection (*i.e.* $\mathbf{q} \neq 0$) has been shown to circumvent the recording of relativistic losses [115, 125]. However, this method consequently measures indirect interband transitions and can not be compared with the results of optical measurements.

The off-line removal method described in [126], iteratively removes the relativistic losses by applying Kröger's equation to a recorded VEEL spectrum. Starting from the recorded spectrum, S_{orig}, the real and imaginary parts of the dielectric function are determined by means of Kramers-Kronig analysis and are used as inputs to Kröger's

equation. The resulting output, $S_{Kröger}$, is a single scattering distribution (SSD) with additional relativistic contributions. The difference between the input and output SSDs is found and subtracted from the input spectrum to give a new SSD, termed S_{new}. S_{new} is then used as the input for the next iteration. The process is repeated until $S_{Kröger}$ converges to S_{orig} within a pre-defined convergence criterion (a 1% difference between S_{orig} and $S_{Kröger}$ was used here).

Fig. 6.11 shows an example of the input and output spectra of the last iteration of the off-line correction process. The relativistic and surface contributions are removed iteratively from the recorded spectrum, S_{orig}, in the spectrum termed corrected. The spectrum produced by applying Kröger's equation to the corrected spectrum, $S_{Kröger}$, converges to S_{orig} within the 1% difference convergence criterion specified.

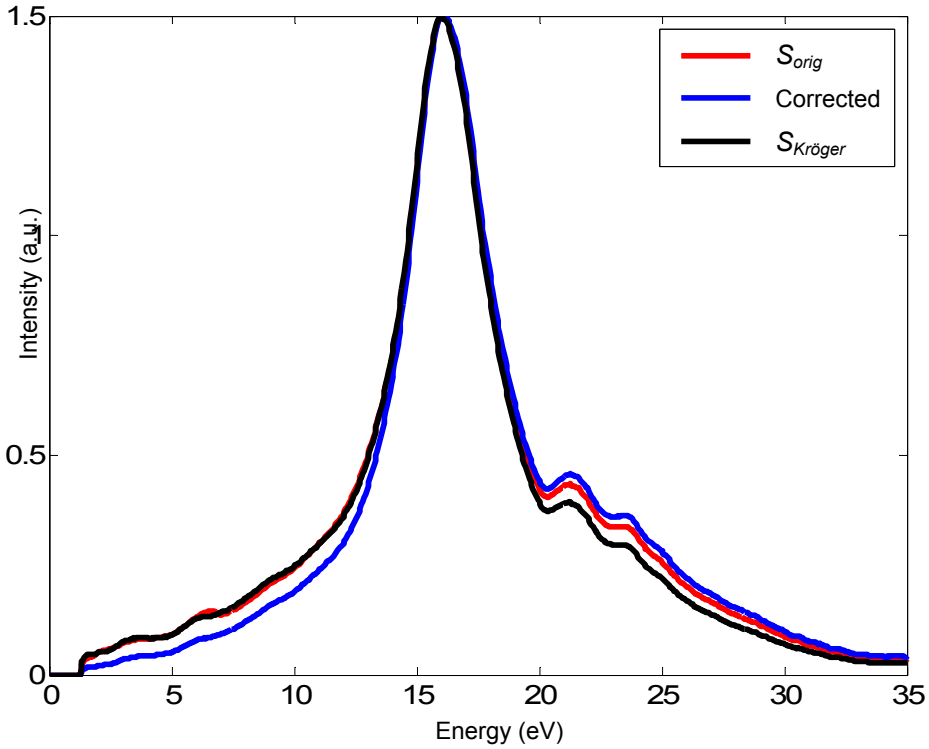


Figure 6.11: The result of the off-line correction method, showing the recorded spectrum, S_{orig}, the output of Kröger's equation in the final iterative step, $S_{Kröger}$ and the result of the off-line correction process, called corrected. The original VEEL spectrum was acquired using the FEI Titan in monochromated TEM mode, 80 kV accelerating voltage and 5 mrad spectrometer collection semi-angle.

6.5 Interband transitions and optical properties of GaAs

6.5.1 Experimental conditions

The VEEL spectra presented in this chapter were acquired using a monochromated FEI Titan (S)TEM instrument with a monochromator excitation of 2 kV. An energy resolution of ~ 0.2 eV (acquisition time << 1 s) was achieved in monochromated mode. All spectra were collected with an energy dispersion of 0.05 eV / channel. Other experimental parameters such as the operating accelerating voltage, probe convergence and spectrometer collection angles are stated for individual spectra. ZLP extraction, removal of plural scattering and Kramers-Kronig analysis were carried out in DigitalMicrograph. The reflected tail deconvolution method was employed for ZLP removal. The relativistic and surface contributions were removed using the off-line correction method described in section 4.

6.5.2 Results and discussion

Fig. 6.12 shows the SSD (prior to the off-line relativistic and surface-loss correction) and the extracted ZLP from a VEEL spectrum of an 80 nm thick GaAs specimen in <110> orientation acquired in TEM mode, using an 80 kV accelerating voltage and 5 mrad spectrometer collection semi-angle. The fine structure present in approximately the first 11 eV of the spectrum corresponds to valence to conduction electron band transitions, relativistic and surface plasmon losses. Despite the presence of the relativistic losses in the spectrum, the results of the simulations in section 3 indicated that their intensity is not expected to dominate the band gap region and cause a shift to the spectrum onset, given the experimental conditions employed here. The bulk plasmon peak occurs at 16.10 ± 0.05 eV (error determined by the energy dispersion of the measurements). The edge seen at ~ 20 eV is the Ga $M_{4,5}$ edge, denoting the 3d – 4p electron transitions in Ga.

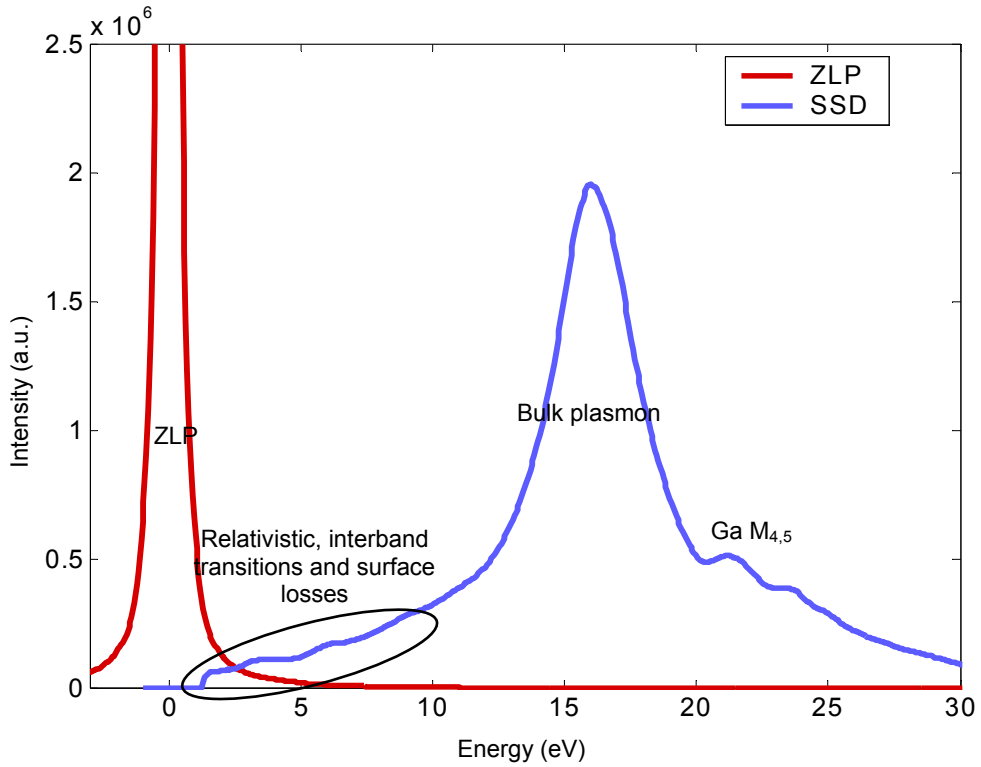


Figure 6.12: VEEL spectrum of GaAs acquired using the FEI Titan in monochromated TEM mode with 80 kV incident electron beam and 5 mrad collection semi-angle.

Fig. 6.13 and 6.14 compare the energy-loss function of GaAs to the results of optical measurements taken from [127]. The VEELS energy-loss function was obtained after removing the surface and relativistic losses, using the off-line correction method described in section 4, and applying the Kramers-Kronig analysis to the corrected spectrum.

The bulk plasmon energies of the optical and VEELS data in Fig. 6.13 occur at 14.9 eV and 16.10 ± 0.05 eV respectively. Given that the results of other experimental measurements and theoretical calculations confirm a bulk plasmon peak of ~ 16 eV [27, 78], this discrepancy is likely to be caused by specimen surface impurities in the optical measurements in which usually the optical reflectivity ($R = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2}$, where n is the refractive index and k the extinction coefficient) is measured [43].

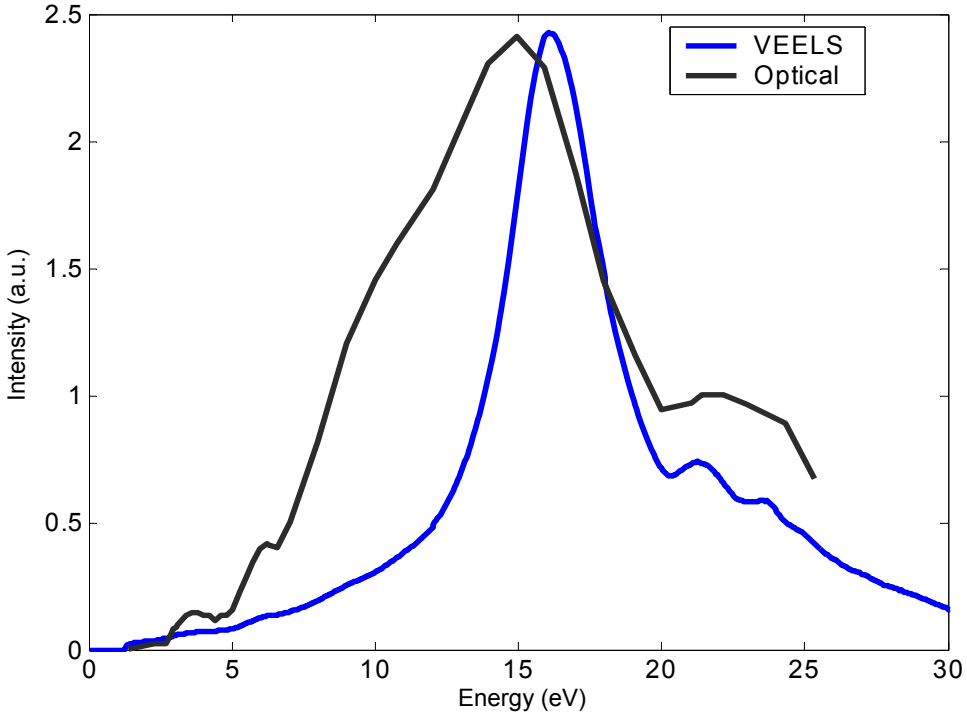


Figure 6.13: The energy-loss function of GaAs obtained by applying the Kramers-Kronig analysis to a VEEL spectrum and that obtained from optical measurements as compiled in [127].

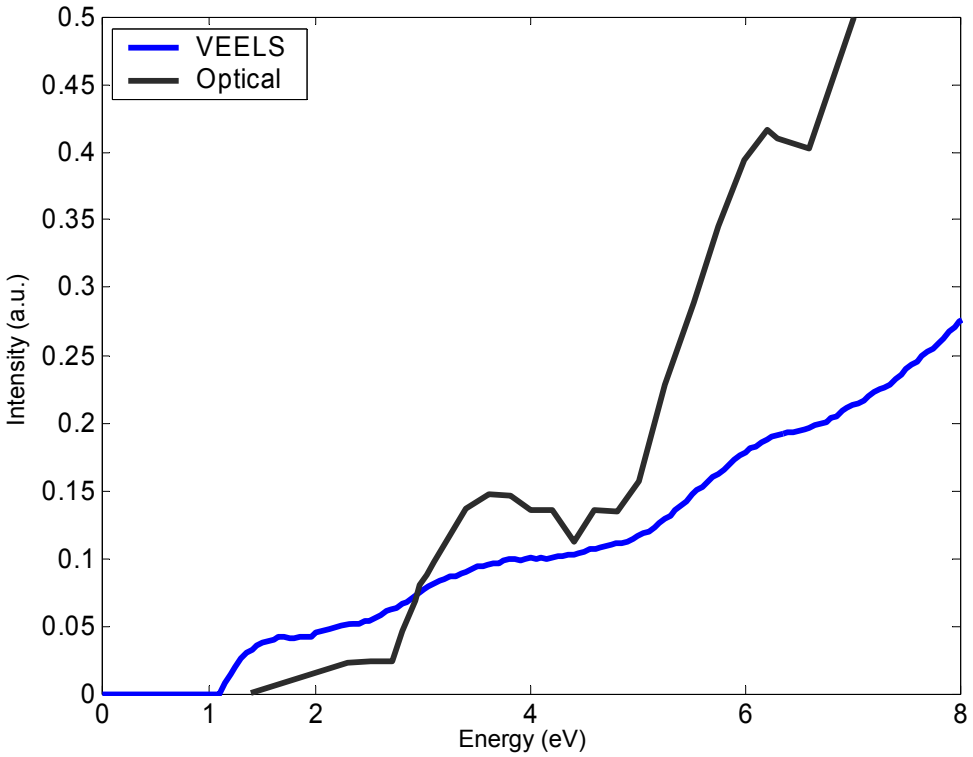
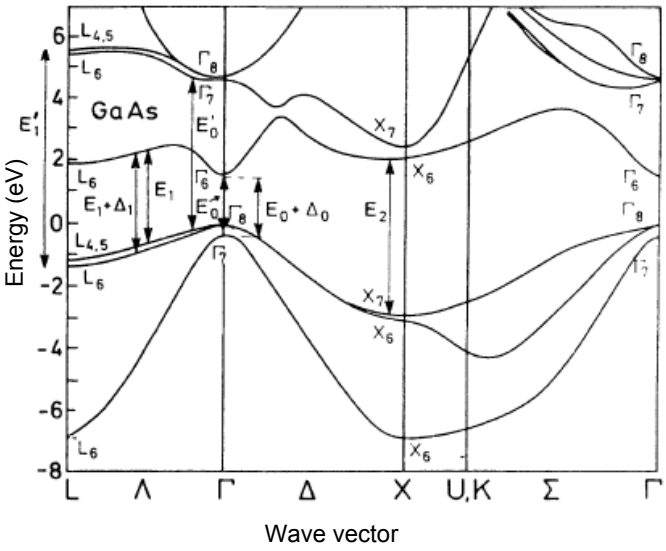


Figure 6.14: The first 8 eV of the energy-loss function of GaAs obtained by VEELS and optical measurements, showing the fine structure due to valence to conduction electron interband transitions, including the band gap transition.

The optical data in Fig. 6.13 exhibits a prominent surface plasmon at ~ 11 eV. Surface contributions are absent in the VEELS data since they were removed from the spectrum by the off-line correction process. The first ~ 8 eV of the VEEL spectrum contains several peaks corresponding to the valence–conduction band transitions, the first of which is the band gap transition. A good agreement exists between the peak position of the interband transitions in the VEEL and optical data, as displayed in Fig. 6.14. The band structure of GaAs and the energy of the corresponding valence – conduction transitions are displayed in Fig. 6.15.



Transition	Energy (eV)
E_0	1.43
$E_0 + \Delta$	1.77
E_1	2.89
$E_1 + \Delta$	3.12
E'_0	4.63
E_2	5.33
E'_1	6.60

Figure 6.15: The theoretical band structure of GaAs from [128] showing the main critical points and their corresponding energies at room temperature taken from [128, 129].

The onset of the energy-loss function in Fig. 6.14 follows a $(E - E_g)^{1/2}$ shape, evidence of the direct band gap transition of GaAs (see section 3.3.1). The band gap of GaAs is well documented as ~ 1.43 eV at room temperature [*e.g.* 128-130]. The VEELS energy-loss function shows an onset energy of 1.15 ± 0.05 eV, a peak energy of 1.5 ± 0.05 eV and a half-maximum energy of 1.3 ± 0.05 eV. The optical data presented in Fig. 6.14 only contained three measurements in the 1.38 eV – 2.29 eV range [127], making it difficult to draw a comparison with the band gap transitions in the VEELS data in this case. However, there are two major differences between the optical and VEELS data that must be taken into account when comparing the results of the two. One is the momentum transferred to the atomic electrons in VEELS,

permitted within the spectrometer collection angle, and the other is the inferior energy resolution of EELS, resulting in the broadening of the features in the energy-loss spectrum. Although the $\mathbf{q}_\perp$ - dependency of the energy-loss function has thus far been ignored, as stated in section 3.3.1, $\mathbf{q}_\perp$ can produce indirect as well as direct transitions at the critical points of the band structure (the Γ – X distance in the band structure of GaAs is equivalent to ~ 8 mrad for 80 kV and ~ 5 mrad for 300 kV accelerating voltages). This, in effect, broadens the interband transition peaks in the VEEL spectrum. A further broadening of the features in the energy-loss spectrum is caused by the limited energy resolution in EELS. This broadening caused by the larger energy spread of the incident beam in EELS relative to that in optical measurements implies that determining the band gap as the onset energy of the VEEL spectrum is not an accurate measure of the band gap energy. A more accurate measure of the band gap can be defined as the energy of the half-maximum intensity of the band gap peak in the spectrum.

The real and imaginary parts of the dielectric function of GaAs obtained from VEELS and optical measurements are displayed in Fig. 6.16 and 6.17, respectively. The main interband transition features that appear in ε_2 are identified in Fig. 6.17. Although the energies at which the main features in the optical and VEEL spectra occur agree, there are differences in the magnitude of these features. This difference can be attributed to the broadening of the interband transition peaks in the VEELS data compared to the optical results, as discussed above. A similar trend was observed in the results of the optical measurements of the dielectric function of GaAs as a function of temperature in [128] in which the magnitude of the peaks in the dielectric function decreased with increasing temperature. In optical measurements an increase in the temperature leads to an increase in the probability of phonon scattering, which can then provide the necessary momentum transfer to the valence electrons for indirect transitions.

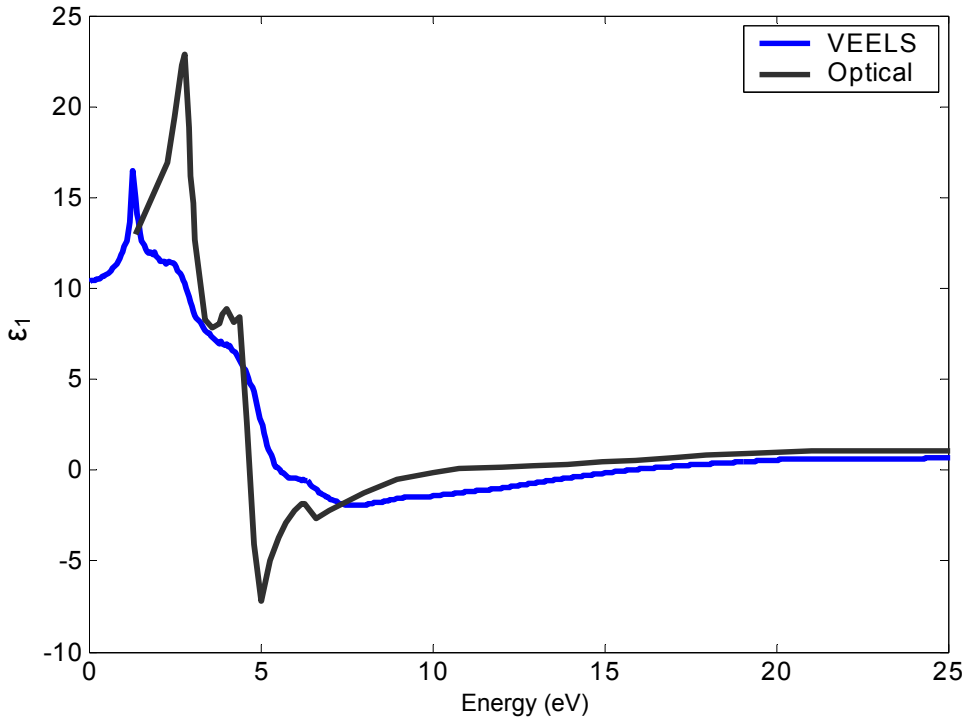
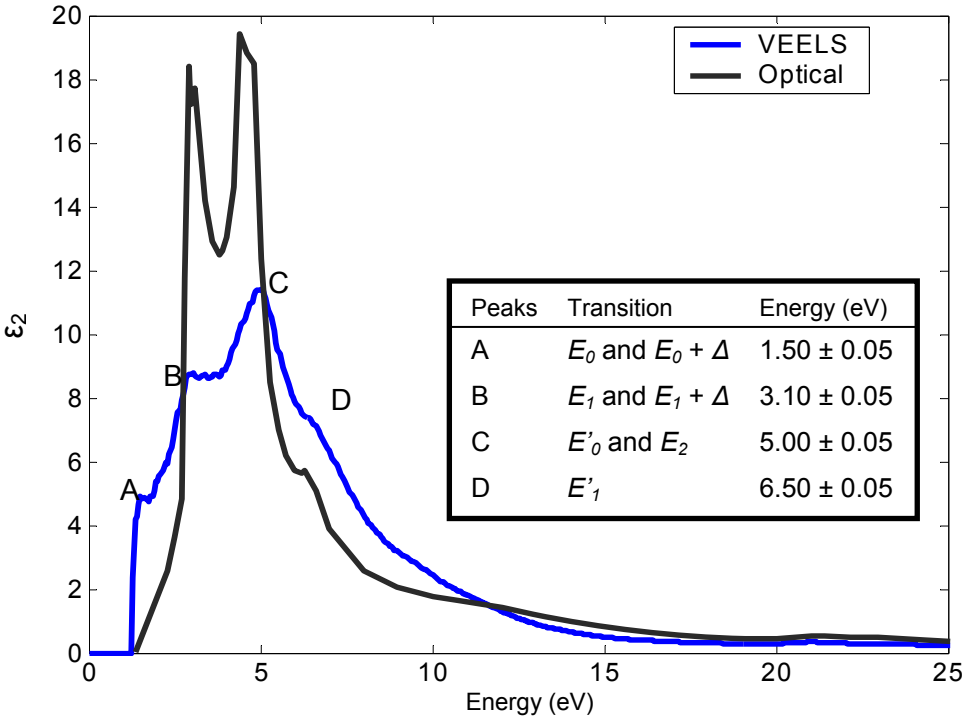


Figure 6.16: The real part of the dielectric function obtained from VEELS compared to that obtained from optical measurements in [127].



Peaks	Transition	Energy (eV)
A	E_0 and $E_0 + \Delta$	1.50 ± 0.05
B	E_1 and $E_1 + \Delta$	3.10 ± 0.05
C	E'_0 and E_2	5.00 ± 0.05
D	E'_1	6.50 ± 0.05

Figure 6.17: The imaginary part of the dielectric function obtained from VEELS compared to that obtained from optical measurements (taken from [127]). The interband transitions corresponding to the peaks in the spectrum and their energies are identified in the table. In the case of points A, B and C, the energy resolution is not sufficient to resolve two different interband transitions.

6.6 Measuring the interband transitions of GaInNAs QWs

6.6.1 Results and discussion

The main advantage of performing band gap measurement by VEELS over the optical techniques is the high spatial resolution achieved by combining STEM and EELS. This presents the prospect of mapping the variations in the band structure of semiconductor quantum structures. The results of the VEELS band gap measurement of an $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$/GaAs QW sample is presented here. The QWs and GaAs spacer layers were grown at 450 °C. The QW layers were 8 nm thick followed by 50 nm thick GaAs spacer layers. Other growth details of this QW sample can be found in section 5.2.

The VEEL spectra presented here were acquired in STEM mode, at an accelerating voltage of 300 kV. The probe diameter and convergence semi-angle were approximately 2 nm and 10 mrad, respectively. The spectrometer collection semi-angle was 11 mrad. A large probe size was selected in order to achieve a higher probe current, and hence, improved signal to noise ratio in the VEELS data. This inevitably degraded the spatial resolution of the measurements.

Although mapping the band structure variations within the QWs was attempted, the results were inhibited by the high noise levels present in the energy-loss spectra collected in STEM mode, which limited the energy and spatial resolution of VEELS. Fig. 6.18 shows the first 5 eV of the SSDs of a series of VEEL spectra recorded at 6 nm intervals across the QW. The increased noise level in these spectra, compared to the spectra collected in TEM – EELS mode (*e.g.* Fig. 6.14), makes it difficult to resolve the peak due to the band gap transition, E_0. However, a redshift of 0.15 ± 0.05 eV is seen in the onset of the spectrum collected on the QW compared to the spectra collected on GaAs. A redshift of 0.15 eV is also observed in the onset of the peak due to the E'_1 interband transition in the QW spectrum, as displayed in Fig. 6.19.

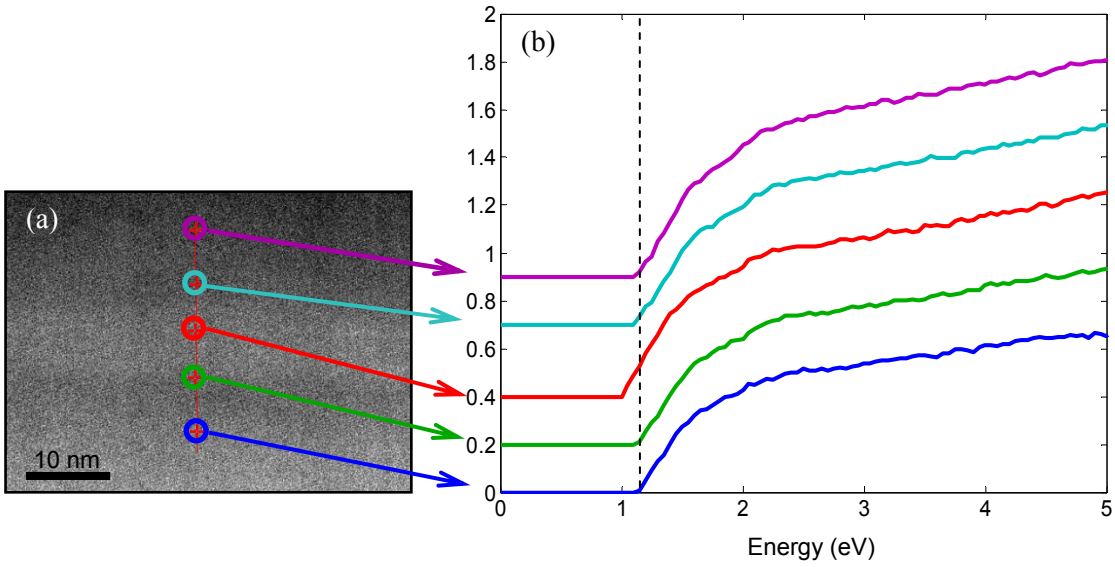


Figure 6.18: (a) HAADF image and (b) a series of SSDs of VEEL spectra obtained across a $Ga_{0.7}In_{0.3}N_{0.05}As_{0.95}$/GaAs QW. The VEEL spectra were acquired at 6 nm intervals. The signal to noise ratio is not sufficient to identify the peak due to the E_0 transition in the spectra. However, a shift can be seen in the onset of the spectrum collected from the QW.

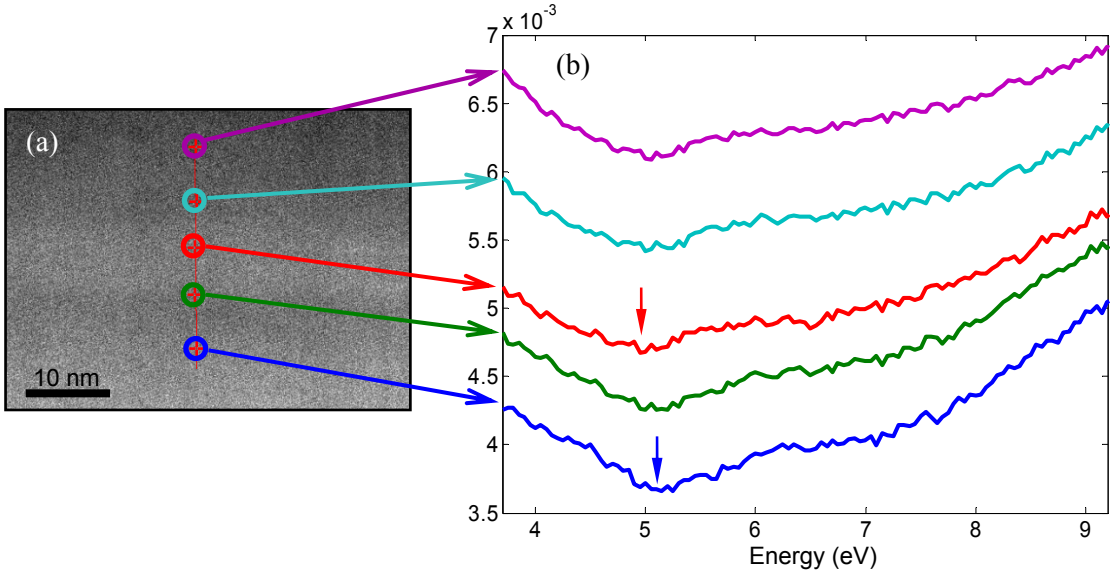


Figure 6.19: (a) The HAADF image and (b) the VEEL spectra of Fig. 6.18 (prior to ZLP removal), showing the redshift in the onset of the E'_1 interband transition of the spectrum collected from the QW.

The results of the photoluminescence spectroscopy measurements on the same sample published in [109] indicated an emission wavelength of 1.585 μm for the QWs. Using the relationship $E_g = \frac{h\,c}{\lambda}$, where h is Plank's constant, c the speed of light, E_g band gap energy and λ the emission wavelength, the QWs are expected to have a band gap energy of ~ 0.78 eV. It is clear that the band gap measured by VEELS here is larger than that obtained from photoluminescence measurements. One source of error in the VEELS measurement can arise from the large electron probe diameter employed here, relative to the thickness of the QW layer. This means that a large volume of the specimen is probed at each acquisition point, within which the band structure may vary. Although, it was possible to improve the spatial resolution of the VEELS measurements by changing the condenser system settings to form a smaller probe diameter (see Ch. 2 section 6.2), the higher noise levels present in the spectra acquired with smaller probes made it difficult to resolve any of the interband transition features.

The limitation encountered here in combining STEM and monochromated EELS reveals an additional challenge in mapping band gap variations in semiconductor nanostructures. The signal to noise ratio of the recorded spectra must be improved without compromising the spatial resolution of the measurements by choosing a large probe diameter. One solution to this problem can be provided by increasing the probe current by increasing the convergence angle of the probe. However, in the absence of aberration correctors, the spherical aberration of the probe forming lenses degrades the spatial resolution by increasing the diameter of the focused probe (see section 2.3.4) and thereby, limits the optimal angular range of the probe.

6.7 Summary

The advantages and challenges of utilising VEELS for measuring the interband transitions and the dielectric response function of semiconductors were discussed in this chapter. The improved energy resolution of monochromated TEM instruments permits probing the local band structure and optical properties of semiconductors with a high spatial resolution. However, the information obtained in VEELS can still be limited by factors such as relativistic losses, the available energy resolution or data acquisition methods. The effect of relativistic losses on VEELS was investigated, followed by methods of minimising or removing their contribution to the energy-loss signal. The results of measuring the interband transitions and dielectric function of GaAs were presented and compared with the results of optical measurements. Finally, the results of measuring the band gap of GaInAs/GaAs QWs using STEM – VEELS was provided and discussed.

7 Conclusions and future work

This work was concerned with the characterisation of the structural properties of InAs/GaAs QDs and GaInNAs/GaAs QWs using analytical TEM. In the case of the InAs/GaAs QDs, a good understanding of the morphology and composition of the QDs after the capping step is still lacking in the literature. *Z*-contrast STEM imaging was utilised in this work to study the morphology and composition of bilayer InAs/GaAs QDs. The effect of the GaAs spacer layer thickness on the nucleation position of the second layer QDs was investigated. The results were in agreement with the existing results in the literature in which for GaAs spacer layer thicknesses less than ~ 20 nm a correlation between the vertical position of the QDs in the first and second layers is observed [85, 86]. Obtaining information on the morphology and composition of the QDs is complicated by the 3D geometry of them. This is because unless tomography is carried out, the information obtained from (S)TEM is projected through the thickness of the thin specimen. Due to time limitation it was not possible to perform tomography on the samples. The analysis carried out on the size of the QDs viewed along the $[1\bar{1}0]$ and $[\bar{1}10]$ directions revealed a smaller height / base length aspect ratio for the QDs along the $[\bar{1}10]$ direction, confirming a larger degree of material intermixing along this direction as previously reported [69, 70]. The capping step is expected to cause changes to the morphology and composition of the QDs. *Z*-contrast STEM imaging was used to examine the composition profile of the QDs. The composition profile of the QDs capped with GaAs along the growth direction was shown to be highly inhomogeneous, with evidence of strong In

segregation towards the top of the QDs. However, a lower degree of compositional inhomogeneity was observed in QDs capped with InGaAs instead of GaAs. This can indicate that smaller morphological changes occur to the QDs during the capping step in samples capped with InGaAs than GaAs due to the smaller lattice mismatch between InAs and InGaAs.

In the case of the GaInNAs/GaAs QWs, increasing the In- or N-content of the QWs or increasing the growth temperature is known to cause modifications in the morphology and composition of the QWs [91-94]. The effects of the growth temperature and In-content on the morphology and composition of GaInNAs/GaAs QWs were investigated in this work using *Z*-contrast STEM imaging and EELS. Comparison between the GaInAs and GaInNAs QWs grown under the same conditions indicated that addition of as little as 1% N leads to a 3D morphology in the form of undulations at the top interface of the QWs. *Z*-contrast STEM images of the QWs indicated neighbouring regions of high and low In concentration along the lateral direction of the QWs with higher In content and grown at higher temperatures. Quantitative EELS analysis carried out using the intensity of the In $M_{4,5}$ signal suggested a similar neighbouring In-rich and In-poor regions along the lateral direction of the QWs. No similar direct evidence of this phase separation of the QWs has been previously reported.

The growth mechanism proposed in [92] was used to explain the changes observed in the morphology and composition of higher growth temperature and In content samples. The metastability of GaInNAs alloy causes phase separation into neighbouring In-rich and N-rich regions, creating compressively strained and tensile strained regions along the QWs respectively (oscillating strain field). This difference in the local strain energy of the QWs leads to a 3D growth mode beyond a critical thickness of the QWs (determined by the growth temperature, N and In contents of the QWs). Different values for the average wavelengths of the interface undulations and the In composition modulation of the QWs were found, with the composition modulations exhibiting smaller wavelength than the interface undulations. This can be interpreted as the effective length of the phase separation of the QWs and the strain field induced in the QWs due to this phase separation respectively. However, fully understanding this effect requires further experimental and theoretical studies. For

example, the exact effect of the oscillating strain field on the growth mode of the QWs and any morphological changes introduced by the overgrowth of the QWs need to be investigated. Furthermore, the results obtained in this work are based on the analyses carried out on the QWs along the <110> orientations and any possible orientation dependence of the observed interface undulations and composition modulations were not taken into account.

In addition to the above results, advantages and challenges of using monochromated VEELS to measure the band gap and other interband transitions in semiconductor structures were discussed. VEELS carried out in a monochromated TEM instrument allows probing the band structure of semiconductors with a high spatial resolution. However, despite the improved energy resolution of monochromated TEM instruments, VEELS band gap measurement of semiconductors was shown to be limited by factors such as relativistic losses and the obtainable energy resolution. Experimental parameters such as the accelerating voltage of the electron beam and the spectrometer collection angle were demonstrated to have a significant effect on the probability of energy-loss by relativistic effects. Reducing the accelerating voltage of the beam, using thinner specimens and larger spectrometer collection angles were shown to suppress the relativistic losses. An off-line method of removing the relativistic losses was also introduced. The complex dielectric function of GaAs obtained by VEELS and optical measurements were compared and a good agreement between the two was found. The STEM – VEELS band gap measurement of GaInNAs/GaAs QWs was found to be constrained by the attainable signal to noise ratio of the acquisitions, with the signal to noise ratio of the data obtained using typical experimental conditions not sufficient to determine the band gap. The signal to noise ratio of acquisitions can be improved by increasing the convergence angle of the electron probe. In the absence of aberration correctors, this leads to the degradation of spatial resolution in STEM. However, this problem does not arise in aberration-corrected STEM instruments and hence, it is of interest to investigate the possibility of probing the band gap of the QWs using aberration-corrected and monochromated instruments.

Appendix A

The following Matlab code was written in order to simulate the probability of energy-loss by an electron beam at a single planar interface, using Eqn. 3.15. The complex dielectric function of the two media, the accelerating voltage of the electron beam, the distance of the beam from the planar interface and the EELS spectrometer collection angle are input into the program.

The integral term in Eqn. 3.14 is replaced by a modified Bessel function of second kind. This Bessel function provides the solutions to differential equations of the general form, $x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} + \left(x^2 - \alpha^2\right) y = 0$, where α is an integer and it determines the order of the Bessel function [131]. The probability of energy-loss is found by evaluating Eqn. 3.15 for the input parameters.

```
function Int = SingleInterface(eps_A1, eps_A2, eps_B1, eps_B2,
eps_E, V, x0, beta);

%% Complex dielectric function of the two media eps_A and
eps_B, energy range of the dielectric functions eps_E,
accelerating voltage V, distance from the interface x0 and
collection semi-angle beta

%% constants: electron charge e, permittivity of vacuum eps0,
Plank's constant h, rest mass of electrons m0, speed of light
c

e = 1.6*10^-19;
eps0 = 8.854*10^-12;
```

```
h_bar = 1.055*10^-34;
m0 = 9.11*10^-31;
h = 6.63*10^-34;
c = 3*10^8;

E = e*V;
v2 = (c^2)*(1 - 1/((E/(m0*c^2) + 1)^2));     %% Finding the
electron velocity
v = sqrt(v2);
l = length(eps_E);
w = eps_E*e/h_bar;
b = beta/1000;
km = m0*v/h_bar * b;

K0 = besselk(0, (2*x0*w/v));          %% zeroth order Bessel
function

for i = 1:l
    P(i) = (eps_B2(i)/(eps_B1(i)^2 +
eps_B2(i)^2)*(log(km*v/w(i)) - K0(i)) +
2*(eps_A2(i)+eps_B2(i))/((eps_A1(i)+eps_B1(i))^2 +
(eps_A2(i)+eps_B2(i))^2)*K0(i));
    Int(i) = e^2/(2*pi^2*eps0*h_bar*v^2) * P(i);     %% The
probability of energy-loss
end;

mm = max(Int);
Int = Int/mm;
figure; plot(eps_E, Int);
```

Appendix B

The following Matlab code simulates the probability of energy-loss due to the bulk, surface, Čerenkov and guided light mode effects. It implements the Kröger equation introduced in section 6.2.4.1 (Eqn. 6.2). The accelerating voltage of the beam, the complex dielectric function of the specimen, the specimen thickness and the spectrometer collection semi-angle are input into the program. The probability of energy-loss is found by integrating Eqn. 6.2 numerically over the wave-vector range of the scattering (determined by the spectrometer collection angle). The numerical integration was implemented using Simpson's rule [see *e.g.* 131]. The output of the programme is a normalised energy-loss spectrum.

```
function SSDs = Kroger(V, eps1, eps2, w, D, b);

%% Accelerating voltage V, complex dielectric function eps1 +
ieps2, energy
%% w, specimen thickness D and collection semi-angle b

%% constants: electron charge e, speed of light c,
permittivity of vacuum eps0, Plank's constant h, rest mass of
electrons m0

 e = 1.6*10^-19;
 c = 3*10^8;
 eps0 = 1;
 h_bar = 6.63*10^-34 / (2*pi);
 m0 = 9.11*10^-31;

 E = e*V;
 v2 = (c^2)*(1 - 1/((E/(m0*c^2) + 1)^2));
 v = sqrt(v2);
```

```
 beta = v/c;
 w = w*e/h_bar;
 EE = w*h_bar/e;
 L_eps = length(eps1);

 t_max = b/1000;
 t_min = 0;
 dt = (t_max - t_min)/L_eps;
 t = t_min:dt:t_max;
 t = t';
 k = m0*v*t*2*pi/(h_bar);

 L = length(k);
 LL = length(w);

eps = complex(eps1, -eps2);

 thetaE = EE/((V + m0*c^2/e)*beta^2);
 Eloss = eps2./(eps1.^2 + eps2.^2);
 mu2 = (1 - eps*beta^2);
 mu02 = (1 - eps0*beta^2);

 for m = 1:LL
     for n = 1:L
%% Evaluating the abbreviated parameters used in Eqn. 6.2
(Kröger's Eqn)
         lamda(m, n) = sqrt(k(n)^2 - eps(m)*w(m)^2/c^2);
         lamda0(m, n) = sqrt(k(n)^2 - eps0*w(m)^2/c^2);
         phi2(m, n) = (lamda(m, n)^2 + w(m)^2/v^2);
         phi02(m ,n) = (lamda0(m, n)^2 + w(m)^2/v^2);
         phi012(m, n) = (k(n)^2 + w(m)^2/v^2 - (eps(m) +
eps0)*w(m)^2/c^2);
         Lp(m, n) = lamda0(m, n)*eps(m) + lamda(m,
n)*eps0*tanh(lamda(m, n)*D/2);
         Ln(m ,n) = lamda0(m, n)*eps(m) + lamda(m,
n)*eps0*coth(lamda(m, n)*D/2);

         P1(m, n) = (mu2(m))*D/(eps(m)*phi2(m, n));

         P2(m, n) = ((phi012(m,
n)^2/(eps(m)*eps0))*((sin(w(m)*D/(2*v)))^2/Lp(m, n) +
(cos(w(m)*D/(2*v)))^2/Ln(m, n)));

         P3(m, n) = beta^2 * (lamda0(m, n)*w(m)*phi012(m,
n)/(eps0*v)) * (1/Lp(m, n) - 1/Ln(m, n)) * sin(w(m)*D/v);

         P4(m, n) = beta^4 * w(m)^2/v^2 * lamda0(m,
n)*lamda(m, n) * ((cos(w(m)*D/(2*v)))^2 * tanh(lamda(m,
n)*D/2)/Lp(m, n) + (sin(w(m)*D/(2*v)))^2 * coth(lamda(m,
n)*D/2)/Ln(m, n));

         P(m, n) = k(n)*(e^2/(pi^2*h_bar*v^2)) * imag(P1(m, n)
- (2*k(n)^2 * (eps(m) - eps0)^2/(phi2(m, n)*phi02(m, n))^2) *
(P2(m, n) + P3(m, n) - P4(m, n)));
     end;
```

```
 end;

h = m0*v/h_bar*dt;

 for x = 1:LL                %% Evaluating the integral
numerically using Simpson's rule
    sum(x) = 0;
    for j = 1:L
        if j == 1
            sum(x) = sum(x) + P(x, 1);
        end;
        if j == L
            sum(x) = sum(x) + P(x, L);
        end;
        if (mod(j, 2) == 0 && j ~= L)
            sum(x) = sum(x) + 4*P(x, j);
        end;
        if (mod(j, 2) == 1 && j ~= L)
            sum(x) = sum(x) + 2*P(x, j);
        end;
    end;
    Integ(x) = h/3 * sum(x);
end;

maxInt = max(Integ);
normInteg = 1.5*Integ/max(Integ);
%% normalising the calculated spectrum
neps = complex(eps1, eps2);
Eloss = imag(-1./neps);
Eloss = Eloss';
mm = max(Eloss); Eloss = 1.5*Eloss/mm;
SSDs = [normInteg; Eloss];
SSDs = SSDs';
figure;
plot(EE, normInteg, 'r', EE, Eloss, 'b');
legend('Kroger', 'no CLs');
end;
```

References

[1] A. R. West, *Basic solid state chemistry*, 2nd Edition, John Wiley and sons Ltd: Chichester (1999)

[2] C. Kittel, *Introduction to solid state physics*, 7th Edition, John Wiley and sons Ltd: New York (1996)

[3] R. R. A. Syms and J. R. Cozens, *Optical guided waves and devices*, McGraw-Hill (1992)

[4] G. W. Hanson, *Fundamentals of nanoelectronics*, Prentice Hall (2007)

[5] J. Singh, *Electronic and optoelectronic properties of semiconductor structures*, Cambridge University Press (2003)

[6] P. Howe, PhD thesis, *InAs/GaAs quantum dots for long wavelength applications*, Imperial College London (2003)

[7] H. Luth, *Surfaces and interfaces of solid materials*, 3rd Edition, Springer (1998)

[8] Y. Arakawa and H. Sakaki, *Appl. Phys. Lett*. **40**, 939 (1982)

[9] D. Bimberg, M. Grundmann and N. N. Ledentsov, *Quantum dot heterostructures*, John Wiley and sons Ltd (1999)

[10] Q. Xie, A. Madhukar, P. Chen and N. Kobayashi, *Phys. Rev. Lett*. **75**, 2542 (1995)

[11] M. Weyers, M. Sato and H. Ando, *Jpn. J. Appl. Phys*. **31**, L853 (1992)

[12] W. G. Bi and C. W. Tu, *Appl. Phys. Lett.* **70**, 1608 (1997)

[13] J. D. Perkins, A. Mascarenhas, Y. Zhang, J. F. Geisz, D. J. Friedman, J. M. Olsen and S. R. Kurtz, *Phys. Rev. Lett.* **82**, 3312 (1999)

[14] C. Skierbiszewski, P. Perlin, P. Wisniewski, W. Knap, T. Suski, W. Walukiewicz, W. Shan, K. M. Yu, J. W. Ager, E. E. Haller, J. F. Giesz and J. M. Olsen, *Appl. Phys. Lett.* **76**, 2409 (2000)

[15] A. R. Denton, N. W. Ashcroft, *Phys. Rev. A* **43**, 3161 (1991)

[16] M. Kondow, T. Kitatani, S. Nakatsuka, M. C. Larson, K. Nakahara, Y. Yazawa, M. Okai and K. Uomi, *IEEE. J. Select. Top. Quant. Elect*. **3**, 1077 (1997)

[17] P. Klar, *Prog. Sol. Stat. Chem*. **31**, 301 (2003)

[18] W. Shan, W. Walukiewicz, J. W. Ager, E. E. Haller, J. F. Giesz, D. J. Friedman, J. M. Olsen and S. R. Kurtz, *Phys. Rev. Lett.* **82**, 1221 (1999)

[19] M. Kondow, K. Uomi, K. Hosomi and T. Mozume, *Jpn. J. Appl. Phys*. **33**, L1056 (1994)

[20] M. Kondow and T. Kitatani, *Semicond. Sci. Technol*. **17**, 746 (2002)

[21] N. Tansu and L. J. Mawst, *IEEE Photon. Tech. Lett*. **14**, 444 (2002)

[22] D. B. Williams, C. B. Carter, *Transmission electron microscopy*, Springer: New York (1996)

[23] M. H. Loretto, *Electron beam analysis of materials*, 2nd Edition, Chapman & Hall: London (1994)

[24] P. J. Goodhew, F. J. Humphreys, *Electron microscopy and analysis*, 2nd Edition, Taylor & Francis: London (1988)

[25] M. Raether, *Springer tracks in modern physics*, Vol. 38, Springer: Berlin (1965)

[26] FEI company, Monochromator users' manual for Titan 80 – 300 kV

[27] R. F. Egerton, *Electron energy-loss spectroscopy in the electron microscope*, 2nd Edition, Plenum press: New York (1996)

[28] R. F. Egerton, M. Malac, *J. Elect. Spectroscopy and Related Phenomena* **143**, 43 (2005)

[29] V. J. Keast, M. Bosman, *Microsc. Res. Tech.* **70**, 211 (2007)

[30] E. M. James, N. D. Browning, A. W. Nicholls, M. Kawasaki, Y. Xin and S. Stemmer, *J. Elect. Microsc.* **47**, 561 (1998)

[31] E. J. Kirkland, R. F. Loane and J. Silcox, *Ultramicrosc.* **23**, 77 (1987)

[32] S. J. Pennycook, D. E. Jesson, A. J. McGibbon and P. D. Nellist, *J. E. Microsc.* **45**, 36 (1996)

[33] B. Rafferty, P. D. Nellist, S. J. Pennycook, *J. Elect. Microsc.* **50**, 227 (2001)

[34] D. O. Klenov, S. Stemmer, *Ultramicrosc.* **106**, 889 (2006)

[35] FEI company, Condenser system users' manual for Titan 80 – 300 kV

[36] H. A. Brink, M. M. G. Barfels, R. P. Burgner and B. N. Edwards, *Ultramicroscopy* **96**, 367 (2003)

[37] S. Nakahara, *Surf. Coat. Tech.* **169-170**, 721 (2003)

[38] S. B. Newcomb, C. B. Boothroyd, W. M. Stobbs, *J. Microsc*. **140**, 195 (1985)

[39] B. Rafferty and L. M. Brown, *Phys. Rev. B.* **58**, 10326 (1998)

[40] M. Inokuti, *Rev. Mod. Phys.* **43**, 297 (1971)

[41] R. H. Ritchie, *Phys. Rev.* **106**, 874 (1957)

[42] V. J. Keast, A. J. Scott, R. Brydson, D. B. Williams and J. Bruley, *J. Microsc.* **203**, 135 (2001)

[43] D. L. Greenaway and G. Harbeke, *Optical properties and band structure of semiconductors*, International Series of Monographs in the Science of Solid State **1**, 1st Edition, Pegamon Press: London (1968)

[44] H. Raether, *Excitations of Plasmons and Interband Transitions by Electrons*, Springer Tracks in Modern Physics, Springer-Verlag: New York **88**, (1979)

[45] R. Garcia-Molina, A. Gras-Marti, A. Howie, R. H. Ritchie, *J. Phys. C: Solid Stat. Phys.* **18**, 5335 (1985)

[46] D. W. Johnson and J. C. H. Spence, *J. Phys.* D **7**, 771 (1974)

[47] D. W. Johnson, *J. Phys. A* **8**, 490 (1975)

[48] P. Nozieres, D. Pines, *Phys Rev.* **113**, 1254 (1959)

[49] R. F. Egerton, *J. Elect. Microsc.* **48**, 711 (1999)

[50] N. D. Browning, M. F. Chisholm and S. J. Pennycook, *Nature* **366**, 143 (1993)

[51] R. H. Ritchie and A. Howie, *Phil. Mag. A* **58**, 753 (1988)

[52] R. F. Egerton, *Micron* **34**, 127 (2003)

[53] R. F. Egerton, *Ultramicrosc.* **107**, 575 (2007)

[54] P. W. Atkins and R. S. Friedman, *Molecular quantum mechanics*, 3rd Edition, Oxford University Press: New York (1997)

[55] J. Marquez, L. Geelhaar and K. Jacobi, *Appl. Phys. Lett.* **78**, 2309 (2001)

[56] P. B. Joyce, T. J. Krzyzewski, G. R. Bell, T. S. Jones, S. Malik, D. Childs and R. Murray, *Phys. Rev. B* **62**, 10891 (2000)

[57] N. Kobayashi, T. R. Ramachandran, P. Chen and A. Madhukar, Appl. *Phys. Lett.* **68**, 3299 (1996)

[58] H. Saito, K. Nishi and S. Sugou, *Appl. Phys. Lett.* **74**, 1224 (1999)

[59] C. P. Liu, P. D. Miller, W. L. Henstrom and J. M. Gibson, *J. Micros.* **199**, 130 (2000)

[60] J. P. McCaffrey, M. D. Robertson, S. Fafard, Z. R. Wasilewski, E. M. Griswold, L. D. Madsen, *J. Appl. Phys.* **88**, 2272 (2000)

[61] N. N. Ledentsov, V. A. Shchukin, M. Grundmann, N. Kirstaedter, J. Bohrer, O. Schmidt, D. Bimberg, V. M., Ustinov, A. Yu. Egorov, A. E. Zhukov, P.

S. Kop'ev, S. V. Zaitsev, N. Yu. Gordeev, Z. I. Alferov, A. I. Borovkov, A. O. Kosogov, S. S. Ruvimov, P. Werner, U. Gosele and J. Heydenreich, *Phys. Rev. B* **54**, 8743 (1996)

[62] M. A. Migliorato, L. R. Wilson, D. J. Mowbray, M. S, Skolnick, M. Al-Khafaji, A. G. Cullis and M. Hopkinson, *J. Appl. Phys.* **90**, 6384 (2001)

[63] C. Heyn, *Phys. Rev B* **64**, 165306 (2001)

[64] D. Leonard, K. Pond and P. M. Petroff, *Phys. Rev B* **50**, 11687 (1994)

[65] J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. Andre and O. Vatel, Appl. Phys. Lett. 64, 196 (1994)

[66] A. Madhukar, Q. Xie, P.Chen and A. Konkar, *Appl. Phys. Lett*. **64**, 2727 (1994)

[67] I. Daruka, J. Tersoff and A. L. Barabasi, *Phys. Rev. Lett*. **82**, 2753 (1999)

[68] G. Costantini, A. Rastelli, C. Manzano, R. Songmuang, O. G. Schmidt, K. Kern and H. von Kanel, *Appl. Phys. Lett.* **85**, 5673 (2004)

[69] G. Constantini, A. Rastelli, C. Manzano, P. Acosta-Diaz, R. Songmuang, G. Katsaros, O. G. Schmidt and K. Kern, *Phys. Rev. Lett.* **96**, 226106 (2006)

[70] P. B. Joyce, T. J. Krzyzewski, P. H. Steans, G. R. Bell, J. H. Neave and T. S. Jones, *Surf. Sci.* **492**, 345 (2001)

[71] K. Eberl, M. O. Lipinski, Y. M. Manz, N. Y. Jin-Phillipp and O. G. Schmidt, *Physica E* **9**, 164 (2001)

[72] R. Songmuang, S. Kiravittaya and O. G. Schmidt, *J. Cryst. Growth* **249**, 416 (2003)

[73] E. Penev, S. Stojkovic, P. Kratzer and M. Sheffler, *Phys. Rev. B* **69**, 115335 (2004)

[74] T. Inoue, T. Kita, O. Wada, M. Konno, T. Yaguchi and T. Kamino, *Appl. Phys. Lett.* **92**, 031902 (2008)

[75] C. Kubel, A. Voigt, R. Schoenmakers, M. Otten, D. Su, T. C. Lee, A. Carlsson and J. Bradley, *Microsc. Micro. Anal*. **11**, 378 (2005)

[76] S. I. Molina, A. M. Sanchez, A. M. Beltran, D. L. Sales, T. Ben, M. F. Chisholm, M. Varela, S. J. Pennycook, P. L. Galindo, A. J. Papworth, P. J. Goodhew and J. M. Ripalda, *Appl. Phys. Lett.* **91**, 263105 (2007)

[77] A. M. Sanchez, R. Beanland, M. H. Gass, A. J. Papworth, P. J. Goodhew and M. Hopkinson, *Phys. Rev. B* **72**, 075339 (2005)

[78] A. Sanchez, R. Beanland, A. J. Papworth, P. J. Goodhew and M. H. Gass, *Appl. Phys. Lett.* **88**, 51917 (2006)

[79] D. Zhi, M. Wei, R. E. Dunin-Borowski, P. A. Midgley, D. W. Pashley, T. S. Jones, B. A. Joyce, P. F. Fewster and P. J. Goodhew, *Microelect. Eng.* **73-74**, 604 (2004)

[80] P. Wang, A. L. Bleloch, M. Faluk, P. J. Goodhew, J. Ng and M. Missous, *Appl. Phys. Lett.* **89**, 72111 (2006)

[81] A. Migliorato, A. G. Cullis, M. Fearn and J. H. Jefferson, *Phys. Rev. B* **65**, 115316 (2002)

[82] D. Zhi, H. Davock, R. Murray, C. Roberts, T. S. Jones, D. W. Pashley, P. J. Goodhew and B. A. Jones, *J. Appl. Phys.* **89**, 2079 (2001)

[83] D. E. Jesson and S. J. Pennycook, *Proc. R. Soc. Lond. A* **441**, 261 (1993)

[84] S. J. Pennycook, D. E. Jesson, A. J. McGibbon, P.D. Nellist, *J. Elect. Microsc.* **45**, 36 (1996)

[85] E. Penev, P. Kratzer and M. Scheffler, *Phys. Rev. B* **64**, 85401 (2001)

[86] P. Howe, E. C. Le Ru, E. Clarke, B. Abbey, R. Murray and T. S. Jones, *J. Appl. Phys.* **95**, 2998 (2004)

[87] P. Howe, E. C. Le Ru, E. Clarke, R. Murray and T. S. Jones, *J. Apl. Phys.* **98**, 113511 (2005)

[88] S. R. Sternberg, *IEEE Computer* **16**, 22 (1983)

[89] N. Tansu and L. J. Mawst, *IEEE Photon. Tech. Lett*. **14**, 444 (2002)

[90] R. Fehse, S. Tomic, A. R. Adams, S. J. Sweeney, E. P. O'Reilly, A. Andreev and H. Riechert, *IEEE J. Select. Top. Quantum. Elect.* **8**, 801 (2002)

[91] P. R. Chalker, T. J.Bullough, M. Gass, S. Thomas and T. B. Joyce, *J. Phys.: Cond. Matter* **16**, S3161 (2004)

[92] W. M. McGee, R. S. Williams, M. J. Ashwin and T. S. Jones, *Surf. Sci.* **600**, L194 (2006)

[93] M. Herrera, D. Gonzalez, M. Hopkinson, M. Gutierrez, P.Navaretti, H. Y. Liu and R. Garcia, *J. Appl. Phys*. **97**, 073705 (2005)

[94] J. M. Ulloa, L.Hierro, M. Montes, B. Damilano, M. Hugues, J.Barjon, J. Y. Duboz and J. Massies, *Appl. Phys. Lett.* **87**, 251109 (2005)

[95] H. Y. Liu, C. M. Tey, C. Y. Jin, S. L. Liew, P. Navaretti, M. Hopkinson and A. G. Cullis, *Appl. Phys. Lett.* **88**, 191907 (2006)

[96] K. Volz, T. Torunski and W. Stolz, *J. Appl. Phys.* **97**, 014306 (2005)

[97] H. F. Liu, N. Xiang and S. J.Chua, *Appl. Phys. Lett*. **88**, 191907 (2006)

[98] H. F. Liu, N. Xiang and S. J. Chua, *Appl. Phys. Lett*. **89**, 71905 (2006)

[99] M. Herrera, D. Gonzalez, M. Hopkinson, P. Navaretti, M. Gutierrez, H. Y. Liu and R. Garcia, *Semicond. Sci. Technol*. **19**, 813 (2004)

[100] J. Neugebauer and C. G. Van de Walle, *Phys. Rev. B* **51**, 10568 (1995)

[101] I. P. Ipatova, V. G. Malyshkin and V. A. Shchukin, *J. Appl. Phys*. **74**, 7198 (1993)

[102] J. W. Cahn, *Acta. Metall.* **9**, 795 (1961)

[103] J. S. Langer, M. Bar-On and H. D. Miller, *Phys. Rev. A* **11**, 1417 (1975)

[104] F. Glas, *Phys. Rev. B* **62**, 7393 (2000)

[105] M. Herrera, D. Gonzalez, M. Hopkinson, M. Gutierrez, P. Navaretti, H. Y. Liu and R. Garcia, *Phys. Stat. Sol. C* **2**, 1292 (2005)

[106] J. M. Moison, C. Guille, F. Houzay, F. Barthe and M. Van Rompay, *Phys. Rev. B* **40**, 6149 (1989)

[107] X. Kong, A. Trampert, E. Tournie and K.H. Ploog, *Appl. Phys. Lett*. **87**, 171901 (2005)

[108] X. Kong, A. Trampert and K.H. Ploog, *Micron* **37**, 465 (2006)

[109] W. M. McGee, R. S. Williams, M. J. Ashwin, T. S. Jones, E. Clarke, J. Zhang and S. Tomic, *Phys. Rev. B* **76**, 85309 (2007)

[110] D. Bohm and D. Pines, *Phys. Rev.* **82**, 625 (1951)

[111] D. Pines and D. Bohm, *Phys. Rev.* **85**, 338 (1952)

[112] P. Specht, X. Xu, R. Armitage, E. R. Weber, R. Erni and C. Kisielowski, *Physica B* **376-377**, 552 (2006)

[113] R. Erni and N. D. Browning, *Ultramicrosc.* **104**, 176 (2005)

[114] L. Gu, V. Srot, W. Sigle, C. Koch, P. van Aken, F. Scholz, S. B. Thapa, C. Kirchner, M. Jetter and M. Ruhle, *Phys. Rev. B* **75**, 195214 (2007)

[115] M. Stoger-Pollach and P. Schattschneider, *Ultramicrosc.* **107**, 1178 (2007)

[116] W. J. De Ruijter, *Micron* **26**, 247 (1995)

[117] H. A. Brink, M. M. G. Barfels, R. P. Burgner and B. N. Edwards, *Ultramicroscopy* **96**, 367 (2003)

[118] B. Rafferty, S. J. Pennycook and L. M. Brown, *J. Elect. Microsc.* **49**, 517 (2000)

[119] M. Stoger-Pollach, *Micron* **39**, 1092 (2008)

[120] R. Erni and N. D. Browning, *Ultramicrosc.* **108**, 84 (2008)

[121] E. Kröger, *Z. Phys.* **216**, 115 (1968)

[122] E. Kröger, *Z. Phys.* **235**, 403 (1970)

[123] F. J. Garcia de Abajo, A. Rivacoba, N. Zabala and N. Namamoto, *Phys. Rev. B* **69**, 155420 (2004)

[124] C. von Festenberg and E. Kröger, *Phys. Lett.* **26A**, 339 (1968)

[125] M. Stoger-Pollach, H. Franco, P. Schattschneider, S. Lazar, B. Schaffer, W. Grogger and H. W. Zandbergen, *Micron* **37**, 396 (2006)

[126] M. Stoger-Pollach, A. Laister and P. Schattschneider, *Ultramicrosc.* **108**, 439 (2008)

[127] E. D. Palik, *Handbook of Optical Constants of Solids*, 1st Edition, Academic Press (1997)

[128] P. Lautenschlager, M. Gariggar, S. Logothetdis and M. Cardona, *Phys. Rev. B* **35**, 9174 (1987)

[129] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*, 3rd Edition, Springer: New York (2005)

[130] M. S. Moreno, K. Jorrisen and J. J. Rehr, *Micron* **38**, 1 (2007)

[131] E. Kreyszig, *Advanced Engineering Mathematics*, 8th Edition, Wiley: New York (1998)