Analysis of the Temperature Dependent Band-Edge Distribution as a Function of Annealing in ZnO and ZnO-Based Thin Film Alloys

Presented in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy

with a
Major in Physics

in the

College of Graduate Studies

University of Idaho

by

Amrah Canul, M.S.

Major Professor
Leah Bergman, Ph.D.

Committee
Ruprecht Machleidt, Ph.D.
M. Grant Norton, Ph.D.

Michael J. Anderson, Ph.D.

Department Administrator
John Hiller, Ph.D.

May 2020

Authorization to Submit Dissertation

This dissertation of Amrah Canul, submitted for the degree of Doctor of Philosophy with a Major in Physics and titled "Analysis of the Temperature Dependent Band-Edge Distribution as a Function of Annealing in ZnO & ZnO-Based Thin Film Alloys," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor:	Leah Bergman, Ph.D	Date
Committee Members:	Ruprecht Machleidt, Ph.D	Date
	M. Grant Norton, Ph.D	Date
	Michael J. Anderson, Ph.D	Date
Department Administrator:	John Hiller, Ph.D	Date

Abstract

An analytical approach based on the derivative of the absorption spectra was used to explore the nature of the band-edge of ZnO and MgZnO thin films grown via the sputtering technique. It was found that the band-edge is composed of a Gaussian where the width and peak position were employed to model the electron-phonon (e-p) interaction and defect characteristics of the film. These characteristics were studied via transmission experiments in the temperature range of 77 K to 532 K. The as-grown film was found to exhibit weak e-p coupling relative to the static contribution of defects similar to the behavior of amorphous-like semiconductors. Upon successive controlled annealing of the ZnO film, up to 800 ^{o}C, the defect component diminished and the phonon contribution became dominant. Specifically, the e-p interaction was similar to semiconductors with long range order. X-ray diffraction and imaging studies agree with these results. The defects in the ZnO films are discussed in terms of structural inhomogeneities and Zn interstitials prevalent in ZnO films grown via sputtering. MgZnO alloys show similar defect and e-p coupling behavior. Furthermore, building on the derivative analysis, we show an analytical technique capable of separating signal/fringe-pattern convolution at the band-edge. Finally, we develop an analytic set of models that can be applied to the full range of nonlinear band-edges. We show that, in the limiting case of linearity, our model is consistent with and produces the same results as the linear Urbach model. Our new model is capable of bypassing "tailing" as an indirect measure of the spatially varying local potential and gives a direct measure of this quantity.

Acknowledgements

A million thanks to all the beautiful souls who removed the barriers along the long path to this success. Thanks to Mrs. Darlene Peters for introducing me to the wonderfully unforgiving world of Physics in high school at St. John's College in Belize City. Thanks to Gavin Courtenay, who took me traveling across my home country, ultimately to meet Lisel Alamilla. Thank you to Lisel Alamilla for introducing me to Bill and Deb McLaughlin who recruited me to the University of Idaho. Thank you to Carmella LeBlanc who gave selflessly to make the United States feel like my home away from home. Thanks to Dr. John Crepeau and the University of Idaho Department of Mechanical Engineering for financial support and believing in me in my darkest academic times. Thank you to my aunts and uncles, Magda Alamilla, Lenie and Lester Hulse and Adelita and Khalid Ghazy for their generous support. Thanks to the University of Idaho International Programs Office for their financial support and welcoming attitude. Thank you to Dr. David McIlroy for believing in my aptitude for Physics and recommending my acceptance into (and helping to facilitate funding for) the graduate program in the Department of Physics at the UI. My deepest gratitude is offered to Dr. Leah Bergman. Your guidance and insight has helped me to cultivate the skills necessary to make a meaningful contribution to the scientific community. Thank you to my committee members, Dr. Machleidt, Dr. Norton and Dr. Anderson, for your efforts and expertise in overseeing this work and my defense. Thanks to my colleagues, Dr. Jesse Huso, Dr. Dinesh Thapa and Jeffrey Lapp for providing stimulating discussion and debate in the pursuit of uncovering scientific truth. I extend my deepest appreciation to Caleb Robinson for his enthusiastic and invigorating discussions. You always provide the catalyst needed to connect key ideas extending the breadth and depth of my understanding of the natural sciences at large. Special thanks to my parents, Maritta and Jose Luis Canul. Not a day passes that I do not acknowledge and appreciate the sacrifices you have made that allow me to do the work I am doing today. To all my family and friends, you have provided me over the years with nothing but love and support and I am thankful, and so very lucky, to have people like you in my life.

Finally, I gratefully acknowledge the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Science and Engineering, Award No. DE-FG02-07ER46386.

Dedication

To my dearest Caleb. You are the spirit of the age. I love you and thank you for your tireless efforts in keeping me on-task, fit, happy and healthy. You do more for me selflessy than I have ever deserved. Your love has made this work possible.

Contents

Authorization to Submit Dissertation . . . ii
Abstract . . . iii
Acknowledgments . . . iv
Dedication . . . vi
Contents . . . vii
List of Tables . . . viii
List of Figures . . . ix
1 Introduction. . . . 1
1.1 Multidisciplinary Approach . . . 3
References . . . 5
2 The Electron-Phonon Interaction. . . . 6
2.1 The Electron . . . 6
2.2 Energy Levels . . . 7
2.2.1 Spherical Harmonics . . . 7
2.3 Chemical Bonds . . . 8
2.4 Bonding in Solids . . . 15
2.5 Classification of Solids . . . 20
2.6 Excitation . . . 21
2.6.1 Exciton Localization . . . 24
2.7 Excited States and Vibrational Modes . . . 26
2.8 Momentum and Bloch Waves . . . 28
2.8.1 Thermal Transport and Phonons . . . 33
2.9 Phonons in Real Solids . . . 35
2.9.1 Type of Phonons . . . 36
2.10 Electron-Phonon Interaction . . . 37
2.11 Thermal Broadening from Electron-Phonon Interaction . . . 39
References . . . 41
3 Thin Film Synthesis and Methodology . . . 42
3.1 Physical Vapor Deposition & Sputtering . . . 42
3.2 Sputtering Gases . . . 44
3.3 Arc Suppression . . . 45
3.4 Substrates . . . 46
3.5 Magnetic Shielding . . . 48
3.6 Magnetron Sputtering . . . 48
3.7 Sputtering System in This Work . . . 49

3.8 Sputtering Targets . . . 51
3.9 Methodology . . . 52
3.9.1 ZnO Thin Films . . . 53
3.9.2 Wurtzite MgZnO Thin Films . . . 53
3.9.3 Wurtzite-Cubic Mixed-Phase MgZnO Thin Films . . . 53
References . . . 56

4 Semiconductors & Bandgap Engineering . . . 57
4.1 Alloying—ZnO & MgO . . . 57
4.1.1 Solubility of Crystal Structures . . . 58
4.2 Defects . . . 61
References . . . 63

5 Optical Spectroscopy and Imaging. . . . 64
5.1 Optical Spectroscopy . . . 64
5.2 Electromagnetic Radiation . . . 64
5.3 Transmission of Radiation . . . 65
5.4 Absorption & The Significance of Resonance . . . 67
5.5 General Spectroscopic Techniques . . . 68
5.6 Transmittance & Absorbance . . . 68
5.7 Scanning Electron Microscopy (SEM) . . . 70
5.8 X-Ray Diffraction (XRD) . . . 72
References . . . 76

6 The Temperature Dependent Band-Edge as a Function of Annealing 77
6.1 Introduction . . . 77
6.2 Defect Distribution . . . 78
6.3 Annealing Experimental Results . . . 83
6.4 Solid-State Phase Transition . . . 84
6.5 Defect-Mediated Electron-Phonon Coupling . . . 87
6.6 Model of Temperature-Dependent Band-Edge Distribution Width . . . 89
6.7 Evidence for Zn_i Out-Diffusion/Lattice Incorporation . . . 91
References . . . 95

7 Analysis of Interference Fringe Patterns in Thin Film Absorption Spectra . . . 97
7.1 Introduction . . . 97
7.2 Overview of Analytical Approach . . . 101
7.3 Analysis . . . 102
7.4 Qualitative Interpretation of Analysis . . . 110
7.5 Temperature Dependent Interference . . . 110
7.6 Modulating Interference Fringes . . . 113
7.7 Interference Pattern Removal . . . 119
7.8 Conclusion . . . 119
References . . . 121

8 Analysis and Modeling of the Band-Edge 122
8.1 History of Urbach Theory 122
8.2 The Basic Urbach Theory 125
8.3 Urbach Analysis in the Literature 128
8.4 Differential Approach 129
8.5 Model Basis 129
8.6 Model Development 132
8.6.1 Differential Analysis & Deconvolution of the Absorption-edge . 132
8.6.2 Piece-wise Reconstruction of the Complex band-edge 143
8.7 Limiting and Special Cases of Nonlinear Model 150
8.7.1 A Limiting Case: The Basic Urbach Model 150
8.7.2 A Special Case: Appearance of a Linear Band-Edge 152
8.7.3 The Urbach Focus and the Unperturbed Band-Edge 154
8.8 Inhomogeneous Broadening 161
8.9 Further Discussion & Conclusion 165
References 169

List of Tables

Table 4.1 Material Properties of Mg & ZnO.. 59

Table 6.1 Gaussian Model Parameters and XRD 86

Table 7.1 Fringe Analysis Parameters 104
Table 7.2 Calculated Interference Maxima/Minima 105
Table 7.3 Predicted Interference Maxima/Minima 105

Table 8.1 Logistic Model Parameters 139
Table 8.2 Logistic Model Parameters Cont'd 140
Table 8.3 Electron-Phonon Interaction Parameters 153
Table 8.4 Annealing Experiment Parameters 160

List of Figures

Figure 1.1 Semiconductor Material Properties 1

Figure 2.1 Spherical Harmonics 9
Figure 2.2 Bonding Schematic 10
Figure 2.3 Overlap Integral 12
Figure 2.4 Morse Potential 13
Figure 2.5 Bonding in Ti_2 15
Figure 2.6 Bonding in Metallic Solids 17
Figure 2.7 Energy Bands in a Solid. 18
Figure 2.8 Electron Density of ZnO 19
Figure 2.9 Classification of Solids 21
Figure 2.10 Isodensity Plots of the Frontier Orbitals of Wurtzite Type ZnO . . 22
Figure 2.11 Hole Mobility. 23
Figure 2.12 Free Exciton 24
Figure 2.13 Localized Exciton 25
Figure 2.14 Vibrational Modes of Electron States. 26
Figure 2.15 The Fourier Transform 27
Figure 2.16 Bloch Wave. 29
Figure 2.17 Dispersion 32
Figure 2.18 Bloch Wave Schematic 33
Figure 2.19 Wave Packet 34
Figure 2.20 Phonon dispersion of ZnO 36
Figure 2.21 Optical Phonons in ZnO Quantum Dots 37
Figure 2.22 Vibronic Transitions in Exciton Peaks 38

Figure 3.1 Sputtering Schematic 43
Figure 3.2 Magnetic Substrate Shielding 47
Figure 3.3 Magnetron Sputtering 49
Figure 3.4 Sputtered Samples 50
Figure 3.5 Sputtering System 51
Figure 3.6 Custom $Mg_{0.07}Zn_{0.93}O$ Target 52

Figure 4.1 Solubility of ZnO & MgO 58
Figure 4.2 $Mg_xZn_{1-x}O$ Alloy Composition 61

Figure 5.1 Electromagnetic Spectrum. 65
Figure 5.2 Refraction of Radiation 66
Figure 5.3 Transmittance to Absorbance 69
Figure 5.4 SEM Image 71
Figure 5.5 XRD Schematic 74

Figure 6.1 Defect Concentration and Phonon Occupation 79

FIGURE 6.2 Derivative Spectrum 80
FIGURE 6.3 Urbach Analysis & Derivative Analysis 81
FIGURE 6.4 Annealing Effects on Differential Band-Edge & XRD 82
FIGURE 6.5 XRD Correlation and SEM 85
FIGURE 6.6 Thermal Broadening as a Function of Annealing 88
FIGURE 6.7 Temperature-Dependent Electron-Phonon Coupling 90
FIGURE 6.8 Defect-mediated Electron-Phonon Coupling 92
FIGURE 6.9 PL for As-grown vs. Annealed ZnO 93

FIGURE 7.1 Transmission Spectra Regions 98
FIGURE 7.2 Bulk vs. Thin Film Interference 100
FIGURE 7.3 Envelope Method 103
FIGURE 7.4 Multiple Internal Reflections 106
FIGURE 7.5 Fringe Order 107
FIGURE 7.6 Deviation from Ideal 108
FIGURE 7.7 Energy Location of Interference Fringes 109
FIGURE 7.8 Analytical Pipeline 111
FIGURE 7.9 Ideal Function Analysis 112
FIGURE 7.10 Temperature Average 114
FIGURE 7.11 Band-Edge Temperature Shift 115
FIGURE 7.12 Interference Analysis Regions 116
FIGURE 7.13 Modulated Interference Fringes 117
FIGURE 7.14 Interference Pattern Removal 118

FIGURE 8.1 Urbach Theory 127
FIGURE 8.2 Logistic Functions 131
FIGURE 8.3 ZnO Temperature Stack 133
FIGURE 8.4 Urbach Focus 134
FIGURE 8.5 77 K ZnO Stack 135
FIGURE 8.6 Phase Separated MgZnO 144
FIGURE 8.7 Wurtzite Mixed-Phase MgZnO 145
FIGURE 8.8 Cubic Mixed-Phase MgZnO 146
FIGURE 8.9 c-Si Band-Edge 155
FIGURE 8.10 Urbach Focus 156
FIGURE 8.11 Linear Band-Edge from Interference 157
FIGURE 8.12 ZnO Step-wise Annealing 159
FIGURE 8.13 XRD Correlation 162

CHAPTER 1

Introduction

Semiconductors have been the foundation of modern technology. First generation semiconductors such as Si and GaAs are still the most prolific semiconductor materials in use today. While first generation semiconductor technology is mature and very well established, these materials have presented some important limitations in the advancement of the technology. As shown in Fig. 1.1, the limitations in application parameters mainly include high voltage operation, high temperature applications and high frequency switching.[1] High voltage operation requires a large energy gap and high critical electric field (related to the breakdown or blocking voltage capacity). High frequency switching also requires a large energy gap but requires high electron velocity or mobility. High temperature applications require high thermal conductivity and a high melting point. Some semiconductor devices include BJTs (bipolar junction transistors), HEMTs (High Electron Mobility Transistors), IGBTs (Insulated Gate Bipolar Transistors), JFETs (Junction Gate Field Effect Transistors), MOSFETs (Metal-Oxide-Semicondcutor Field Effect Transistor) and thyristors.

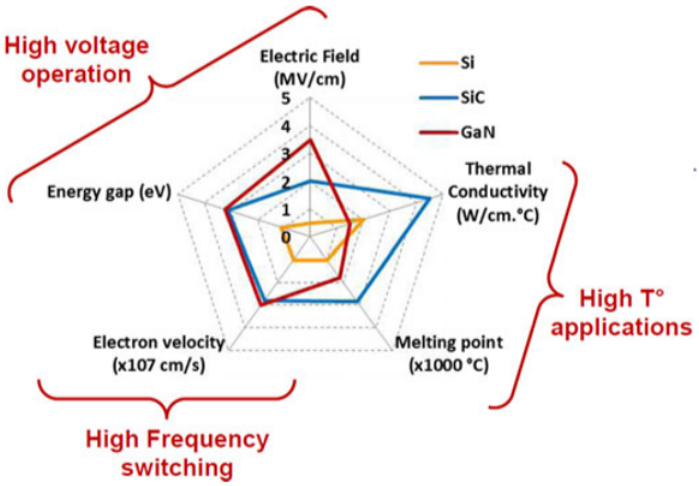


FIGURE 1.1: Semiconductor material properties important for applications mainly include high voltage, high temperature and high frequency switching.[1]

The advent of wide-bandgap (WBG) semiconductors, such as GaN, SiC and ZnO, have addressed some of these shortcomings. While GaN and SiC are relatively equal in high voltage and high frequency applications, GaN provides a higher critical electric field, while SiC provides higher thermal conductivity. ZnO also provides high electron mobility.[1]

The United States has led the way on WBG research. In 2002, DARPA initiated the Wide Bandgap Semiconductor Technology Initiative (WBGSTI) that lasted from 2002-2009 and included phases of material development, device development and reliability and performance optimization.[2] This initiative has effected advances in technology such as high RF wireless communication, next generation jammers, air-missile defense radars, Active Electronically Scanned Array detection, and long-distance, small target detection. The enhanced detection is attributed to the ability of WBG semicondutors to amplify signals without adding noise. Replacing Si with SiC, for example cuts system power consumption to 1/25 of original levels. It has also enhanced switching speeds by up to 200 times. WBG semiconductors have also assisted in carbon emission reduction due to enhanced efficiency of DC to AC power conversion in hybrid electric vehicles.

Ultimately, the improved efficiency of WBG semiconductors markedly reduce heat dissipation which make cooling systems smaller, lighter and simpler. In addition to extending the range of vehicles and aircraft from these improved systems, there is additional reduction in production costs.[2]

Recently, semiconductors with bandgaps significantly wider than 3.4 eV have attracted attention due to their potential to access previously inaccessible levels of performance.[2] The bandgap scales nonlinearly with device performance and Ultra-Wide Bandgap (UWBG) semiconductors offers higher breakdown fields, higher temperature stability and higher irradiation robustness than currently available WBG technology. Recent advances in growth and substrate development have opened a

new era of UWBG semiconductor materials. Some UWBG materials include AlGaN, β-Ga_2O_3 and boron nitride.

Future applications in WBG and UWBG semiconductors will be in satellite communications, high speed computers, precision-guided early warning detection systems, intelligence and reconnaissance, electronic warfare and intelligent fire control systems.

1.1 MULTIDISCIPLINARY APPROACH

It is the intention of the author to draw information from a wide array of knowledge bases, including solid-state physics and chemistry, electromagnetism, optics, and material science, in hopes of more clearly resolving the complexity of their interaction in the field of semiconductor study, if only in a qualitative sense. History shows that scientists have long used heuristic models to better communicate the abstract qualities of the physical phenomena they wish to describe, usually independent of mathematical treatment. These models serve as a method of communicating large amounts of information in a condensed fashion in an attempt to accelerate scientific thought. While mathematics is important in Science and provides a rigorous framework for physics, it represents operational or computational knowledge of subjects that can also be communicated qualitatively. For this reason, this work will attempt to use models that do not operationally describe all aspects of this work, but which have some pedagogical value and lend qualitative perspective on the issues at hand. The models should provide self-consistency throughout a wide array of disciplines and promote meaningful and insightful interpretation of experimental data. The majority of data analysis in this work is based in statistics and treats the bulk behavior of electrons in semiconductors. Statistical treatments are useful as a starting point to bridging explanatory gaps between emergent phenomena in semiconductors and fundamental rules that govern individual electrons, however, this function is performed best when

coupled with our most current models of electrons. As such, a large portion of this work will attempt to interpret the statistical behavior of electrons in semiconductors through the pedagogical lens of molecular orbital theory and electron wave functions, which offers some insight into the physical forces that give rise to various selection rules observed in transitions at the band-edge.

REFERENCES

[1] J. Millan, P. Godignon, X. Perpina, A. Perez-Tomas, and J. Rebollo, IEEE Transactions on Power Electronics **29**, 2155 (2014).

[2] H. Jin, L. Qin, L. Zhang, X. Zeng, and R. Yang, MATEC Web of Conferences **40**, 01006 (2016), publisher: EDP Sciences.

[3] A. Canul, *Analysis of Band-Edge Dynamics in ZnO and MgZnO Thin Films*, Ph.D. thesis, University of Idaho (2015).

[4] G. D. Cody, B. Abeles, B. Brooks, P. Persans, C. Roxlo, A. Ruppert, and C. Wronski, Journal of Non Crystalline Solids **59-60**, 325 (1983).

[5] S. M. Wasim, C. Rincón, G. Marín, P. Bocarand a, E. Hernández, I. Bonalde, and E. Medina, **64**, 195101 (2001).

[6] H. Che, *Band gaps by design: Tailoring ZnO based semiconductor alloy films*, Ph.D. thesis, University of Idaho (2014).

[7] E. Economou, C. Soukoulis, M. H. Cohen, and S. John, in *Disordered Semiconductors* (Springer, 1987) pp. 681–695.

[8] D. Thapa, J. Huso, J. L. Morrison, C. D. Corolewski, M. D. McCluskey, and L. Bergman, Optical Materials **58**, 382 (2016).

[9] M. Vishwas, K. N. Rao, K. A. Gowda, and R. Chakradhar, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **77**, 330 (2010).

CHAPTER 2

The Electron-Phonon Interaction

Abstract — This chapter is dedicated to the keystone advances and developments in the field of solid-state physics which underlie the electron-phonon interaction in semiconductor materials. The chapter draws from fundamental texts in the fields of solid-state physics and molecular orbital theory of bonding from physical chemistry. Sec. 2.1 begins with a modern understanding of the electron from quantum wave mechanics. Sec. 2.2 discusses the orbital energy levels associated with electrons. Sec. 2.3 describe how electrons become associated through wavefunction overlap. Sec. 2.4 outlines how solids are formed through the association of atoms. Sec. 2.5 describes the three classes into which solids are grouped. The physics of excited states and their interaction with vibrational modes are described in Sec. 2.6 and Sec. 2.7, respectively. Momentum and k vectors are introduced in Sec. 2.8. Phonons in real (imperfect) solids are discussed in Sec. 2.9. Finally, the electron-phonon interaction is introduced in Sec. 2.10.

2.1 THE ELECTRON

Electrons are small and light and thus show properties of both particles and waves. In 1923, Louis de Broglie suggested that the properties of electrons in atoms are better explained by treating the electrons as waves rather than as particles.

Electrons that are bound to nuclei occupy an area of space referred to as an orbital. The Heisenberg uncertainty principle states that we can never determine exactly where the electron is. We can, however, determine the electron density—the probability of finding the electron in a particular part of the orbital. The orbital is thus defined as an *allowed energy state* for an electron, with an associated probability function that defines the distribution of electron density in space.[1]

2.2 ENERGY LEVELS

Atomic orbitals are grouped into different "shells" at different distances from the nucleus. Each shell is identified by a principal quantum number n, with $n = 1$ for the lowest-energy shell closest to the nucleus. As n increases, the shells are farther away from the nucleus—higher in energy—and can hold more electrons.

The first electron shell contains just the 1s orbital which is spherically symmetric. The electron density decays exponentially as a function of the distance from the nucleus. The second electron shell consists of the 2s and 2p orbitals. The 2s orbital is also spherically symmetric—although not simply an exponential function. Because most of the 2s electron density will be farther from the nucleus than that of the 1s, the 2s orbital will be higher in energy. In addition to the 2s orbital, the second shell also contains three 2p atomic orbitals, one oriented in each of the three spatial directions. These orbitals are called $2p_x$, $2p_y$ and $2p_z$. The 2p orbitals are higher in energy than the 2s, because the average location of the electron in a 2p orbital is farther from the nucleus.

2.2.1 *Spherical Harmonics*

The electron in an atomic orbital is like a stationary, bound, vibration—a standing wave in three dimensions—and spherical harmonics are used in the computation of their configurations.[1] In essence, an electron's energy in an orbital must be described by an integer number of standing waves to be stable, otherwise some of the waves would be radiated as light or lost as heat.[2]

FUNDAMENTAL FREQUENCY–1S — The waveform of a 1s orbital is like a vibrating guitar string at its fundamental frequency[1]), except in three-dimensions. The time-dependent wave function, ψ, is a mathematical description of the shape of the

[1]The fundamental frequency of a guitar string is a standing wave with the entire string alternately displaced upward and downward

wave as it vibrates and the electron density at any point is given by ψ^2, the square of the wave function at that point.

Blue and yellow regions of the wave functions in Fig. 2.1 indicate the instantaneous phase of the constantly changing wave function.

First Harmonic–2p — If you place a finger at the center of a guitar string, while plucking the string, your finger keeps the midpoint of the string from moving. The displacement at the midpoint is always zero; this point is a node. The string now vibrates in two parts, with the two halves vibrating in opposite directions—the two halves of the string are out of phase. When one is displaced upward, the other is displaced downward.

The first harmonic of the guitar string resembles the higher energy $2p$ orbital. The $2p$ orbitals are represented as two "lobes" separated by a node. The two lobes of the p orbital are out of phase with each other. Whenever the wave function is positive in one lobe, it is negative in the other.

Fig. 2.1 shows the forms of the first few real spherical harmonic functions.

2.3 CHEMICAL BONDS

In 1915, G.N. Lewis proposed several new theories describing how atoms bond.[3] One of these theories states that a filled shell of electrons is especially stable and atoms transfer or share electrons in such a way as to attain a filled shell of electrons. This principle has become known as the octet rule because a filled shell implies eight valence electrons for the elements in the second row of the periodic table.

It is known today that atomic orbitals can combine and overlap to give more complex standing waves. We can add and subtract their wavefunctions to give the wave function of new molecular orbitals (MO). This process is described in one way

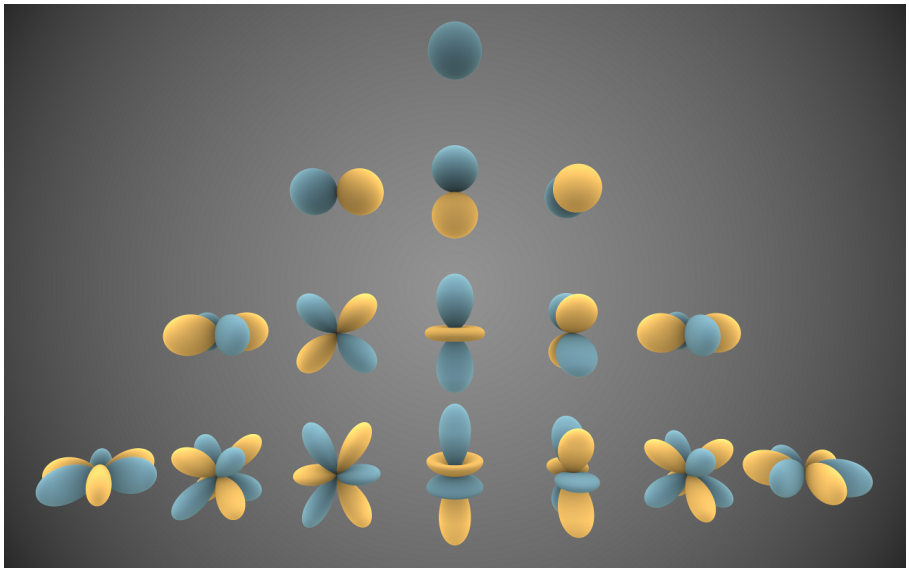


FIGURE 2.1: Visual representations of the first four bands of real spherical harmonic functions. Blue portions are regions where the function is positive, and yellow portions represent regions where it is negative. (Inigo Quilez 2014)

by the linear combination of atomic orbitals (LCAO), developed in 1929 by Sir John Lennard-Jones. [2]

Valence Bond Theory (VB), largely replaced today by LCAO, is also founded in quantum mechanics and demonstrated that electrons hold nuclei together, that is, forms bonds when shared by two nuclei. In 1927, calculations by W. Heitler and F. London established this fact.[4] The results showed that an energy minimum occurs at a certain internuclear distance if the electrons are free to associate with either nucleus. There is always an equilibrium distance for the two bonded nuclei—if they are too close together, their electrostatic repulsion pushes them apart; if they are too far apart, their attraction for the bonding electrons is diminished. The internuclear distance where attraction and repulsion are balanced, which also gives the minimum energy—and the strongest bond—is the bond length. Electron density accumulates the area in space between the two nuclei, creating a bond orbital. Non-bonding orbitals are typically more diffuse than bonding ones.[5]

[2] It is worth noting that orbitals between different atoms can interact as well as orbitals within the same atom.

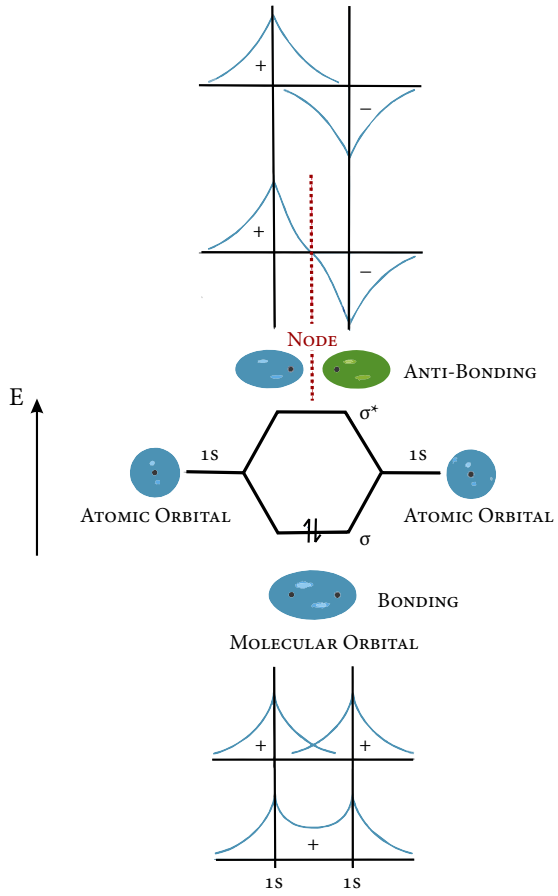


FIGURE 2.2: Schematic of the bonding and anti-bonding orbitals formed by overlapping 1s orbitals.[1]

The hydrogen molecule is the simplest example of covalent bonding in LCAO. As two hydrogen atoms approach each other, their 1*s* wavefunctions can add constructively or destructively where they overlap, depending on the phase. This is illustrated in the schematic in Fig. 2.2. When they are in phase in the region between the nuclei, the wave functions "reinforce" each other and increase the electron density in this bonding region. The result is a σ bonding molecular orbital. When two hydrogen 1*s* orbitals overlap out of phase, a σ^* anti-bonding molecular orbital results. The two 1*s* wave functions have opposite signs and cancel out where they overlap. The result is a nodal plane separating the two atoms.

The function which describes the new electron distribution, the molecular orbital, is called σ, and ψ_1 and ψ_2 are the atomic 1*s* wave functions on atoms 1 and 2.

$$\sigma = c1\psi_1 + c2\psi_2 \tag{2.1}$$

The coefficients, c_1 and c_2, are a measure of the contribution which the atomic orbital is making to the molecular orbital. In this case, they are of course equal in

magnitude, but they may be positive or negative to indicate the phase. To obtain the electron density, we square the function in Eq. 2.1, which gives:

$$\sigma^2 = (c1\psi_1)^2 + (c2\psi_2)^2 + 2c_1\psi_1 c_2\psi_2 \tag{2.2}$$

We see that σ^2 differs from the superposition of the two atomic orbitals $(c1\psi_1)^2 + (c2\psi_2)^2$ by the term $2c_1\psi_1 c_2\psi_2$. This leads to the two solutions shown in Fig. 2.2 The case where both c_1 and c_2 are positive correspond to the electron density increasing between the two atoms, and the formation of a bonding orbital. The case where c_1 and c_2 are of opposite sign models the decreasing of electron density between the two atoms and the formation of an anti-bonding orbital.

The physical interpretation of orbital phase is illustrated by recalling that wave-functions describe the time-dependent nature of electron density in space. When electrons in orbitals are in phase they move in-sync and alternately occupy the same region in space. The electrons form a molecular orbital by keeping out of each others' way with one electron on one side of the orbital, while the other is on the opposite side most of the time. Thus, we are allowed to put two electrons into one orbital if they have opposite spins, even if they still repel each other from having the same charge provided that the electron motion is synchronized. This is referred to as electron correlation.[6]

Conversely, when out-of-phase, they move out-of-sync and try to occupy the same region in space simultaneously. The electrostatic repulsion results in a nodal plane and a significant reduction in electron density in this region. Furthermore, the reduction in electron density allows the nuclei to repel each other ever more strongly. As will be shown in detail shortly, σ bonding molecular orbitals are lower in energy than that of σ^* anti-bonding molecular orbitals.

The force holding the two atoms together is dependent on the extent of the overlap in the bonding orbital. If we bring the two 1s orbitals from a position where there

is essentially no overlap through the bonding arrangement to superimposition, the extent of overlap steadily increases. The mathematical description is called the overlap integral:

$$S_{12} = \int \psi_1 \psi_2 d\tau \tag{2.3}$$

For a pair of 1s orbitals, the overlap integral rises from zero at infinite separation to 1 at superimposition as shown in Fig. 2.3.

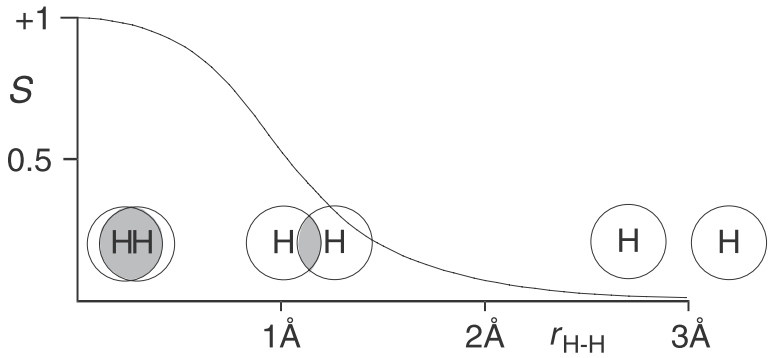


FIGURE 2.3: *The overlap integral S for two* $1s_H$ *orbitals as a function of internuclear distance.*[6]

The energy E of an electron in a bonding and anti-bonding orbital is given by:

$$E = \frac{\alpha + \beta}{1 + S} \tag{2.4}$$

and

$$E = \frac{\alpha - \beta}{1 - S} \tag{2.5}$$

respectively.

α represents the energy of an electron in an isolated atomic orbital, and is called a Coulomb integral. The function represented by the β contributes to the energy of an

electron in the field of both nuclei, and may be called the delocalization integral. It is roughly proportional to the overlap integral so it appears in the equations twice. The function β is negative, lowering the value of E in Eq. 2.4 and raising it in Eq. 2.5.

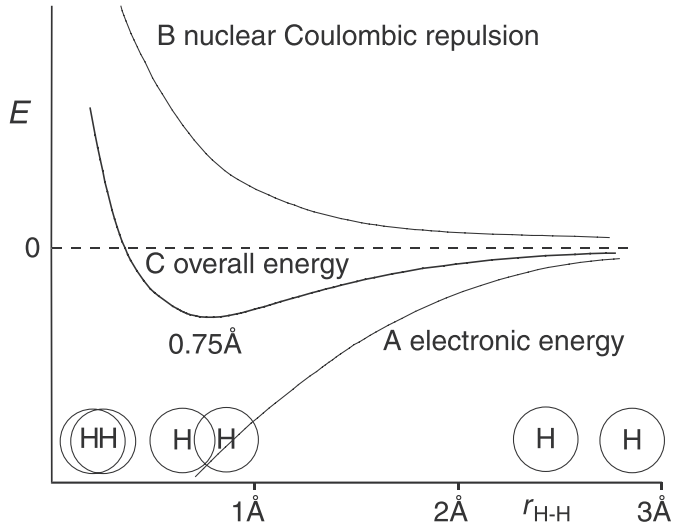


FIGURE 2.4: *Electronic attraction, nuclear repulsion and the overall effect as a function of internuclear distance for two* $1s_H$ *atoms.*[6]

The overall effect on the energy of the hydrogen molecule relative to that of two separate hydrogen atoms as a function of the internuclear distance is given in Fig. 2.4. If the bonding orbital is filled, the energy derived from the electronic contribution in Eq. 2.4 steadily falls as the two hydrogen atoms are moved from infinity towards one another, shown by curve "A." At the same time, the nuclei repel each other and the nuclear contribution to the energy goes steadily up, shown by the curve "B." The sum of these two is the familiar Morse plot "C" for the relationship between internuclear distance and energy, with a minimum at the bond length. If we had filled the antibonding orbital instead, the resultant curve would be a steady increase in energy as the nuclei are pushed together.

The detailed form of α and β is where mathematical complexity appears, even in the simple Hydrogen molecule. They come from the Schrödinger equation and they are integrals over all coordinates and are given by:

$$\alpha = \int \psi_1 H \psi_1 d\tau \tag{2.6}$$

$$\beta = \int \psi_1 H \psi_2 d\tau \tag{2.7}$$

where H is the Hamiltonian, or energy operator written as:

$$\begin{aligned} H = &-\frac{\hbar^2 \nabla_1^2}{2m_e} - \frac{\hbar^2 \nabla_2^2}{2m_e} + \frac{e^2}{4\pi\epsilon_o}\left(\frac{1}{R} + \frac{1}{|\vec{r}_1 - \vec{r}_2|}\right. \\ &\left. - \frac{1}{|\vec{R}_1 - \vec{r}_1|} - \frac{1}{|\vec{R}_2 - \vec{r}_2|} - \frac{1}{|\vec{R}_1 - \vec{r}_2|} - \frac{1}{|\vec{R}_2 - \vec{r}_1|}\right) \end{aligned} \tag{2.8}$$

$\vec{R}_1$ and $\vec{R}_2$ are locations of the two nuclei, $|\vec{R}_2 - \vec{R}_1| = R$, $\vec{r}_1$ and $\vec{r}_2$ are the coordinates of the electrons belonging to the first and second nucleus, respectively. The first two terms refer to the kinetic energy of the two electrons. The operators ∇_1^2 and ∇_2^2 act only on the coordinates $\vec{r}_1$ and $\vec{r}_2$, respectively. The electrostatic term contains the repulsion between the two nuclei and the repulsion between the two electrons as well as the attraction of each electron to each nucleus. It is easy to see how the solution to this problem is not trivial.

Higher energy orbitals overlap in the much the same way as s orbitals to yield bonding and anti-bonding orbitals. Interestingly, bonding and anti-bonding interactions have been observed in isolated dielectric spheres which support resonant electromagnetic modes analogous to electronic orbitals.[7]

The standing-wave model of electrons are useful in describing resonance events that are characteristic of the absorption of radiation.

2.4 BONDING IN SOLIDS

Now that the reader has been introduced to LCAO, we can use this model in shaping the way we envision many atoms coming together to form a solid. Bonding in solids is more complex than in diatomic molecules, but the fundamental principles in LCAO still hold. For instance, the overlap for metals containing d-orbitals is more complex, but gives rise to σ, π, and δ bonding/anti-bonding molecular orbitals as shown in Fig. 2.5.

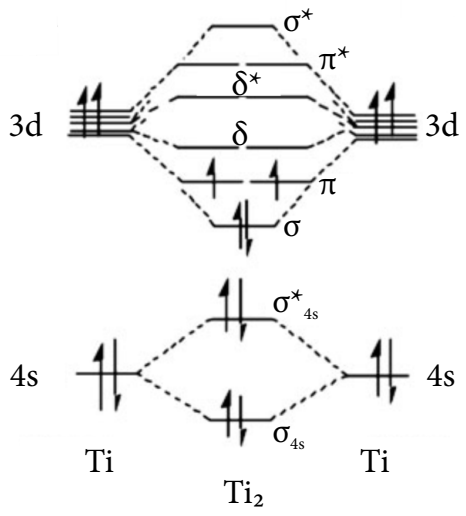


FIGURE 2.5: Molecular orbital diagram for the homonuclear diatomic molecule, Ti_2, showing the discrete energy gap between filled and empty molecular orbitals.

It can be expected that finding the quantum mechanical eigenstates (energy states) of all the electrons in the solid is a formidable problem. We would have to construct a wave function that depends on the coordinates of all the electrons and also of all the nuclei, which make up the positive part of the potential. The first approximation we can make is to ignore the motion of the ions by "freezing" them into their equilibrium position. We are allowed to make this assumption based on the relative mass between electrons and nuclei. When nuclei move out of position, their motion is so slow that the electrons readjust their distribution such that they stay in a modified ground state,

but still in the ground state. When the nuclei move back, the electrons readjust once more. This separation of nuclear motion from electronic motion is known as the Born-Oppenheimer approximation and will be useful in later discussion. The other simplification made in treating a solid system is that we do not consider the correlated motion of all electrons. We only calculate the electronic states for one electron moving in an effective potential given by all nuclei and other electrons.

The stationary Schrödinger equation for the one-electron states then becomes:

$$-\frac{\hbar^2\nabla^2}{2m_e}\psi(r) + U(r)\psi(r) = E\psi(r) \tag{2.9}$$

One great help in finding the eingenstates is the symmetry of the lattice. No matter how complicated the potential $U(r)$ is, we know that is must be lattice periodic, that is:

$$U(r) = U(\vec{r} + \vec{R}) \tag{2.10}$$

Finally, when we have found the eigenstates in the one-electron picture, we fill them with all the electrons according to the Pauli principle. This gives the correct occupation of the states *but only for zero temperature.*

A key concept in LCAO molecular orbital theory is the formation of the same number of molecular orbitals as the number of atomic orbitals that are combined, e.g. there are 12 molecular orbitals formed when $4s$ and $3d$ orbitals combine in the Ti_2 molecule, as shown in Fig. 2.5.[8] Thus, we end up with the same number of molecular orbitals as atomic orbitals, but due to spin-pairing, not all the molecular orbitals will be occupied.

As the number of atoms increases to infinity within a crystal lattice, the difference in energy between bonding and antibonding states goes approximately to zero, resulting in a continuum of states as shown in Fig. 2.6 for a conducting solid.

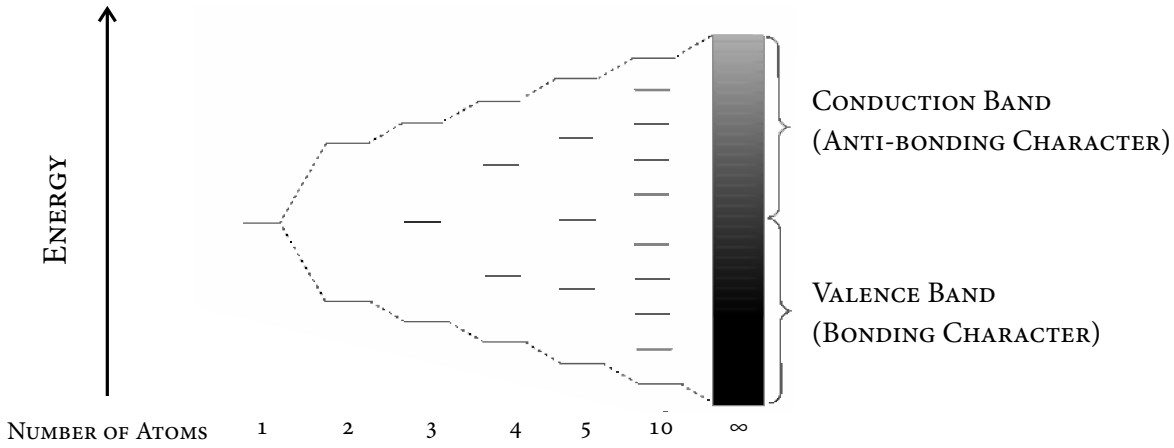
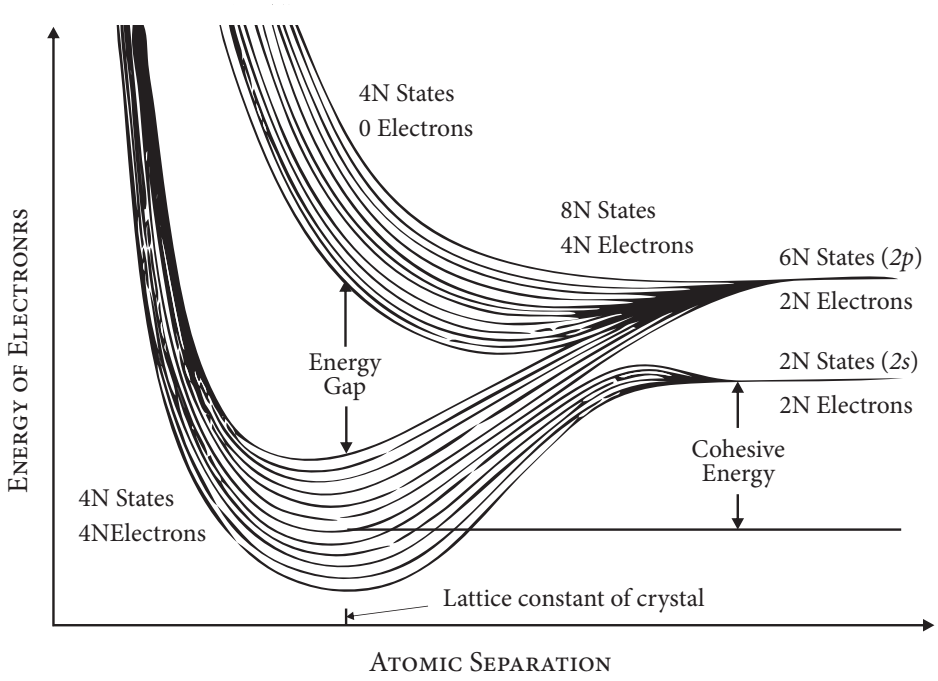
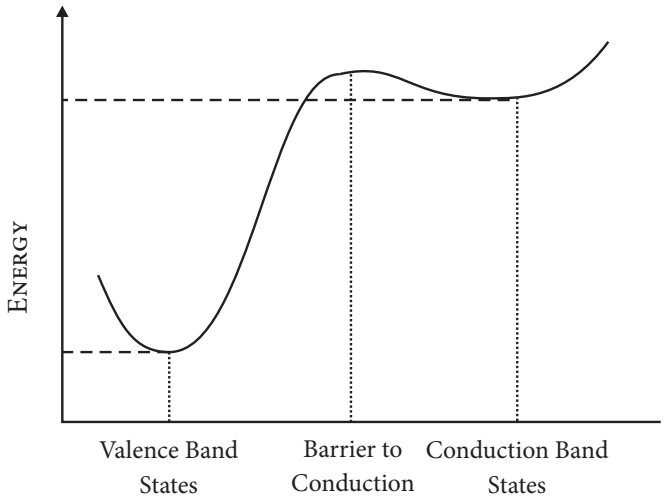


Figure 2.6: Band diagram for a metallic solid illustrating the continuum between valence and conduction bands, i.e. no bandgap, for increasing numbers of atoms.

In semiconducting solids, the formation of a distinct energy gap appears between the continuum of bonding and antibonding states as shown in Fig. 2.7. The figure depicts a band diagram and can be thought of as being constructed of many such Morse potentials as that in Fig. 2.4, with all the energies slightly adjusted to accommodate the Pauli exclusion principle. In semiconductors, the exclusion principle leaves us with all electrons occupying the bonding states while the antibonding states above the gap remain unfilled. The electron occupied band of states is known as the valence band and comprises the lowest allowed energy states. The unfilled band at higher energies constitute the conduction band. The energy gap, E_g, between these levels is known as the bandgap. In essence, the bandgap is the energetic barrier to conduction and delocalized excitation. In the absence of the energy described by this barrier, one would not expect to find an electron within the energy range described. For similar reasons, one would not expect to find an electron in the nodal plane of an orbital due to the massive (infinite) energetic requirement of this state.



a)



b)

Figure 2.7: *Energy bands in diamond as a function of atomic separation. The lattice constant of the crystal gives the lowest energy states.[9] The covalent bonds in diamond are constructed from the linear combination of the* 2*s orbital and three* 2*p orbitals. This results in four* sp^3 *orbitals that stick out in a tetrahedral configuration from the carbon atoms.[9] (b) Schematic illustrating the energetic barrier between valence and conduction states in semiconductors.*

When a band is filled, and all electron states are occupied, the solid cannot carry a current. However, the conduction band contains no electrons, so an electron must assume energies available here in order to traverse the crystal and produce a net current.

It is not instructive to think of solid structures as materials in which there are singly definable molecules.[10] Instead, molecular orbital theory concludes that the orbitals used to describe bonds can encompass many atoms of the structure and are not necessarily associated with particular bonds between any atom pairs.[5] This leads to a model of solids as comprising nuclei periodically embedded in widely distributed electronic clouds. We show the electron density plots for ZnO in Fig. 2.8.

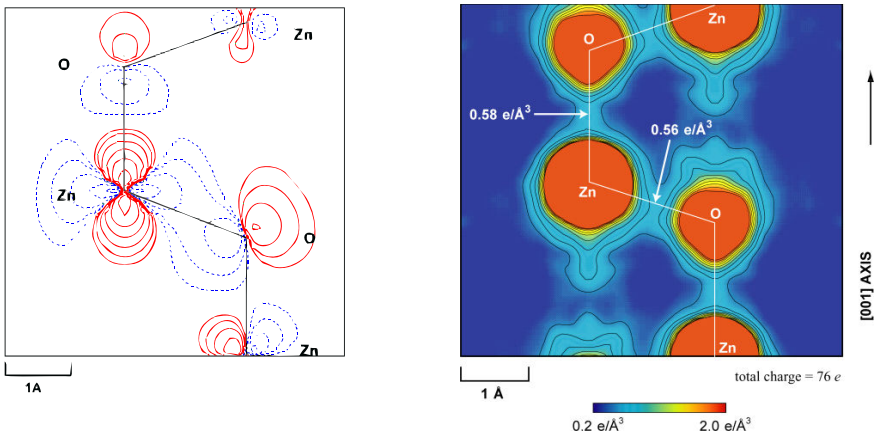


FIGURE 2.8: *Electron wavefunctions of Zn 3d- and O 2p-orbitals via Dv-xα (Discrete Variational Xα) method. Blue indicates the negative region and red indicates the positive region of the wavefunction. (b) The electron density in ZnO by MEM (Maximum Entropy Method). Generally in ZnO, it is the $2s^2 2p^4$ electrons in O and the $3d^{10} 4s^2$ electrons in Zn that are relevant in chemical bonding. In the semiconductor, it is believed that hybridization of localized Zn 3d electron orbitals with O 2p electron orbitals plays a crucial role in energy band formation formation in ZnO.*

Generally in ZnO, it is the $2s^2 2p^4$ electrons in O and the $3d^{10} 4s^2$ electrons in Zn that are relevant in chemical bonding. In the semiconductor, it is believed that hybridization of localized Zn $3d$ electron orbitals with O $2p$ electron orbitals plays a crucial role in energy band formation formation in ZnO.[11]

As the number of atoms increases to infinity within a crystal lattice, the change in energy between energy levels within bonding and anti-bonding regions goes approximately to zero, resulting in a continuum of states. This is an application of the Pauli exclusion principle, which states for N electrons, there must be $N/2$ available states to distribute electron density.

The electron occupied band of states is known as the valence band and comprises the lowest allowed energy states. The unfilled bands at higher energies constitute the conduction band. If present, the energy gap, E_g, between these levels is known as the bandgap as shown in Fig. 2.7a. In essence, the bandgap is the energetic barrier to conduction as illustrated in the schematic in Fig. 2.7b. In the absence of the energy described by this barrier, one would not expect to find an electron within the range described therein. For similar reasons, one would not expect to find an electron in the nodal plane of an orbital due to the massive (infinite) energetic requirement of this state.

When a band is filled, and all electron states are occupied, the solid cannot carry a current. However, the conduction band contains no electrons, so an electron must assume energies available here in order to delocalize and facilitate energy transfer.

2.5 CLASSIFICATION OF SOLIDS

Solids are generally classified as conductors (or metals), semiconductors, or insulators depending on the range of the energy gap, if present, and the relative room-temperature population of conduction electrons.

In semiconductors, the energy gap may extend over a range of about three electron-volts (eV) and the number of electrons per cubic centimeter in the upper band, having enough energy for mobility, is less than 10^{20} cm^{-3}. In a conductor, the upper band is populated on the order of about 10^{23} cm^{-3} at energies significantly higher than that

of the maximum gap energy.[9] Lastly, insulators are characterized as having a large energy gap, greater than 3 eV, and negligible electron concentration at energies in the conduction band. However, with the rise of UWBG semiconductors and higher enery devices, energy gaps greater than 3 eV are still attributed to semiconductor materials.

The energy diagram model, shown in Fig. 2.9, highlights the differences among the three primary classes of solids.

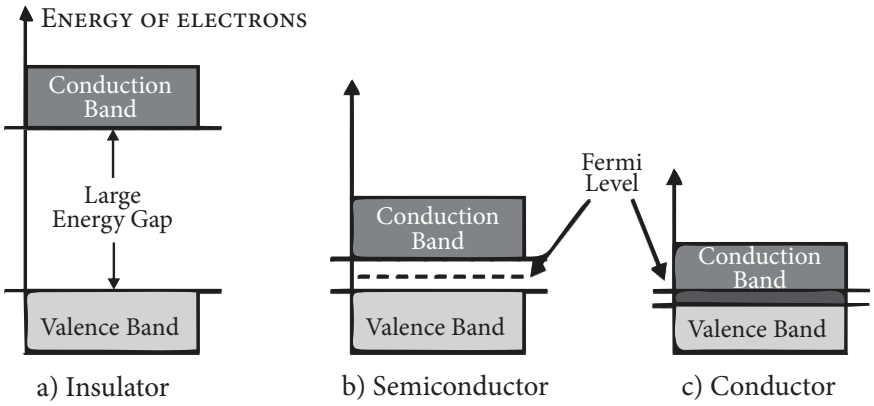
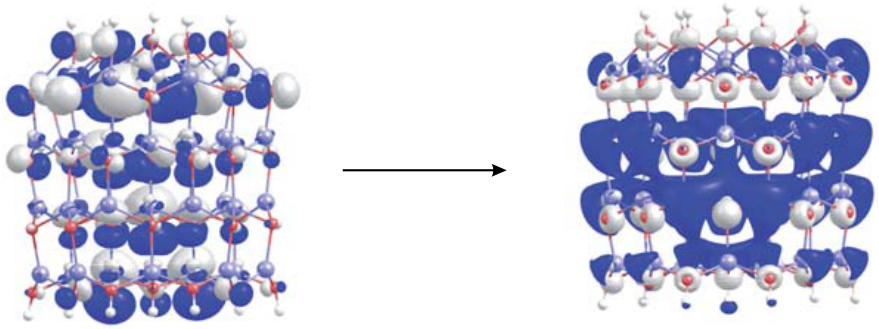


Figure 2.9: Solids are generally classified as conductors, semiconductors or insulators depending on the range of the energy gap and the relative availability of conduction electrons.[12]

2.6 Excitation

As discussed, electrons with sufficient energy occupy higher energy orbitals at average distances farther away from the nucleus. Fig. 2.10 shows electron transitions from the highest occupied molecular orbital (HOMO) transition to the lowest unoccupied molecular orbital (LUMO) from density functional theory calculations by Azpiroz et al.[14] Generally, it is the valence electrons involved in bonding which undergo excitation. The excitation of core electrons is possible, but requires much higher energies, generally corresponding to X-rays.

Electron excitation into conduction can be visualized as electrons possessing enough energy to traverse the crystal largely unimpeded by the attraction to the positive charge in atomic nuclei. It is worth noting that the Coulomb force of nuclei on electrons are always present in a solid, but the energy of electrons determines—and is inversely proportional to—the effect of this attraction.



HOMO LUMO

FIGURE 2.10: *Isodensity Plots showing the highest occupied molecular orbital (HOMO) and the lowest unoccupied mocular orbital (LUMO) in wurtzite type ZnO. Light blue = Zn, red = O, and white = H atoms.*[14]

In inorganic as well as organic solids, energy states are generally available below those required for conduction. *These states are said to occupy the band-edge region and correspond to electrons, due to their lower energy, still remains largely coupled to atomic nuclei and are therfore localized.* These electronic states may include those introduced by defects, discussed in detail in following sections, and states commonly referred to as excitonic states.[12] Generally, excitonic states are more stable than defect states and can thus contribute to energetic transfer processes in semiconductors as well be outlined presently.[13]

During excitation into an excitonic state, the electron density decreases at the valence level. It is well known that electron density screens nuclear charge[1,5,15],

thus the decrease in nuclear screening corresponds to a positive charge and the introduction of a Coulomb potential given by:[16]

$$U(\vec{r_1}) = \int \psi_c^*(\vec{r_2}) \frac{e^2}{|\vec{r_1} - \vec{r_2}|} \psi_c(\vec{r_2}) dr_2 \tag{2.11}$$

which may interact with other valence electrons in the crystal. As a result, a valence electron and the electron in the excitonic energy level may undergo nuclear "swapping" and the excitation can be imagined as moving through the solid as a wave. Due to symmetry of the lattice, the de-screened nucleus *appears* to move through the lattice in company with the excitation wave as shown in Fig. 2.11. The associated positive charge is commonly referred to as a "hole."

These coupled, electron-hole pairs are called excitons and are treated, in practice, as a *quasiparticle*. These quasiparticles, being neutral as a unit, cannot transport charge but do transport energy by the mechanism described above.

Excitons may be tightly bound (localized) and involve coupling to nearby nuclei or weakly bound (delocalized). The following sections illustrate the qualitative difference using an example in quartz.

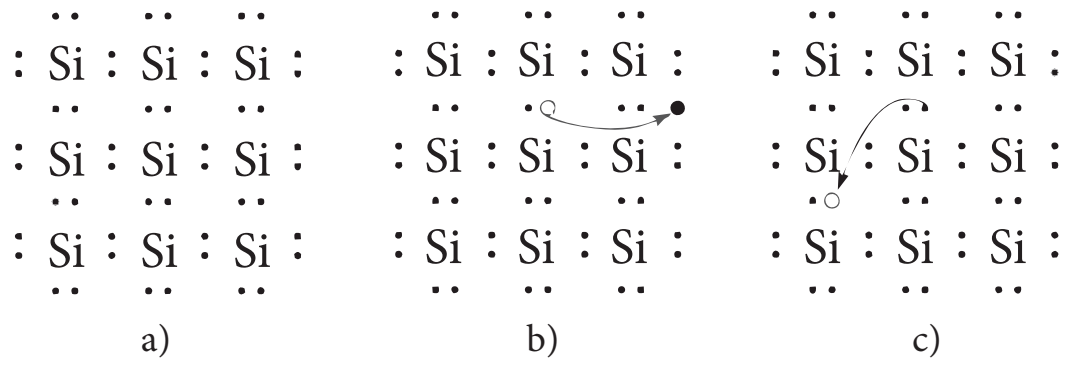
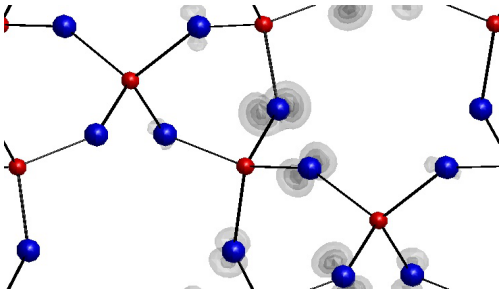


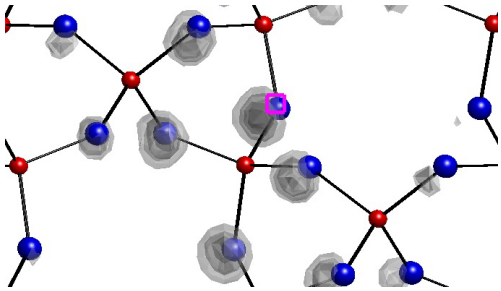
FIGURE 2.11: *(a)* A silicon lattice with valence electrons in the ground state. *(b)* When a valence electron in a Silicon lattice is excited to an excitonic level, it becomes mobile and can conduct current but only while paired with the "hole." *(c)* The electron and hole are free to move through the lattice.[13]

2.6.1 *Exciton Localization*

When an exciton is created in an ideal crystal, it can propagate freely and is thus quite delocalized. However, there are situations in which the exciton can be highly localized. The following Figs.—2.12 & 2.13—shows delocalized and localized excitons in the ideal quartz crystal structure created via density functional theory calculations. The silicon atoms have a tetrahedral bonding configuration with four oxygen atoms lying at each corner of the tetrahedron. Red indicates silicon, while blue indicates oxygen. The figures were generated via density functional theory calculations for a self-trapping exciton mechanism by Ismail-Beigi et al.[17]



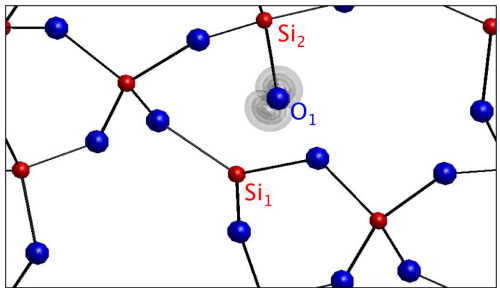
a)



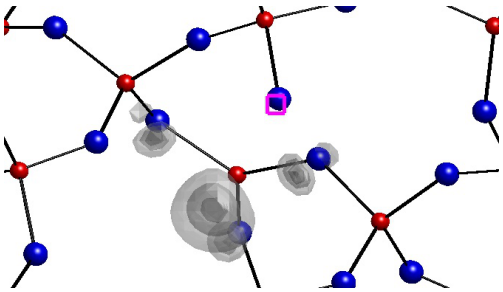
a)

FIGURE 2.12: *Free exciton delocalization in SiO_2. (a) The probability contours for finding the hole-component of a free exciton. As seen, the hole is quite delocalized over different O atoms. (b) we select one location from the set in (a) for the hole location and calculate the probability density of finding the electron-component of the exciton. As seen, the distribution is quite delocalized as well.*[17]

In Fig. 2.12a, we see the probability contours for finding the hole-component of a free exciton. As seen, the hole is quite delocalized over different O atoms. In (b) we select one location from the set in (a) for the hole location and calculate the probability density of finding the electron-component of the exciton. As seen, the distribution is quite delocalized as well. In Fig. 2.13, when the exciton is trapped, due to an unsaturated bond between Si & O, the hole is now completely localized on the O atom as shown in (a). Now, the electron-component probability density is highly localized around the unsaturated bond, thus leading to a highly localized excited state, with both electron- and hole-components localized around the unsaturated bond. This qualitative picture for free and localized excitons will be important in later discussion.



a)



b)

FIGURE 2.13: *Exciton localization near a SiO_2 unsaturated bond. The exciton is trapped, due to a broken bond between Si & O. The hole is now completely localized on the O atom. Now, the electron-component probability density is highly localized around the unsaturated bond, thus leading to a highly localized excited state, with both electron- and hole-components localized around the unsaturated bond.*[17]

2.7 EXCITED STATES AND VIBRATIONAL MODES

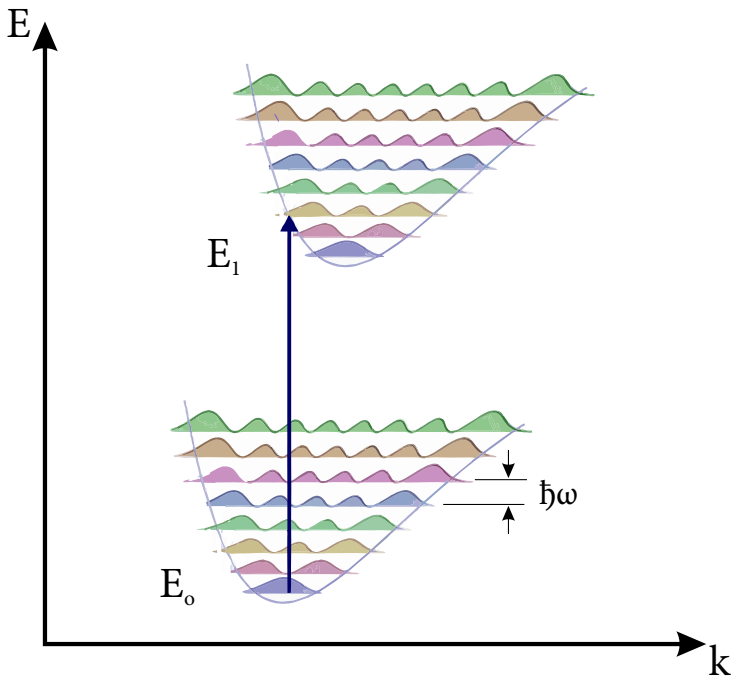


FIGURE 2.14: *A transition from a valence band state to and excited state. Each electron energy level is associated with a spectrum of vibrational modes.*

Electrons in solids exist in a lattice periodic potential. It is therefore customary to consider only a single Morse curve (associated with any unit cell of the lattice) in modeling electron dynamics as shown in Fig. 2.14. The lower curve corresponds to an electronic ground or valence state and the upper curve corresponds to an excited state. It is possible to vibrationally excite the lattice structure, so the energy and the overall wavefunction of electrons can be divided approximately into vibrational and electronic contributions. These vibrational modes can be modeled by wavefunctions described by Felix Bloch, to be detailed shortly. One interesting note is that the Heisenberg uncertainty principle tells us that the energy of a system described by a quantum harmonic oscillator potential cannot have zero energy. Thus, physical systems such as atoms in a solid lattice cannot have zero energy even at absolute

zero temperature. The energy of the ground vibrational state is often referred to as "zero-point vibration." One concept that will be useful in the upcoming section is the concept of the Fourier transform. Much in the way that a waveform in one dimension can consist of a superposition of many frequencies, so too can three dimensional vibrations of orbitals consist of a superposition of many frequencies. This concept is illustrated in Fig. 2.15.

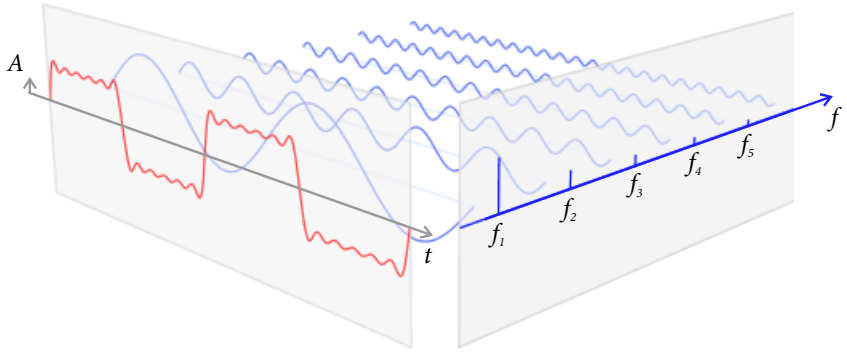


FIGURE 2.15: *A function f (red) is resolved into a Fourier series: a linear combination of sines and cosines (in blue). The component frequencies of these sines and cosines spread across the frequency spectrum and are represented as Dirac delta functions in the frequency domain. The original function can be thought of as being the "amplitude at a given time, t", and the Fourier transform of the function is the "amplitude at given frequency, f". (Lucas V. Barbosa 2013)*

2.8 MOMENTUM AND BLOCH WAVES

In 1928, Bloch showed that electron motion may be modeled with plane waves modulated by a lattice periodic function. The form of the wavefunction showed that the electronic states are spread out over the entire crystal as plane waves. In other words, the wavefunction is a stationary solution for a perfectly periodic potential in some agreement with our previous model in LCAO. Thus, Bloch showed that electrons traveled through the entire crystal without bouncing into lattice nuclei at all. We can then envision vibrational modes as perturbations of electronic wavefunctions. This allows us to relate the momentum possessed by electrons to the wavevector, k, of the constantly vibrating field.

MOMENTUM SPACE — Momentum space is analogous to real or coordinate space. In coordinate space, each position is fully described by a single 3-component vector, while in momentum space, each value of momentum is fully described by a single 3-component wavevector. The real lattice in coordinate space corresponds to the reciprocal lattice in momentum space. Classical particles may be described by momentum p, whereas wavefunctions may be described by the wavevector k. When a particle is described by a quantum state, it is worth remembering that this quantum state can be represented as a superposition (i.e. a linear combination) of basis states. If one chooses the eigenfunctions of the position operator as a set of basis functions, one speaks of a state as a wavefunction $\psi(\vec{r})$ in position space. If one chooses the eigenfunctions of the momentum operator as a set of basis functions, the resulting wavefunction $\phi(\vec{k})$ is said to be the wavefunction in momentum space. Thus, the wavefunctions $\psi(\vec{r})$ and $\phi(\vec{k})$ both describe the behavior of a particle in different ways. Due to the periodic nature of the lattice, and consequently, of momentum space, the forms of the Bloch waves need only be found for a single unit cell. (A single unit cell in momentum space is commonly referred to as the Brillouin zone.) Bloch showed

that the general form of the wavefunction describing electron motion has the very simple form of:[18]

$$\psi_k(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_k(\vec{r}) \tag{2.12}$$

where $u_k(r)$ takes advantage of the periodicity of the lattice as:

$$u_k(\vec{r}) = u_k(\vec{r} + \vec{R}) \tag{2.13}$$

The index k refers to the fact that the function $u_k(r)$ and $\psi_k(r)$ can vary depending on the wavevector k. Fig. 2.16 shows the isosurface of the square modulus of the Bloch wave in Silicon corresponding to the lower valence band in the Brillouin zone.

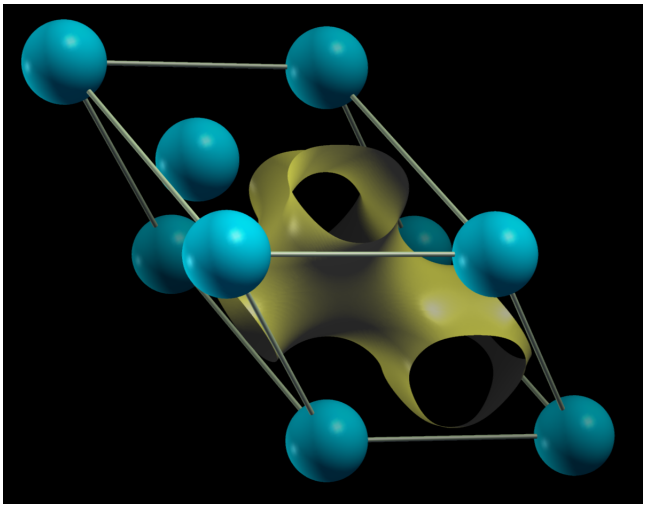


FIGURE 2.16: *Isosurface of the square modulus of the Bloch wave in Silicon corresponding to the lower valence band in the Brillouin zone. (Lorenzo Paulatto)*

To illustrate Bloch's theorem, we use a cubic crystal of side, L and Born-von Kármán boundary conditions. The Born-von Kármán boundary conditions are cyclic boundary conditions which model a finite chain of atoms in which the "end" is tied to the "beginning" and states that $u_{N+n}(t) = u_n(t)$. Every solution of the Schrödinger

equation given in Eq. 2.9 consistent with these boundary conditions can be written as a sum of plane waves:

$$\psi(\vec{r}) = \sum_k c_k e^{i\vec{k}\cdot\vec{r}} \tag{2.14}$$

The lattice periodic potential is written as a Fourier series, using the reciprocal lattice vectors $\vec{G}$:

$$U(\vec{r}) = \sum_G U_g e^{i\vec{G}\cdot\vec{r}} \tag{2.15}$$

where we require $U_{-G} = U_G^*$

Therefore,

$$-\frac{\hbar^2\nabla^2}{2m_e}\psi(r) = \sum_k \frac{\hbar^2 k^2}{2m_e} c_k e^{i\vec{k}\cdot\vec{r}} \tag{2.16}$$

and

$$U(r)\psi(r) = \sum_{k',G} U_G c_{k'-G} e^{i\vec{k'}\cdot\vec{r}} \tag{2.17}$$

Thus, the Schrödinger equation is reduced to a set of equations:

$$\left(\frac{\hbar^2 k^2}{2m_e} - E\right) c_k + \sum_G U_G c_{k-G} = 0 \tag{2.18}$$

We can view this equation as a set of equations for every allowed k in the first Brillouin zone.

Technically, the set consists of infinitely many equations but in practice, it is often sufficient to take only a few coefficients to be different from zero. If we take one such equation for a certain k it contains the coefficients for k itself and those for k plus all

possible reciprocal lattice vectors. This implies that our original wavefunction given in Eq. 2.14 has non-vanishing terms with these coefficients and can be written as:

$$\psi(r) = \sum_G c_{k-G} e^{i(\vec{k}-\vec{G})\cdot\vec{r}} \tag{2.19}$$

or

$$\psi(r) = e^{i\vec{k}\vec{r}} \left(\sum_G c_{k-G} e^{i\vec{G}\cdot\vec{r}} \right) \tag{2.20}$$

The term in brackets is a *Fourier series* over the reciprocal lattice vectors and therefore is a lattice periodic function, which is the essence of Bloch's theorem.

For a one-dimensional solid with lattice constant a, the reciprocal lattice is spanned by "vectors" of length $g = 2\pi/a$ and the potential can be written as a Fourier series:

$$U(x) = \sum_n U_n e^{ingx} \tag{2.21}$$

The only coefficients we are going to keep are $U_1 = U_{-1}$ and we call them simply U. For a given k, we can write equations of the type given in Eq. 2.18 for c_k, $c_k - g$, $c_k + g$. We get:

$$\left(\frac{(\hbar - g)^2 k^2}{2m_e} - E \right) c_{k-g} + U c_k = 0 \tag{2.22}$$

$$\left(\frac{\hbar^2 k^2}{2m_e} - E \right) c_k + U c_{k-g} + U c_{k+g} = 0 \tag{2.23}$$

$$\left(\frac{(\hbar + g)^2 k^2}{2m_e} - E \right) c_{k+g} + U c_k = 0 \tag{2.24}$$

This is a linear system of equations that has three solutions, E_1, E_2 and E_3 for every value of k. The solutions are shown in Fig. 2.17 with higher order contributions of

the potential. Note the range of energies for which there are no states at all, which correspond to bandgaps.

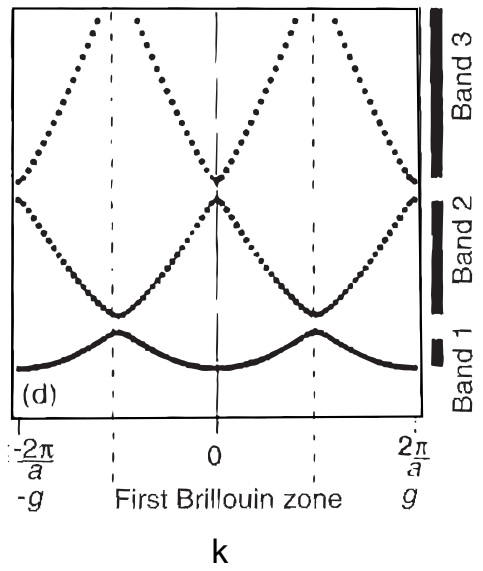


FIGURE 2.17: *Solution of Bloch waves for a one-dimensional solid and higher order contributions of the lattice potential. This plot is commonly referred to as a dispersion relation.*[18]

The relationship between n energy eigeinvalues and allowed values of k is the dispersion relation for the electronic states and is usually called the electronic band structure of the solid. In this context, k can be interpreted as either the wavevector of the electron wavefunction and can also be viewed as a quantum number of the electronic states. In comparison to the atomic analogue, n is the main quantum number specifying the shell containing the electrons and l and m are the parameters of the spherical harmonic functions. In this sense, l and m are quantum numbers that are related to the spherical symmetry of the atom. *In the solid, the symmetry of the wave solution is given by the periodic lattice structure and k can be viewed as the quantum number related to this symmetry.*

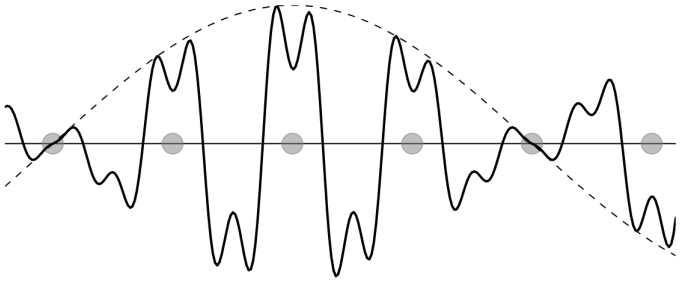


FIGURE 2.18: *Solid line: A schematic of a typical Bloch wave in one dimension. (The actual wave is complex; this is the real part.) The dotted line is from the $e^{i\vec{k}\cdot\vec{r}}$ factor. The light circles represent atoms.(Nicoguaro)*

2.8.1 *Thermal Transport and Phonons*

The solution of Bloch waves given by Eq. 2.20 describes a set of modulated plane waves that is delocalized over the whole crystal. Such solutions in and of themselves cannot be used to describe energy transport throughout a crystal. In Bloch waves, the phase velocity does not describe the *velocity of energy* in the solid.[3] "Monochromatic" Bloch waves as shown in Fig. 2.18 are distorted as they travel through the medium and this distortion makes it difficult to define their average velocity. For

[3]This is a general statement of wave mechanics, the exception being waves in vacuum.

this reason, we define a group of waves or *wave packet* of finite length consisting of only a limited number of frequencies. In other words, to describe energy transport, we must consider a superposition of Bloch waves within a certain Δk, close to the k of interest. These wave packets are commonly referred to as *phonons* in analogy to photons. Fig. 2.19 shows three waves (light green) of varying frequencies and phase velocities superposed to give a wave packet (dark green).

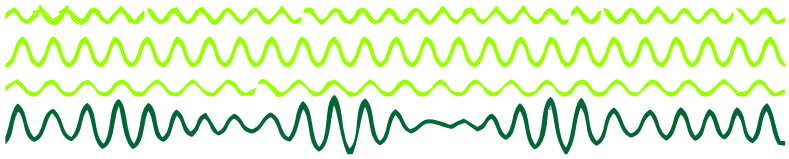


FIGURE 2.19: *Three waves (light green) of varying frequencies and phase velocities superposed to give a wave packet (dark green).*

The energy velocity is proportional to the velocity of this wave packet and in the limit that the frequencies become equal, is known as the group velocity, v_g given by:

$$v_g = \frac{d\omega}{dt} = \frac{1}{\hbar}\frac{dE(k)}{dk} \tag{2.25}$$

In words, the group velocity is given by the slope of the bands in Fig. 2.17. We see that the group velocity is zero at the Brillouin zone boundaries, which means that we have standing waves there. It is worth mentioning that zero group velocity does not imply that electrons do not move. If we could measure the velocity of a wave packet at the zone boundary, the expectation value would be zero, so we could not say if it moves to the right or to the left. However, the expectation value for the kinetic energy would not be zero.

2.9 PHONONS IN REAL SOLIDS

Phononic phenomena defines the propagation of energy through condensed material media, as opposed to photonic phenomena, which does not require condensed material media. It is worth repeating that this energy exists as *packets* of vibrational energy in the bond orbitals of the periodic arrangement of atoms (whose solution depend on the lattice periodic potential) wherein the energy is proportional to the frequency of vibration.

The previous derivation of phonons from Bloch waves in solids is based on a perfectly periodic lattice and our example is confined to the one-dimensional case. While this suffices for a fundamental qualitative understanding of these concepts, we must note the differences between the real and ideal cases. For three-dimensional cases, the band energy depends on a three-dimensional $\vec{k}$ and the Brillouin zone is not always easy to identify (without prior labeling of the Γ point). Generally, at low energies, the bands still look like parabolas near the Γ point. Usually the bands will also be given for different high-symmetry directions in the first Brillouin zone as well. The bands are filled up to the Fermi level which usually corresponds to zero energy. For metals, there are many bands which may cross the Fermi level. For semiconductors and insulators, bandgap openings at the Brillioun zone is typical. In both materials, this corresponds to an "absolute bandgap," which defines the gap opening in the entire band structure. In semiconductors and insulators, there will be no bands which cross the Fermi level. No electrons could increase their energy by a small amount in order to become conduction electrons; as seen previously, the energy must correspond to the amount defined by the gap. It is worth noting once more, that at elevated temperatures a small distribution of electrons will manage to accomplish this.

The band structures of materials may be determined experimentally by using angle-resolved photo-emission, based on the photoelectric effect. The sample under-

goes neutron scattering or exposure to UV-radiation and electrons are emitted, which are then sorted by their $\vec{k}$ and energy. From here, it possible to work backward to obtain the band structure of the sample.

It is also possible to theoretically derive the phonon dispersion of real solids by using density functional theory calculations in addition to implementing various approximations. Calzolari et al. have demonstrated accurate calculations of both the electronic and vibrational spectrum of ZnO allowing for the simulation of phonon modes as well as high and low-frequency dielectric constants. Fig. 2.20 shows the phonon dispersion of ZnO used in this work.[19]

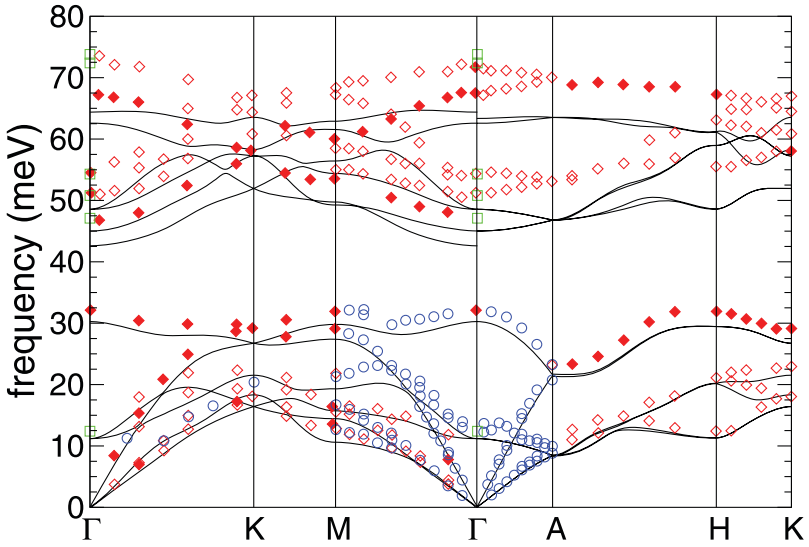


FIGURE 2.20: *Phonon dispersion of ZnO*

2.9.1 *Type of Phonons*

In general, solids may assume two types of phononic behavior—acoustical and optical—which correspond to the lower and upper branches of solutions shown in Fig 2.20, respectively. The solutions that go to zero for small k is called the acoustic branch. It corresponds to the propagation of acoustic waves through the crystal. The solution

that has a finite ω at $k = 0$ is called the optical branch. *Vibrations in the optical branch may coupl to an electromagnetic field.*

Fig. 2.21, based on theoretical calculations, shows the optical phonon modes in wurtzite ZnO quantum dots. This image is a useful basis for envisioning phonon modes in ZnO nanocrystals and, depending on the crystal structure, may give the same frequency spectrum.

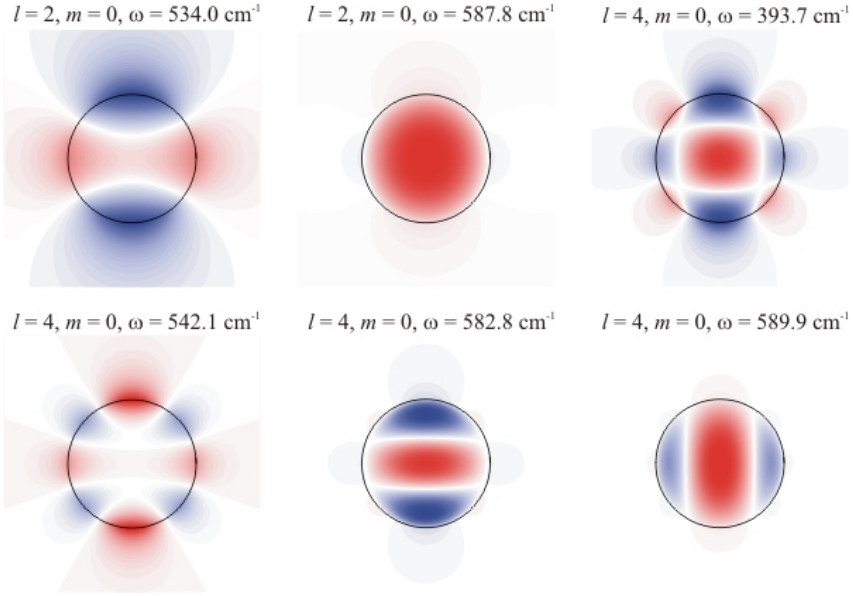


FIGURE 2.21: *Calculated optical phonon modes in wurtzite ZnO quantum dots. The quantity k, related to ω can be interpreted as an additional quantum number of the electron wavefunction. V.A. Fonoberov and A.A. Balandin, Phys. Rev. B (2004)*

2.10 ELECTRON-PHONON INTERACTION

In real solids, Bloch waves see some interaction with core ions that lead may lead to the scattering of Bloch wave packets. If the ions move, then the electrons must move in a changing potential field. Generally speaking, all the explanations come down to the fact that real lattices are not perfect. A thermally excited environment can lead to the destruction of the perfect translational symmetry in the lattice and cause scattering of the Bloch electrons. The interaction between Bloch electrons and vibrations of

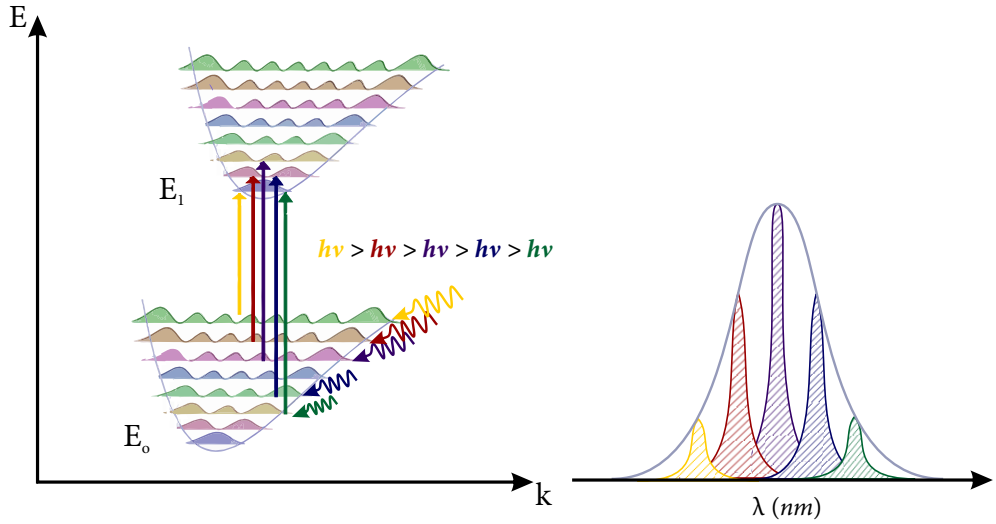


FIGURE 2.22: *The schematic highlights available vibronic transitions into excitonic levels. The probability of vibronic transitions is proportional to the overlapping phase of initial and final states. The electron–phonon interaction is proportional to the range of vibronic transitions available (base width of unresolved peak (light blue)).*

core ions is called the electron-phonon interaction. The underlying mechanism can be understood as potential fluctuations in the Bloch wave environment that destroy their symmetry. By definition, when Bloch waves are scattered, this generally leads to inhibition of their propagation throughout the lattice. Although scattering from lattice vibrations are important, scattering can occur at very low temperatures due to remaining imperfections in the crystal. It is worth repeating that Bloch wave packets are the main mechanism being used herein to describe thermal transport (as well as electrical transport) in a solid. *Consequently, the electron-phonon coupling defines the* ***efficiency*** *with which the energy is transferred between Bloch electrons and vibrations of the lattice.*

2.11 THERMAL BROADENING FROM ELECTRON-PHONON INTERACTION

In absorption spectra of semiconductors, phonon broadening is caused by the mechanisms outlined in Sec. 2.10 wherein absorbing species participate in lattice vibration. Since resonance is a requirement for absorption, at increased temperatures, the fluctuations in orbital configurations result in a wider variance of probability of resonance with radiation. The Franck-Condon principle states that phase will be conserved for these transitions, therefore, transitions can occur into a wide range of excited-state vibrational modes as well, so long as this requirement is met. In semiconductors, phonon broadening can be observed even at room temperature.

CONCLUSION — We have seen that the rules governing individual electron interactions remain in effect in solid structures, in an admittedly more complex configuration. The correlated motion of electrons in a solid is constrained by both electronic and vibrational wavefunctions, which lead to distinct energetic pathways within the solid. We have seen that in ZnO, the proposed hybridization of a localized Zn 3*d*

electron with an O 2*p* electron is thought to play a crucial role in giving ZnO its electrical properties—leading most notably to the formation of the energy bands and the band gap.

REFERENCES

[1] L. G. W. Jr, *Organic Chemistry*, 7th ed. (Prentice Hall, Upper Saddle River, NJ, 2009).

[2] M. Nelkon and P. Parker, *Advanced Level Physics*, 7th ed. (Heinemann International Literature & Textbooks, 1995).

[3] G. N. Lewis, Journal of the American Chemical Society **38**, 762 (1916).

[4] W. Heitler and F. London, Zeitschrift für Physik **44**, 455 (1927).

[5] F. Carey and R. Sundberg, *Advanced Organic Chemistry*, Advanced Organic Chemistry (Springer US, Boston, MA, 2007).

[6] I. Fleming, *Molecular Orbitals and Organic Chemical Reactions*, student edition edition ed. (Wiley, Chichester, West Sussex, U.K, 2009).

[7] M. I. Antonoyiannakis and J. B. Pendry, EPL (Europhysics Letters) **40**, 613 (1997).

[8] B. D. Fahlman, *Materials Chemistry* (Springer Science & Business Media, 2007).

[9] J. I. Pankove, *Optical processes in semiconductors* (Prentice-Hall, englewood Cliffs, NJ, 1971).

[10] E. Smith and G. Dent, *Modern Raman Spectroscopy: A Practical Approach*, 1st ed. (Wiley, Hoboken, NJ, 2005).

[11] Y. Hu, B. Cai, Z. Hu, Y. Liu, S. Zhang, and H. Zeng, Current Applied Physics **15**, 423 (2015).

[12] C. Kittel, *Introduction to Solid State Physics*, 8th ed. (Wiley, Hoboken, NJ, 2004).

[13] C. C. Hu, *Modern Semiconductor Devices for Integrated Circuits*, 1st ed. (Prentice Hall, Upper Saddle River, N.J, 2009).

[14] J. M. Azpiroz, I. Infante, X. Lopez, J. M. Ugalde, and F. D. Angelis, Journal of Materials Chemistry **22**, 21453 (2012).

[15] L. Pauling, Journal of the American Chemical Society **54**, 3570 (1932).

[16] D. C. Reynolds, *Excitons: Their Properties and Uses* (Elsevier, 2012).

[17] S. Ismail-Beigi and S. G. Louie, Physical Review Letters **95**, 156401 (2005).

[18] P. Hofmann, *Solid State Physics: An Introduction* (John Wiley & Sons, 2011).

[19] A. Calzolari and M. B. Nardelli, Scientific Reports **3** (2013), 10.1038/srep02999.

CHAPTER 3

Thin Film Synthesis and Methodology

Abstract — In this chapter, we outline the growth approach used in this work. We specifically discuss the measures implemented in order to ensure the repeatability of growth experiments in addition to ensuring optimum sample quality. The ZnO and $Mg_{0.07}Zn_{0.93}O$ thin films studied in this work were grown via DC magnetron sputtering using custom sputtering targets.

3.1 physical vapor deposition & sputtering

Physical vapor deposition methods comprise techniques such as RF (Radio Frequency) and DC (Direct Current) magnetron sputtering, the latter being used in this work. Aside from common use in the semiconductor industry, sputter deposition has found utility in industrial tool bit coating applications and optical coating for glass and storage media and has been used as long ago as 1877 to coat mirrors.[1–3] The phenomenon of sputtering was first described over 150 years ago by William Robert Grove in 1852 and independently by Julius Plücker in 1858 who reported vaporization and film formation of metal films by sputtering.[4,5]

Formerly known as cathode disintegration, sputtering, as its former name implies, fundamentally involves the disintegration of a material and subsequent deposition of this material onto a substrate.[6] The process involves several components including electrodes, sputtering targets or coating material, sputtering gases (that may be inert or reactive) and substrates.

The working gas is introduced into the sputtering chamber, initially placed under vacuum. This gas is then ionized in the chamber via an applied electric field. A cascade of positively charged ions is then initiated in the plasma; these ions are accelerated toward the anodic sputtering target and collide with target surface molecules.

This bombardment results in molecular ejection via the transfer of momentum. Upon displacement, the displaced target material falls and deposits uniformly onto the substrate aligned below. Generally, the dislodged target material is electrically neutral and is able to pass through the applied electric field uninhibited. This process is illustrated in Fig. 3.1.

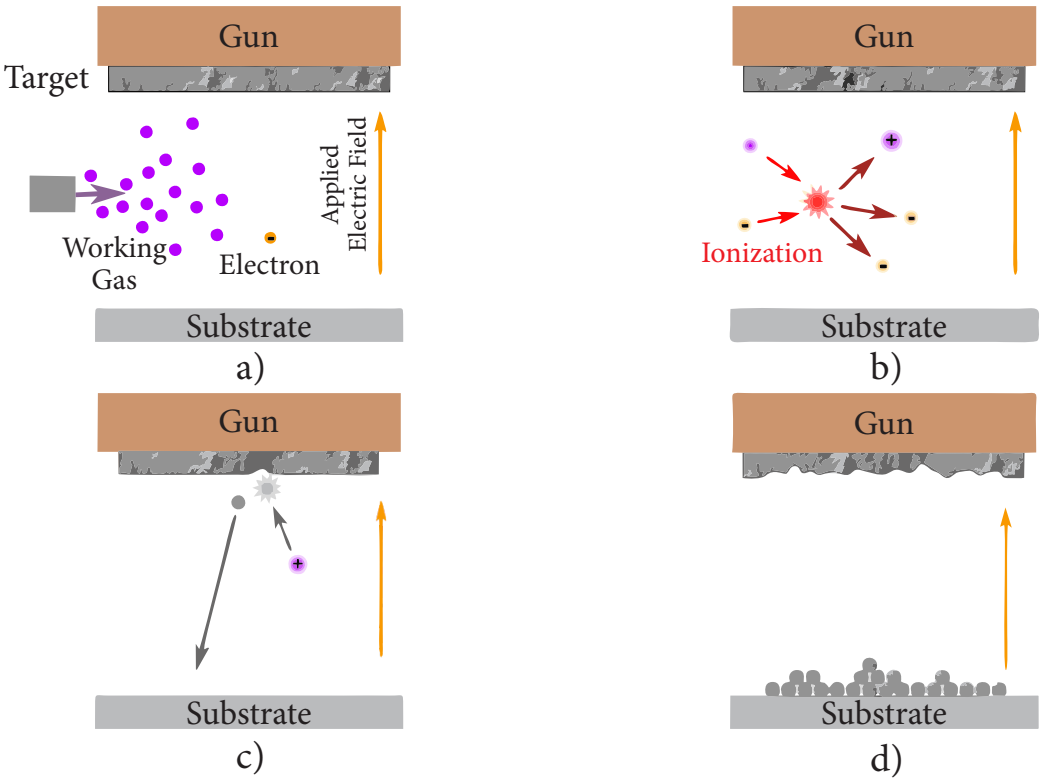


FIGURE 3.1: *A schematic illustrating the process of DC sputtering. (a) The chamber is filled with the working gas and electrons are accelerated by the applied electric field. (b) The electron collides with a working gas atom—an ion and more electrons are generated. (c) The ion is accelerated toward and bombard the target, ejecting a small portion of the target. The ejected material travels and deposits onto the substrate. (d) This process is repeated over time until the substrate is coated with a thing layer of the target material.*

Reactive sputtering occurs when the sputtered material undergoes a reaction with a reactive gas on the path from target to substrate. In this work, Zn and Zn–Mg sputtering targets are bombarded by Argon ions which act as the inert gas and Oxygen gas acts as the reactive agent. This combination results in the deposition of ZnO

and $Mg_{0.07}Zn_{0.93}O$ type thin film alloys, respectively. Sputtering, despite the many parameters present, provide a great degree of control over the microstructure and growth of thin films.

These parameters include, but are not limited to, those mentioned previously—target material, working and reactive gases—in addition to the flow rate of both gasses, chamber temperature and pressure, target power source, sputtering power, magnetron configuration of the sputtering gun, and magnetic shielding of the substrate. Each of the aforementioned parameters and their effect on thin film growth will be discussed herein.

The average number of target atoms released per incident ion is known as the sputtering yield. Simply put, this yield is a function of both ion and target atom(s) masses, ion incident angle and energy, and the surface binding energy of atoms on the target surface.[7] In addition to these considerations, in crystalline targets, the orientation of crystal axes can be modified with respect to the surface orientation, thereby affecting the surface binding energy, albeit usually negligibly so.

3.2 SPUTTERING GASES

Working and reactive gases play a crucial role in controlling thin-film properties. As mentioned previously, the working gas atom transfers momentum to the target atom and thus, if both species possess a mass of similar order, this transfer takes place most efficiently. When working with gases, pressure becomes a significant variable to consider. The chamber pressure throughout sputtering comprises a preliminary base pressure, P_{base}, the plasma strike pressure, P_{strike}, initial working gas pressure, P_{Ar}, and a working pressure, P_{total}. *In this work, Argon is used as the working gas.*

The chamber is first pumped down to its minimum base pressure before growth. Argon is then introduced into the chamber to strike the plasma. Following this, the

Argon flow is adjusted and pressure is lowered in preparation for growing. If a reactive gas, P_{reactive} is required, it is now introduced to achieve the total, or working, pressure needed for growing.

Dalton's law of partial pressure states that the total pressuring of the mixture of non-reactive gases is equal to the sum of the partial pressure of individual gases. Therefore, the total pressure in the sputtering chamber during the sputtering process is given by the following:

$$P_{\text{total}} = P_{\text{base}} + P_{\text{working}} + P_{\text{reactive}} \tag{3.1}$$

Sputtering power should ensure that ions posses enough momentum for collision, while partial pressures and flow rate ideally ensure the minimization of mean free path, allowing Ar ions to travel to the target efficiently.

In general there is an optimum total operating pressure to maximize deposition rate. The trick to picking an optimum pressure is a compromise in the interplay between increasing the number of Ar ions available for sputtering and increasing the scattering events (reduced mean free path) between Ar ions and neutral Ar atoms.

3.3 ARC SUPPRESSION

When using a conductive target, a DC power source is preferable because it increases the rate of sputtering and in turn reduces total sputtering time. DC power sources, however, are more prone to arcing, which is a source of several unwanted, and sometimes also dangerous, issues in sputtering. When the uniform glow of a plasma develops intense, localized concentrations of energy, usually due to the formation of insulating regions, arcing results—a phenomenon similar to lightning.[8] The intense energy evolved near arcing sites can result in explosive, and potentially harmful, ejection of macro-particles. There are several methods to combat arcing when sputtering

in the DC regime. Reverse-voltage pulsing has been shown to dramatically reduce arcing, while other electronics, equipped for arc detection and suppression, attenuate this phenomenon, but unfortunately do not prevent it completely.[9] Sputtering with highly pure targets is required since small impurities on the surface of the target can lead to arcing.

When using an insulating target, consistent bombardment of the insulator results in significant build-up of positive charge, which in turn repels incoming ions of the working gas. As this phenomenon continues over time, an increased sputtering power is required to maintain sputtering rate, in some cases up to 10^{12} volts. In many instances sputtering rate is dramatically reduced and possibly halted altogether. One method of circumventing this problem is in using RF sputtering.

In RF sputtering, the polarity of the applied electric field is alternated by driving the sputtering gun at a particular frequency, unlike the fixed voltage in DC sputtering. For the first half of the alternating cycle, positively charged ions are driven toward the target and electrons are driven toward the substrate. This portion of the cycle is similar to that of DC sputtering and thus the target begins to be affected by charging. However, the alternating cycle reverses the polarity of the applied field and drives electrons back toward the target, neutralizing it. This cycle of brief charging and subsequent neutralization significantly reduces the occurrence of ion repulsion, unfortunately, at the cost of sputtering rate—a direct result of only sputtering over half-cycles.

3.4 SUBSTRATES

The substrate chosen for sputtering has the ability to affect the structural properties of deposited thin films. Generally, growing on material that has a similar if not the same lattice parameters (and crystal orientation) as the film is beneficial. In cases where

lattice parameters are not similar enough between substrate and alloy, a buffer layer could first be introduced which serves as an average to "bridge" lattice parameters. Differently cut substrates will have different lattice parameters, namely, z, x, y, m and n-cut.

Potential interference of substrates with spectroscopic techniques also bears considering. For example, when a chemical element is common to both substrate and sample, EDS alloy compositional readings may detect elemental traces from the substrate. Another example of possible instrumental interference is a substrate with high potential for photoluminescence, which can interfere with UV-Photoluminescence measurements.

If crystal orientation is not a concern, i.e. in polycrystalline materials, mechanical properties of the substrate come into consideration. Quartz and Sapphire may be chosen for their high melting point, transparency to UV or relative durability while fluorinated ethylene propylene (FEP) may be chosen for its flexibility.

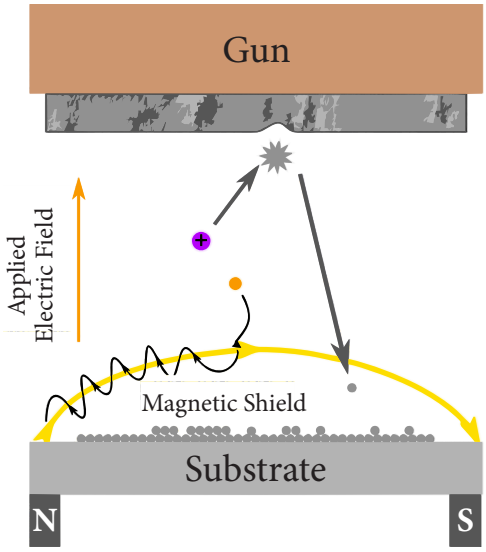


FIGURE 3.2: *During sputtering, a substrate is magnetically shielded. A magnetic field (yellow) is created directly beneath the substrate and the resulting field captures charged particles, particularly electrons, preventing them from damaging the substrate and sputtered coating.*

3.5 MAGNETIC SHIELDING

Due to the high energy environment of the sputtering chamber, safety measures must be implemented to protect substrates and deposited material. As an example of why these measures are sometimes necessary, consider the alternating polarity of the applied electric field in RF sputtering. This reversed polarity not only reduces charge buildup on the target, but redirects ions from the working gas toward the substrates, potentially reversing deposition achieved heretofore, if not damaging the substrates completely. Polymer substrates are notably susceptible to damage in this fashion. Upon collision with the substrate, even in DC sputtering, the mean free path of electrons is so confined to the thin surface layer that all the kinetic energy is directed into virtually boiling this layer of the substrate. To combat collisions of charged particles with the substrate, we introduce magnetic shielding based on the Lorentz Force Law:

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{3.2}$$

where q represents the charge of the particle, E is the electric field, v is the velocity of the charged particle and B is the magnetic field.

As seen in the Fig. 3.2, a magnet placed beneath the substrates generates the field that protects the substrate from damage by electrons and ions, in both DC and RF sputtering, but allow through sputtered material.

3.6 MAGNETRON SPUTTERING

Further adjustments may be implemented to the sputtering system in order to increase the sputtering rate. "Surface magnetron" sputtering configurations were introduced in the 1960s and early 1970s after the effects of a magnetic field on the trajectories of electrons had been realized.[10] These magnetron sources use a magnetic field that

loop in and out of the target surface in a closed "racetrack" pattern as shown in Fig. 3.3. Electrons are then confined to this "racetrack" thereby resulting in high density plasma near the target surface. This high density plasma area contains more ions and in turn contributes to the increase in sputtering rate. Magnetron configurations can be utilized in conjunction with DC or RF sputtering.

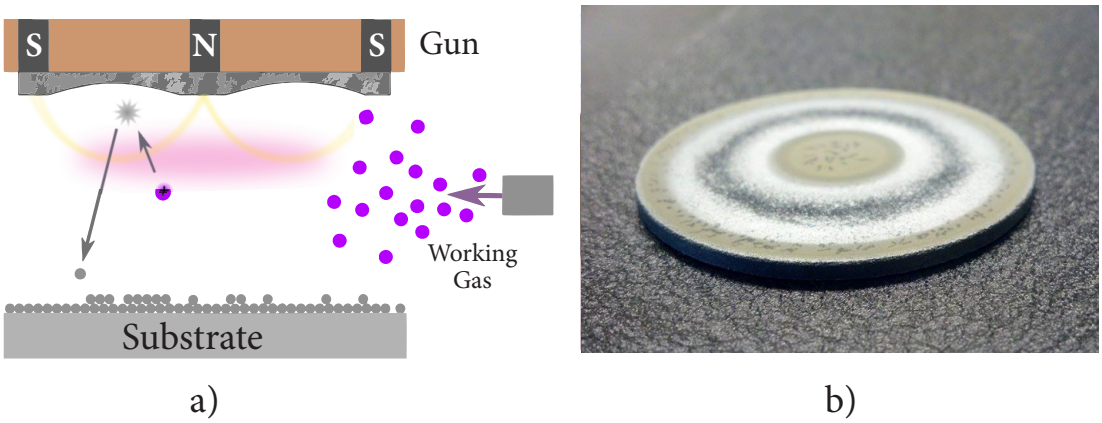


FIGURE 3.3: *(a)*A cross section view of the magnetron configuration used in this work. *(b)* A sputtered Zn-Mg target. The racetrack shaped pattern on the target suggests an area with enhanced sputtering rate by confining electrons within a certain region via the magnetic field.

3.7 SPUTTERING SYSTEM IN THIS WORK

The sputtering chamber utilized in this work is custom built. The following paragraphs outline some of the system's components and capabilities.

The sputtering chamber uses a Varian V200 molecular Turbo pump coupled with a Welch model 1402 DuoSeal belt-drive mechanical roughing pump and interlocking system that monitors flow. The chamber is vacuum sealed using Viton elastomer flanges and copper conflate-style flanges. Pressure in the system is determined via a triple-gauge configuration. A Duniway Convection gauge measures in the range

1000–10^{-4} *Torr*, a capacitive gauge measures in the range 100–10^{-4} *Torr*, and a Bayard-Alpart style hot cathode ion gauge measures in the range 10^{-3}–10^{-10} *Torr*. The preceding configuration takes the chamber to a well-monitored 2.6×10^{-6} *Torr*. Despite the mixture of multiple gases in the chamber, the capacitive gauge monitors pressure during growth independent of chamber gas composition. This feature is most useful during reactive sputtering.

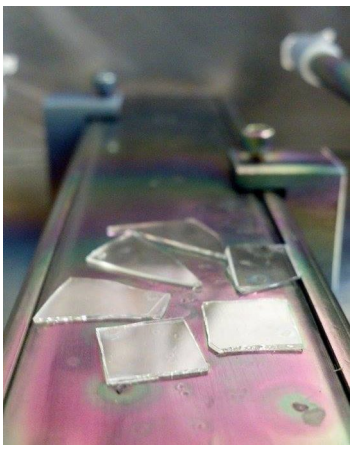


FIGURE 3.4: *A photograph of sputtered $Mg_{0.07}Zn_{0.93}O$ samples studied in this work. The samples were grown on Quartz and Sapphire substrates.*

The sputtering system utilizes a double-gun configuration capable of dual-target sputtering. However, only single-target sputtering is used in this work via a water cooled ST20 magnetron sputtering gun from AJA International. This gun is only compatible with 2–inch diameter disk targets. Flow of both working and reactive gasses, Argon and Oxygen, is monitored independently upstream from the sputtering chamber. Flow is regulated by MKS Type P4B digital Mass Flow Controllers (MFC) based on differential heat transfer between temperature-sensing heater elements attached to the exterior of the sensing tube. Real-time gas flow control is achieved through a computer-based flow control application. The MFC and flow control application work in conjunction to gauge flow with a resolution of 0.07 standard cubic centimeters per minute (*sccm*) and an accuracy of approximately $\pm 1\%$.

FIGURE 3.5: *A photograph of the sputtering system used in this work. On the right is the growth chamber, beneath which can be seen the turbo molecular vacuum pump backed by the mechanical roughing pump. On the left is an equipment rack with all accompanying control electronics and power sources.*

By closely monitoring the introduction of each gas to the sputtering system and the contribution of each gas to Dalton's law of pressure we insure the repeatability of experiments with reasonable accuracy.

3.8 SPUTTERING TARGETS

The targets used in this work were custom-made Zn–Mg and Zn–metal targets. The Zn–Mg targets were made by mixing Mg and Zn metal powders in desired ratios followed by pressing the mixture into a 2–inch die by a Carver hydraulic press, at 50,000 *lbs* for over 24 hours. The Mg and Zn metal powders were purchased from Alfa Aesar and are of 99.8% and 99.9% purity, respectively.

The most significant advantage of pressed metal targets lies in the freedom of controlling the Zn-Mg ratio, as compared to the limited choices of commercially available ones. However, pressed metal targets cannot endure high input power due to its relatively lower density and higher probability for having internal flaws.

For the growth of $Mg_{0.07}Zn_{0.93}O$ and $Mg_{0.15}Zn_{0.85}O$ films, conductive Zn–Mg targets were sputtered and thus a DC power source was utilized. The DC source used in this work is an MDX 500 by Advanced Energy with a maximum power of 500 *W* and built in arc-suppression circuitry.

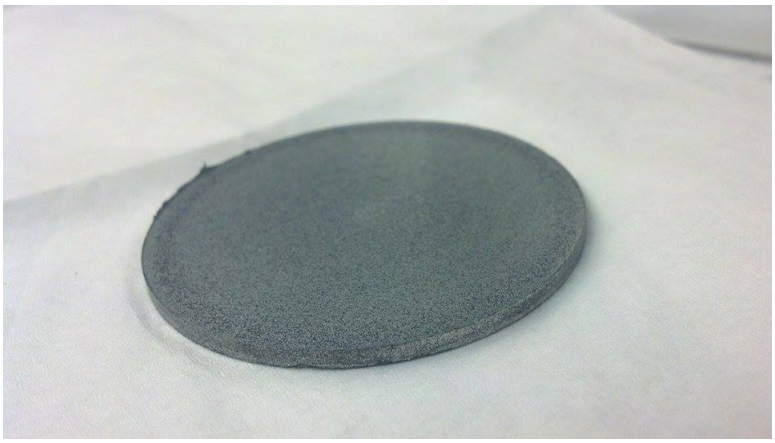


FIGURE 3.6: *A photograph of a custom made $Mg_{0.07}Zn_{0.93}O$ target, made to contain 7% Mg, pressed by a Carver hydraulic press.*

3.9 METHODOLOGY

Cold and room temperature transmission measurements were performed using an Agilent 300 Cary UV-Vis transmission system in double beam mode to remove substrate effects. Cold temperature measurements were performed using an Instec 621 V microcell customized for UV measurements and a sorption pump to keep the sample chamber at a pressure of 10^{-2} Torr.

All analyses were conducted in OriginLab 2018 scientific graphing and data analysis software and PeakFit Peak Analyzer. The data analyzed from the literature

(c-Si in Chap. 8) was digitized via Engauge Digitizer in point match mode. The absorption coefficient was calculated using $\alpha(E) = \frac{-ln(T)}{d}$, where T is the normalized transmission and d is the film thickness.[13] The thicknesses were calculated using a well-established method of interference fringe pattern analysis described in Chap. 7.[11,12]

3.9.1 *ZnO Thin Films*

The ZnO films analyzed in this work were grown on quartz substrates using a DC magnetron sputtering system. The sputtering chamber was evacuated to a base pressure of 10^{-5} Torr and deposition of the films was carried out by sputtering the Zn metal target in an oxygen-argon gas mixture at a pressure of 12 mTorr and a delivered power of 30 W. The films were grown for 90 min. at 250 °C. The thickness of the ZnO films were ≈700-800 nm.

3.9.2 *Wurtzite MgZnO Thin Films*

The $Mg_{0.07}Zn_{0.93}O$ and $Mg_{0.15}Zn_{0.85}O$ films analyzed in Chap. 6 were grown on quartz substrates using a DC magnetron sputtering system. The sputtering chamber was evacuated to a base pressure of 10^{-6} Torr and deposition of the films was carried out by sputtering the Mg-Zn target pf the given stoichiometry in an oxygen-argon gas mixture at a pressure of 11 mTorr and a delivered power of 30 W. The films were grown for 120 min. at 250 °C. .

3.9.3 *Wurtzite-Cubic Mixed-Phase MgZnO Thin Films*

The mixed-phase $Mg_{0.75}Zn_{0.25}O$ film analyzed Chap. 8 were grown on sapphire substrates using a RF magnetron sputtering system. The sputtering chamber was evacuated to a base pressure of 10^{-5} Torr and deposition of the films was carried out

by sputtering the Mg-Zn target in an oxygen-argon gas mixture at a pressure of 17 mTorr and a delivered power of 70 W. The films were grown for 210 min. at 250 °C. The thickness of the $Mg_{0.75}Zn_{0.25}O$ film was ≈300 nm.

The ZnO film discussed in Sec. 8.7.3 and 8.8 was subjected to post-growth annealing treatments at 200, 400, 600, 800 °C, all for 1 h under a controlled Ar environment in a Lindberg furnace. The as-grown ZnO and $Mg_{0.75}Zn_{0.25}O$ films discussed in Chap. 8 were subjected to the same post-growth annealing treatment at 900 °C.

Conclusion — Many variables were shown to influence semiconductor thin-film growth and this chapter has outlined the sputtering growth methods used to synthesize ZnO thin films and MgZnO thin film alloys in addition to outlining the customized features used. We have shown that numerous measures were taken to ensure as much control over the growth environment as possible which in turn ensures minimal variation in experimental samples.

REFERENCES

[1] B. A. Banks, M. J. Mirtich, S. K. Rutledge, and D. M. Swec, Thin solid films **127**, 107 (1985).

[2] J. L. Edwards and R. J. Roedel, Electronics Letters **19**, 962 (1983).

[3] M. J. Ellison, *Advances in Thin Film Coatings for Optical Applications*, Vol. 6286 (Society of Photo Optical, 2006).

[4] W. R. Grove, Philosophical transactions of the Royal Society of London **142**, 87 (1852).

[5] J. T. Merz, *A History of European Thought in the Nineteenth Century* (W. Blackwood, 1912).

[6] L. Holland, *Vacuum deposition of thin films* (Wiley, 1966).

[7] P. Sigmund, Physical Review **184**, 383 (1969).

[8] A. Anders, Thin Solid Films **502**, 22 (2006).

[9] D. C. Carter, R. L. Arent, and D. J. Christie, in *50th Annual Technical Conference Proceedings* (Advanced Energy Industries, Inc., 2007).

[10] D. M. Mattox, *The Foundations of Vacuum Coating Technology* (Springer Science & Business Media, 2004).

[11] J. C. Manifacier, J. Gasiot, and J. P. Fillard, J. Phys. E: Sci. Instrum. **9**, 1002 (1976).

[12] D. Thapa, J. Huso, J. L. Morrison, C. D. Corolewski, M. D. McCluskey, and L. Bergman, Optical Materials **58**, 382 (2016).

[13] M. A. Kastner, S. R. Ovshinsky, and G. A. Thomas, *Disordered Semiconductors*, 1987th ed. (Springer, New York, 1987).

CHAPTER 4

Semiconductors & Bandgap Engineering

Introduction — The bandgap is perhaps the single most useful property of semiconductor materials. The bandgap allows the semiconductor to operate as an energetic switch, facilitating the flow of energy into (or out of) the semiconductor at pre-determined values. The energy of the bandgap is dependent on the chemical interactions between sub-unit atoms in the solid; therefore, we can engineer bandgaps by altering the semiconductor at the atomic level.

4.1 Alloying—ZnO & MgO

Currently in the field of bandgap engineering, there remains a limited number of accessible bandgap energy ranges. If the need arises for a device that operates in a particular energy range that is not available naturally, the appropriate materials must be created. One method of bandgap engineering focuses on introducing a substitutional impurity (or multiple impurities) to create a semiconductor alloy, which, in essence, is a solid solution of multiple materials. This work focuses on the extrinsic semiconductor ZnO and its alloy $Mg_xZn_{1-x}O$. In the $Mg_xZn_{1-x}O$ alloy system, Zn lattice points are substituted with Mg atoms to create the ternary compound.

To understand the interaction between alloyed semiconductors, we must first be familiar with the component parts. ZnO is a semiconductor with a wurtzite crystal structure, a bandgap of approximately 3.3 meV and an exciton binding energy of approximately 60 meV. MgO, more an insulator, has a cubic crystal structure, a bandgap of approximately 7.8 eV and an exciton binding energy of 100 meV.[1]

Substituting Mg for Zn in the ZnO lattice changes the optical properties of the lattice as a whole. The changes are largely based on the ratio of electronegativities of Mg to Zn, the ratios of their covalent radii, and also on the ratio of their mass.

The ratio of the mass of Mg to Zn is considered to be ≈ 1, and the two atoms have comparable covalent radii, although Mg is slightly larger.

As shown in Tab. 4.1, the electronegativity of Mg is less than that of Zn. The ratio of electronegativities of O to Mg is ≈ 2.63. Therefore the electron density of MgO is located around O and MgO is said to be ionically bonded. In the case of ZnO, the ratio of electronegativities of O to Zn is ≈ 2.08—less than that of 2.63 in MgO. Therefore the electron density accumulates between Zn and O more equally in a semi-ionic-covalent bond.

MgO crystal structures exist in the rocksalt cubic phase, whereas ZnO crystal structures exist in the hexagonal wurtzite phase as shown in Fig 4.1.[2] The mismatch in crystal structures of the alloy end-members gives $Mg_xZn_{1-x}O$ varying crystal structure depending on Mg concentration. Additionally, solubility plays a key role in governing the properties of this alloy.

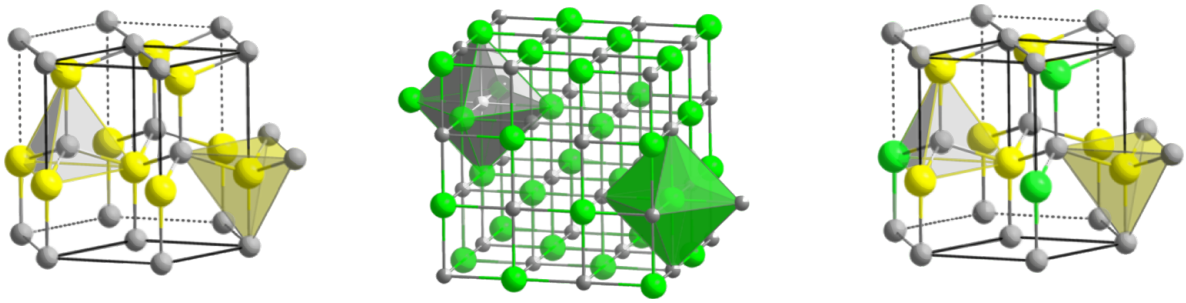


FIGURE 4.1: *(a) The wurtzite hexagonal crystal structure of ZnO. (b) The rocksalt cubic crystal structure of MgO. (c) At low concentrations, $Mg_xZn_{1-x}O$ retains a wurtzite structure.*

4.1.1 *Solubility of Crystal Structures*

The extent of the solubility of a substitutional impurity in a semiconductor is measured by the saturation concentration. Saturation concentration in $Mg_xZn_{1-x}O$ is the

point at which the wurtzite-type alloy can dissolve no more Mg without precipitating cubic-phase MgO. This results in a mixed-phase wurtzite-cubic material and is known as phase segregation. Saturation in $Mg_xZn_{1-x}O$ has been shown to occur at about 40% Mg. This mismatch in crystal structure between ZnO and MgO leads to a limitation in solubility on both ends of the alloy composition spectrum, that is, Mg-doped ZnO and Zn-doped MgO, at equilibrium.[4] Accordingly, extremely high concentrations of Mg lead to cubic structure $Mg_xZn_{1-x}O$.[3]

TABLE 4.1: *Material properties of MgO and ZnO and constituent atoms, Mg, Zn & O.*

Compound	**Structure**	**Lattice Constants (Å)**	**Bandgap *(eV)***
MgO	Rocksalt Cubic	a = 4.2	7.8
ZnO	Hexagonal Wurtzite	a = 3.2 c = 5.2	3.4
Atom	**Covalent Radius (Å)**	**Electronegativity†**	**Oxidation States**
Mg	1.36	1.31	+2
Zn	1.25	1.65	+2
O	0.73	3.44	−2, −1

† Pauling's

After obtaining the ternary compound $Mg_xZn_{1-x}O$, the bandgap settles somewhere between the two bandgaps of the end-members, based on composition of the compound, and lies between 3.3–7.8 eV. The value of the energy gap of a ternary semiconductor alloy is determined by the following equation:[4]

$$E_{g,MgZnO}(x) = (1-x)E_{g,ZnO} + xE_{g,MgO} - bx(1-x) \tag{4.1}$$

where x is the composition fraction, E are the respective bandgap energies, and b is the bowing parameter. The quantity b incorporates the mismatch in electronegativities and covalent radii between the substituting atomic species in the determination of the

[4]Equilibrium refers to a state in which the relative phase composition is stable and the concentration of atoms in any given phase remains constant in time.

alloy bandgap.[5,6] Thus, the bowing parameter describes the deviation from linear behavior of the bandgap with composition in the alloy system.

When $b = 0$, the bandgap of the alloy changes linearly with respect to the relative composition of the two binary compounds. When $b > 0$, the behavior of the resultant bandgap exhibits a bowing effect, which makes it possible to create an alloy with a bandgap lower than that of either end binaries. One of the first reports of the synthesis of $Mg_xZn_{1-x}O$ with a Mg content of up to 33% was by Ohtomo et al.[7] by pulsed laser deposition (PLD). Ohtomo found a linear increase in the bandgap with composition up to 4.15 eV for $0 < x < 0.36$.

Bandgap bowing in $Mg_xZn_{1-x}O$ is therefore not an issue in practice, and this ternary system is created to increase the bandgap in ZnO.[8] Fig. 4.2 shows that bandgap engineering of ZnO with Mg achieves a wide range of bandgap energies from UV well into the visible.

The $Mg_xZn_{1-x}O$ alloy investigated in this work consists of 7%, 15% and 75% Mg by composition. The 7 % and 15 % MgZnO which corresponds to a wurtzite type alloy with a bandgap slightly wider than that of ZnO. The 75 % MgZnO alloy corresponds to a mixed-phase structure with a wide bandgap.

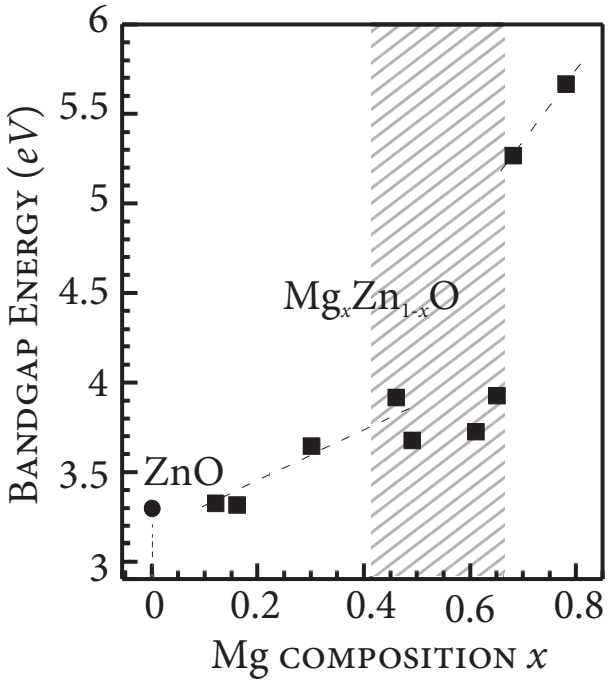


Figure 4.2: *Bandgap engineering of ZnO can achieve a wide range from UV into the visible. Squares represent the bandgap energy of $Mg_xZn_{1-x}O$. The shadowed area indicates the phase segregation region where cubic and wurtzite crystal structures coexist. Figure adapted from Hui (2014).*[9]

4.2 DEFECTS

While alloying offers many benefits to semiconductor technology, impurities form additional energy levels and cause band edges to be distorted via Coulomb interaction. This distortion leads to the "tails" at the band-edge. The nature and source of tailing in semiconductors, and the Urbach model of tailing is the focus of Chap. 8.

In addition to band distortion and tailing, other mechanical changes in doped semiconductors may come about. Since a substitutional impurity would normally possess a covalent radius that is larger or smaller than the host atom, mechanical strain is introduced in the lattice, and can be either compressive or dilatational. An interstitial can also introduce a compressional strain whereas a vacancy can introduce a dilational strain. Compression and dilation can increase and reduce the energy

gap, respectively, due the presence of additional atomic nuclei, or lack thereof. This effect is directly linked to the band edge perturbation previously discussed. Crystallographic defects known as dislocations also contribute to irregularities within a crystal. Dislocations occur at the edge of an extra plane of atoms and are characterized by the mismatch between planes of atoms, especially at grain boundaries. Dislocations produce both compressional and dilational strains which also perturb the band edges.

More information on structural defects and their effects on the band-edge is given in Chap. 6.

Conclusion — We have seen in this chapter that altering the intrinsic nature of semiconductors may result in potentially changing any number of its electrical, optical, and mechanical properties. In $Mg_xZn_{1-x}O$ we widen the bandgap of ZnO but expect to see an increased presence of defects intrinsic in the nature of alloying.

REFERENCES

[1] R. C. Whited, C. J. Flaten, and W. C. Walker, Solid State Communications **13**, 1903 (1973).

[2] A. F. Wells, *Structural Inorganic Chemistry*, 5th ed. (Oxford University Press, Oxford Oxfordshire : New York, 1984).

[3] J. L. Morrison, J. Huso, H. Hoeck, E. Casey, J. Mitchell, L. Bergman, and M. G. Norton, Journal of Applied Physics **104**, 123519 (2008).

[4] J. A. Van Vechten and T. K. Bergstresser, Physical Review B **1**, 3351 (1970).

[5] M. Ferhat, physica status solidi (b) **241**, R38 (2004).

[6] **21**, 10.1088/0953-8984/21/7/075802.

[7] A. Ohtomo, M. Kawasaki, T. Koida, K. Masubuchi, H. Koinuma, Y. Sakurai, Y. Yoshida, T. Yasuda, and Y. Segawa, Applied Physics Letters **72**, 2466 (1998).

[8] A. Zeuner, H. Alves, D. M. Hofmann, B. K. Meyer, M. Heuken, J. Bläsing, and A. Krost, Applied Physics Letters **80**, 2078 (2002).

[9] *Band gaps by design : Tailoring ZnO based semiconductor alloy films*, Ph.D. thesis.

CHAPTER 5

Optical Spectroscopy and Imaging

Abstract — This chapter explores the nature of electromagnetic radiation and provides several mechanisms for the interaction of radiation and semiconducting solids. We will see that the basis for the interaction, and ultimately the transfer of energy, is a coupling between the radiation field and the fields existing within the solid structure. In spectroscopy, we experimentally observe the interaction of radiation with matter, and in this work, our semiconductors, ultimately gaining insight into the band structure of the thin films and thin-film alloys. This chapter focuses on UV-Visible absorption spectroscopy, the relationship between transmission and absorption of radiation, Scanning Electron Microscopy (SEM) and X-ray Diffraction (XRD) in ZnO and $Mg_{0.07}Zn_{0.93}O$.

5.1 OPTICAL SPECTROSCOPY

Spectroscopy is science that deals with the interaction of radiation with matter. Historically, the interactions between electromagnetic radiation and matter were of primary interest, but spectroscopy has expanded to include acoustic waves and particle beams such as ion or electron beams. Still, electromagnetic radiation is most widely used. Electromagnetic radiation comprises a vast spectrum, including visible light, infrared, gamma rays, X-rays, ultraviolet (UV), microwave and radio-frequency radiation.

5.2 ELECTROMAGNETIC RADIATION

Electromagnetic radiation is most easily described by the wave model which involves quantities such as wavelength, frequency, velocity and amplitude. In contrast to phononic phenomenon, electromagnetic radiation propagates readily through the

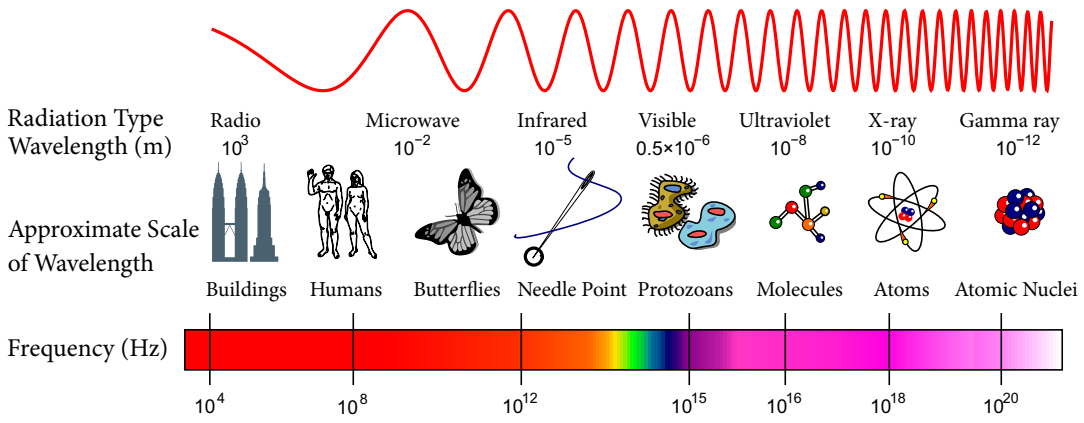


FIGURE 5.1: *The electromagnetic spectrum encompasses a vast range of energies—a range so great that a logarithmic scale is required. This spectrum has been qualitatively divided based on the methods used to generate and detect the different types of radiation. Some overlaps do occur. (Inductiveload, 2007)*

vacuum of space. In a vacuum, the velocity of radiation is independent of wavelength and is at its maximum—denoted by the symbol, c—and has been determined to be $2.998 \times 10^{8}\ m/s$. The velocity of radiation in air differs only slightly from c, and is written, for vacuum or air, as $c = 3 \times 10^{8} m/s$. As the density of matter increases, propagation of radiation through a medium slows by the interaction between the electromagnetic field of the radiation and the elementary particles that constitute the medium. Since frequency is fixed by the source, the wavelength must decrease as radiation passes from vacuum to denser media.

5.3 TRANSMISSION OF RADIATION

The rate at which radiation propagates through a transparent substance, as noted previously, is less than its velocity in a vacuum and depends on the kinds and concentration of atoms, ions or molecules in the medium. Therefore, the radiation

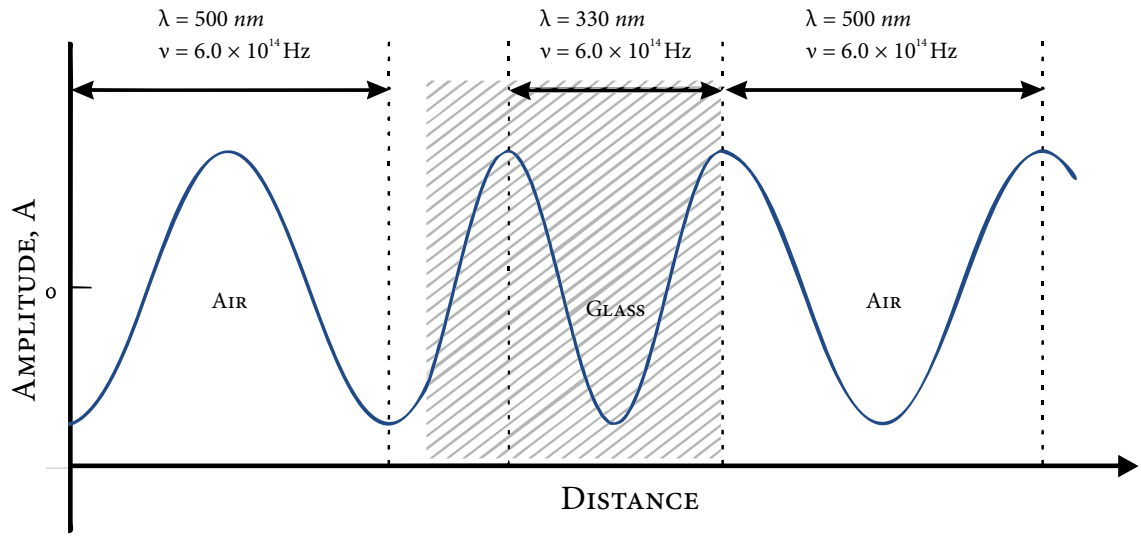


FIGURE 5.2: *As radiation passes from a less dense medium to a more dense medium a change in wavelength occurs. The reverse change occurs as the radiation re-enters the air.*

must be interacting in some way with the matter. A change in frequency is not observed—so, the interaction does not involve a permanent energy transfer.

The interaction involved in transmission can be attributed to periodic polarization of the atoms, ions or molecules that constitute the medium.[1] Polarization in this context is the temporary deformation of the electron clouds associated with atoms by the alternating field of the electromagnetic radiation. The polarizability of the valence shell electrons depend on the strength of their interaction with the nucleus and intervening filled shells. Provided that the radiation is not absorbed, the energy required for polarization is only retained for 10^{-14}–10^{-15} *s* by the species and is re-emitted without alteration as the substance returns to its original state.[5] Since there is no net energy transfer, the frequency of the emitted radiation is unchanged, but the rate of its propagation is slowed by the time it takes for retention and re-emission to occur.

[5]Radiation from polarized particles should be emitted in all directions in a medium. If the particles are small, however, destructive interference prevents the propagation of significant amounts of radiation in any direction other than that of the original path.

The refractive index is related to polarizability and is a measure of a medium's interaction with radiation, and therefore the speed of radiation propagation, as defined by:

$$\eta(\nu) = \frac{c}{v(\nu)} \tag{5.1}$$

where $\eta(\nu)$ is the refractive index at a specific frequency, v is the velocity of the radiation in the medium and c its velocity in vacuum.

The velocity of radiation in matter is wavelength dependent and since c in Eq. 5.1 is independent of wavelength, the refractive index of a substance must also change with wavelength. The variation of refractive index with wavelength or frequency is called dispersion. Dispersion plots exhibit two types of regions, normal and anomalous. In normal dispersion, the slope is gradual, whereas in anomalous dispersion the slope is sharp.

5.4 ABSORPTION & THE SIGNIFICANCE OF RESONANCE

Anomalous dispersion always occurs at frequencies that correspond to the natural harmonic frequency associated with ionic, inter-atomic or orbital perturbation of a substance. At such a frequency, resonance occurs and a permanent energy transfer takes place.

Resonance is a requirement for the absorption of radiation and depends on both the frequency of the driving electromagnetic field and the natural frequency of the oscillator. The absorption of UV-visible radiation by electrons can be described classically by the frequency-dependent dielectric function of a material. Furthermore, light can only couple to optical phonons very close to the center of the Brillouin zone at $k = 0$ and the electromagnetic excitation of such a mode corresponds to the introduction of a time-varying dipole within the crystal.

5.5 GENERAL SPECTROSCOPIC TECHNIQUES

In absorption spectroscopy, we measure the amount of radiation absorbed as a function of wavelength. In semiconducting solids, the onset of the absorption of radiation indicates that the electromagnetic energy meets that required by electrons in overcoming the bandgap energy. In photoluminescence spectroscopy, the emission of photons is measured after absorption which yield information about radiative relaxation pathways in a sample. Furthermore, Raman spectroscopy uses inelastic scattering to glean information about the vibrational modes of a sample. Absorption, photoluminescence and Raman spectroscopy generally yield complementary information about a sample and are usually used in conjunction, however, the primary focus of this work is absorption spectroscopy.

5.6 TRANSMITTANCE & ABSORBANCE

In practice, calculating the energy of radiation absorbed is determined by the experimental measurement of transmittance, T, of samples. The transmittance of a material is given by the following equation which is a ratio of the intensities of incoming, I_o, and outgoing, I', electromagnetic radiation.

$$T = \frac{I'}{I_o} \tag{5.2}$$

Reflection, R, also occurs at air/sample/substrate interfaces during a measurement, along with multiple internal reflections within the film itself. Taking these into consideration, the transmittance is then given by the following:[2]

$$T = \frac{I'}{I_o} = \frac{(1-R)^2 e^{\frac{-4\pi\nu\kappa\chi}{c}}}{(1-R^2)e^{\frac{-8\pi\nu\kappa\chi}{c}}} = \frac{(1-R)^2 e^{\frac{-\alpha t}{c}}}{(1-R^2)e^{\frac{-2\alpha t}{c}}} \tag{5.3}$$

where $\alpha = \frac{4\pi\nu\kappa}{c}$ is equivalent to the absorption coefficient of the film described above.

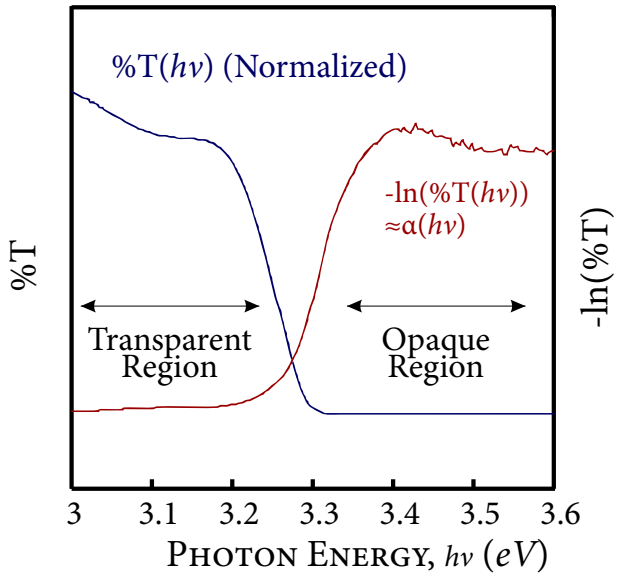


Figure 5.3: *Normalized transmission of electromagnetic radiation in ZnO thin film as a function of photon energy (blue) and corresponding absorption as a function of photon energy (red). The conversion to absorption is a simple negative logarithm. Note the film thickness has not been included in the conversion.*

For semiconductor materials, the reflectivity, R is negligible even near the absorption edge.[3] For the purposes of band edge analysis, the transmittance, T, of a semiconductor with film thickness, t, can be expressed as an approximation of the form:

$$T = e^{-\alpha t} \tag{5.4}$$

Consequently, with measured transmittance and film thickness, t, the absorption coefficient, α, can be obtained:

$$\alpha = -t^{-1} ln T \tag{5.5}$$

5.7 SCANNING ELECTRON MICROSCOPY (SEM)

Scanning electron microscopy (SEM) is a non-destructive surface analyzing technique with unique capabilities. It works in similar fashion to an optical microscope just with a different source of radiation[4]. While an optical microscope forms images from visible light reflected from a sample's surface, an SEM uses a beam of electrons to scan across a specimen's surface in a television-like raster. These electrons interact with atoms in the specimen, inelastically scattering and lose some of their energy to produce various signals such as secondary electrons, X-rays, and Auger electrons.

Secondary electrons contain information about the specimen's surface morphology and is the most common imaging mode. With an appropriate detector, emitted secondary electrons can be captured and analyzed, producing readily available images of the specimen. Compared to visible light, electrons have much shorter wavelengths which are capable of conveying information with much higher resolution and higher magnification can be obtained without loss of detail. The practical maximum resolu-

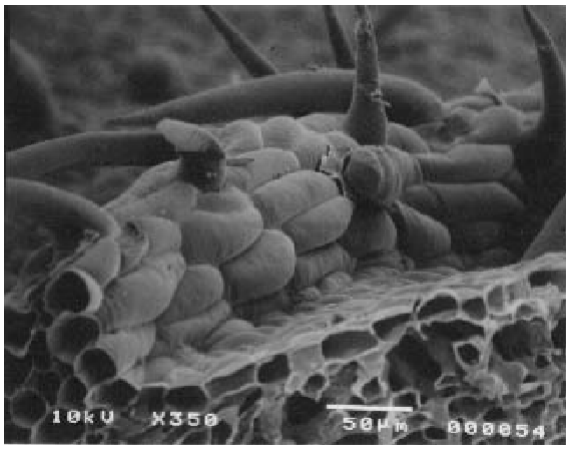


FIGURE 5.4: *An SEM image taken of the back of a leaf showing the raised wall of cells. (Museum of Science, Boston)*

tion of an SEM can reach to 750,000x,[4] which is about 375 times the resolution limit of the best optical microscope.

Secondary electrons are generated by scattering of the primary electron beam by the loosely bound conduction-band electrons of the specimen[5]. Due to the small energy associated with this electronic transition, secondary electrons have a low exit energy of only a few eVs which constrains their trajectory. Although secondary electrons are generated along the entire trajectory of the beam electron within the specimen, only the ones generated within a few nanometers beneath the specimen's surface are likely to escape and be detected. Such shallow escaping depth of secondary electrons makes SEM suitable for imaging the surface of a specimen.

Another area in which SEM is superior to optical microscopy is in the narrow electron probe aperture which leads to a large depth of field, which means the ability of maintaining focus across a larger range of distance. As a result, the three-dimensional details on a specimen's surface will appear sharp, regardless of surface roughness. Fig. 5.4 gives an example SEM image of the back of a bean plant leaf. The texture from raised wall of cells in the vein can been observed clearly under 350x magnification. The combination of a large magnification, small signaling depth, and, a large depth of field makes SEM a favorable technique for studying surface topography of the thin film materials grown for this work.

5.8 X-RAY DIFFRACTION (XRD)

X-ray diffraction (XRD) is a nondestructive analytical technique for the study of crystal structures and atomic spacing in crystalline materials. A crystal consists of a periodic arrangement of atoms into a lattice and such arrangement can be viewed as sets of evenly spaced atomic planes. When irradiated by a monochromatic beam of X-ray, an atomic plane can act as a partially transparent mirror. Part of the X-ray beam will be

reflected and part will be transmitted and may be reflected by another plane of atoms deeper into the material. Fig. 5.5 shows a set of atomic planes with spacing d being irradiated by a beam of X-ray of wavelength λ, at an angle Θ. Since the wavelength of X-rays are comparable to the spacing between atomic planes, constructive interference between two reflected beams can occur when the path difference between the two beams equals a multiple of the incident wavelength. This condition is know as Bragg's law and can be summarized by the equation:[6]

$$n\lambda = 2dsin\theta \tag{5.6}$$

where n is an integer, d is the interatomic spacing, λ is the wavelength and Θ is the angle of incidence.

Typically, the intensity of the scattered X-ray beam is detected as a function of 2Θ and the measurement is based on the first-order reflected beam so that $n = 1$ in Eq. 5.6. For a given wavelength of incident X-rays, narrow and intense peaks can be observed at angles where constructive interference occurs. Each crystalline material usually has its unique set of d-spacings which gives rise to diffraction peaks at specific angles. By comparing the detected pattern of peaks with standard references, the crystal structure of the examined specimen can be identified. The peak position, relative intensity of peaks, and, line shapes also contain information such as strain in the lattice planes and the lattice constants. The peak Full-width half-max (FWHM) also contains information and is inversely proportional to the grain size. Typically, broadened XRD peaks are an indication of small grain size.[6]

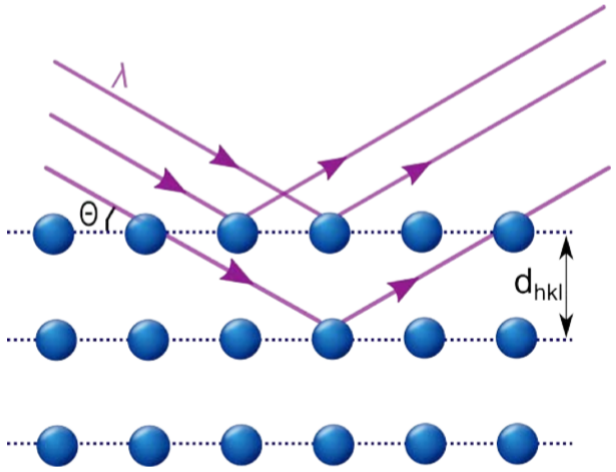


FIGURE 5.5: *A schematic illustrating Bragg's law. An X-ray beam of wavelength λ is incident on a set of atomic planes with spacing d, at an angle Θ. If the path difference between the two reflected beams equal to a multiple of the incident wavelength, constructive interference will occur and cause an enhancement in the intensity.*

Conclusion — In this chapter we have seen that the interaction of radiation or electron beams with semiconductors in controlled experiments can yield valuable information about the sample. In absorption spectroscopy we measure the intensity of radiation transmitted to calculate the amount of radiation absorbed, or absorption coefficient. The absorption coefficient can be used to gain valuable information about the band structure of semiconductors and alloys. SEM can be used to image the surface of samples, while XRD gives valuable information about the structure of the sample.

REFERENCES

[1] D. A. Skoog, F. J. Holler, and S. R. Crouch, *Principles of Instrumental Analysis*, 6th ed. (Brooks Cole, Belmont, CA, 2006).

[2] J. I. Pankove, *Optical processes in semiconductors* (Prentice-Hall, Englewood Cliffs, NJ, 1971).

[3] S. T. Tan, B. J. Chen, X. W. Sun, X. Hu, X. H. Zhang, and S. J. Chua, Journal of Crystal Growth **281**, 571 (2005).

[4] J. I. Goldstein and D. B. Williams, Journal of Electron Microscopy Technique **5**, 105 (1987), _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jemt.1060050111.

[5] D. E. Newbury, *Advanced scanning electron microscopy and X-ray microanalysis* (1986) oCLC: 861706289.

[6] C. Suryanarayana and M. G. Norton, *X-Ray diffraction: a practical approach* (1998) oCLC: 861706202.

CHAPTER 6

The Temperature Dependent Band-Edge as a Function of Annealing

---

Abstract — In this chapter we demonstrate the relationship between the band-edge defect distribution, controlled *via* annealing, and the of degree electron-phonon (e-p) coupling in ZnO, $Mg_{0.07}Zn_{0.93}O$ and $Mg_{0.15}Zn_{0.85}O$. We demonstrate that the band-edge in derivative spectra is composed of a Gaussian related to the random distribution of defects in the material. We show, based on UV-Vis absorption spectra at 77 K, that the band-edge distribution of defects was decreased as a function of increasing annealing temperature. Furthermore, we show that the broadening of the Gaussian obeys a one-to-one correlation with the XRD Full-Width Half Max (FWHM) indicating that the dominant defects leading to band-edge broadening are structural inhomogeneities such as grain boundaries.[1] We show that ZnO undergoes an amorphous-to-polycrystalline phase transition at 400 °C. In the amorphous-like morphology, ZnO exhibits none-to-weak e-p coupling. Following the phase transition, ZnO exhibits traditional electron-phonon coupling typical of more crystalline structures. $Mg_{0.07}Zn_{0.93}O$ and $Mg_{0.15}Zn_{0.85}O$ demonstrate similar behavior to ZnO in the as-grown/900 °C annealing temperature limits, respectively.

6.1 INTRODUCTION

There are trends in the literature which indicate that defects, specifically cation-cation disorder in semiconductor ternary alloy systems such as CuInTe and CuInSe, significantly impact the e-p interaction of the material.[2–8] The relationship between the temperature broaadening term and the defect broadening term of the band-edge *slope* for these materials, in comparison to ZnO and MgZnO, is shown in Fig. 6.1. The effects of e-p interaction include thermal broadening and thermal redshift of the

optical gap. Exploring this relationship is critical to the continued rational application of semiconductor technology in variable temperature environments such as in space exploration.[9] Rincón proposes that, in general, in the presence of defects, only a fraction of the phonon modes excited at a given temperature can interact with electrons, thereby reducing the effect of thermal broadening.[5] We set out to more generally establish this dependence by investigating the e-p interaction *via* thermal broadening of the optical gap in ZnO thin film as function of annealing. As will be shown, this investigation also gives insight into the defect types present in the as-grown and annealed films.

Annealing thin films at elevated temperatures allows the evolution into a more stable state, removing defects and contributing to more ordered morphology.[10] The samples were annealed following the procedure described in Chap. 3.

6.2 DEFECT DISTRIBUTION

To uncover the defect distribution in our samples, we analyzed the log of the energy derivative of $\alpha(E)$, shown in Fig. 6.2(a). The rationale for this is treated in detail in Chap. 8.

In Fig. 6.3 (b) we show that the form of $d(ln\alpha(E))/dE$ can be described with good accuracy by a Gaussian. Fig. 6.3 (b) shows the data after the removal of interference patterns. The features at the band-edge comprise (i) an artefact from interference and (ii) the actual absorption signal of the material. Analysis described in detail in Chap. 7 confirms that the former is an artefact from interference.[11] The present chapter focuses on the extracted absorption signal, modeled excellently by the following Gaussian form:

$$\frac{d(ln\alpha(E))}{dE} = \frac{1}{E_u} e^{\frac{-(E-E_0)^2}{2w^2}} \tag{6.1}$$

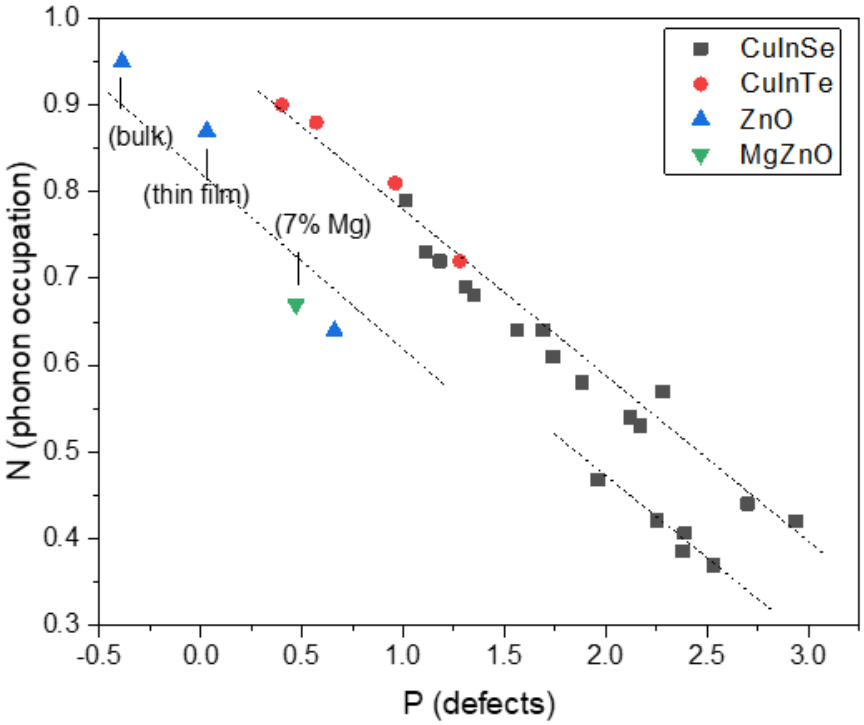


FIGURE 6.1: *The relationship between the temperature broadening term and the defect broadening term of the band-edge slope for various stoichiometries of CuInTe and CuInSe, ZnO and MgZnO.*[2–8]

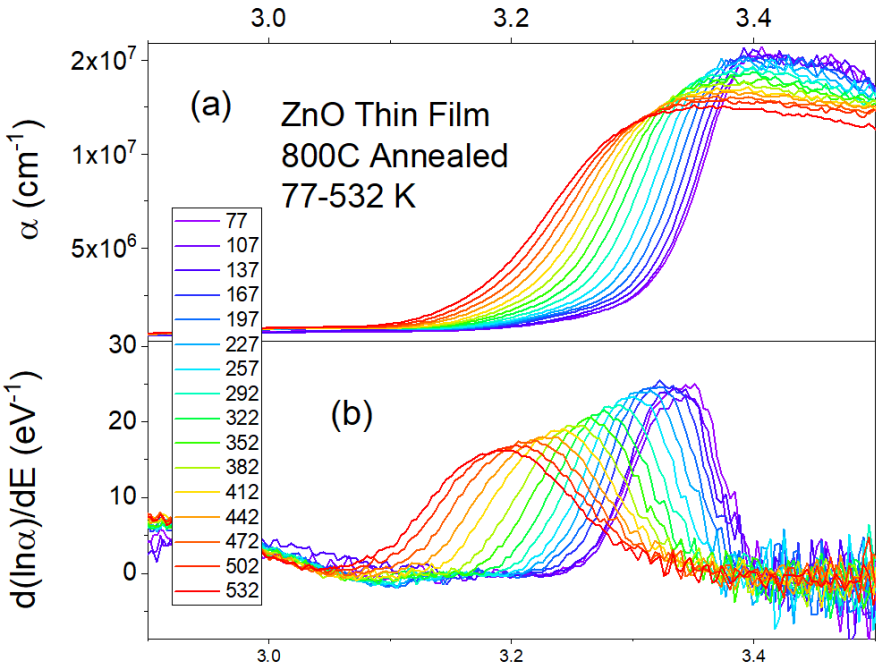


FIGURE 6.2: *(a) The absorption coefficient (b) energy derivative of the logarithm of the absorption coefficient of ZnO thin film annealed at* 800 °*C acquired at temperatures from* $77 - 532$ *K. The interference artefacts have been removed from the plots of* $d(ln\alpha(E))/dE$.

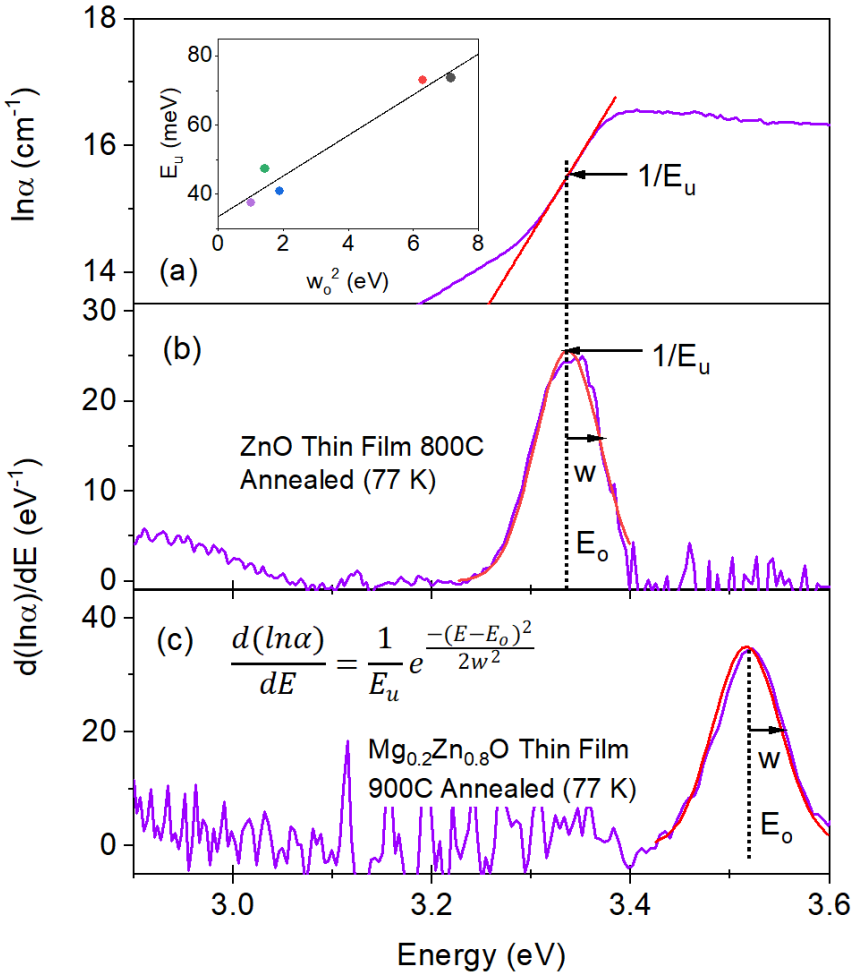


FIGURE 6.3: *(a) Semi-logarithmic plot of $\alpha(E)$, $ln\alpha(E)$, and (b) energy derivative of $ln\alpha(E)$ for ZnO thin film annealed at 800 °C acquired at 77 K. (c) Energy derivative of $ln\alpha(E)$ for $Mg_{0.2}Zn_{0.8}O$ thin film annealed at 900 °C acquired at 77 K. The interference artefacts have been removed from the plots of $d(ln\alpha(E))/dE$. The parameters of the Gaussian model in Eq. 6.1 are shown. The inset in (a) shows E_u as a function of the variance w_o^2 of the defect distribution. The line is a fit to Eq. 6.2.*

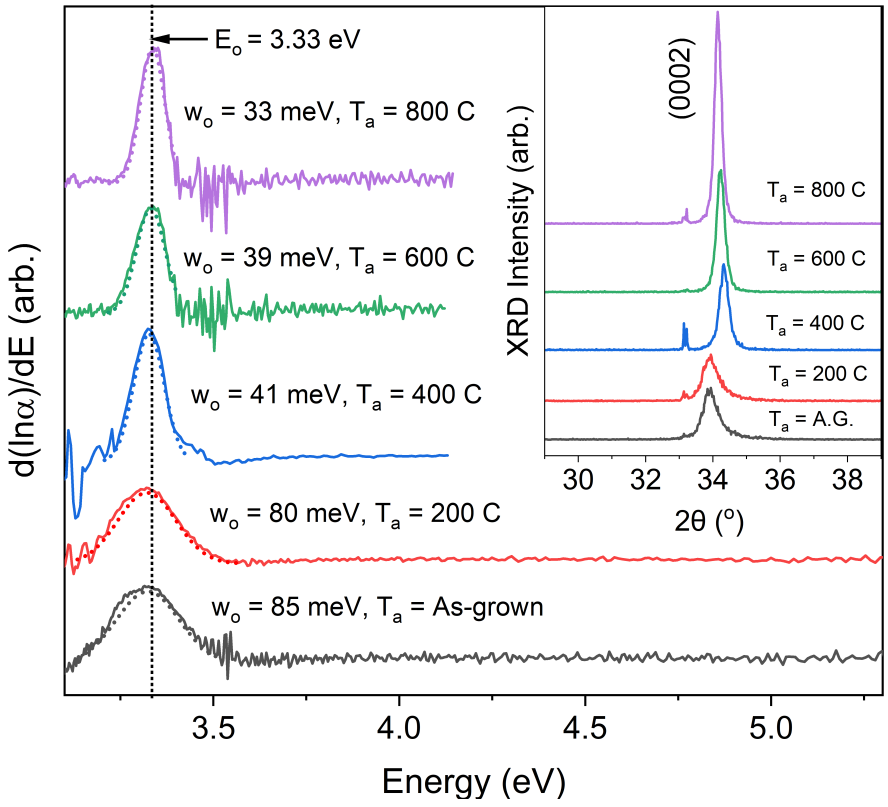


FIGURE 6.4: *(a) The energy derivative of $ln\alpha(E)$ for the ZnO thin film sample as-grown and annealed (T_a) at* 200, 400, 600, *and* 800 °*C after removal of the interference artifact. The standard deviation of the Gaussians (dotted fits to data) are shown to narrow with increased annealing temperature while the UBE, E_o, shown at* 3.33 *eV by the vertical dotted line, remains invariant to defects. The inset shows the XRD diffraction pattern for ZnO thin film as-grown and annealed at* 200, 400, 600, *and* 800 °*C.*

where, E_u is the Urbach energy, E_o is the UBE and w is the Gaussian standard deviation, (where w^2 is the variance of the distribution).

Urbach theory asserts that defects alter the gap value locally through local band deformation.[12,13] Defects are assumed to be randomly distributed in a material, so the distribution of energy at the band-edge can be described by a normal distribution.[12,14] Our results show that derivative analysis is required to experimentally uncover the normal distribution of defects. The thoery supporting this transformation is described in Chap. 8.

Economou shows by first-principles calculation that E_u is proportional to the variance, w^2, in the bandgap energy distribution at the band-edge due to the defect distribution in the material:[12]

$$E_u = 2Aw^2 \tag{6.2}$$

To analyze our relationship between w^2 and E_u, it was first necessary to separate defect broadening from thermal broadening. We analyzed the spectra at 77 K where thermal contributions to band-edge broadening are negligible.[5,8,14]

Fig. 6.3 shows our plot of E_u vs. ${w_o}^2$ for the ZnO thin film at 77 K for each annealing condition. This allows the analytical extraction of A which was found to be 0.003 eV^{-1}, within the range of values reported by Bacalis *et al* for various materials.[15]

6.3 ANNEALING EXPERIMENTAL RESULTS

Analysis of the defects at 77 K for each annealing stage—starting at 200 $^{\circ}C$ and increasing by 200 $^{\circ}C$ up to 800 $^{\circ}C$ —revealed that the UBE, E_o, is invariant to the defect concentration in the material, measuring between $\approx 3.33 - 3.34\ eV$ throughout all stages of annealing treatment as shown in Fig. 6.4. Values of E_o are shown in Table 6.1. Regardless of defect concentration, broadening is centered about E_o. In

agreement with our results, Economou predicts that E_o represents the average value of the fluctuating local energy environment.[15]

Furthermore, the defect broadening parameter, w_o, decreases with each increase in annealing temperature, as shown in Fig. 6.4. XRD measured at each annealing temperature, as shown in the inset of Fig. 6.4, agreed with this finding. As expected, the XRD FWHM (Full-width Half-Max) also decreased with each increase in annealing temperature. A key finding of this dissertation, as shown in Fig. 6.5, is that the defect parameter w_o obeys a one-to-one correlation (Pearson's r = 0.99) with the XRD FWHM. This suggests that the dominant defects measured by w_o are structural inhomogeneities such as grain boundaries. The results are compiled in Table 6.1.

6.4 SOLID-STATE PHASE TRANSITION

The analysis of w_o and the XRD FWHM as a function of annealing in Fig. 6.5 shows that both parameters are drastically reduced at 400 $^{\circ}C$, corresponding to a 39% decrease for the XRD FWHM and 49 % for w_o. This indicates the activation of an amorphous-to-polycrystalline solid-state phase transition corresponding to grain growth activation. This agrees with the results of Vishwas *et al.* who also found this phase transition at 400 $^{\circ}C$ *via* XRD for ZnO thin film grown by the sol-gel method.[16] Grain-growth activation is further supported by the SEM images shown in Fig. 6.5 which compares the morphology of the as-grown sample and the sample annealed at 800 $^{\circ}C$. The as-grown sample shows clustering and no well-defined grains while the sample annealed at 800 $^{\circ}C$ shows well defined grains. In contrast, the XRD intensity shows a gradual improvement over all stages of annealing as shown in Fig. 6.5.

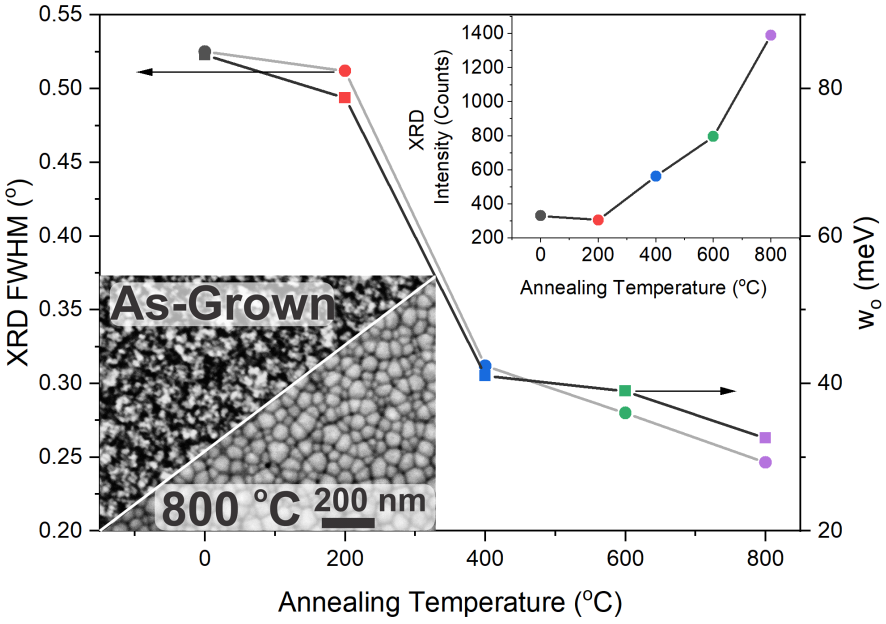


FIGURE 6.5: *The defect parameter, w_o plotted with the XRD FWHM for ZnO thin film as-grown and annealed at* 200, 400, 600, 800 °*C. The Pearson's correlation is* 0.99 *between these two parameters. The top inset shows the XRD intensity and the bottom inset shows SEM images of ZnO thin-film as-grown (top) and annealed at 800 °C.*

TABLE 6.1: *E_o $(0K)$, w_o and E_u obtained by fitting the absorption signal of $d(ln\alpha(E))/dE$ at 77 K to Eq. 6.1, the XRD FWHM obtained by XRD peak analysis, the e-p coupling constant (α_o), phonon temperature (Θ) and phonon frequency ($h\nu_p$) from fitting the BET model to $w(T)$ for the ZnO film as-grown (AG) and annealed successively by 200 $^\circ C$ up to 800 $^\circ C$, and for the $Mg_{0.2}Zn_{0.8}O$ film as-grown and annealed at 900$^\circ C$*

ZnO	$E_o(0K)$ (eV)	w_o (meV)	XRD FWHM (°)	E_u (meV)	α_o $(10^{-3}K^{-1})$	$\Theta(K)$	$h\nu_p(meV)$
AG	3.32	85	0.52	73.64	0	0	0
200 $^\circ C$	3.32	80	0.51	72.94	1.41	1277	110
400 $^\circ C$	3.33	41	0.31	40.90	2.70	850	73
600 $^\circ C$	3.33	39	0.28	47.37	3.29	893	77
800 $^\circ C$	3.34	33	0.25	37.37	3.16	857	74
MgZnO							
AG(15% Mg)	3.65	136.7	-	113.38	0	0	0
900 $^\circ C$ (7% Mg)	3.52	33.74	-	28.53	6.49	840.25	72

6.5 DEFECT-MEDIATED ELECTRON-PHONON COUPLING

We now consider the effect of defects such as grain boundaries on the thermal broadening of the band-edge. w for the as-grown sample (Fig. 6.6) shows no temperature dependence, and in the sample annealed at 200 $^{\circ}C$, w shows negligible temperature dependence. Studies on the thermal contribution to band-edge broadening of amorphous materials using Urbach analysis indicates a constant slope, or E_u, or little to no thermal response of the band-edge slope, for a broad range of temperatures in amorphous materials.[17–20] Some amorphous materials that display this behavior include α-Ge[20] (amorphous Ge), hydrogenated α-Si[21], α-As_2S_3[22] and α-GeTe alloys.[23] These results further support the conclusion that the structure of our as-grown ZnO (and the sample annealed at 200 $^{\circ}C$ to a lesser degree) exists mainly as disordered networks. Wasim suggests that in Cu ternaries, maximum structural disorder corresponds to no thermal broadening at the band-edge and asserts that near this threshold, a structural phase transition in these materials is expected to occur.[7] This correspondence of a structural phase transition with enhanced thermal broadening is in agreement with our results.

Following the structural phase transition at 400 $^{\circ}C$ annealing, and for 600 $^{\circ}C$ and 800 $^{\circ}C$ annealing, a significant thermal broadening response is seen (Fig. 6.6). The prevalence of thermal broadening as opposed to defect broadening is characteristic of more crystalline solids.[18] The broadening as a function of temperature, $w(T)$, shows a trend similar to that predicted by the Wasim model, which is reasonable as we show in Chap. 8, E_u is proportional to w^2. To our knowledge, no model has yet been developed for the temperature dependence of $w(T)$ extracted *via derivative absorption spectra*.

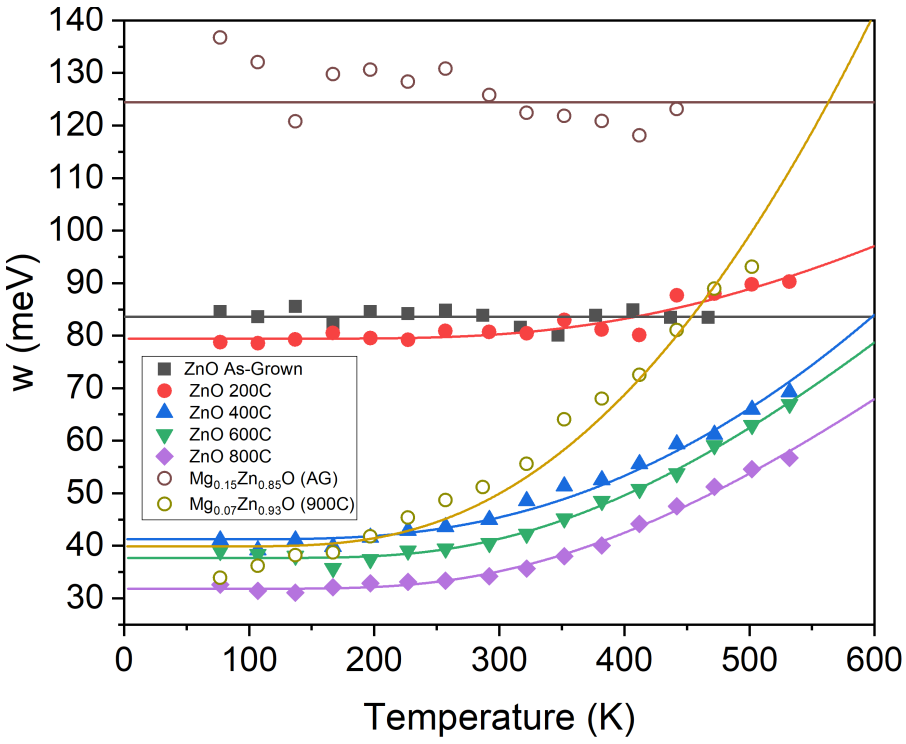


FIGURE 6.6: *The Gaussian standard deviation as a function of sample temperature for the ZnO film as-grown and annealed at* 200, 400, 600 *and* 800 °*C, and for the* $Mg_{0.15}Zn_{0.85}O$ *film as-grown and the* $Mg_{0.07}Zn_{0.93}O$ *film annealed at* 900 °*C. The solid lines are the fits to the BET model in Eq. 6.3.*

6.6 MODEL OF TEMPERATURE-DEPENDENT BAND-EDGE DISTRIBUTION WIDTH

There are many models that describe the temperature-dependent width of broadening in photo-luminescence (PL), ellipsometry and the absorption spectra. We applied a Bose-Einstein-type (BET) model to the temperature dependence of $w(T)$ extracted *via* derivative absorption spectra:[24–27]

$$w(T) = w_o \left(1 + \frac{2\alpha(T)}{e^{\Theta/T} - 1}\right) \tag{6.3}$$

where w_o quantifies the defect baseline. Θ is the phonon temperature of the material involved in broadening and corresponds to the average phonon frequency in the material *via* $k_b\Theta = h\nu_p$. $\alpha(T)$ quantifies the degree of e-p coupling and is a weight-factor on the Bose-Einstein term. $\alpha(T)$ is usually regarded in the literature as having no temperature dependence, however, assuming this factor is independent of temperature did not yield physically reasonable results. As shown in Fig. 6.7, our analysis shows that $\alpha(T)$ is approximately linearly temperature dependent and can be approximated by $\alpha(T) = \alpha_o T$, where α_o is defined as the e-p coupling constant.

The results of fitting Eq. 6.3 to the data in Fig. 6.6 are shown in Table 6.1. The results of $h\nu_p$ for the sample annealed at 400, 600, and 800 $^\circ C$ indicate phonon frequencies of 73, 77 and 74 meV, respectively. These are in good agreement with measured values of the $A_1(LO)$ mode in ZnO ($\approx$ 550 cm^{-1}), $\approx$ 70 meV.[28–30] In contrast, assuming $\alpha(T)$ has no temperature dependence yields phonon frequencies of 52, 121 and 119 meV, respectively, which is not physically reasonable since the $A_1(LO)$ mode is the highest available optical mode in ZnO as shown in Fig. 2.20.

We also note that values for Θ for the sample annealed at 200 $^\circ C$ give a value for $h\nu_p$ higher than the the $A_1(LO)$ mode in ZnO. Wasim proposes that unusually high values for Θ may indicate higher-than usual frequency modes at localized sites[5] which are more prevalent in disordered networks.[31]

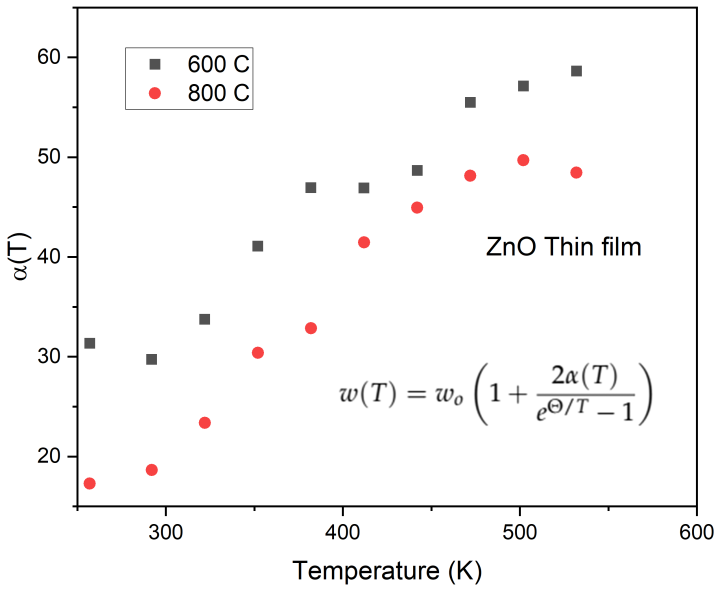


FIGURE 6.7: *Solving for $\alpha(T)$ in Eq. 6.3 by substituting experimental values for w_o, T and a known value for Θ for ZnO yields the following form.*

Fig. 7.5 summarizes the results of applying the model in Eq. 6.3 to the temperature-dependent data shown in Fig. 6.6. This more generally establishes the trend found in the literature.

6.7 EVIDENCE FOR Zn_i OUT-DIFFUSION / LATTICE INCORPORATION

At 200 $^{\circ}C$ annealing, minor thermal broadening is activated. We assert that this annealing temperature, at which there is insufficient energy to induce a structural change—as shown by XRD—causes the out-diffusion and lattice incorporation of Zn_i. Shown in Fig. 6.9, our previous work confirmed *via* visible PL that Zn_i are the dominant native defect in sputtered as-grown ZnO.[10] Theoretical predictions by Kohan *et al.* also indicate that Zn_i are highly unstable and as a result are expected to be thermally active at this relatively low annealing temperature.[32]

The presence of Zn_i in the thin flim studied here can be attributed to the DC sputtering technique used. This growth technique is commonly reported in literature to lead to the growth of Zn-rich films, thereby favoring the formation of Zn_i-related defects.[33–35] This can be understood by noting that sputtering involves the competing processes of target oxidation and removal of target material. To achieve sputtering growth, target material must be removed at a sufficient rate to prevent oxide overcoat formation. However, by removing target material at a higher rate than the oxidation process occurs, the sputtered material may not completely oxidize. The net result of these competing processes is a Zn-rich film.

CONCLUSION — In summary, we observe an amorphous-to-polycrystalline phase transition in the ZnO thin film at 400 $^{\circ}C$. The e-p coupling constant, α_o, shows an inverse relationship to the defect parameter, w_o. With each successive annealing treatment, following the removal of Zn_i at 200 $^{\circ}C$ and grain growth at 400 $^{\circ}C$, an increase in the degree of the band-edge thermal broadening was observed. Paired

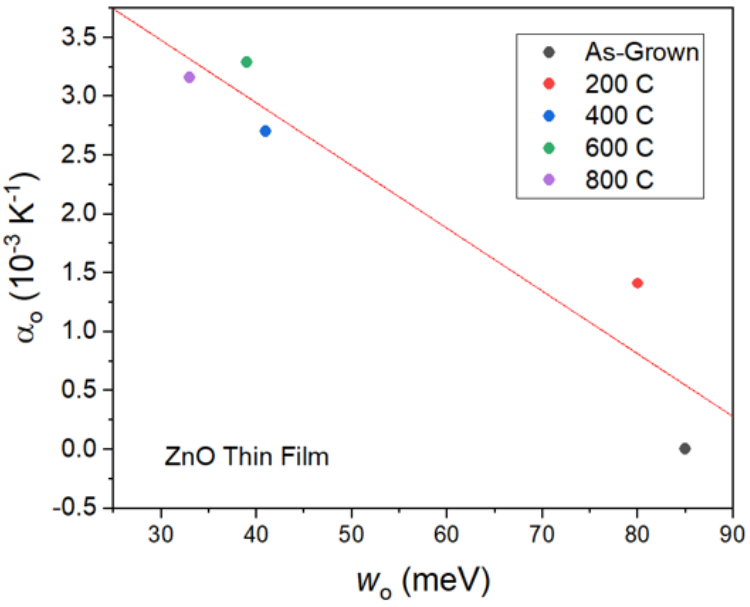


FIGURE 6.8: *α_o, the electron-phonon coupling constant, as a function of w_o, the defect concentration, from applying the model in Eq. 6.3 to the temperature-dependent data shown in Fig. 6.6*

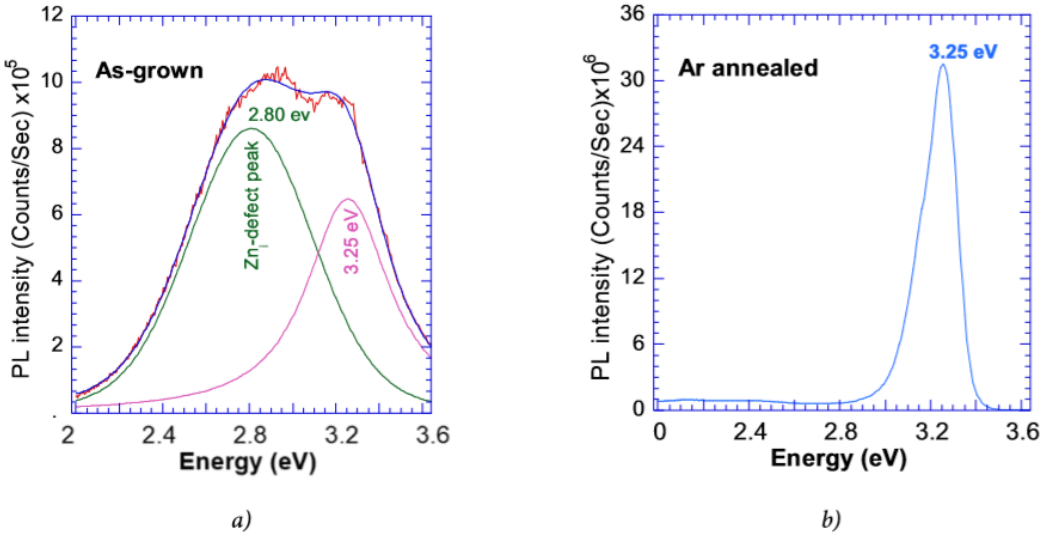


FIGURE 6.9: *PL spectra of the ZnO films: (a) as grown film showing two peaks at 2.80 and 3.25 eV, were obtained with a voigt profile fitting to the experimental curve. The 3.25 eV is attributed to UV band-edge emission and 2.80 eV is due to Zni related defects. (b) Ar annealed ZnO film showing the significantly enhanced UV PL with nearly quenched visible emission peak.*[10]

with Wasim's investigation of this effect for disorder *via* alloy inhomogeneity, this suggests that *any* defect mechanism which interrupts lattice periodicity is sufficient, though perhaps to varying degrees, in reducing the e-p coupling properties of a material. Further investigations into the inverse relationship between defects and e-p interaction of other semiconductor systems are underway. While the prediction of this trend by Wasim is further confirmed through our experimental work, investigation into the physical mechanism *via* localization is required.

REFERENCES

[1] C. Suryanarayana and M. G. Norton, *X-Ray Diffraction: A Practical Approach* (Springer US, 1998).

[2] H. Nakanishi, T. Sawaya, S. Endo, and T. Irie, Japanese Journal of Applied Physics **32**, 200 (1993).

[3] S. M. Wasim, G. Marın, C. Rincón, and G. Sánchez Pérez, Journal of Applied Physics **84**, 5823 (1998).

[4] T. Shioda, S. Chichibu, T. Irie, H. Nakanishi, and T. Kariya, Journal of Applied Physics **80**, 1106 (1996).

[5] C. Rincón, S. M. Wasim, G. Marín, R. Márquez, L. Nieves, G. S. Pérez, and E. Medina, Journal of Applied Physics **90**, 4423 (2001).

[6] E. M. Ismardo Bonalde, Phys. Rev. B **69** (2004), 10.1103/PhysRevB.69.195201.

[7] S. M. Wasim, C. Rincón, G. Marín, P. Bocaranda, E. Hernández, I. Bonalde, and E. Medina, Phys. Rev. B **64**, 195101 (2001).

[8] R. C. Rai, Journal of Applied Physics **113**, 153508 (2013).

[9] H. Liu, Y. Peng, J. Long, W. Lv, K. Liang, R. Yang, and D. Han, in *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)* (2019) pp. 1–3, iSSN: 2577-0829.

[10] D. Thapa, J. Huso, J. L. Morrison, C. D. Corolewski, M. D. McCluskey, and L. Bergman, Optical Materials **58**, 382 (2016).

[11] J. C. Manifacier, J. Gasiot, and J. P. Fillard, J. Phys. E: Sci. Instrum. **9**, 1002 (1976).

[12] M. A. Kastner, S. R. Ovshinsky, and G. A. Thomas, *Disordered Semiconductors*, 1987th ed. (Springer, New York, 1987).

[13] J. I. Pankove, *Optical Processes in Semiconductors* (Courier Corporation, 1975).

[14] G. D. Cody, T. Tiedje, B. Abeles, B. Brooks, and Y. Goldstein, Phys. Rev. Lett. **47**, 1480 (1981).

[15] N. Bacalis, E. N. Economou, and M. H. Cohen, Phys. Rev. B **37**, 2714 (1988).

[16] M. Vishwas, K. N. Rao, K. V. A. Gowda, and R. P. S. Chakradhar, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **77**, 330 (2010).

[17] S. Abe and Y. Toyozawa, J. Phys. Soc. Jpn. **50**, 2185 (1981).

[18] A. F. Zatsepin, Y. A. Kuznetsova, V. N. Rychkov, and V. I. Sokolov, Journal of Nanoparticle Research **19**, 111 (2017).

[19] I. Z. Kostadinov, J. Phys. C: Solid State Phys. **10**, L263 (1977).

[20] N. Connell, physica status solidi (b) **53**, 213 (1972).

[21] E. C. Freeman and W. Paul, Phys. Rev. B **20**, 716 (1979).

[22] F. Kosek and J. Tauc, Czech J Phys **20**, 94 (1970).

[23] A. Deneuville, A. Mini, and B. K. Chakraverty, Phys. Rev. Lett. **37**, 295 (1976).

[24] C. Rincón, S. M. Wasim, G. Marín, G. Sánchez Pérez, and G. Bacquet, Journal of Applied Physics **82**, 4500 (1997).

[25] L. Malikova, W. Krystek, F. H. Pollak, N. Dai, A. Cavus, and M. C. Tamargo, Phys. Rev. B **54**, 1819 (1996).

[26] P. Lautenschlager, Phys. Rev. B **35**, 9174 (1987).

[27] A. K. Viswanath, J. I. Lee, D. Kim, C. R. Lee, and J. Y. Leem, Phys. Rev. B **58**, 16333 (1998).

[28] D. C. Reynolds, D. C. Look, B. Jogai, R. L. Jones, C. W. Litton, W. Harsch, and G. Cantwell, Journal of Luminescence **82**, 173 (1999).

[29] D. W. Hamby, D. A. Lucca, M. J. Klopfstein, and G. Cantwell, Journal of Applied Physics **93**, 3214 (2003).

[30] R. Cuscó, E. Alarcón-Lladó, J. Ibáñez, L. Artús, J. Jiménez, B. Wang, and M. J. Callahan, Phys. Rev. B **75**, 165202 (2007).

[31] P. DEAN, Rev. Mod. Phys. **44**, 127 (1972).

[32] A. F. Kohan, G. Ceder, D. Morgan, and C. G. Van de Walle, Phys. Rev. B **61**, 15019 (2000).

[33] W. S. Shi, O. Agyeman, and C. N. Xu, Journal of Applied Physics **91**, 5640 (2002).

[34] V. V. Khomyak, M. M. Slyotov, I. I. Shtepliuk, G. V. Lashkarev, O. M. Slyotov, P. D. Marianchuk, and V. V. Kosolovskiy, Journal of Physics and Chemistry of Solids **74**, 291 (2013).

[35] S.-Y. Chu, W. Water, and J.-T. Liaw, Journal of the European Ceramic Society **23**, 1593 (2003).

CHAPTER 7

Analysis of Interference Fringe Patterns in Thin Film Absorption Spectra

---

Abstract — In this chapter we discuss the physical origin of interference patterns in the transmission (and absorption) spectra of thin films. We show, based on the theory, that it is possible to predict where these features will arise in the spectra. Consequently, we show that especially in the derivative spectra—which magnifies fine features —interference fringes persist deeply into the band-edge region. Following the confirmation of interference fringe features at the band-edge, we demonstrate an analytical technique which deconvolves the interference fringe from the band-edge signal.

7.1 INTRODUCTION

Experimental measurements can be decomposed into two parts:

- *Signal*, which carries information about the sample that is of specific interest to the scientist
- *Noise*, which is made up of unwanted extraneous information which degrades the accuracy and precision of analysis.[1]

In optical spectroscopy, anomalies, artefacts and interference are common sources of noise. At times, these may only be a nuisance and do not undermine quantitative analysis but at other times can be catastrophic and lead to incorrect conclusions if not properly mitigated.

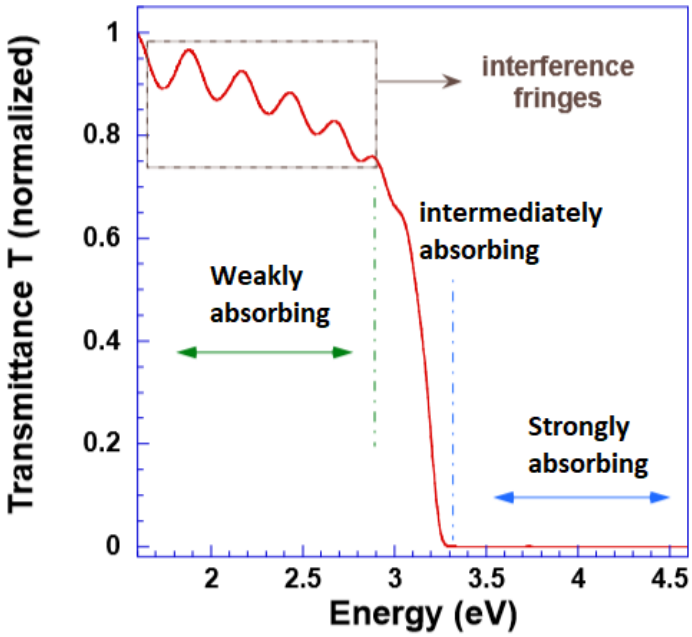


FIGURE 7.1: *As shown by the vertically dotted lines, the normalized transmission spectrum can be broadly categorized into (from left to right) a weakly absorbing, an intermediately absorbing and strongly absorbing region.*

The transmission spectrum of thin, non-scattering films of uniform thickness deposited on substrates has an undulating sinusoidal baseline originating from multiple internal reflections within the film; these features are commonly referred to as “thin-film interference" or simply "interference fringes."[2] As shown in Fig. 7.1, the normalized transmission spectrum can be divided into three distinct regions:

- a weakly absorbing region (T>0.6)
- an intermediately absorbing region (0.6> T > 0.4)
- a strongly absorbing region (T < 0.4)

Interference fringes usually diminish gradually from the weakly absorbing to the strongly absorbing regions. However, in the intermediately absorbing region—the band-edge—interference fringes are convolved with spectral features of interest and the extracted quantitative parameters have the potential to be inaccurate. To current knowledge of the author, fringe patterns in thin-film absorption spectroscopy have not been reported to significantly affect the determination of parameters such as the optical gap or Urbach energy.

Interference fringes can in fact be useful and are used in determining the refractive index or thickness of a thin film. In this chapter, we show that although interference patterns gradually diminish in intensity, remnants of their features extend well into the strongly absorbing region (T<0.1). This is especially significant when analyzing derivative spectra as a means of magnifying fine-features. In fact, it is through the derivative spectra that the extent of interference fringes in affecting the lineshape of the band-edge becomes clear. Here, fringe patterns have the potential to significantly affect the accuracy of quantitative analysis. Derivative spectra is analyzed here in order to determine the temperature- and energy-dependent density-of-states at the band-edge as discussed in Chap. 6. The quantitative usefulness of extracting precision density-of-state measurements from differential UV-Vis absorption spectra rests

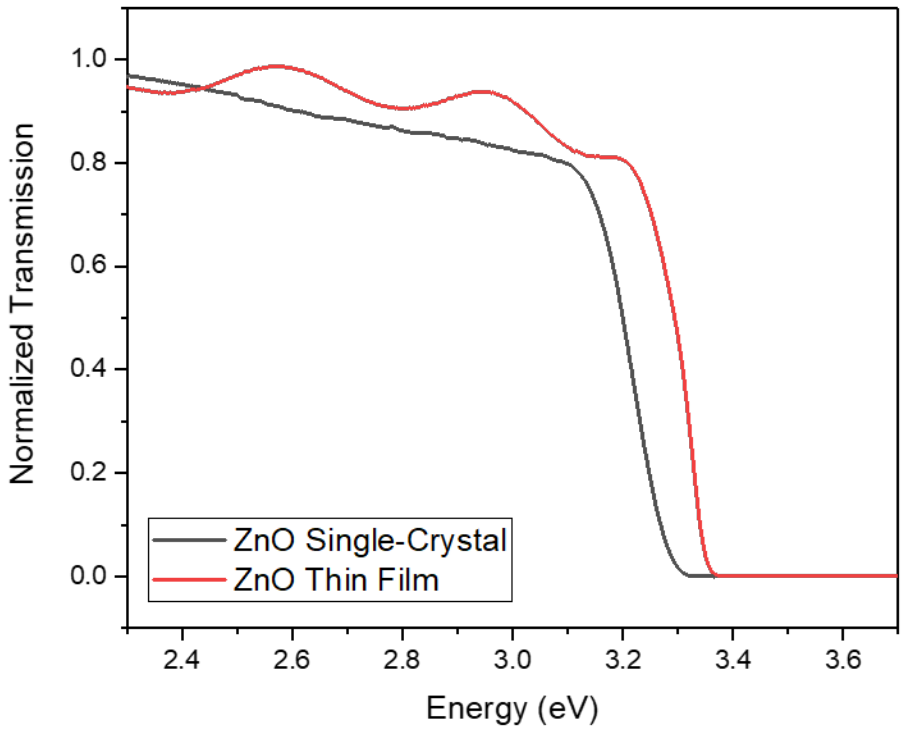


FIGURE 7.2: *ZnO thin film sample showing interference fringes while the bulk ZnO sample does not.*

largely on accurate fringe-baseline corrections. It is worth noting that in samples where interference is not present, such as in bulk materials, analysis of derivative spectra is relatively straightforward. Fig. 7.2 illustrates the difference between bulk and thin-film ZnO.

7.2 OVERVIEW OF ANALYTICAL APPROACH

There are several ways of accounting for or eliminating fringe patterns. Some physical approaches include alteration of the sample, involving roughening of the surface or crushing or wedging the sample, or measuring absorption at non-orthogonal angles of incidence, specifically at or near Brewster's angle.[3] Some computational approaches include performing a sinusoidal baseline correction or more complex corrections from explicit optical theory.[4]

This chapter outline the initial steps in accounting for interference, i.e. determining where in the spectra we can expect to find higher order interference maxima/minima and whether or not these are responsible for spectral features at the band-edge in the derivative spectra. When present, Sec. 7.7 describes the removal of interference from the signal. This analysis follows that outlined in "The Envelope Method" by Manifacier *et al*.[5] Using this method, the refractive index (n), extinction coefficient (k) and film thickness (d) can be determined from interference fringes in the transmission spectrum alone. It is worth nothing that the envelope method is applicable only in the weak and intermediately absorbing regions.[5] Here, the envelope method is applied within these regions to determine the film thickness. The film thickness is then used to predict higher order interference maxima and minima in the strongly absorbing region.

7.3 ANALYSIS

The analysis begins with the equation:

$$d_{n,m} = \frac{\lambda_n \lambda_m}{2(n(\lambda_n)\lambda_m - n(\lambda_m)\lambda_n)} \tag{7.1}$$

where n(λ_n) and n(λ_m) are the refractive indices at two adjacent maxima or minima at λ_n or λ_m. Values of the wavelength-dependent refractive index of ZnO thin film (determined via ellipsometry measurements) were referenced from Stelling *et al.*[6] By using the an "envelope" drawn onto the fringe patters, as shown in Fig. 7.3, the film thickness is calculated for each pair of adjacent maxima and minima and an average of these values is taken as the final measured film thickness, as shown in Table 7.2.

The relationship between fringe order (m), wavelength (λ) and film thickness (d) is given by:

$$2n(\lambda)d = m\lambda \tag{7.2}$$

Eq. (7.2) is derived from the optical geometry of multiple internal reflections within a single film for an orthogonally incident beam of light.[7] This is illustrated in Fig. 7.4.[8] Here, m is either an integer or half-integer for maxima or minima. By solving for m, this equation is used to predict the "location" of the higher order maxima and minima in the spectra. Using the average thickness of the film and the wavelength value at the maxima (or minima) within the weakly and intermediately absorbing regions, we can determine their assigned "order" m as shown in Table 3. Next, knowing the sequence of m, we can work in the inverse direction and plug in values for the next integer up (m+1) (or half-integer up (m+$\frac{1}{2}$) to determine its wavelength value, i.e. location on the spectra in the strongly absorbing region where our interest lies. It should be noted that the value of the wavelength-dependent refractive index, n(λ), was not known *a priori*, so the value of n(λ) from the previous

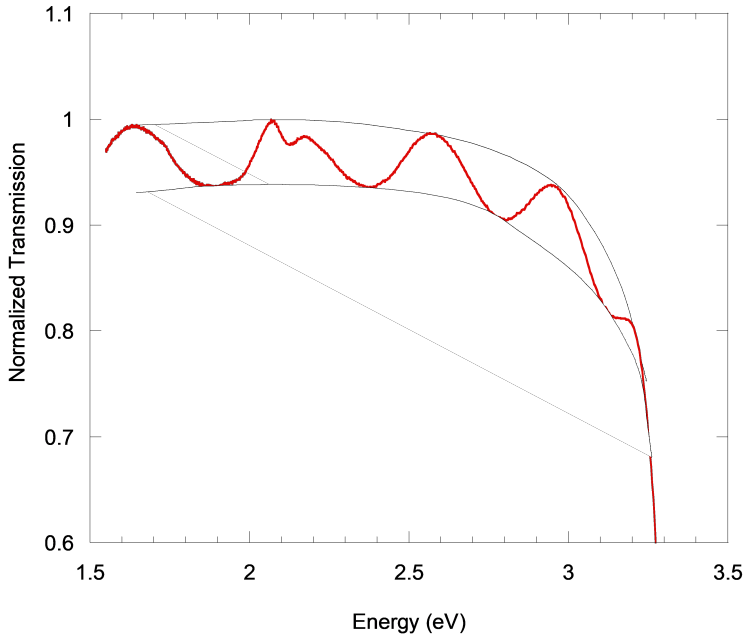


FIGURE 7.3: *The envelope functions wrapping the maxima and minima points of the interference fringe patterns in the weakly and intermediately absorbing regions.*

TABLE 7.1: *Values of λ, $n(\lambda)$, $T_{max}(\lambda)$, $T_{min}(\lambda)$ and d for adjacent minima for Zno thin film at 77 K.*

	λ (nm)	$n(\lambda)$ (-)	$T_{max}(\lambda)$ (-)	$T_{min}(\lambda)$ (-)	d (nm)
			Maxima		
1	759.73	1.6055	0.9946	0.9342	694.81
2	570.41	1.6159	0.9994	0.9325	882.91
3	482.53	1.6402	0.9866	0.9242	788.88
			Minima		
1	759.73	1.6055	0.9946	0.9342	694.81
2	570.41	1.6159	0.9994	0.9325	882.91
3	482.53	1.6402	0.9866	0.9242	788.88

fringe order was used as an initial guess and an iterative process in n and λ was performed until n converged to the 4th decimal place. The final row in Table 7.2 and 7.3 show the conversion from wavelength to energy (eV).

TABLE 7.2: *Calculated values of order m (up to 6.0) from the known maxima/minima locations, λ.*

m	3.5	4.0	4.5	5.0	5.5	6.0
λ (nm)	759.73	656.39	570.41	522.04	482.53	443.02
Energy (eV)	1.632	1.889	2.174	2.375	2.570	2.799

TABLE 7.3: *Predicted values of λ for $m > 6.0$.*

m	6.5	7.0	7.5	8.0	8.5
λ (nm)	408.17	385.46	367.68	335.42	313.88
Energy (eV)	3.038	3.217	3.372	3.696	3.950

The data in Table 7.2 and 7.3 was fitted to a power function in Figure 7.5 indicating that this relationship is $m \propto 1/\lambda^{0.91}$. Since we expect $m \propto 1/\lambda^{1.0}$ from theory, it follows that our analysis has accumulated error. These errors stem from many possible sources including the deviation from ideal sinusoidal behavior of the fringe patterns, illustrated in Fig. 7.6, the free-hand drawing of the envelope functions, averaging of the film thicknesses from multiple fringe-order pairs (average standard deviation of 9%), rounding errors to obtain integer and half-integer values from calculated m, and lastly, the error in using the refractive index of a different ZnO thin-film sample.

The results of the analysis indicate that the location of the 7.0 and 7.5 fringe orders overlap the band-edgeas shown in Fig. 7.7. We can conclude from a theoretical analysis alone that at least one (m = 7.0) fringe maximum may be significantly convolved with the band-edge. Sec. 7.6 provides experimental support or this conclusion. First, the following section will outline a comparison of observed fringe-patterns with an idealized case study.

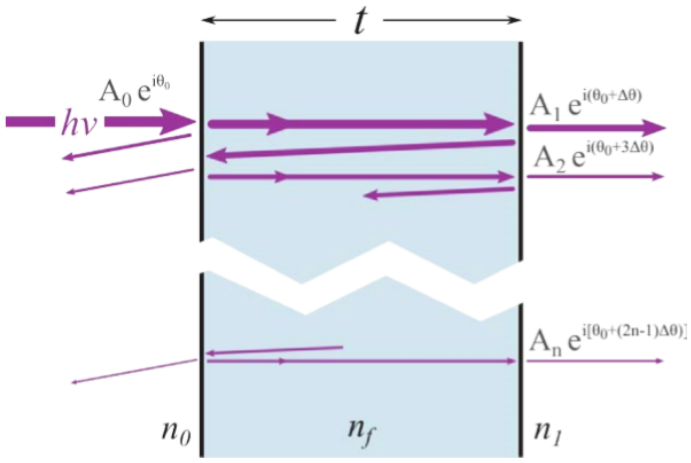


FIGURE 7.4: *When incident light travels through a film, reflections occur on the two interfaces and multiple internal reflections occur within the film. The change in the phase angle each time the light wave propagates in the film will cause constructive and destructive interference between different orders of transmitted light wave when certain conditions are met, resulting in interference fringes in the transmittance spectrum.*[8]

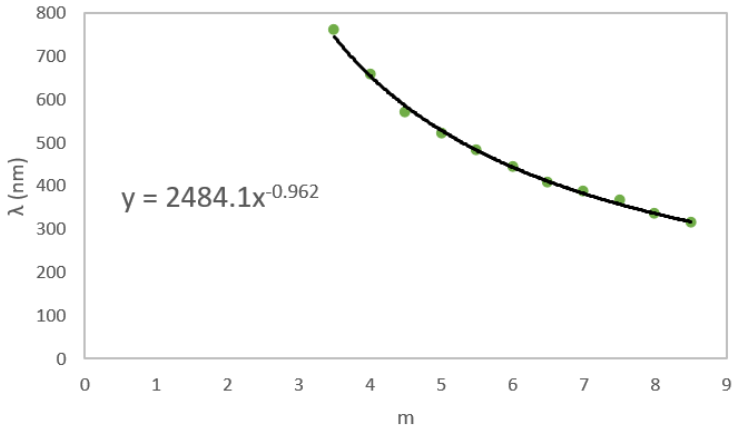


FIGURE 7.5: *The relationship between fringe order, m, and wavelength, λ. An inverse relationship is expected from theory.*

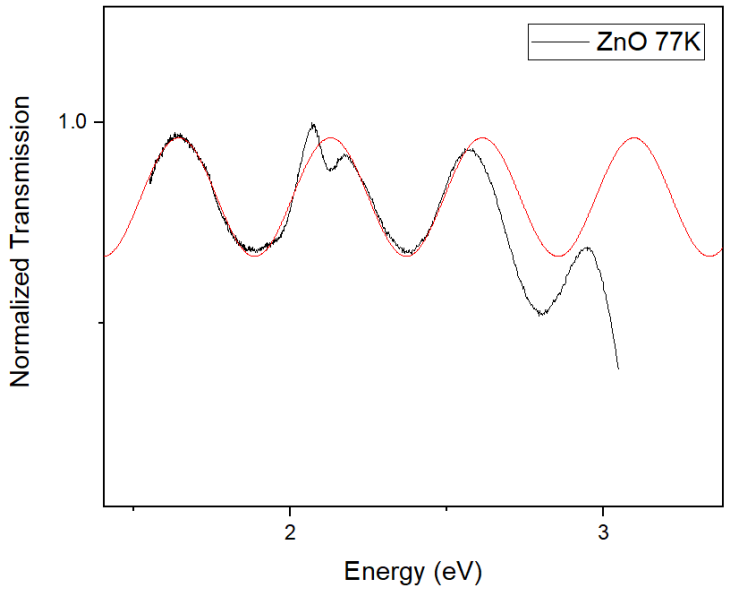


FIGURE 7.6: *A sine curve overlaid onto the interference fringe pattern illustrate the deviation of the spectra from this ideal.*

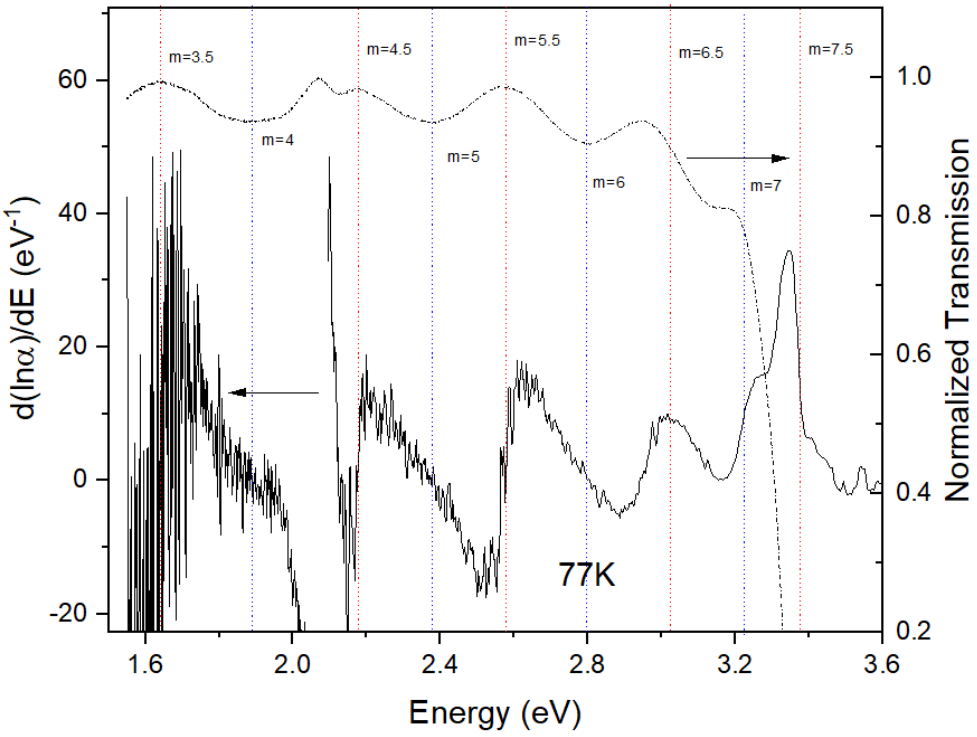


FIGURE 7.7: *The transmission spectra (upper) alongside the derivative spectra (lower) and indicates the energy location of fringe maxima and minima as shown in Table 3.*

7.4 QUALITATIVE INTERPRETATION OF ANALYSIS

It is useful to get a more qualitative understanding of the form of interference fringes as we manipulate the spectra down the analytical pipeline, i.e. as we transform fringes in the transmission spectra to the (logarithmic) absorption spectra, and as we transform that to the derivative spectra. We have performed similar transformations on an ideal sinusoidal function. Figure 7.9 shows the original sine function (analogous to interference in Transmission), the conversion to $ln(A)$ or $ln(-ln(T))$, and $d(ln(A))/dE$.

Figure 7.8 shows the sequential transformation of the analyzed spectra. Notice similar forms to that in Figure 7.9. The purpose of this exercise is to get perspective on the 'first order' behavior of the interference baseline pattern. As is the case with any scientific investigation, understanding spectra from a qualitative perspective is the first step in opening up new lines of inquiry for quantitative analysis. The derivative represents the interference as sharp inflections or pulses. Seen in all three forms of the spectra, there is a diminishing of the intensity of the interference pattern as we approach the strongly absorbing regime—this is most noticeable in the derivative spectra, which again, magnifies fine features. The sharp inflections become attenuated and gradually become somewhat saw-tooth in form.

7.5 TEMPERATURE DEPENDENT INTERFERENCE

In addition to spectra taken at 77 K, the previous analysis was performed for the additional temperatures 287 K, 437 K, and 527 K to investigate the temperature dependence of interference. Since features at the band-edge have a significant dependence on temperature it is reasonable to expect that the spectrum at each temperature will require its own special correction. As shown in Figure 7.10, the average relationship between wavelength, λ, and fringe-order, m, did not vary for spectra over a wide

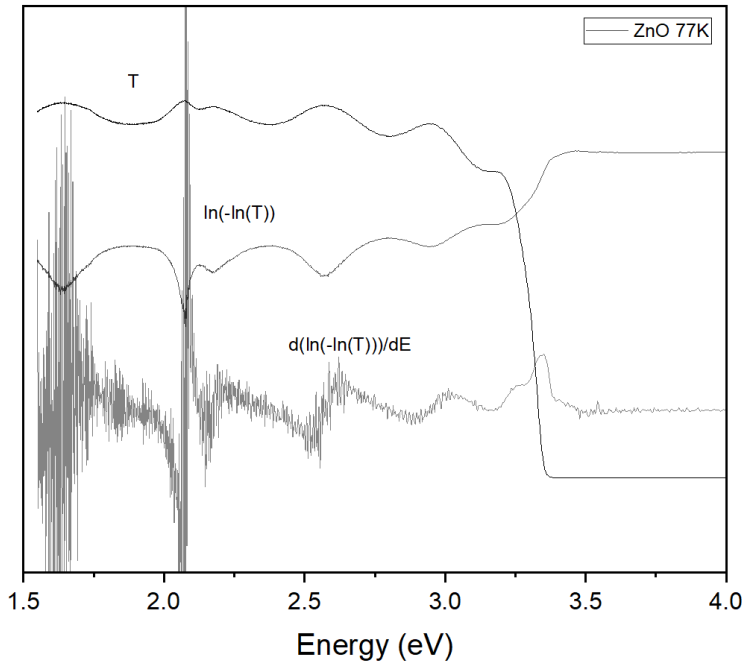


FIGURE 7.8: *The sequential transformation of the analysed spectra. Notice similar forms to that in Figure 7.9*

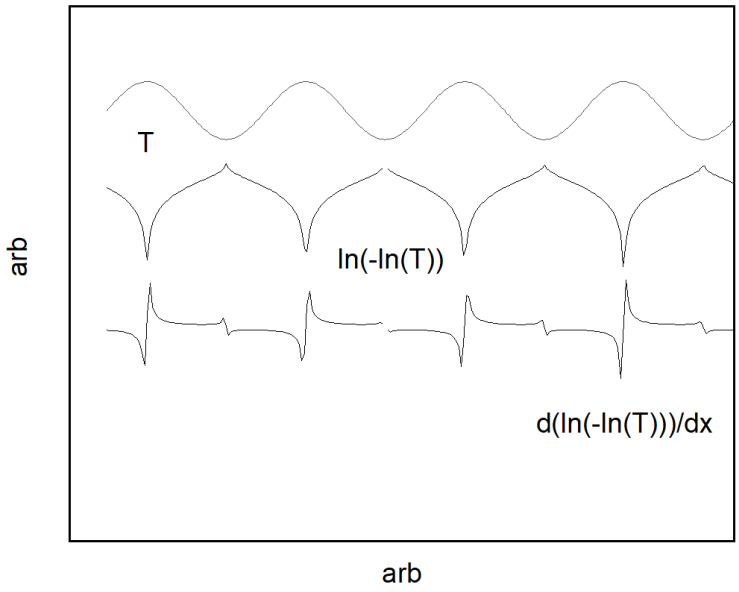


FIGURE 7.9: *(a) an ideal sinusoidal function (top) representing fringe patterns in transmission spectra (T). The change in the functional form is illustrated as we move down the analytical chain: (b) conversion to absorption* $(-ln(T))$ *and subsequent logarithmic representation* $ln(-ln(T))$*, and finally (c) the first derivative of this quantity.*

temperature range; specifically at 77, 287, 437 and 527 K. The data was fitted to a power function which yielded the relationship $m \propto 1/\lambda^{0.92}$. Note the line which separates the experimentally determined location of fringe-order and those predicted from theory.

The average standard deviation of λ (for all m) across this wide temperature range is small, which comes to $\pm$ 2.66 nm. This implies that for any fringe order m, across the range 77—527 K, we observe no temperature dependence (within error) of its energy location. In short, interference fringe patterns are invariant to temperature in the sample studied here.

Figure 7.12 shows the same with special interest at a) the error region and b) the Urbach region. The error region has been named as such because fringe maxima/minima in this region, the intermediately absorbing region, will yield a large error for their energy assignment and should not be used in experimental analysis. This is not to say that the fringes themselves have a dependence on temperature but that the strong shift from transmission to absorption dominates the behavior in this region and takes precedent over the fringe pattern. The Urbach region highlights the strong temperature shift of this region for the four spectra analyzed.

7.6 MODULATING INTERFERENCE FRINGES

As a final check for confirming the presence of interference fringe features at the band-edge, we conducted a controlled condensation experiment. The surface roughness leading to surface scattering of incident light has been shown to affect the response of interference fringe patterns in thin films.[9] We utilized condensation as a non-destructive means of increasing the scattering at the surface of the film. The ability of condensed thin ice films on a substrate to affect interference fringes was demonstrated by Teolis *et al.* and investigated *via* reflectance spectrometry.[3] We first acquired an

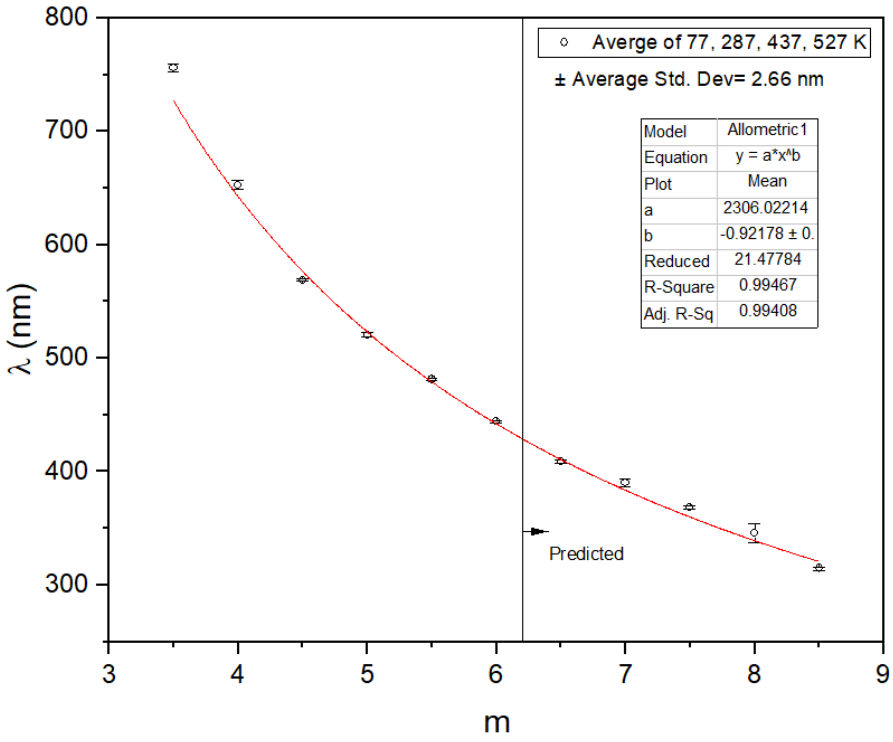


FIGURE 7.10: *Averaged relationship between wavelength, λ, and fringe-order, m, for spectra over a wide temperature range, specifically 77, 287, 437 and 527 K. The data was fitted to a power function which yielded the relationship* $m \propto 1/\lambda^{0.92}$. *Note the line which separates experimentally determined fringe-order location and those predicted from theory. The error bars indicate the standard deviation in λ from the average of all temperatures. The average standard deviation for all m was 2.66 nm.*

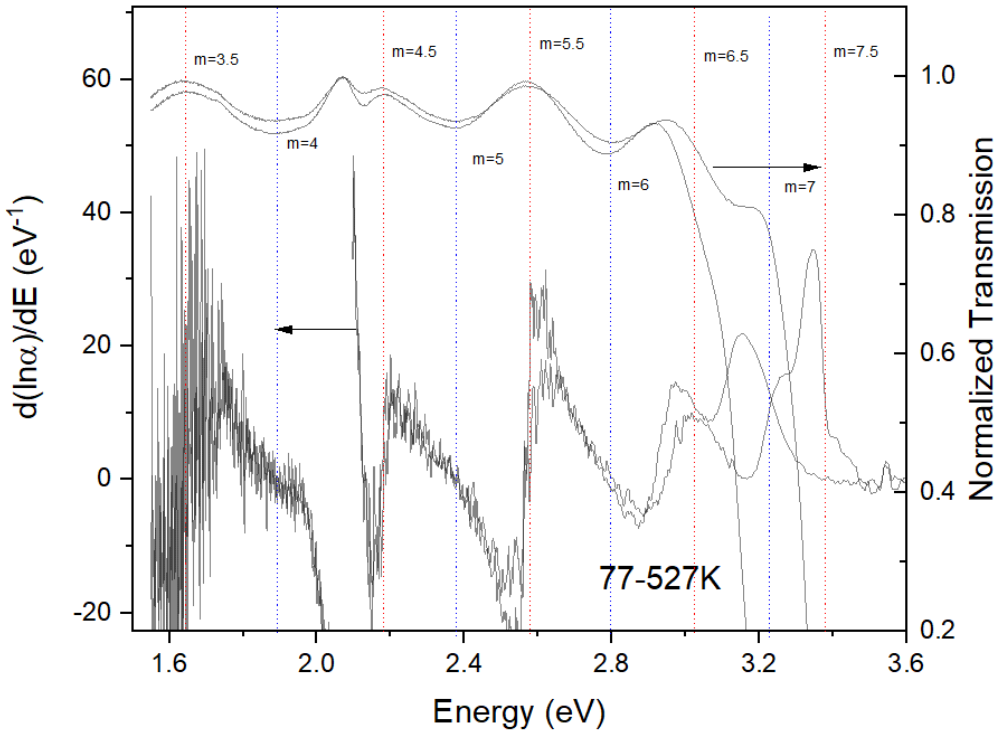


FIGURE 7.11: *Transmission spectra (upper) and corresponding derivative spectra (lower) for 77 and 527K. This highlights the significant temperature shift of band-edge features and conversely the non-temperature-dependence of interference maxima/minima. The white and red dotted lines represent the expected location of fringe maxima and minima in spectra at any temperature, respectively, obtained from data in Figure 7.10.*

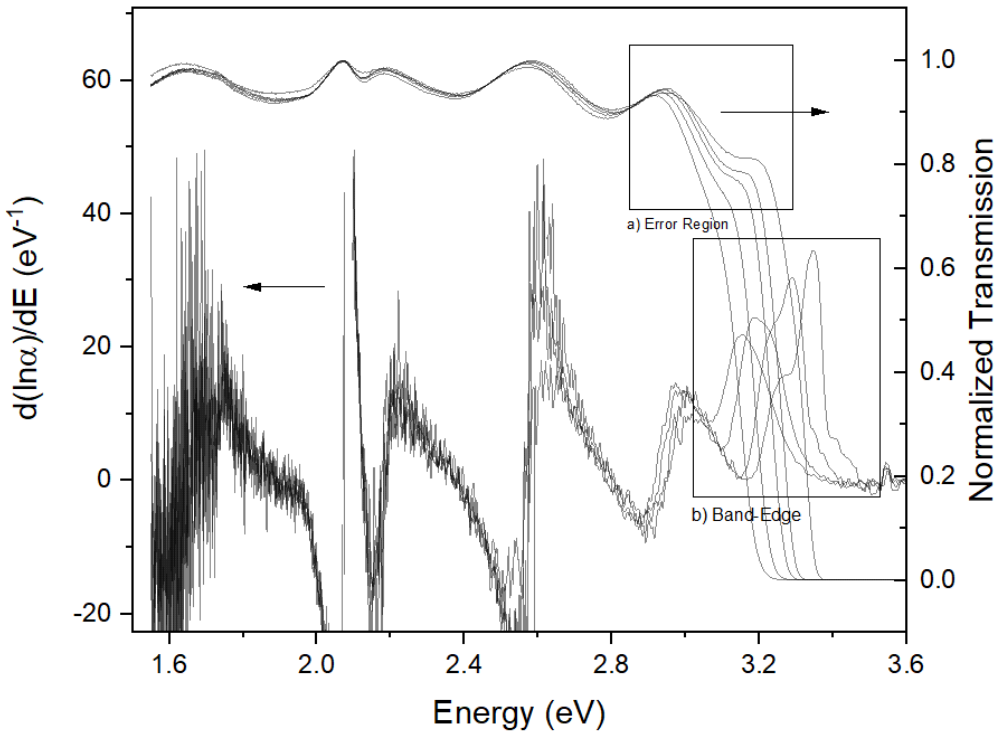


FIGURE 7.12: *The transmission spectra (upper) and corresponding derivative spectra (lower) for 77, 287, 437 and 527 K. Of special interest is a) the error region and b) the Urbach region.*

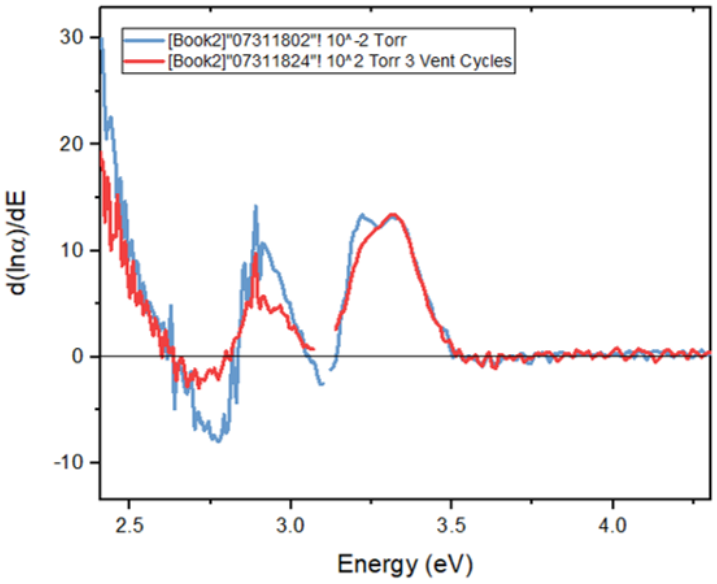


FIGURE 7.13: *Derivative absorption spectra of ZnO with (red) and without (blue) condensation and increased surface scattering. The red curve shows a clear modulation and attenuation of the interference pattern intensity, while the band-edge signal remains unchanged.*

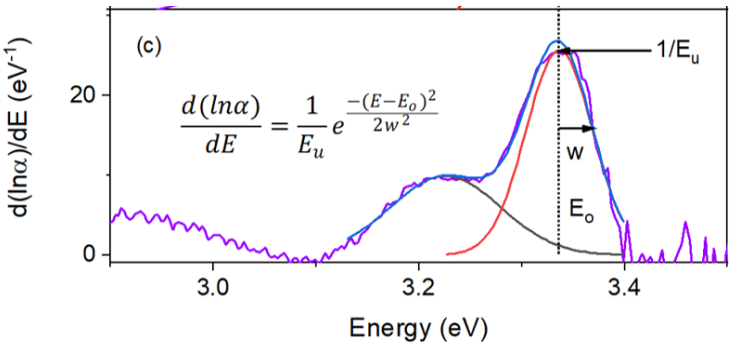


FIGURE 7.14: *Derivative absorption spectra of ZnO with the interference pattern shown in grey and and the band-edge signal shown in red.*

absoprtion spectrum under vacuum at pressure at 10^{-2} Torr at 77 K. Following this, we vented the sorption pump 3 times for equal amounts of time to allow controlled condensation on the thin film then returned the chamber back to pressure at 10^{-2} Torr. We then acquired another spectrum with condensation and increased surface scattering. The results, shown in Fig. 7.13, shows a clear modulation and attenuation of the interference pattern intensity, while the band-edge signal remains unchanged.

7.7 INTERFERENCE PATTERN REMOVAL

The convolution of interference patterns with the spectroscopic signal of semiconducting thin films is a rich and ongoing area of research.[2,4] As shown these patterns are *roughly* described by a sinusoidal function in energy.[5] Thus, taking the energy derivative of a logarithmic sinusoidal function yields the roughly cotangent baseline seen in Fig. 7.8. However, approaching the absorbing region of the band-edge reduces the "contrast"[10] of the interference pattern resulting in an attenuated cotangent. Ultimately this is reasonably approximated by a Gaussians shown in Fig. 7.14. Owing to the reasonable fit of this approximation, the interference pattern can be removed *via* deconvolution analysis, which treats the fringe pattern as a unique and separate signal.

7.8 CONCLUSION

In conclusion, it was determined that the features at the band-edge in the derivative spectra do not originate from interference. The interference pattern, magnified in the derivative spectra, are instead a baseline and are convolved with these features. The nature of the convolution is yet to be characterized. It was also determined that the interference patterns are invariant to temperature across the range 77-527 K. Future work will seek to quantify the effect of and deconvolve interference from spectral fea-

tures in order to accurately extract the Urbach energies and energy-dependent density-of-states from the band-edge. In addition, future work should seek to determine the index of refraction from interference fringes (or extinction coefficient from absorption spectra) and checked against externally referenced indices of refraction. This analytic approach will become important for novel materials wherein the energy-dependent index of refraction is not yet known.

REFERENCES

[1] D. A. Skoog, F. J. Holler, and S. R. Crouch, *Principles of Instrumental Analysis*, 6th ed. (Brooks Cole, Belmont, CA, 2006).

[2] N. J. Harrick, Appl. Spectrosc., AS **31**, 548 (1977).

[3] B. D. Teolis, M. J. Loeffler, U. Raut, M. Famá, and R. A. Baragiola (2007).

[4] M. Milosevic and S. W. King, Journal of Applied Physics **112**, 093514 (2012).

[5] J. C. Manifacier, J. Gasiot, and J. P. Fillard, J. Phys. E: Sci. Instrum. **9**, 1002 (1976).

[6] C. Stelling, C. R. Singh, M. Karg, T. A. F. König, M. Thelakkat, and M. Retsch, Scientific Reports **7**, 1 (2017), number: 1 Publisher: Nature Publishing Group.

[7] F. L. Pedrotti, L. M. Pedrotti, and L. S. Pedrotti, *Introduction to Optics*, 3rd ed. (Pearson, Upper Saddle River, N.J, 2006).

[8] H. Che, *Band gaps by design : Tailoring ZnO based semiconductor alloy films*, Ph.D. thesis (2014).

[9] H. Nagata, A. Yamaguchi, and A. Kawai, Jpn. J. Appl. Phys. **34**, 3754 (1995).

[10] M. V. Klein, *Optics* (Wiley, 1970).

chapter 8

Analysis and Modeling of the Band-Edge

8.1 history of urbach theory

Einstein's general theory of relativity was first proven by measurements of the deflection of light around the Sun (gravitational lensing) during the solar eclipse of May 29, 1916. The deflection of light by the warped space of the sun was to be imprinted on photographic plates. These plates used to document the images played one of many crucial roles in marking the success of the experiment. The deflection was to be measured as accurately as possible by calipers and compared to the predicted deflection made by the relativistic theory.

As research into photographic plates progressed into the 1950s, Franz Urbach discovered in 1953 that a region in the absorption spectra of AgBr had a familiar shape. Urbach noticed that the band-edge, where the material began to absorb light, could be approximately described by an exponential function. The exponential function was known earlier in the optics of photographic layers in 1952[1] and was later applied in 1955 to the analysis of the spectral distribution of photo-plate sensitivity.[2] Instead of the ideal "step-function" form of absorption from hitting the band gap energy, there was instead a softer and more gradual increase. He referred to this general feature as "tailing," since he imagined the step-function as acquiring a "tail."

Urbach further observed that this tailing effect was proportional to the material temperature. At higher temperatures the tailing effect became more pronounced and "stretched" the band-edge over a wider range of energies. When the material is thermally excited, a broader range of energies of light was observe to be absorbed approaching the band-gap. Inversely, when the material is thermally quenched, a

narrower range of energies of light was observed to be absorbed approaching the gap.

Many models have been developed to describe the exponential dependence of absorption on photon energy. Many of the first theoretical papers on the shape of the tail were developed for the alkali-halides. The first theoretical interpretation of the exponential model was attempted by Dexter who proposed that the tail in the absorption spectra was due to the energy of the excited electron state changing due to the lattice deformation by thermal vibrations.[2] At high temperatures the vibrating lattice atoms can be approximated by independent harmonic oscillators which modifies the energy by a linear term. This model can obtain a curve that resembles the Urbach rule but cannot describe the behavior within a wide energy of absorption and temperature ranges.

Building on the work of Dexter, in 1959, Toyazawa[3] proposed that the energy of absorption be modified by a nonlinear deformation term. As the absorption increases with temperature at the low-energy tail, one can assume that the main contribution in the tail formation comes from regions in the crystal that are significantly deformed. One difficulty with this model is connected to the linear term. It gives the main contribution to the center of the band-edge and will give some contribution to the band tail. Toyazawa introduced another assumption that the vibrations responsible for the band tail differ from those contributing to the width of the central part of the band-edge. This was known as the "two-modes" model. The two modes model was further developed in the one-dimensional case considering the Condon approximation. As mentioned in Chap. 2, the approximation states that during an electronic transition, a change from one vibrational energy level to another will be more likely to happen if the two vibrational wave functions overlap more significantly. For these conditions the absorption band shape was modeled as Gaussian at the top and exponential at lower energy assuming that two types of vibrations take part in the formation of

the tail. The main theoretical developments by Toyazawa do not take into account the electronic structure and phonon dispersion in the crystals when considering the shape of the band tail. Therefore, as cited by Kurik,[2] the sensibility of the parameters in the Urbach model to these characteristics remain unclear. However, under different approximations, specific relations between the Urbach parameters and the interaction of excitation with phonons are obtained.

In 1968, Davydov proposed that the tail of the absorption band are due to the quantum transitions from the vibrational sublevels of the crystal lattice to the first electron excitation level.[4] Davydov shows that the absorption shape, both exponential at the tail and Gaussian at the top can be obtained for a single type of lattice vibration. He used a Franck-Condon type model to describe the overlapping vibrations between ground and excited states. Davydov first introduced the idea that the Urbach rule arises from a large number of phonon states filled according the Boltzmann law at a given temperature and interacting with electron excitation. It is assumed in these works that the Urbach rule is seen largely at higher temperatures due to electron coupling to the Boltzmann distribution of phonons. This was the first divergence of the temperature dependent theory between low and high temperatures because of possible absorption structure observed at lower temperatures. The works of Toyazawa and other theoretical treatments such as by Seagall[2] differ in that Toyazawa assumes the Urbach rule is due to the self-trapping of excitons while Seagall does not assume such self-trapping. Still, both theories offer insight into the Urbach rule.

Since the 1950s, many refinements have been made to Urbach's theory, indeed due to its investigation in a host of many diverse semiconductor systems. On the whole, all have built on the fundamental assumption of an approximately exponential absorption edge. The theory which underlies the experimental investigation in this dissertation builds on Urbach's theory of the absorption edge in semiconductors but from a different direction than most. Instead of building on Urbach's theory by

accepting this fundamental assumption as true, we inspect the assumption which underlie Urbach's theory and seek an improvement. In this we offer new grounds on which Urbach theory can rest and, in so doing, allow the theory to capture even more within its scope.

8.2 THE BASIC URBACH THEORY

The Urbach rule models an exponential dependence of the absorption coefficient, $\alpha(E)$, on the photon energy, E, at the band-edge.

$$\alpha(E) = \alpha_o e^{\frac{E-E_o}{E_u}} \tag{8.1}$$

Here, the Urbach energy, E_u, is a "rate parameter" in the exponential increase. α_o is a scaling factor and E_o is the Urbach focus. This parameter in the basic Urbach model is described by Economou[5] as corresponding to the "ideal" or "unperturbed" band-edge (UBE). In essence, E_o is defined as the energy of the band-edge in an ideal structural environment. Following conventional Urbach analysis, E_o is reported to be observed in experiment as the Urbach focus in temperature-dependent absorption spectra and found at higher energies relative to the band-edge as shown in Fig. 8.4 (top).

Urbach analysis requires plotting the absorption coefficient on a logarithmic scale to linearize the Urbach rule:

$$ln\alpha(E) = \frac{1}{E_u}(E - E_o) + ln(\alpha_o) \tag{8.2}$$

This models a linear band-edge with the constant $1/E_u$ as the slope:

$$d(ln\alpha(E))/dE = \frac{1}{E_u} \tag{8.3}$$

Since E_u and the slope are inversely related, a sharp slope at the band-edge corresponds to a small E_u and *vice versa*.

It has been shown by experiment that disorder of many kinds in a material lead to a greater degree of tailing. This is seen as broadening in the band-edge and "softening" or reduction in the slope.

In other words, E_u measures tailing at the band-edge in absorption spectra of semiconductors and quantifies two closely related phenomena:

1. the general degree of disorder in the material, and

2. the variability in the local potential.

As mentioned in Chap. 1, the development of the Urbach rule and subsequent Urbach analysis stems from the assumption of the exponential dependence of the absorption coefficient on photon energy at the band-edge. This trend has been reportedly identified in direct- and indirect-gap semiconductors, crystalline and amorphous materials, in and nanodimensional structures, ferroics, superionic conductors and organics.[6–34] *The Urbach rule has been applied in so many instances that it is generally accepted as a universal rule in semiconductor physics.*

Despite this widely accepted universality, the following reasons indicate that the Urbach rule only *approximately* describes the band-edge.

If the band-edge is exponential, we expect to see strictly linear band-edges in logarithmic plots. In general, however, the linearity of the band-edge in logarithmic absorption plots is not universal. It was noted by Dow and Redfield in 1972 that experimental observations of the full Urbach rule are very rare.[35] As mentioned

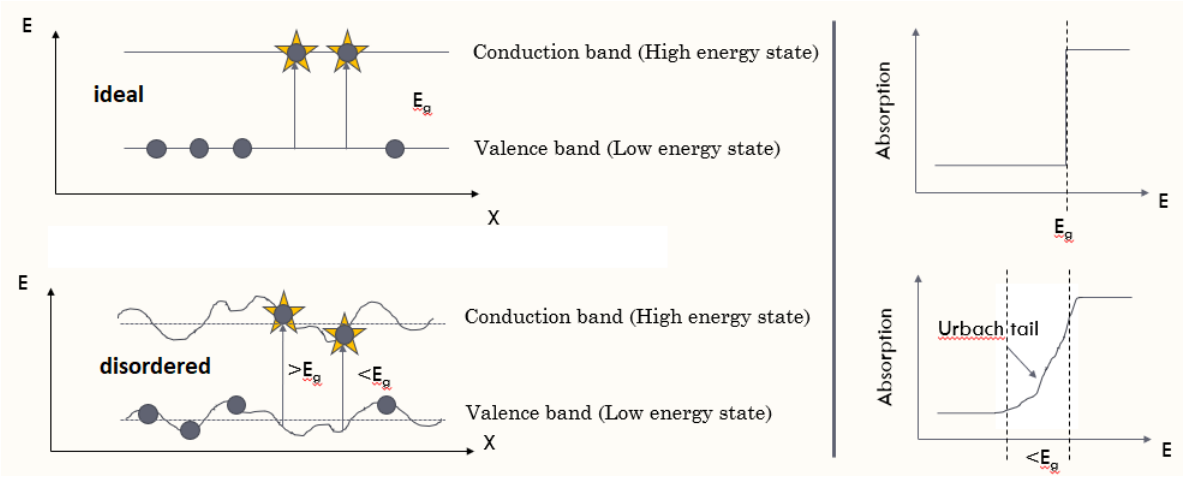


FIGURE 8.1: *(Top left) Transitions from the valence band to the conduction band in ideal semiconductor systems occur at the same energy in a perfectly crystalline lattice. (Top right) This appears as a sharp transition to absorption at the bandgap, E_g, in absorption spectra. (Bottom left) Transitions from the valence band to the conduction band in disordered semiconductor systems occur at variable energies in due the fluctuating local energy environment caused by defects. (Bottom Right) This appears as a soft transition to absorption at the band-edge, in absorption spectra.*

in Chap. 1, in many cases, especially at low temperature, the band-edges have structure—they are nonlinear.

In his original letter, Urbach states that the logarithmic plot only "*approaches* a straight line." He states further, "In the region of the band-edge, the relation $dlog\epsilon/dv = -1/kT$ is *closely approached*."

8.3 URBACH ANALYSIS IN THE LITERATURE

The Urbach rule has been developed extensively through experiment and theory. Building on this substantial progress, we fulfill the need for a nonlinear model evident in the literature.[36–42] A survey of the literature indicates challenges to application of the linear model and analysis of the Urbach energy parameter from experimental spectra. There are many examples throughout the literature of very loosely defined Urbach fitting regions. In some cases, only a small range of the band-edge is used to obtain Urbach slopes.[38,39,42] An example of this is shown in Fig. 8.4 (top). In one case, Rai explicitly notes that the absorption data is complex[43] and, like other authors, select a small but consistent range of the absorption coefficients to obtain the Urbach slopes for spectra taken at different temperatures.[36,38] In other cases, the fitting ranges or regions change as a function of temperature or alloy concentration.[40,41,44] In fact, Sadigh *et al.* showed, using density functional theory (DFT) calculations of the Urbach tail in silica glass, that the upper and lower limits of the Urbach region in the band-edge shows significant temperature dependence.[45] Finally, in the most extreme fitting cases, the model is applied to highly nonlinear band-edges as in the case of crystalline Si (c-Si).[8] To our knowledge, only one instance of quantifying linearity at the logarithmic band-edge has been stated when applying the linear model.[27]

Challenges to standardization of Urbach analysis is likely rooted in the fact that the Urbach rule does not describe the full range of the logarithmic band-edge in

many cases. In these cases, the Urbach rule only approximates an energy dependent exponential increase in the absorption coefficient at the band-edge and that the actual functional form is more complicated. This dissertation was in part motivated from a survey of the current literature on Urbach analysis and in part by the need for a more complete model of tailing at the band-edge.

8.4 DIFFERENTIAL APPROACH

Using differential and deconvolution analysis, we uncover the detailed form of the components of the complex nonlinear band-edge. From the results of this analysis, we develop a new model of the band-edge using a few key physical principles and four free parameters. We show that nonlinear band-edges arise as a result of the inherent nature of broadening at the band-edge. Nonlinear band-edges may acquire additional fine structure from a convolution of the absorption edge signal with artifacts such as other absorption edge signals or interference fringe patterns. The convolution of multiple absorption edges will be illustrated using the cubic $Mg_{0.75}Zn_{0.25}O$ (cubic and wurtzite mixed-phase) thin film alloy at 292 K and c-Si at 299 K while the convolution of absorption edges with interference fringe patterns will be illustrated using ZnO thin film at 77 K and 377 K and in the wurtzite $Mg_{0.75}Zn_{0.25}O$ (cubic and wurtzite mixed-phase) thin film alloy at 292 K.

Furthermore, we will show that the basic Urbach model is a limiting case of our model, giving insight into the theory of normally distributed varying local potentials.

8.5 MODEL BASIS

The nucleus of model development is the deconvolved peaks of $d(ln\alpha(E))/dE$, which were modeled accurately by Gaussian distributions. We have already discussed in Chap. 7 that at least one feature in the derivative spectra originates from thin film

interference. Additionally, we showed that the fringe feature could be approximated well by a Gaussian. For the band-ege signal, the probability density function (PDF) of the normal distribution was approximated by the PDF of the logistic distribution given generally by the following:

$$f(x) = \frac{ae^{\frac{-(x-x_o)}{s}}}{s(1+e^{\frac{-(x-x_o)}{s}})^2} \tag{8.4}$$

Eq. 8.4 is plotted in Fig. 8.2b) for various values of x_o and s, $a = 1$.

a is the integrated area of the distribution, x_o is the center location of the distribution on the x-axis and s determines the spread of the distribution. Eq. 4 is plotted in Fig. 4(b) for various values of x_o and s; $a = 1$.

Since $ln\alpha(E)$ is the integral of $d(ln\alpha(E))/dE$, $ln\alpha(E)$ was found to be modeled by the integral of the logistic PDF. The integral of the PDF is known as the cumulative distribution function or CDF and is given generally by the following:

$$\int f(x)dx = F(x) = \frac{a}{1+e^{-\frac{(x-x_o)}{s}}} + F_o \tag{8.5}$$

Eq. 8.5 is plotted in Fig. 8.2a) for various values of x_o and s, $a = 1$.

Here the parameters are the same but display graphically in a different way since we have taken the integral. a is the maximum value of the curve, x_o is the inflection point of the curve and s is inversely proportional to the steepness of the curve. F_o is an arbitrary constant of integration and corresponds to a shift on the y-axis. Eq. 5 is plotted in Fig. 4(a) for various values of x_o and s; $a = 1$, $F_o = 0$.

Having a model for $ln\alpha(E)$, obtaining a model for $\alpha(E)$ is simple, it is given generally by the following:

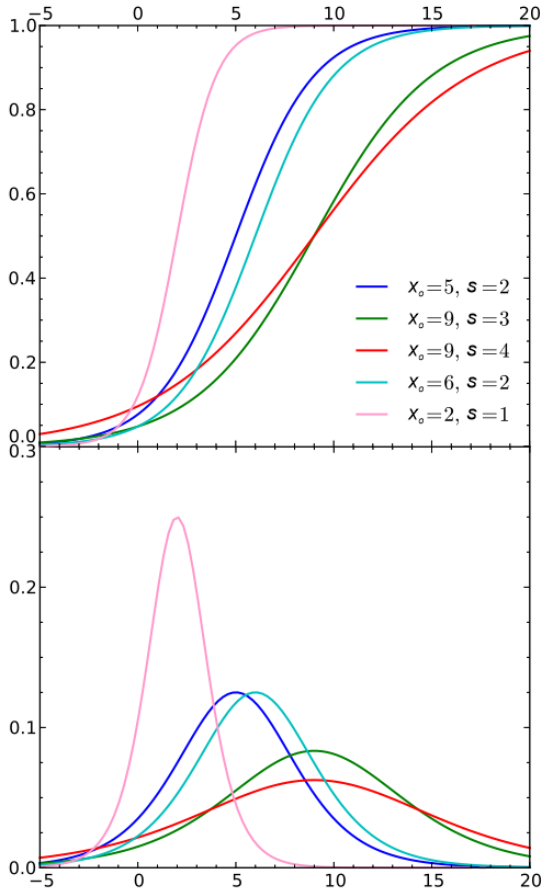


FIGURE 8.2: *a) General forms of Cumulative Distribution Functions as given by Eq. 8.5 for various values of x_o and s, $a = 1$. b) General forms of Probability Density Functions as given by Eq. 8.4 for various values of x_o and s, $a = 1$.*

$$G(x) = e^{F(x)} = F_o e^{\frac{a}{1+e^{-\frac{(x-x_0)}{s}}}} \tag{8.6}$$

All parameters here are as previously described even though the shape of the curve is different.

8.6 MODEL DEVELOPMENT

8.6.1 *Differential Analysis & Deconvolution of the Absorption-edge*

The band-edge is often modeled as a complicated function of many absorption regimes.[5] In general, different energy ranges of the band-edge are described piece-wise and mutually exclusive by different functional forms. In order of increasing energy some notable regions are: the Urbach region, Halperin-Lax region, Anderson region and the Tauc region.[5]

In order to accurately study and separate band-edge regions, we have employed differential analysis on band-edges with fine structure in $ln(\alpha(E)/\alpha_o)$ to obtain $d(ln\alpha(E))/dE$ ($ln\alpha_o$ drops out as a constant). The analysis includes a ZnO thin film at 77 K shown in Fig. 8.5, $Mg_{0.75}Zn_{0.25}O$ (wurtzite and cubic mixed-phase) thin film alloy shown in Figs. 8.6, 8.7 and 8.8, and c-Si shown in Fig. 8.9. The results are illustrated by the green solid curves in Figs. 8.5(c), 8.7(c), 8.8(c) and 8.9(b). We have followed the analysis of Cody[8] who used $d(ln\alpha(E))/dE$ to analyze the fine structure of the band-edge in c-Si. Fig. 8.9(b) includes our replication of the results of Cody's differential analysis.

Differential analysis is widely employed as an analytical tool in spectroscopic techniques such as absorption and reflectance.[46–50] The reference text Spectrochemical Analysis gives an extensive overview of its many applications.[51] Here, differential analysis is useful for quantitatively unmasking the characteristic slope behavior of

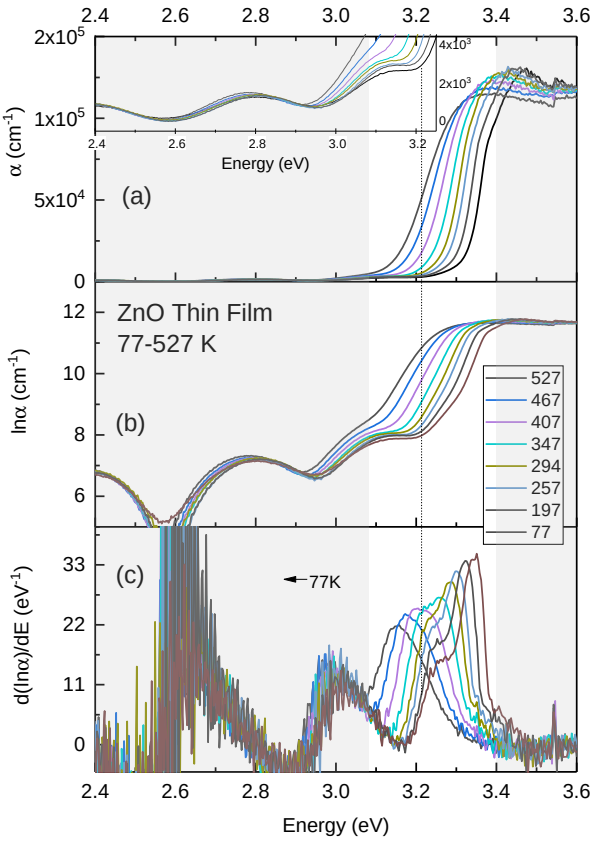


FIGURE 8.3: *(a) Absorption coefficient, $\alpha(E)$ (b) logarithmic absorption coefficient, $ln(\alpha(E)/\alpha_o)$ and (c) the energy derivative of the logarithmic absorption coefficient, $d(ln\alpha(E))/dE$ at the band-edge for ZnO thin film from 77-527 K. The expansion in (a) and the vertical dotted line qualitatively show the interference fringe pattern persisting into the band-edge. The non-shaded region indicates the general energy range used for analysis in this work ($\approx 3.10 - 3.40$ eV).*

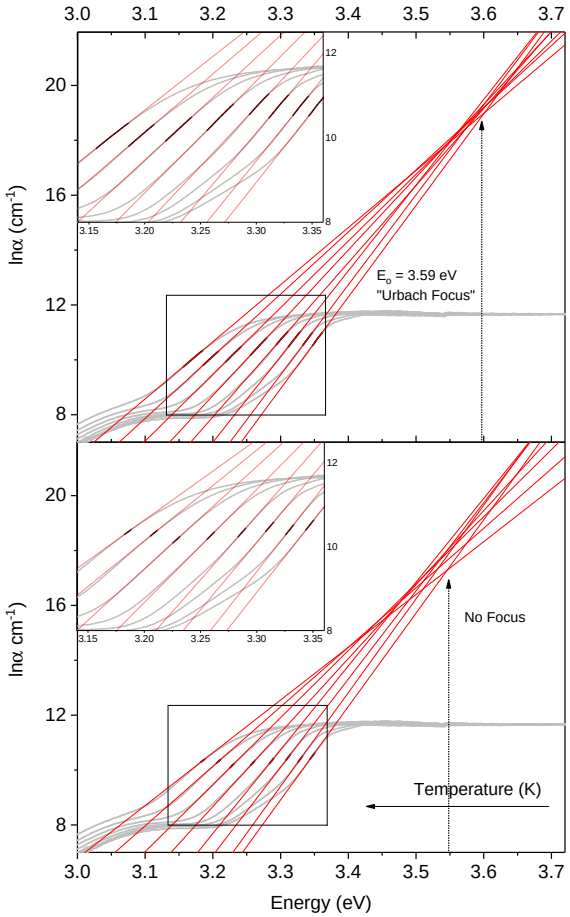


FIGURE 8.4: *(Top) Conventional Urbach analysis of temperature-dependent UV-Vis absorption spectra for ZnO thin film from 77 – 527 K. Shown is the "Urbach focus" at 3.59 eV which indicates the point of convergence of the Urbach slopes (red). The expansion shows the limited range and varying regions of the band-edge analyzed (shown by the bold line sections) due to prioritizing the Urbach focus in the basic Urbach model. (Bottom) Differential Urbach Analysis on temperature-dependent absorption spectra for ZnO thin film from 77 – 527 K. The tangents do not converge. The expansion shows more clearly the tangents at the inflection points of the curves obtained from the derivative spectra shown in Fig. 8.3(c).*

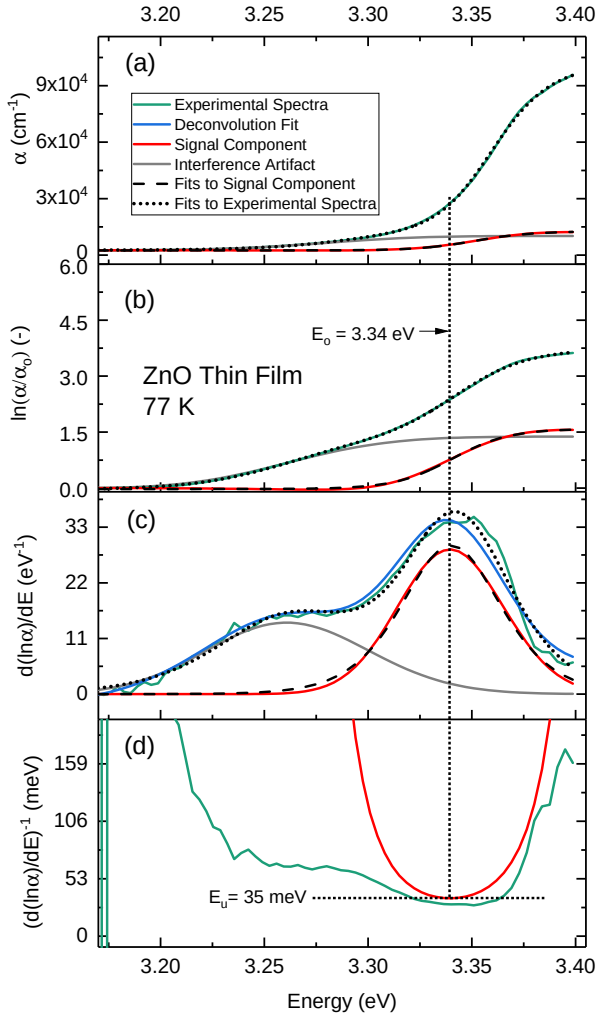


FIGURE 8.5: *(a) Absorption coefficient, $\alpha(E)$,(b) $ln(\alpha(E)/\alpha_o)$, (c) energy derivative of $ln(\alpha(E)/\alpha_o)$ and (d) inverse plot of $d(ln\alpha(E))/dE$ for a ZnO thin film at 77 K. The green solid lines show the experimental spectra. The grey and red solid lines show the analyzed components of the spectra. The black dashed lines are the fits to Eqs. (8.9), (8.8) and (8.7), respectively, and the black dotted lines are the fits to Eqs. (8.15), (8.14) and (8.13), respectively. The fitted parameters are shown in Table 8.1. The solid blue line in (c) is the fit from deconvolution analysis. The Urbach energy, E_u, is shown in (d) corresponding to the value of the inverse slope at the inflection point, E_o. Comparison of this value with E_u calculated from Eq. (8.17) is shown in Table 8.1.*

the components of the complex band-edge function. To this end, we further applied deconvolution analysis to the differential spectra in order to fully separate the components. The results of deconvolution are shown by the grey, red and blue solid curves in Figs. 8.5(c), 8.7(c), 8.8(c) and 8.9(b).

As will be treated in depth shortly, the results indicate that

- the differential absorption signal is well modeled by a Gaussian in all cases and
- this signal may be convolved with other artifacts.

In order to successfully apply deconvolution analysis, we found that the artifacts are best treated on a case-by-case basis. Analysis beyond the scope of this work indicated that the low-energy artifact of ZnO at 77 K in Fig. 8.5(c) and wurtzite $Mg_{0.75}Zn_{0.25}O$ in Fig. 8.7(c) originate from thin film interference. In cubic $Mg_{0.75}Zn_{0.25}O$ shown in Fig. 8.8(c), the convolving artifact (grey solid curve) is the lower-energy wurtzite absorption baseline. In contrast, as shown in Fig. 8.9(b), the fine structure in c-Si does not originate from artifacts but, as stated by Cody,[8] from the two indirect edges of c-Si.

The interference artifacts shown by the grey solid curve in Fig. 8.5(c) for ZnO and Fig. 8.7(c) for wurtzite $Mg_{0.75}Zn_{0.25}O$ were approximated by Gaussians. It must be stressed that, unlike the absorption signal wherein the Gaussian has physical meaning, there is no physical significance, only a good approximation of form.

Conversely, the artifact in the cubic $Mg_{0.75}Zn_{0.25}O$ spectra shown in Fig. 8.8(c) does not originate from thin film interference because the cubic signal lies well beyond the energy range of interference effects in the deeply absorbing region. This is highlighted in Fig. 8.6. This presents the unique opportunity to test our model against an already interference-free band-edge. Nevertheless, this signal was found to be shifted by the absorption baseline from the wurtzite signal as shown in Fig. 6(c) by the grey solid curve.

Deconvolution of the absorption signal in the derivative spectra highlights a key point. The results show that the step-wise, mutually exclusive picture of the band-edge can be more simply represented by a single function—the Gaussian. Furthermore, the convolution of the Gaussian signal—hereafter denoted by $d(ln\alpha_s(E))/dE$—with artifacts of varying origins fits the experimental data in $d(ln\alpha(E))/dE$ to very good accuracy as shown by the solid blue curves in Figs. 8.5(c), 8.7(c), 8.8(c) and 8.9(b).

While unexpected, it is not too surprising that statistical theory can be applied to the derivative of the band-edge. In fact, Pankove[52] states that the absorption coefficient is proportional to the cumulative or integral probability of transitions between states. It is known in statistics that the derivative of a cumulative probability function (CDF) is a probability density function (PDF). Therefore, we propose that the signal component $d(ln\alpha_s(E))/dE$ represents a probability density and is the experimental observation of the normally distributed variability in the local potential due to the random distribution of material defects as predicted by semiconductor theory.[5,35,53–56] One instance in the literature has recently appeared simulating the band-edge using a Gaussian PDF.[57] Another has also suggested a band-edge modeled by a Gaussian distribution of band gap energies centered at a mean value and characterized by a standard deviation.[58] We further propose that the probability density function represented here corresponds to the density of states (DOS) at the band-edge. The units of the derivative plots are (eV^{-1}), consistent with units of DOS.

Continuing with model development, if $d(ln\alpha_s(E))/dE$ is Gaussian, its integral $ln\alpha_s(E)$ is given by an error function. This, however, cannot be represented using elementary functions. We instead represent the normal distribution of the analyzed signal component $d(ln\alpha_s(E))/dE$ to good approximation using the simpler form of the probability density function (PDF) of the logistic distribution:

$$\frac{d(ln\alpha_s(E))}{dE} = \frac{a_c e^{\frac{-(E-E_o)}{\gamma}}}{\gamma\left(1+e^{\frac{-(E-E_o)}{\gamma}}\right)^2} \tag{8.7}$$

In Eq. (8.7), a_c represents the total population of photons absorbed or relative *photon capacity* and corresponds to the integral area of the signal peaks shown in $d(ln\alpha(E))/dE$ of Fig. 8.5(c), 8.7(c), 8.8(c) and 8.9(b). E_o is the energy of the UBE and is found at the peak center. This corresponds to the photon energy with the maximum probability density of absorption. Finally, γ represents the spread or variability in the energy about the UBE as a result of the variability in the local potential. This is reflected in the width of the distribution. The results of fitting Eq. (8.7) to the component data is shown by the black dashed line(s) in $d(ln\alpha(E))/dE$ of Figs. 8.5(c), 8.7(c), 8.8(c) and 8.9(b). The fitted parameters are given in Table 8.1.

As early as 1845, logistic distributions were employed as a model for normal distributions in science and statistics specifically in fields such as growth curve analysis and population dynamics.[59] The logistic function now finds applications in a wide range of fields, including biology, ecology, chemistry, economics, geoscience, mathematical psychology, sociology, political science, financial mathematics, hydrology, bioassay and quantal response data problems.[60–68]

Again, the central assumption in all these applications is that the phenomenon modeled is a normally distributed random variable. In our case, the normally distributed random variable is the absorption of a photon of energy, E, by an electron (corresponding to the local potential) and subsequent occupation of an excited state. In physics, we model the statistical distribution of fermions occupying the energy states of a system in thermal equilibrium via the Fermi-Dirac distribution. In fact, the Fermi-Dirac distribution is also of the logistic form.

TABLE 8.1: *Fitted parameters to Eq. (8.7), Eq. (8.8) and Eq. (8.9) for analyzed band-edge signal components and Eq. (8.13), Eq. (8.14) and Eq. (8.15) for as-is experimental spectra of ZnO thin film, mixed-phase $Mg_{0.75}Zn_{0.25}O$ thin film and c-Si (Cody). $\langle x \rangle$ and $\pm\sigma$ indicate the mean value and standard deviation of parameters obtained for all analyzed signal components and as-is experimental spectra. The final two columns compare E_u derived from our model parameters and E_u^* extracted analytically from the spectra via the inverse slope method.*

Material	Eq.	Signal	α_o (cm $^{-1}$)	a_c (-)	E_o (eV)	γ (meV)	E_u (meV)	E_u^* (meV)
	8.7	1	-	1.88	3.34	16.04	34.13	34.97
	8.8	1	-	1.58	3.34	13.11	33.19	34.97
ZnO (77 K)	8.9	1	$2.6x10^3$	1.62	3.34	14.18	35.01	34.97
	8.13	1	-	2.37	3.34	17.31	29.21	34.97
	8.14	1	-	2.57	3.34	18.23	28.37	34.97
	8.15	1	$2.7x10^3$	2.08	3.35	14.84	28.60	34.97
	$\langle x \rangle$			2.02	3.34	15.62	31.42	
	$\pm\sigma$			0.37	0.004	1.78	2.75	
	8.7	1	-	1.75	3.25	21.30	48.69	50.00
	8.8	1	-	1.70	3.26	20.13	47.36	50.00
ZnO (377 K)	8.9	1	$3.4x10^3$	1.74	3.26	21.15	48.62	50.00
	8.13	1	-	2.27	3.25	25.22	44.44	50.00
	8.14	1	-	2.29	3.25	24.11	42.11	50.00
	8.15	1	$3.4x10^3$	2.48	3.25	25.72	41.48	50.00
	$\langle x \rangle$			2.04	3.25	22.94	45.45	
	$\pm\sigma$			0.32	0.005	2.16	2.95	
	8.7	1	-	1.42	3.43	25.58	72.06	74.18
	8.8	1	-	1.39	3.44	24.33	70.01	74.18
MgZnO (75%) (292 K)	8.9	1	$1.5x10^5$	1.40	3.44	24.80	70.86	74.18
	8.13	1	-	1.11	3.44	23.75	85.58	74.18
	8.14	1	-	1.31	3.44	23.78	72.61	74.18
	8.15	1	$1.5x10^5$	0.95	3.45	21.76	91.62	74.18
	$\langle x \rangle$			1.26	3.44	24.00	77.12	
	$\pm\sigma$			0.17	0.006	1.18	8.34	
	8.7	2	-	1.72	6.04	113.29	263.00	274.00
	8.8	2	-	1.67	6.04	108.22	259.19	274.00

TABLE 8.2: *Table 8.1 continued.*

Material	Eq.	Signal	α_o (cm $^{-1}$)	a_c (-)	E_o (eV)	γ (meV)	E_u (meV)	E_u^* (meV)
MgZnO (75%) (292 K)	8.9	2	2.0x10^6	1.76	6.04	117.67	267.43	274.00
	7	2	-	1.93	6.04	125.17	259.42	274.00
	8	2	-	3.07	6.1	164.82	257.40	215.08
	9	2	2.0x10^6	2.02	6.04	129.00	255.45	274.00
	$\langle x \rangle$			2.03	6.05	126.36	260.32	
	$\pm\sigma$			0.48	0.022	18.54	3.92	
	8.7	1	-	2.07	1.03	3.38	6.53	6.70
	8.8	1	-	1.85	1.03	2.85	6.16	6.70
c-Si (299 K)	8.13	1	-	2.02	1.03	3.33	6.59	6.70
	8.14	1	-	1.84	1.03	2.72	5.91	6.70
	$\langle x \rangle$			1.95	1.03	3.07	6.30	
	$\pm\sigma$			0.1	0	0.29	0.28	
	8.7	2	-	2.22	1.05	3.68	6.63	6.60
	8.8	2	-	2.09	1.05	3.38	6.47	6.60[6]
c-Si (299 K)	8.13	2	-	2.19	1.05	3.63	6.63	6.60
	8.14	2	-	2.24	1.05	3.53	6.30	6.60
	$\langle x \rangle$			2.19	1.05	3.56	6.51	
	$\pm\sigma$			0.06	0	0.11	0.14	

Since the phenomenon being modeled obeys Fermi-Dirac statistics, it is fitting that we also use the logistic distribution as the basis for our model.

To obtain the more familiar form of this distribution from the PDF in $d(ln\alpha(E))/dE$, we take the integral of Eq. (8.7) to obtain the CDF. This models $ln\alpha_s(E)$ as:

$$ln(\alpha_s(E)/\alpha_o) = \frac{a_c}{1+e^{-\frac{(E-E_o)}{\gamma}}} = f_s(E) \tag{8.8}$$

Here the parameters are the same as in Eq. (8.7) but manifest graphically in a different way since we have taken the integral. α_o is introduced as a constant of integration and functions as a normalization factor. It is obtained from the initial value of the absorption coefficient in the range selected for analysis. It is useful to define the normalized signal function in Eq. (8.8) $\left[a_c/1+e^{-\frac{(E-E_o)}{\gamma}}\right]$ as $f_s(E)$. a_c again represents the total population of photons absorbed or relative *photon capacity* and corresponds to the maximum of the red (and grey for c-Si) solid curves shown in $ln(\alpha(E)/\alpha_o)$ in Figs. 8.5(b), 8.7(b), 8.8(b) and 8.9(a). E_o is the energy of the UBE and corresponds to the greatest rate of increase in absorption. E_o is found at the inflection point of the red (and grey for c-Si) solid curves in $ln(\alpha(E)/\alpha_o)$, marked by the vertical dotted line(s) of Figs. 8.5, 8.7, 8.8 and 8.9. Finally, γ represents the spread or variability in the energy about the UBE as a result of the variability in the local potential. This is reflected in the *steepness* of the red (and grey for c-Si) solid curves in $ln(\alpha(E)/\alpha_o)$ of Figs. 8.5(b), 8.7(b), 8.8(b) and 8.9(a). Here, the parameter γ makes a more concrete connection between the variability in the local potential illustrated by the spread of $d(ln\alpha_s(E))/dE$ and the phenomenon of “tailing" in the absorption spectra. The results of fitting Eq. (8.8) to the signal component data is shown by the black dashed lines in $ln(\alpha(E)/\alpha_o)$ of Figs. 8.5(b), 8.7(b), 8.8(b) and 8.9(a). The fitted parameters are given in Table 8.1.

In the following three paragraphs we discuss in more detail the physical significance of applying statistical theory to the modeling of the experimental data.

The logistic CDF $ln\alpha_s(E)$ models the probability of absorbing a photon of a given energy. *Cumulative* in "CDF" implies that the value of the absorption coefficient at a given energy is a cumulative measurement of transitions occurring both *at and less than* the given energy. Energy bands in semiconducting solids perhaps allow for this cumulative effect in contrast to systems such as absorbing molecules which do not form energy bands.

To gain perspective on the probability of absorption, we consider the derivative of the CDF, or the *rate of change in the probability of absorption*. This yields the PDF, which shows the probability per unit energy of absorption. In essence, the PDF reveals the discrete DOS contributions to the cumulative value measured in the absorption coefficient. It must be stressed that, unlike the CDF, the PDF itself does not correspond to a physically "real" probability. Integrating the PDF gives the physically observable probability of absorption.

We see in the PDFs of $d(ln\alpha(E))/dE$ in Figs. 8.5(c), 8.7(c), 8.8(c) and 8.9(b), the energy of maximum probability density of absorption lies at the center of the normal distribution, corresponding to the UBE, E_o. Integrating the PDF yields the maximum rate of increase in the probability of absorption approaching E_o, found at the inflection point in the CDF. In general, the integral population of photons absorbed is at first low preceding E_o, increases significantly approaching E_o, then gradually saturates as energy increases past E_o.

Now that functional forms for $d(ln\alpha_s(E))/dE$ and $ln(\alpha_s(E)/\alpha_o)$ have been described, obtaining the functional form of the absorption coefficient signal $\alpha_s(E)$ is relatively straightforward. Using Eq. (8.8), we raise $ln(\alpha_s(E)/\alpha_o)$ to the exponent of e and obtain:

$$\alpha_s(E) = \alpha_o e^{a_c / \left(1+e^{-\frac{(E-E_o)}{\gamma}}\right)} = \alpha_o F_s(E) \tag{8.9}$$

All parameters in Eq. (8.9) are the same as previously described. It is useful to define the normalized signal function in Eq. (8.9) $\left[e^{a_c / \left(1+e^{-\frac{(E-E_o)}{\gamma}}\right)}\right]$ as $F_s(E)$.

We now summarize the premises on which the model stands:

1. The absorption coefficient $\alpha_s(E)$ (or $ln(\alpha_s(E)/\alpha_o)$ on a log scale) is proportional to the cumulative probability of absorption.

2. The energy derivative of the cumulative probability, $d(ln\alpha_s(E))/dE$, represents the probability density.

3. The UBE corresponds to the maximum probability density of absorption in the PDF or the maximum rate of change in the probability of absorption in the CDF.

4. The variability in the energy about the UBE is a result of the variability in the local potential due to the random distribution of defects in the material.

It must be stressed in closing that the model presently discussed was derived for the signal components of the absorption spectra analyzed via deconvolution. As shown, the results of analysis illustrate that the experimental spectra may represent a convolution of this signal with any arbitrary artifact. To model the experimental data *as is* (meaning not analyzed via deconvolution) requires rigorous treatment, discussed in the following section.

8.6.2 *Piece-wise Reconstruction of the Complex band-edge*

So far, we have only discussed fitting the deconvolved signal components of the experimental spectra. It was shown in Sec. 8.6.1 that the analyzed signal, $d(ln\alpha_s(E))/dE$,

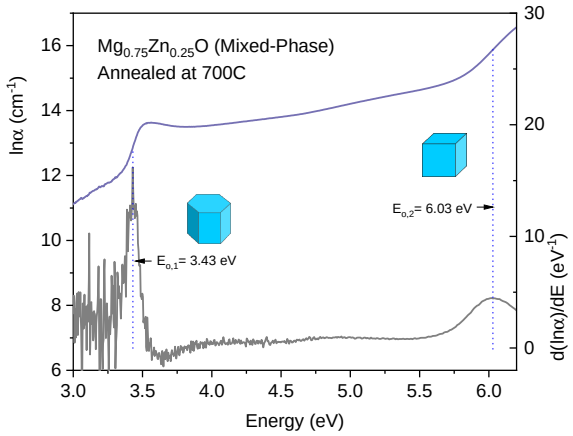


FIGURE 8.6: *(a) $ln(\alpha(E)$ (top) and $d(ln\alpha(E))/dE$ (bottom) for mixed-phase $Mg_{0.75}Zn_{0.25}O$ thin film at 292 K. The wurtzite band-edge is shown by the vertical dotted line at 3.43 eV and the cubic band-edge is shown by the vertical dotted line at 6.03 eV. In contrast to the cubic band-edge, the wurtzite band-edge is seen to be convolved with the interference pattern.*

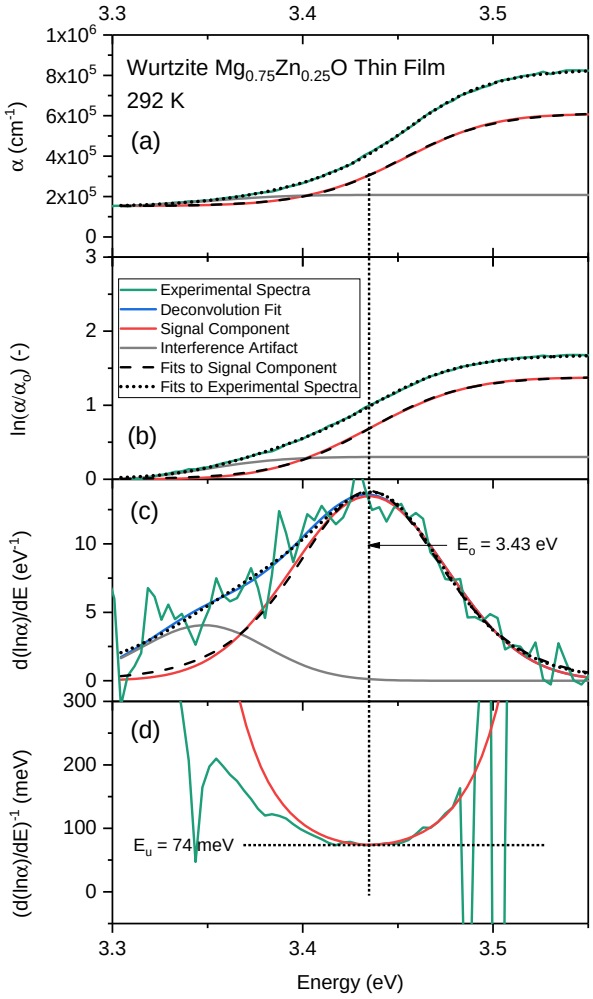


FIGURE 8.7: *(a) Absorption coefficient, $\alpha(E)$, (b) $ln(\alpha(E)/\alpha_o)$, (c) energy derivative of $ln(\alpha(E)/\alpha_o)$ and (d) inverse plot of $d(ln\alpha(E))/dE$ for wurtzite $Mg_{0.75}Zn_{0.25}O$ (cubic and wurtzite mixed-phase) thin film at 292 K. The green solid lines show the experimental spectra. The grey and red solid lines show the analyzed components of the spectra. The black dashed lines are the fits to Eqs. (8.9), (8.8) and (8.7), respectively, and the black dotted lines are the fits to Eqs. (8.15), (8.14) and (8.13), respectively. The fitted parameters are shown in Table 8.1. The solid blue line in (c) is the fit from deconvolution analysis. The Urbach energy, E_u, is shown in (d) corresponding to the value of the inverse slope at the inflection point, E_o. Comparison of this value with E_u calculated from Eq. (8.17) is shown in Table 8.1.*

was fitted to good accuracy by the normal distribution represented by the logistic PDF. Furthermore, the integrated signal, $ln(\alpha_s(E)/\alpha_o)$, was fitted to good accuracy by the logistic CDF.

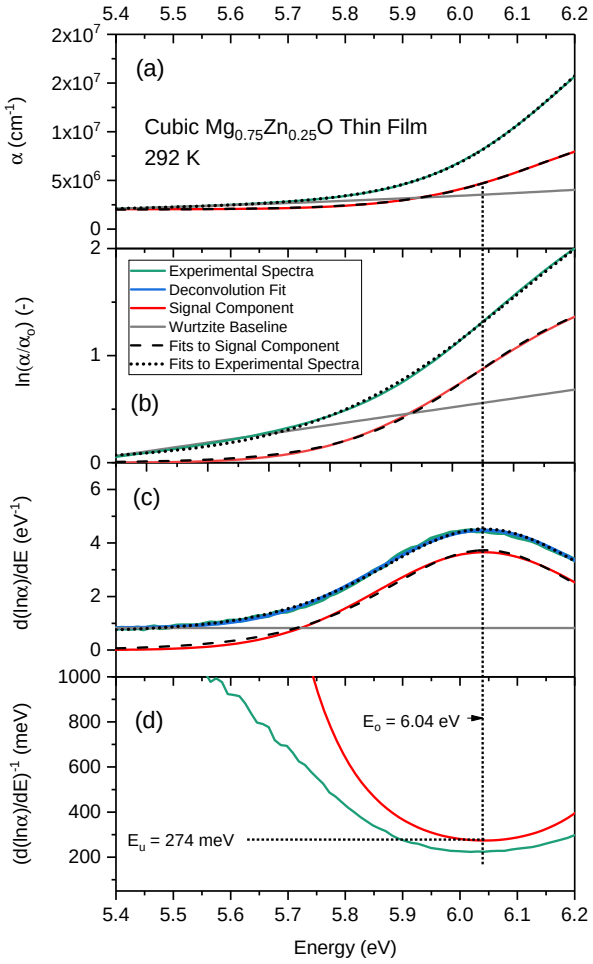


FIGURE 8.8: *(a) Absorption coefficient, $\alpha(E)$, (b) $ln(\alpha(E)/\alpha_o)$, (c) energy derivative of $ln(\alpha(E)/\alpha_o)$ and (d) inverse plot of $d(ln\alpha(E))/dE$ for cubic $Mg_{0.75}Zn_{0.25}O$ (cubic and wurtzite mixed-phase) thin film at 292 K. The green solid lines show the experimental spectra. The grey and red solid lines show the analyzed components of the spectra. The black dashed lines are the fits to Eqs. (8.9), (8.8) and (8.7), respectively, and the black dotted lines are the fits to the experimental spectra. The fitted parameters are shown in Table 8.1. The solid blue line in (c) is the fit from deconvolution analysis. The Urbach energy, E_u, is shown in (d) corresponding to the value of the inverse slope at the inflection point, E_o. Comparison of this value with E_u calculated from Eq. (8.17) is shown in Table 8.1.*

As discussed in Sec. 8.6.1, the signals at the band-edge were found to be convolved with artifacts of varying origins. Further analysis indicated that the as-is experimental spectra in $d(ln\alpha(E))/dE$ and $ln(\alpha(E)/\alpha_o)$ can be fitted from a linear combination of the signal(s) and artifact. This is described by Eq. (8.10) and Eq. (8.11) below. This surprisingly simple result is facilitated by logarithmic and subsequent derivative transformation of the data. Inversely, this implies that the absorption coefficient can be modeled as a product of the signal(s) and artifact, described by Eq. (8.12). This section will expand on the mathematical details of this general result.

To summarize the crucial Eqs. of piece-wise reconstruction:

$$\frac{d(ln\alpha(E))}{dE} = \sum_{s,a} \frac{d(f_{s,a}(E))}{dE} \tag{8.10}$$

$$ln(\alpha(E)/\alpha_o) = \sum_{s,a} f_{s,a}(E) \tag{8.11}$$

$$\alpha(E) = \alpha_o \prod_{s,a} F_{s,a}(E) \tag{8.12}$$

In the equations above, $d(f_s(E))/dE$ is given by Eq. (8.7) and $f_s(E)$ and $F_s(E)$ are given by Eqs. (8.8) and (8.9), respectively. These represent the signal components of the experimental spectra.

Furthermore, $d((f_a(E))/dE$, $f_a(E)$ and $F_a(E)$ represent signal modification functions arising from the effects of artifacts.

As shown in Fig. 8.5(c) for ZnO and Fig. 8.7(c) for wurtzite $Mg_{0.75}Zn_{0.25}O$, the form of both $d(f_s(E))/dE$ and $d(f_a(E))/dE$ were represented by Eq. (8.7). Additionally,

while no artifacts were found in c-Si, both indirect edges in $d(ln\alpha(E))/dE$ shown in Fig. 8.9(b) were found to be represented by Eq. (8.7). Therefore, in these cases the as-is experimental spectra is given generally by the following forms for any arbitrary number, n, of band-edge components corresponding to PDFs of the logistic function or interference patterns approximated by the same.

$$\frac{d(ln\alpha(E))}{dE} = \sum_{i=1}^{n} \frac{a_{c,i} e^{\frac{-(E-E_{o,i})}{\gamma_i}}}{\gamma_i \left(1 + e^{\frac{-(E-E_{o,i})}{\gamma_i}}\right)^2} \tag{8.13}$$

$$ln(\alpha(E)/\alpha_o) = \sum_{i=1}^{n} \frac{a_{c,i}}{1 + e^{-\frac{(E-E_{o,i})}{\gamma_i}}} \tag{8.14}$$

$$\alpha(E) = \alpha_o \prod_{i=1}^{n} e^{a_{c,i} / \left(1+e^{-\frac{(E-E_{o,i})}{\gamma_i}}\right)} \tag{8.15}$$

Fits of the as-is experimental spectra to Eqs. (8.13),(8.14) and (8.15) are shown by the black dotted lines of $d(ln\alpha(E))/dE$, $ln(\alpha(E)/\alpha_o)$ and $\alpha(E)$, respectively, in Figs. 8.5, 8.7 and 8.9. The results are given in Table 8.1 and are in excellent agreement with the data.

Exact modeling the interference pattern is the subject of future work and can be easily substituted for the Gaussian approximation made here. This is the reason we have included the interference pattern in our model. Even though it can be considered as extrinsic to the physics of “tailing," it is complementary to the band-edge signal and is required for fitting purposes. *Improving the accuracy of the interference pattern function will result in more accurate extraction of the band-edge signal and parameters.*

As discussed in Sec. 8.6.1, the origin of the artifact in cubic $Mg_{0.75}Zn_{0.25}O$ is hypothesized to be the absorption signal of the lower-energy wurtzite band-edge. For this special case, in Eq. (8.10), $d(f_a(E))/dE = c$, a constant, while $d(f_s(E))/dE$ is given by the logistic PDF in Eq. (8.7). Furthermore, in Eq. (8.11), $f_a(E) = cE$ and $f_s(E)$ is given by Eq. (8.8). Finally, in Eq. (8.12), $F_a(E) = e^{cE}$ and $F_s(E)$ is given by Eq. (8.9).

Fits of the as-is experimental spectra to Eq. (8.10), Eq. (8.11) and Eq. (8.12) for the case described above are shown by the black dotted lines of $d(ln\alpha(E))/dE$, $ln(\alpha(E)/\alpha_o)$ and $\alpha(E)$, respectively, in Fig. 8.8. The results are given in Table 8.1. The fits are in excellent agreement with the data.

In general, there is a large number of free parameters involved in fitting the as-is experimental spectra with the generalized forms (Eq. (8.10), Eq. (8.11) and Eq. (8.12)). As can be seen, depending on which functional form of the band-edge is used, the extracted parameters vary for the same experimental data. This variation is however negligible as shown in Table 8.1, which shows the mean value and standard deviation of the band-edge parameters extracted from both the as-is experimental spectra (using Eqs. (8.10), (8.11) and (8.12)) and from the spectra analyzed via deconvolution (using Eqs. (8.7), (8.8) and (8.9)). This analysis illustrates the precision of both approaches. Despite this close agreement, for simplicity and ease-of-use, it is recommended to first analyze the signal components via deconvolution and fit these individually to the functions given in Sec. 8.6.1.

This concludes the segment of this work relating the development of the new model and its testing against the complex band-edges of ZnO at 77 K, $Mg_{0.75}Zn_{0.25}O$ (cubic and wurtzite mixed-phase) at 292 K and c-Si at 299 K. The following section will illustrate that our nonlinear model and the linear Urbach model converge to the same results.

8.7 LIMITING AND SPECIAL CASES OF NONLINEAR MODEL

8.7.1 *A Limiting Case: The Basic Urbach Model*

As mentioned previously, surveying the literature indicates a challenge in standardization of Urbach analysis in many instances. A limit analysis of our model gives some insight into this variability. In parallel, this limit analysis shows that our nonlinear model and the basic Urbach model converge to the same results when both are applied to approximately linear regions in the spectra.

We begin by taking the classical definition of the Urbach energy as *the inverse slope of the "linear" band-edge.* A glance back at Fig. 8.3 shows that the logarithmic band-edge is not linear when the energy derivative (the *slope*) is not a constant. The best we can say for a nonlinear band-edge is that it is "practically linear" at the *inflection point*, corresponding to the peak centers of the deconvolved signals. An example of this can be found at the vertical dotted line of Fig. 8.5(b) and (c). If the location of the extraction of the Urbach energy parameter is vanishingly small (an inflection point), this would explain the trend in the variable linear analysis in the literature.To illustrate this, we take the limit of Eq. (8.7) as $E \rightarrow E_o$, which models the *slope* of $ln\alpha_s(E)$ approaching the inflection point. We obtain the following:

$$\lim_{E \rightarrow E_o} \frac{d(ln\alpha_s(E))}{dE} = \frac{1}{4}\frac{a_c}{\gamma} \tag{8.16}$$

Finally, taking the inverse of Eq. (8.16) gives the *inverse slope at the inflection point* as:

$$\left(\lim_{E \rightarrow E_o} \frac{d(ln\alpha_s(E))}{dE} \right)^{-1} = 4\frac{\gamma}{a_c} = E_u \tag{8.17}$$

This result illustrates three key points:

1. the Urbach energy corresponds to an inflection point in nonlinear band-edge signals,

2. the Urbach energy, represented by the spread per unit area, is physically defined as the spread in the local potential per unit photon capacity and (

3. while the Urbach energy is proportional to the spread in the local potential, γ gives a direct measure of this quantity. Table 8.1 compares analytical values of E_u with values calculated from the relation in Eq. (8.17).

The analytical values were extracted from the inverse slope at the inflection point in the spectra for ZnO, wurtzite and cubic $Mg_{0.75}Zn_{0.25}O$ and c-Si shown by the horizontal dotted lines in Figs. 8.5(d), 8.7(d), 8.8(d) and 8.9(c), respectively. The results of this comparison are in excellent agreement.

We reiterate that this limit analysis shows the Urbach energy is found in the curve, $ln(\alpha(E)/\alpha_o)$, only in the place that is approximately linear, i.e. the inflection point at E_o in nonlinear band-edges. This approximation of the sigmoidal curve in $ln(\alpha(E)/\alpha_o)$ as linear is the basis of the basic Urbach model. Here we demonstrate that it is a *limiting case* of a more complete model [Eq. (8.7), Eq. (8.8) and Eq. (8.9)] of tailing at the band-edge.

Utilizing the energy derivative of $ln(\alpha(E)/\alpha_o)$ is advantageous by providing an entirely experimental determination of the Urbach energy, removing all ambiguity in locating and determining the slopes of interest and most critically, bypassing any assumptions “built-in" to the basic Urbach model, such as linearity or the Urbach focus. It must be emphasized that this approach also offers consistent criteria for directly measuring the variability in the local potential in addition to the corresponding UBE and can do so for multi-component band-edges.

Cody utilized the derivative spectrum $d(ln\alpha(E))/dE$ but only to locate the energy position of the indirect gaps of Si.[8] Cody did not mention the connection between the values of the slope on the abscissa and the definition of the Urbach energy—the value of the inverse slope. This is understandable as the author was primarily arguing for the existence of an Urbach edge in c-Si in the first place, which was not clearly understood at the time. This was done by drawing a straight line through the highly nonlinear edge shown in Fig. 8.9(a). A value of 9.6 meV was obtained, which is close to the values of $E_{u,1}$ and $E_{u,2}$ we obtained—6.70 meV and 6.60 meV, respectively. We briefly note that at least one instance in the literature reports multiple Urbach energies at the band-edge corresponding to different structural phases in the sample.[23]

In this section we have shown that in the limit of linearity the new and basic models converge to the same results. In the following section we show that linear band-edges may also represent a special case.

8.7.2 *A Special Case: Appearance of a Linear Band-Edge*

It is not the intention of this work to downplay the appearance of linear band-edges in the literature. A significant number of band-edges are (or least appear to be) linear. Our analysis shows that linear band-edges are compatible with our model but may represent a special case as illustrated by the following.

Urbach observed an approximately linear band-edge in $ln(\alpha(E)/\alpha_o)$ for AgBr, however, this was a classic example of a material with fine structure at low-temperature as noted by Dow and Redfield.[35] Structure in the band-edge is mostly observed in materials at low temperature—here we have shown that this can be due to distinct band-edge components corresponding to direct or indirect gaps, or in the case of thin films, artifacts of interference.

This section focuses on the spectra of ZnO thin film affected by interference artifacts. This can be seen in Fig. 8.3(b) and (c). As shown, we found that as sample

TABLE 8.3: *Fitted parameters of the electron-phonon interaction model applied to the data in the inset of Fig. 8.10(a), showing the temperature dependence of E_o.*

E (oK) (eV)	Θ (K)	$\hbar\omega$ (meV)	α (meV)
3.34	511	44	0.12

temperature is increased from low-temperature (77 K) the fine structure is diminished. This is highlighted in the inset of Fig. 8.10(b) showing the same dataset presented in Fig. 8.3(c). In essence, the interference artifact and signal become increasingly convolved due to thermal broadening. This subject is treated in depth in Chap. 7. At some temperature, the form of the components are such that their superposition yields a practically constant slope, or "plateau" in $d(ln\alpha(E))/dE$, which gives a linear band-edge in $ln(\alpha(E)/\alpha_o)$. This occurs near 377 K for the dataset shown in the inset of Fig. 8.10(b) and was chosen for further analysis shown in Fig. 8.11.

The results of this analysis indicate that the interference pattern may be convolved with the signal in such a way as to give the *appearance* of a linear band-edge in the experimental data of $ln(\alpha(E)/\alpha_o)$. This would lead to the erroneous conclusion that the Urbach rule applies to the entire linear region when, in fact, it is applied to the effects of an artifact. While this may not be the case in all instances, care should be taken when analyzing the spectra of thin films so that accurate data may be extracted.

8.7.3 *The Urbach Focus and the Unperturbed Band-Edge*

As mentioned in the introduction, E_o is reportedly observed in experiment as the Urbach focus shown in Fig. 8.4 (top). Analysis using the inflection point from the as-is derivative spectra in Fig 8.3(c) shows the absence of an Urbach focus in temperature-dependent spectra for ZnO as shown in Fig. 8.4 (bottom) for the same data as in Fig. 8.4 (top) which shows a focus. The slopes in Fig. 8.4 (top) were acquired by assuming the presence of the Urbach focus in the basic Urbach model. We also investigated this trend on the deconvolved signal of the differential band-edge in ZnO as shown in Fig. 8.10. These tangents in $ln(\alpha_s(E)/\alpha_o)$ were acquired from the maximum value at the inflection point in $d(ln\alpha_s(E))/dE$. The results also indicate no Urbach focus. An Urbach focus is acquired from a constant rate of change of the slope with temperature. Instead, via derivative analysis, we see an *increase* in the rate of change of the slope

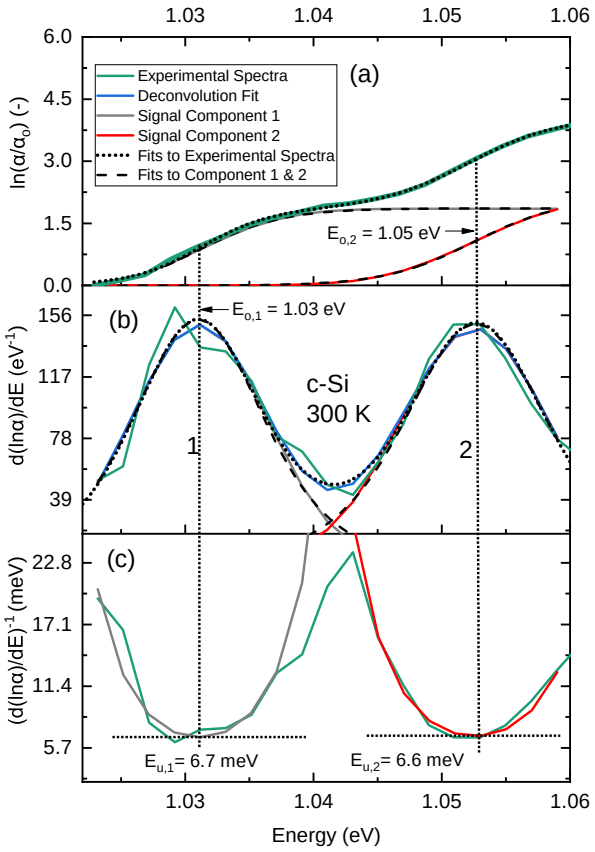


FIGURE 8.9: *(a) $ln(\alpha(E)/\alpha_o)$ for Si single crystal (c-Si) at 299 K (Cody[8], reprinted with permission from Elsevier.) (b) Energy derivative of $ln(\alpha(E)/\alpha_o)$ and (c) inverse plot of $d(ln\alpha(E))/dE$. The green solid lines show the experimental spectra. The grey and red solid lines show the analyzed components of the spectra. The black dashed lines are the fits to Eqs. (8.8) and (8.7), respectively, and the black dotted lines are the fits to Eqs. (8.14) and (8.13), respectively. The fitted parameters are shown in Table 8.1. The solid blue line in (c) is the fit from deconvolution analysis. The Urbach energies, E_u, are shown in (d) corresponding to the value of the inverse slope at the inflection point, E_o. Comparison of this value with E_u calculated from Eq. (8.17) is shown in Table 8.1.*

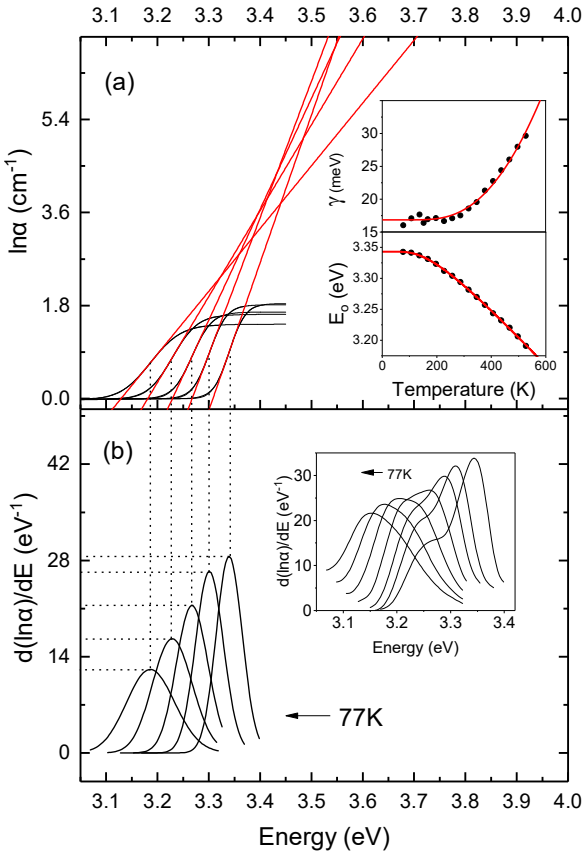


FIGURE 8.10: *(a) Logistic CDF corresponding to the band-edge signal in Fig. 8.5(c) as a function of temperature from 77-527 K and fitted slopes extracted at the inflection point via derivative analysis shown in (b). The results indicate no Urbach focus. A focus is obtained from a constant rate of change of the slopes, but here we see an increase in the rate of change of the slope with temperature. The top inset in (a) shows γ and the bottom inset in (a) shows E_o, both as a function of temperature. The fit to $\gamma(T)$ is a guide to the eye and $E_o(T)$ is fitted to the electron-phonon interaction model. The parameters are given in Table 8.3. (b) The analyzed band-edge signal shown in Fig. 8.5(c) as a function of temperature from 77-527 K. The inset shows the unresolved features at the band-edge for $d(ln\alpha(E))/dE$ for representative spectra from 77-527 K.*

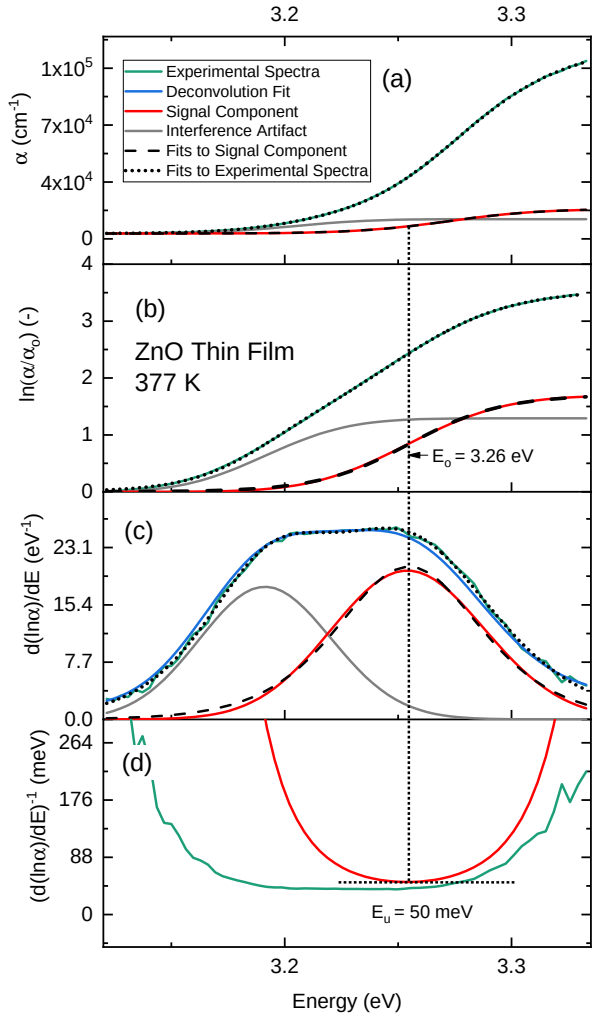


FIGURE 8.11: *(a) Absorption coefficient, $\alpha(E)$, (b) $ln(\alpha(E)/\alpha_o)$, (c) energy derivative of $ln(\alpha(E)/\alpha_o)$ and (d) inverse plot of $d(ln\alpha(E))/dE$ for ZnO thin film at 377 K. The green solid lines show the experimental spectra. The grey and red solid lines show the analyzed components of the spectra. The black dashed lines are the fits to Eqs. (8.9), (8.8) and (8.7), respectively, and the black dotted lines are the fits to Eqs. (8.15), (8.14) and (8.13), respectively. The fitted parameters are shown in Table 8.1. The solid blue line in (c) is the fit from deconvolution analysis. The Urbach energy, E_u, is shown in (d) corresponding to the value of the inverse slope at the inflection point, E_o. Comparison of this value with E_u calculated from Eq. (8.17) is shown in Table 8.1.*

with temperature. The inset in Fig. 8.10(a) shows that the temperature dependence of E_o in our model obeys the trend expected from the electron-phonon interaction model:

$$E_o(T) = E_o(0K) - 2\alpha/(e^{\Theta/T} - 1) \tag{8.18}$$

where $\Theta = \hbar\omega$. The results of fitting this model to the data is given in Table 8.3. Note that $\hbar\omega$ for ZnO is significantly underreported at 44 meV. The accepted literature value for this quantity is approximately 70 meV[69?,70], as mentioned in Chap. 6.

Based on the presently refined model, we see that, experimentally, the UBE, E_o, lies at the center of the PDF shown in $d(ln\alpha(E))/dE$. It follows that the spread in the PDF, γ, represents the variability in this ideal energy as a result of the nonuniform local potential.

Theoretically approaching ideal crystal structure with a perfectly uniform local potential, we can write the limit of Eq. (8.8) as $\gamma \to 0$. Here, the mathematics predict that the PDF in Eq. (8.7) will be described by a Dirac delta function, $\delta(E - E_o)$, centered at the energy of the UBE. In statistics, it is known that as the standard deviation of a normally distributed random variable narrows and tends to zero, the normal distribution tends toward a Dirac delta function.[71] In this limiting case, integration of the Dirac delta function gives the CDF in $ln(\alpha(E)/\alpha_o)$ as a Heaviside step-function, $H(E - E_o)$, at the energy of the UBE, E_o, with maximum value of photon absorption, a_c.[72] This is illustrated by the following:

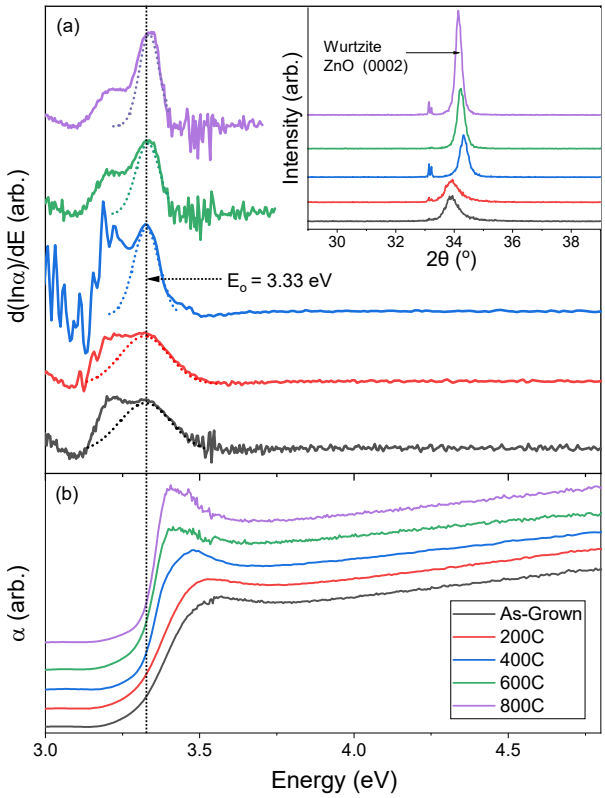


FIGURE 8.12: *(a) $d(ln\alpha(E))/dE$ at 77 K for ZnO thin film as-grown (grey) and annealed at 200 (red), 400 (blue), 600 (green) and 800 (purple) ^oC. The dotted fits to the spectra are the result of fitting Eq. (8.7) to the deconvolved signal component. E_o, shown by the dotted line at 3.33 eV, were given by these fits. The top inset shows the corresponding XRD diffractogram for the ZnO wurtzite peak (0002) for the same succession of annealing treatments. (b) Absorption coefficient plots, $\alpha(E)$, at 77 K for ZnO thin film as-grown and annealed at 200, 400, 600 and 800^oC corresponding to the derivative plots in (a).*

TABLE 8.4: *The unperturbed band-edge, E_o and defect broadening, γ_o, measured at 77 K and FWHM from XRD for ZnO thin film as-grown and annealed at 200^{o}C increments up to 800^{o}C .*

Temp. (^{o}C)	E_o (eV)	γ_o (meV)	XRD FWHM (^{o})
0	3.324	54.0	0.52
200	3.321	49.8	0.51
400	3.326	25.9	0.31
600	3.332	24.9	0.28
800	3.338	18.8	0.25

$$\lim_{\gamma \to 0} ln(\alpha(E)/\alpha_o) = a_c \lim_{\gamma \to 0} \frac{1}{1 + e^{-\frac{(E-E_o)}{\gamma}}} \tag{8.19}$$

$$= a_c(H(E - E_o)) \tag{8.20}$$

$$= \begin{cases} a_c & E > E_o \\ 0 & E < E_o \end{cases} \tag{8.21}$$

To test this mathematical result experimentally we have analyzed $d(ln\alpha(E))/dE$ for ZnO thin film acquired at 77 K for the sample as grown and annealed incrementally by 200 ^{o}C up to 800 ^{o}C. Post-growth annealing treatment has been shown to remove defects and improve the crystal quality of thin films. The results in Fig. 8.12(a) show that, as predicted by the definition of the UBE, E_o, found at 3.33 eV, is invariant to the defect concentration in the sample. This indicates that the UBE can be recovered from the spectra in $d(ln\alpha(E))/dE$ irrespective of the degree of "tailing" in $\alpha(E)$, shown in Fig. 8.12(b). Here, E_o varies by approximately 14 meV over the full range of annealing treatment. Furthermore, we show that the improvement from post-growth annealing is reflected in γ_o (γ at 77 K), which decreases for higher annealing temperatures, shown in Fig. 8.13. At temperatures above 77 K, γ has contributions from both thermal broadening in addition to diverse defect broadening mechanisms. This is discussed further in the following section.

8.8 INHOMOGENEOUS BROADENING

In solid state systems, Gaussian broadening such as that found in this work is indicative of inhomogeneous broadening which is a result of a fluctuating local electric field.[51] This fluctuating local electric field is caused by local perturbations of varying sources such as structural defects, impurities, thermal fluctuations and others.

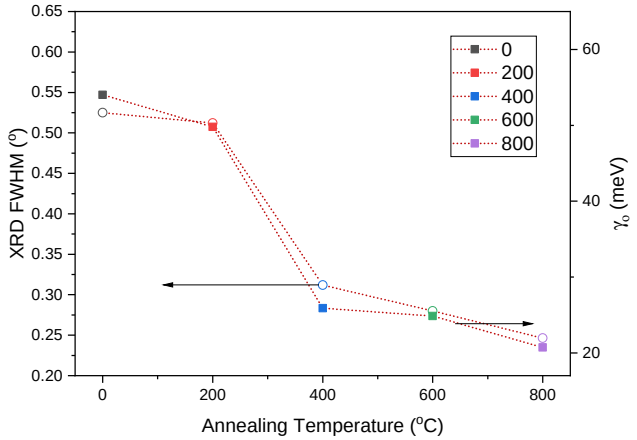


FIGURE 8.13: *The XRD full-width half-max (FWHM) (open circles) and defect broadening parameter γ_0 (solid squares) obtained from fitting Eq. (8.7) to the deconvolved spectra at 77 K for ZnO thin film as-grown and annealed at 200, 400, 600 and 800°C. All parameters are shown in Table. 8.4. The correlation (Pearson's R) between these two physical parameters was found to be 0.998.*

Cody[36] and Wasim[73] have described an equivalent effect on band-edge broadening due to perturbations from thermal excitation and perturbation from defects. This has been studied through the temperature dependence of the band-edge slope and ultimately, E_u. In Fig. 8.10(b), we show the temperature dependence of the Gaussian lineshape and in the top inset of Fig. 8.10(a), the temperature dependence of the broadening parameter, $\gamma(T)$. For the range of temperatures studied, the Gaussian lineshape of the band-edge was found to be conserved. Furthermore, the temperature dependence of $\gamma(T)$ was found to have a similar trend as the temperature dependence of $E_u(T)$ found in the literature.[36] This is reasonable to expect since these parameters are linearly proportional, as shown by Eq (8.17). To our knowledge, there is no model to describe the temperature dependence of the broadening parameter, $\gamma(T)$, in absorption spectra.

Illustrated in the top inset of Fig. 8.10(a), lowering the sample temperature below a certain threshold no longer reduces the broadening parameter, indicating only the presence of defect broadening. The lineshape of the differential band-edge was then analyzed here in the absence of thermal perturbation (electron-phonon interactions) at 77 K. As shown in Fig. 8.12(a), the lineshape was also found to be Gaussian. The broadening effect of thermal perturbation from electron-phonon interaction appears to be indistinguishable from the broadening effect of defects on the band-edge. That is, the Gaussian lineshape is conserved for both, in agreement with the predictions of Cody and Wasim.

The thorough characterization of the temperature dependence of the broadening parameter, $\gamma(T)$, and its relationship to defects in the material is the subject of Chap. **??**. Some results of this chapter are reiterated in brief as follows. Through successive annealing experiments at 200, 400, 600 and 800 ^{0}C, we study the change of the 77 K band-edge distribution as a function of defects as shown in Fig. 8.12(a). At 200 ^{0}C annealing, we found a small reduction to the band-edge broadening. At 400 ^{0}C, we

found significant reduction of the band-edge broadening. Furthermore, at 600 and 800 ^{0}C, we found some small reduction in the broadening. We correlated γ_o with the XRD FWHM and show there is 0.99 Pearson's correlation between these two parameters. All parameters are shown in Table. 8.4. The significant reduction in the XRD FWHM at 400 ^{0}C indicates grain growth activation at this temperature. This agrees with the results of Vishwas *et al*. who also found this activation at 400 ^{0}C via XRD for ZnO thin film grown by the sol-gel method.[74] These facts strongly suggest that the main defect contribution to the band-edge broadening and ultimately to the Urbach tail in the ZnO thin film studied here at 77 K originates from grain boundaries.

While the main contribution to the band-edge tail originates from grain boundaries, there are also likely contributions from Zn interstitial (Zn_i) point defects in the as-grown film shown by the grey curve in Fig. 8.12(a). Our previous work showed that our samples grown via sputtering are rich in Zn_i.[75] While annealing at 200 ^{0}C does not seem to affect grain boundaries, Zn_i are highly unstable and are expected to be thermally active at relatively low annealing temperatures.[75] We conclude that the small improvement to the band-edge broadening parameter at 200 ^{0}C annealing is likely attributed to the migration of Zn_i, readily out-diffused or lattice-incorporated at this temperature.

The key result of this experimental investigation relative to this chapter is the *invariance of the band-edge lineshape to the defect type*, specifically grain boundaries and Zn_i in ZnO and alloy inhomogeneity in MgZnO. As discussed in the previous section, the effect on the band-edge of removing both defect types via annealing is to change the width of the distribution. The distribution itself remains centered about E_o and, more importantly, the normal distribution is changed symmetrically. This remains true for every increase in annealing temperature and at 400 ^{0}C where the most significant changes to the band-edge broadening is observed.

This supports the conclusion that the underlying physical phenomena is, at its core, local electric field fluctuation, and shows an equivalent effect on the band-edge of thermally induced and defect-induced electric field fluctuation and, further, between electric field fluctuation from multiple defect types. Importantly, we found the Gaussian lineshape to be invariant to the source of electric field fluctuation, thus meeting the standard definition of inhomogeneous broadening.

This suggests that, in the ZnO sample studied, the full range of the band-edge uniformly includes broadening contributions from all the sources mentioned here. Specifically, the defect contributions to broadening do not seem to be exclusive to any distinct part of the tail. This is in contrast to theories in which the low-, medium-, and high-energy regions of the tail are hypothesized to correspond to specific defect types.[5] We note that for the data shown in Fig. 8.12, the excitonic region corresponds to the high-energy tail of the Gaussians and likely contribute to small asymmetries here. Due to the approximately one-to-one correlation between γ_o and the XRD FWHM we assert that the effect of excitons is negligible compared to the effect of grain boundaries on the broadening of the band-edge in our ZnO thin film. In conclusion, we uncover the Gaussian lineshape resulting from perturbation by inhomogeneous sources—such as possible Zn_i, grain boundaries and electron-phonon interactions—which are uniformly distributed across the full range of the band-edge tail.

8.9 FURTHER DISCUSSION & CONCLUSION

Many physical mechanisms have experimentally been shown to perturb the band-edge and change the nature of tailing in the absorption spectra of semiconductors. These include but are not limited to alloy inhomogeneity, thermal excitation, impurities, and structural defects. In addition, the band-edge states have been proposed to

comprise bound excitons interacting with lattice vibrations as proposed by Toyozawa (1959, 1964)[76], electric field broadening of an exciton line as proposed by Dow and Redfield (1970)[76], or electric field broadening of direct transitions between parabolic bands.[76] Mott and Davis summarize a short review by Hopfield (1968) on the various models that have been proposed for giving rise to the Urbach edge.[76] These widely varying mechanisms have all been hypothesized to reflect general disordering which varies the local potential in the material. In the words of Economou (1987), this "proliferation of formalisms and physical mechanisms tend to obscure the common feature of a fluctuating potential being the source of the exponential tails."[55] He proposes that the general character of the exponential tails suggests "a quasi-universal mechanism that bypasses the complexity of real materials."[55] John *et al.*,[77] states it is strongly suggested that "the underlying physics is both simple and general."

Advancing in this spirit, we have made no assumptions about the form, exponential or otherwise, of the band-edge.

In his published work of 1953,[5] Economou's theoretical investigations into the tailing in absorption spectra are primarily based on the assumptions that:

1. the tail in the measured absorption coefficient is a result of tailing in the DOS and

2. a Gaussian probability distribution of the local potential is responsible for tailing in the DOS.[5,56]

He proceeds to derive the theoretical (and causal) connection between (2) and (1). That is, he proceeds to show that the nonuniform local potential leads to the exponential DOS, which is then observed in the absorption coefficient.

We base our work largely on Economou's theoretical considerations but make slight variations mostly in the number and nuance of the assumptions being made. We do not seek to show that the nonuniform local potential leads to an exponential

tail in the DOS because, as we have shown, the tail in the absorption spectra is, on the whole, not exponential. We instead assume the validity of Economou's more fundamental assumption of the presence of a Gaussian probability distribution in the local potential and show directly its connection to the observed absorption coefficient without introducing an intermediate assumption about the DOS.

Still, we are able to analytically show the connection between the theoretically predicted nonuniform local potential and the experimentally observed absorption coefficient.

To this end, we show that the form of the nonuniform local potential can be experimentally uncovered in derivative spectra and described by a normally distributed probability density corresponding to the DOS at the band-edge. Due to the cumulative probabilistic observation of absorption in semiconductors, the probability densities integrate directly to the form of the tail in the absorption spectra. The form was found instead to be exponential-logistic. As we show, this form can be approximated under certain conditions—and for limited ranges—by the simple exponential form given by the linear Urbach model.

In summary, we have built conceptually on the foundation of basic Urbach theory and, by examining the assumption of band-edge exponentiality, have taken one step further toward a more complete and experimentally accurate model of tailing at the band-edge. While being consistent with the linear Urbach model, our generalized model fits complex, nonlinear band-edges and is capable of extracting much more quantitative data from absorption spectra than previously anticipated.

Furthermore, by outlining a procedure for utilizing derivative spectra, we have provided a simple solution to two problems:

1. the convolution of artifacts with the band-edge signal and
2. the variability in Urbach analysis from application of the linear Urbach model to nonlinear band-edges.

Coupled with this more standardized analytical approach, our model is capable of extracting band-edge parameters in a more objective and universally consistent way as shown from successful applications to the nonlinear band-edges of ZnO thin film (at various temperatures), mixed-phase $Mg_{0.75}Zn_{0.25}O$ thin film, and c-Si. Additionally, in the limit of zero variability in the local potential, our model predicts the ideal form of absorption, a step function, at the energy of the UBE, E_o, corresponding to the maximum probability density of absorption. As mentioned before, transitions at the band-edge may be defect-related, exciton-related, related to direct or indirect transitions between parabolic bands. It must be stressed that it is not the intention of this work to distinguish or characterize the type of transitions occurring at the band-edge but to show that our model may be capable of being applied to many such cases.

Finally, several limitations may arise in the application of our current model, mostly associated with the limitations of deconvolution analysis. Here we have investigated easily identified normal distributions in the derivative spectra, however, there may be cases where the convolution of band-edge components is highly complex. Additionally, further sensitivity analyses of the fitting parameters of this model is required. This becomes especially important when fitting linear combinations of the model functions (given in Sec. 8.6.2) as this significantly increases the number of free parameters with each additional band-edge component.

Application to a wider variety of semiconducting materials is necessary to test the model further and to explore its strengths and limitations. It is anticipated that the current development will contribute to the study of band-edges in novel systems of soft matter, organics, bioinorganics and nanobioconjugated systems.

REFERENCES

[1] A. Tudorovskii, AN SSSR, Moscow-Leningrad (1948).

[2] M. V. Kurik, physica status solidi (a) **8**, 9 (1971).

[3] Y. Toyozawa, Progress of Theoretical Physics **22**, 455 (1959).

[4] A. S. Davydov and A. Lubchenko, in *Doklady Akademii Nauk*, Vol. 179 (Russian Academy of Sciences, 1968) pp. 1301–1303.

[5] M. A. Kastner, S. R. Ovshinsky, and G. A. Thomas, *Disordered Semiconductors*, 1987th ed. (Springer, New York, 1987).

[6] W. Martienssen, Journal of Physics and Chemistry of Solids **2**, 257 (1957).

[7] M. V. Kurik, physica status solidi (a) **8**, 9.

[8] G. D. Cody, Journal of Non-Crystalline Solids **141**, 11 (1992), *Reprinted with permission from Elsevier.*

[9] M. Kranjčec, I. P. Studenyak, and M. V. Kurik, Journal of Non-Crystalline Solids **355**, 54 (2009).

[10] D. A. Drabold, Y. Li, B. Cai, and M. Zhang, Phys. Rev. B **83**, 045201 (2011).

[11] S. Duman, B. Gurbulak, S. Dogan, and F. S. Ozcelik, J. Phys.: Conf. Ser. **707**, 012027 (2016).

[12] M. E. Kassem, A. A. Higazy, R. R. Zahran, and A. Y. Kandeil, Materials Letters **16**, 236 (1993).

[13] M. Kranjčec, I. D. Desnica, I. P. Studenyak, G. S. Kovacs, M. V. Potory, Y. V. Voroshilov, and V. Y. Gebesh, Materials Research Bulletin **34**, 2297 (1999).

[14] B. Abay, H. S. Güder, H. Efeoğlu, and Y. K. Yoğurtçu, Journal of Physics and Chemistry of Solids **62**, 747 (2001).

[15] K. Zickus, A. Audzijonis, J. Batarunas, and A. Sileika, physica status solidi (b) **125**, 645.

[16] Q. Shen and T. Toyoda, Jpn. J. Appl. Phys. **38**, 3163 (1999).

[17] V. P. Kunets, N. R. Kulish, V. P. Kunets, and M. P. Lisitsa, (2002).

[18] B. I. Turko, V. B. Kapustyanyk, V. P. Rudyk, M. V. Partika, M. V. Kvasnytsya, and A. P. Vas'kiv, J Appl Spectrosc **74**, 310 (2007).

[19] P. P. Guranich, E. I. Gerzanich, A. G. Slivka, V. S. Shusta, and V. A. Bobela, Ferroelectrics **132**, 173 (1992).

[20] K.-i. Noba and Y. Kayanuma, Phys. Rev. B **60**, 4418 (1999).

[21] I. P. Studenyak, M. Kranjčec, G. S. Kovács, V. V. Pan'ko, I. D. Desnica, A. G. Slivka, and P. P. Guranich, Journal of Physics and Chemistry of Solids **60**, 1897 (1999).

[22] I. P. Studenyak, M. Kranjčec, G. S. Kovacs, I. D. Desnica-Franković, V. V. Panko, and V. Y. Slivka, Journal of Materials Research **16**, 1600 (2001).

[23] V. Mishra, A. Sagdeo, V. Kumar, M. K. Warshi, H. M. Rai, S. K. Saxena, D. R. Roy, V. Mishra, R. Kumar, and P. R. Sagdeo, Journal of Applied Physics **122**, 065105 (2017).

[24] I. P. Studenyak, O. A. Mykajlo, Y. M. Vysochanskii, and V. B. Cajipe, J. Phys.: Condens. Matter **15**, 6773 (2003).

[25] M. Kranjčec, I. P. Studenyak, and M. V. Kurik, Journal of Physics and Chemistry of Solids **67**, 807 (2006).

[26] J. Szczyrbowski, physica status solidi (b) **105**, 515.

[27] L. Nadjia, E. Abdelkader, B. Naceur, and B. Ahmed, Journal of Rare Earths **36**, 575 (2018).

[28] Ng Tse Nga, Wong William S., Lujan Rene A., and Street Robert A., Advanced Materials **21**, 1855 (2009).

[29] Z. N. Kayani, S. Iram, R. Rafi, S. Riaz, and S. Naseem, Appl. Phys. A **124**, 468 (2018).

[30] A. Kaphle and P. Hari, Journal of Applied Physics **122**, 165304 (2017).

[31] H. Ertap, Optical Materials **83**, 99 (2018).

[32] F. Chandoul, A. Boukhachem, F. Hosni, H. Moussa, M. S. Fayache, M. Amlouk, and R. Schneider, Ceramics International **44**, 12483 (2018).

[33] D. A. Jones and J. U. Lee, Nano Lett. **11**, 4176 (2011).

[34] I. Studenyak, M. Kranjčec, M. Kurik, and I. Studenyak, International Journal of Optics and Applications **4**, 76 (2014).

[35] J. D. Dow and D. Redfield, Phys. Rev. B **5**, 594 (1972).

[36] G. D. Cody, T. Tiedje, B. Abeles, B. Brooks, and Y. Goldstein, Phys. Rev. Lett. **47**, 1480 (1981).

[37] D. J. Dunstan, J. Phys. C: Solid State Phys. **15**, L419 (1982).

[38] S. R. Johnson and T. Tiedje, Journal of Applied Physics **78**, 5609 (1995).

[39] A. E. Rakhshani, J. Phys.: Condens. Matter **12**, 4391 (2000).

[40] A. M. Bakry and A. H. El-Naggar, Thin Solid Films **360**, 293 (2000).

[41] C. Rincón, S. M. Wasim, G. Marín, R. Márquez, L. Nieves, G. S. Pérez, and E. Medina, Journal of Applied Physics **90**, 4423 (2001).

[42] R. Bhattacharya, R. Mondal, P. Khatua, A. Rudra, E. Kapon, S. Malzer, G. Döhler, B. Pal, and B. Bansal, Phys. Rev. Lett. **114**, 047402 (2015).

[43] R. C. Rai, Journal of Applied Physics **113**, 153508 (2013).

[44] T. Shioda, S. Chichibu, T. Irie, H. Nakanishi, and T. Kariya, Journal of Applied Physics **80**, 1106 (1996).

[45] B. Sadigh, P. Erhart, D. Åberg, A. Trave, E. Schwegler, and J. Bude, Phys. Rev. Lett. **106**, 027401 (2011).

[46] A. Emel'yanov, Physics of the Solid State **58**, 1081 (2016).

[47] K. Zhang, C. Zhen, W. Wei, W. Guo, G. Tang, L. Ma, D. Hou, and X. Wu, RSC Adv. **7**, 36026 (2017).

[48] L. C. Xue, L. Q. Wu, S. Q. Li, Z. Z. Li, G. D. Tang, W. H. Qi, X. S. Ge, and L. L. Ding, Physica B: Condensed Matter **492**, 61 (2016).

[49] T. C. O'Haver, Anal. Chem. **51**, 91A (1979).

[50] L. H. Bennett, *Electronic Density of States: Based on Invited and Contributed Papers and Discussion* (U.S. National Bureau of Standards, 1971).

[51] J. D. Ingle and S. R. Crouch, *Spectrochemical Analysis*, 1st ed. (Prentice Hall, Englewood Cliffs, N.J, 1988).

[52] J. I. Pankove, Phys. Rev. **140**, A2059 (1965).

[53] S. John, M. Y. Chou, M. H. Cohen, and C. M. Soukoulis, Phys. Rev. B **37**, 6963 (1988).

[54] C. T. Chan, S. G. Louie, and J. C. Phillips, Phys. Rev. B **35**, 2744 (1987).

[55] E. N. Economou, N. Bacalis, and M. H. Cohen, Journal of Non-Crystalline Solids Amorphous and liquid semiconductors, **97-98**, 101 (1987).

[56] N. Bacalis, E. N. Economou, and M. H. Cohen, Phys. Rev. B **37**, 2714 (1988).

[57] J. F. Wager, AIP Advances **7**, 125321 (2017).

[58] G. Rey, G. Larramona, S. Bourdais, C. Choné, B. Delatouche, A. Jacob, G. Dennler, and S. Siebentritt, Solar Energy Materials and Solar Cells **179**, 142 (2018).

[59] E. O. George and G. S. Mudholkar, Metrika **30**, 1 (1983).

[60] J. Berkson, Journal of the American Statistical Association **39**, 357 (1944).

[61] D. Costarelli, Neural Netw. **67**, 28 (2015).

[62] Z. Chen, F. Cao, and J. Hu, Applied Mathematics and Computation **256**, 565 (2015).

[63] I. A. Basheer and M. Hajmeer, Journal of Microbiological Methods Neural Computting in Micrbiology, **43**, 3 (2000).

[64] J. Han and C. Moraga, in *From Natural to Artificial Neural Computation*, Lecture Notes in Computer Science (Springer, Berlin, Heidelberg, 1995) pp. 195–201.

[65] D. Costarelli and R. Spigler, Neural Networks **44**, 101 (2013).

[66] J. Dombi and Z. Gera, Fuzzy Sets and Systems **154**, 275 (2005).

[67] Y. Chalco-Cano, H. Román-Flores, and F. Gomide, Fuzzy Sets and Systems Theme: Fuzzy Interval Analysis, **159**, 1376 (2008).

[68] H. P. Ritzema, ed., *Drainage Principles and Applications*, revised edition ed. (Water Resources Pubns, Wageningen, 1994).

[69] D. W. Hamby, D. A. Lucca, M. J. Klopfstein, and G. Cantwell, Journal of Applied Physics **93**, 3214 (2003).

[70] R. Cuscó, E. Alarcón-Lladó, J. Ibáñez, L. Artús, J. Jiménez, B. Wang, and M. J. Callahan, Phys. Rev. B **75**, 165202 (2007).

[71] in *Stochastic Dynamics of Structures* (Wiley-Blackwell, 2010) pp. 343–348.

[72] R. P. Kanwal, *Generalized Functions: Theory and Technique*, 2nd ed. (Birkhäuser Boston, Boston, 1998).

[73] S. M. Wasim, C. Rincón, G. Marín, P. Bocaranda, E. Hernández, I. Bonalde, and E. Medina, Phys. Rev. B **64**, 195101 (2001).

[74] M. Vishwas, K. N. Rao, K. V. A. Gowda, and R. P. S. Chakradhar, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **77**, 330 (2010).

[75] D. Thapa, J. Huso, J. L. Morrison, C. D. Corolewski, M. D. McCluskey, and L. Bergman, Optical Materials **58**, 382 (2016).

[76] N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (OUP Oxford, 2012).

[77] S. John, C. Soukoulis, M. H. Cohen, and E. N. Economou, Phys. Rev. Lett. **57**, 1777 (1986).