ABSTRACT

LONG, JOSEPH PRESTON. Spectroscopic and Electrical Studies of Hafnium-Based High-κ Thin Film Dielectrics on Germanium Surfaces. (Under the direction of Gerald Lucovsky.)

The research discussed here has been carried out in order to advance the basic understanding of the compatibility between germanium surfaces and hafnium-based high-κ dielectric materials with a particular emphasis on their potential for microelectronic applications. To this end, spectroscopic studies were carried out to determine the physical and electronic properties of Ge/high-κ structures, and MOS capacitors were fabricated to study their electrical characteristics. Crystallinity, thermal stability, electronic defect levels, and Hf d-state degeneracy removal in this material system were studied via x-ray absorption spectroscopy (XAS). The presence of conduction band edge defect states was confirmed by spectroscopic ellipsometry (SE), and medium energy ion spectroscopy (MEIS) was employed to investigate chemical distributions within these structures. Capacitance-voltage (C-V) measurements made on related MOS structures provided valuable insight into the nature of carrier trapping and charged traps in these devices, and lastly, current-voltage (I-V) measurements revealed information about the integrity of these gate materials while in contact with germanium surfaces. While the nitridation of germanium surfaces was shown to mitigate atomic migration and diffusion to a degree during device processing, both hafnia and hafnium nitro-silicate films demonstrated a large degree of physical and chemical instability when in contact with germanium. Moreover, these difficulties were found to be correlated with electrically active defects which make this technology unsuitable for CMOS applications at present.
Spectroscopic and Electrical Studies of Hafnium-Based High-κ Thin Film Dielectrics on Germanium Surfaces

by
Joseph Preston Long

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Physics

Raleigh, North Carolina

2008

APPROVED BY:

________________________  __________________________
Jack Rowe               Robert Kolbas

________________________  __________________________
David Aspnes            Gerald Lucovsky
Chair of Advisory Committee
Dedication

With love to all of my family and friends who supported me throughout this endeavor.
Biography

Joseph Preston (J.P.) Long was born in Durham, NC on January 9, 1979 to Robert Delaney and Beth Amos Long. He spent most of his formative years in Charlotte, NC, where he attended East Mecklenburg High School. Following his graduation from high school in 1997, J.P. attended the University of North Carolina at Chapel Hill on a James M. Johnston scholarship, earning Bachelor of Science degrees in both physics and mathematics in 2001. He decided for personal reasons to remain in the Raleigh/Durham area, accepting an attractive offer from Professor Jan F. Schetzina of North Carolina State University to study wide bandgap III-nitride semiconductor materials and devices. Under Dr. Schetzina’s supervision, J.P. earned a Master’s of Science degree in physics in 2004. Professor Schetzina retired shortly thereafter, and J.P. decided to complete his graduate studies, transferring to Professor Gerald Lucovsky’s group at North Carolina State University in pursuit of a Doctor of Philosophy degree in physics. Under Dr. Lucovsky’s supervision, J.P. investigated complex oxides, particularly group IV-B transition metal oxides, silicates, and nitro-silicates. With his Ph.D now near completion, J.P. plans to attend law school and earn the degree of Juris Doctor, focusing on intellectual properties and patents. It is in this field that he plans to make his career, and he looks forward to this new set of challenges with great anticipation.
Acknowledgements

I would like to thank everyone who has made this work possible. This includes, first and foremost, Dr. Gerald Lucovsky whose advice and support have been vital in this endeavor. I would like also to thank Dr. Jan Schetzina for his encouragement and teachings during my graduate studies and Mr. Joe Matthews, lab technician extraordinaire, without whom I would have been useless as an experimentalist. He taught me virtually everything I know about working in a lab, and I thank him, especially, for all of the hands-on knowledge he imparted to me. Without the help of the PAMS machinists, Tim Harvell, Len Smith, Frank Milkowski, and Chris Hewett, I never would have made it this far either. I would also like to acknowledge the various graduate students, post doctoral researchers, and professors who helped me along the way. This list includes Sean Washburn, Mark Johnson, Doug Barlage, John Muth, Yoganand Saripalli, Cristiano Krug, Sanghyun Lee, Hyungtak Seo, Chris Hinkle, and Kwun-Bum Chung. Special thanks are in order as well for my advisory committee, Drs. David Aspnes, Jack Rowe, Robert Kolbas, and again, Gerald Lucovsky. Lastly, but certainly not least, I thank wholeheartedly all of those who gave me their love and support during this often very difficult and challenging time. This includes most notably my parents, Bob and Beth Long; my sister, Leigh Kugler; and my dear friend, Mayely Boyce. Thank you all.
Table of Contents

LIST OF TABLES .......................................................................................................................... viii
LIST OF FIGURES ...................................................................................................................... xi

CHAPTER 1: INTRODUCTION ................................................................................................. 1
  1.1 – Semiconductor Roadmap .............................................................................................. 1
  1.2 – Possible Benefits and Drawbacks of Using Germanium Substrates ...................... 5
  1.3 – High-κ Transition Metal Oxides as Gate Dielectrics on Si ................................... 8
  1.4 – The Possibility High-κ Transition Metal Oxides on Ge Surfaces ....................... 16
  1.5 – Overview of the Dissertation ..................................................................................... 17
  References .............................................................................................................................. 29

CHAPTER 2: THIN FILM DEPOSITION AND PROCESSING .................................................. 34
  2.1 – RPECVD Process Overview ..................................................................................... 34
  2.2 – Hafnium-Based High-κ Films by RPECVD ............................................................. 37
  2.3 – Germanium and Germanium-Containing Films by RPECVD ............................. 40
  2.4 – Thin Film Deposition and Processing: Summary and Conclusions .................. 43
  References .............................................................................................................................. 50

CHAPTER 3: HAFNIUM BASED HIGH-κ MATERIALS ON GERMANIUM SUBSTRATES ....... 53
  3.1 – Restatement of the Motivation and Proposed Solutions ........................................ 53
  3.2 – Germanium Substrate Cleaning ............................................................................. 56
  3.3 – Sample Preparation ................................................................................................. 63
  3.4 – Spectroscopic Results for Ge/High-κ Stack Structures ...................................... 66
  3.5 – Electrical Results for Ge/High-κ Stack Structures ............................................. 73
  3.6 – Summary of Results for High-κ Thin Films on Ge Substrates ............................. 80
  References .............................................................................................................................. 105

CHAPTER 4: HAFNIUM BASED HIGH-κ MATERIALS ON PSEUDO-MORPHIC GERMANIUM SURFACES GROWN ON SILICON ......................................................... 111
  4.1 – Brief Restatement of the Motivation .................................................................... 111
  4.2 – Sample Preparation ................................................................................................. 111
  4.3 – Spectroscopic Properties of Si/PM-Ge/High-κ Structures .................................. 115
  4.4 – C-V Properties of Si/PM-Ge/High-κ Structures .................................................. 118
  4.5 – Gate Leakage Properties of Si/PM-Ge/High-κ Structures ................................... 125
  4.6 – Summary of Results for Si/PM-Ge/High-κ Structures ......................................... 129
CHAPTER 5: SUMMARY, FUTURE WORK, AND CONCLUDING REMARKS ........................................ 152
   5.1 – Summary of Experiments and Results ................................................................. 152
   5.2 – Recommendations for Future Work ............................................................... 153
   5.3 – Concluding Remarks .................................................................................. 156
References ........................................................................................................ 158

APPENDIX 161

APPENDIX A: THE THEORY OF POINT GROUPS ......................................................... 162
   A.1 – Definition of a Group ..................................................................................... 162
   A.2 – Symmetry Operations .................................................................................. 163
   A.3 – Point Groups .............................................................................................. 163
   A.4 – Representations of Point Groups ................................................................ 165
   A.5 – The Reduction of Reducible Representations ............................................. 167
   A.6 – The Applications of Point Group Theory ................................................... 171
References ........................................................................................................ 178

APPENDIX B: MOLECULAR ORBITAL THEORY ......................................................... 179
   B.1 – Molecular Orbitals ...................................................................................... 179
   B.2 – Atomic Orbitals and Their Transformation Properties ................................ 180
   B.3 – Bond Hybridization .................................................................................... 182
   B.4 – Allowed and Forbidden Transitions ............................................................ 183
   B.5 – Symmetry Adapted Linear Combinations of Atomic Orbitals ................. 186
References ........................................................................................................ 193

APPENDIX C: CRYSTAL FIELD SPLITTING AND JAHN-TELLER DISTORTIONS ................. 194
   C.1 – Crystal Field Splitting .................................................................................. 194
   C.2 – Jahn-Teller Distortions ............................................................................... 197
References ........................................................................................................ 201

APPENDIX D: THE STRUCTURE AND POLYMORPHISMS OF GROUP IV–B TRANSITION METAL OXIDES AND RELATED ALLOYS ................................................................. 203
   D.1 – The Symmetry of Rutile TiO$_2$ ...................................................................... 203
   D.2 – The Symmetry of Baddeleyite ZrO$_2$ and HfO$_2$ ...................................... 208
   D.3 – The Symmetry of ([Hf, Ti, Zr]O$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ Compounds ........ 213
References ........................................................................................................ 230
APPENDIX E: A DESCRIPTION OF INTERMEDIATE PHASES AND SELF-ORGANIZATION IN NON-CRYSTALLINE SOLIDS ........232
   E.1 – Short Range Order: Bond Constraint Theory ............................................................. 232
   E.2 – Long Range Order: Intermediate Phases and Percolation Theory ...............240
   References .......................................................................................................................... 251

APPENDIX F: AUGER ELECTRON SPECTROSCOPY (AES) ........................................... 252
   F.2 – Elemental Analysis by AES ................................................................................. 254
   F.3 – Thickness Estimations by AES ......................................................................... 255
   References .......................................................................................................................... 262

APPENDIX G: X-RAY ABSORPTION SPECTROSCOPY (XAS) ...................................... 263
   G.1 – The X-Ray Absorption Process and Experiment ........................................... 263
   G.2 – XAS Spectra ....................................................................................................... 265
   References .......................................................................................................................... 276

APPENDIX H: SPECTROSCOPIC ELLIPSOMETRY (SE) ............................................. 278
   H.1 – The Theory and Experimental Set-Up for SE ............................................. 278
   H.2 – SE as it Relates to Transition Metal Oxides and Related Alloys .......... 283
   References .......................................................................................................................... 290

APPENDIX I: MEDIUM ENERGY ION SPECTROSCOPY (MEIS) .................................. 291
   I.1 – The Theory and Experimental Set-Up for MEIS ........................................ 291
   References .......................................................................................................................... 299

APPENDIX J: CAPACITANCE-VOLTAGE (C-V) MEASUREMENTS ......................... 300
   J.1 – The Theory of Capacitance in MOS Structures ........................................ 300
   J.2 – C-V Modeling and the Analysis of MOS Device Structures ............... 302
   References .......................................................................................................................... 313

APPENDIX K: CURRENT-VOLTAGE (I-V) MEASUREMENTS .................................... 314
   K.1 – The Theory of Leakage Current in MOS Structures ......................... 314
   References .......................................................................................................................... 320
List of Tables

Chapter 1: Introduction

NONE

Chapter 2: Thin Film Deposition and Processing

Table 2.1: The differences between direct and remote plasma CVD processes (after Lucovsky and Tsu [6]). The left side depicts the progression of a direct process; whereas, the right side depicts the steps in a remote process ..............................................46

Chapter 3: Hafnium-Based High-κ Materials on Germanium Substrates

Table 3.1: The three germanium chemical surface treatments investigated using real-time SE and AES .........................................................85

Table 3.2: A summary of the approximate chemical make-up of the films shown in figure 3.8 .................................................................92

Chapter 4: Hafnium Based High-κ Materials on Pseudomorphic Germanium Surfaces Grown on Silicon

NONE

Chapter 5: Summary, Future Work, and Concluding Remarks

NONE
Appendix A: The Theory of Point Groups

Table A.1: The multiplication table for the group $C_{3v}$. The operation in the left column is applied first, followed by the operation in the top row .................................................................175

Table A.2: The three irreducible representations for $C_{3v}$: $A_1$, $A_2$, and $E$. $A_1$ and $A_2$ are both one-dimensional, and $E$ is two-dimensional (i.e. a $2 \times 2$ matrix). Note that all rows and columns of this table are orthogonal as they must be, due to the Great Orthogonality Theorem. (This data can be found in most texts on point groups, e.g. R. McWeeny [2].) .........................176

Table A.3: The character table for the group, $C_{3v}$. Note that table A.3 is simply a compact way of writing the traces of the three representations in table A.2. (This data can be found in most texts on point groups, e.g. R. McWeeny [2] or D.M. Bishop [3].) ........................................................................................................177

Appendix B: Molecular Orbital Theory

Table B.1: The symmetry of the atomic orbitals commonly used in molecular orbital theory .................................................................191

Appendix C: Crystal Field Splitting and Jahn-Teller Distortions

Table C.1: The character table for the cubic point group, $O_h$.................................200

Appendix D: Crystal Field Splitting and Jahn-Teller Distortions

NONE

Appendix E: The Structure and Polymorphisms of Group IV–B Transition Metal Oxides and Related Alloys

NONE
Appendix F: The Structure and Polymorphisms of Group IV–B Transition Metal Oxides and Related Alloys

NONE

Appendix G: X-Ray Absorption Spectroscopy (XAS)

NONE

Appendix H: Spectroscopic Ellipsometry (SE)

NONE

Appendix I: Medium Energy Ion Scattering (MEIS)

NONE

Appendix J: Capacitance-Voltage (C-V) Measurements

NONE

Appendix K: Current-Voltage (I-V) Measurements

NONE
Chapter 1: Introduction

Figure 1.1: Scaling in the EOT for CMOS devices, both historically and as projected by the ITRS (after Wong and Iwai [5]). Note that alternative high-κ dielectrics are now the only available option for future scaling ....................................................20

Figure 1.2: Measured channel mobilities in silicon CMOS devices (after Frank, et al. [3] and references therein). Note the decline in mobility for channels with thickness less than approximately 10 nm ........................................................................21

Figure 1.3: A comparison of electron and hole mobilities in both silicon and germanium (data from refs. [9-12]).................................22

Figure 1.4: X-ray photoemission spectra showing the thermal decomposition of GeO₂ into gaseous GeO (from Prabhakaran, et al. [15], © 2000 Elsevier Science S.A. All rights reserved.).................................................................23

Figure 1.5: X-ray photoemission spectra showing the thermal decomposition of pure Ge₃N₄ formed by plasma nitridation of the Ge surface. (from Kutsuki, et al. [16], © 2008 The Japan Society of Applied Physics)..................................................................................24

Figure 1.6: Core level XPS spectra showing the structure of the suboxide region at the Si/SiO₂ interface on both Si (100) and Si (111) substrates (from Himpsel, et al., © 1988 The American Physical Society).................................................................................25

Figure 1.7: Cross-section HRTEM micrographs showing the existence of the ZrSi₂ silicide phase in ZrO₂/Si interfaces after annealing at 1000ºC (from Stemmer [24], © 2004 American Vacuum Society)............................................................................26

Figure 1.8: The thermal stability of zirconium silicate films as demonstrated by (a) an FTIR spectrum showing crystallization in a (ZrO₂)₀.₆(SiO₂)₀.₄ film after a 900ºC
RTA, (b) an XRD spectrum showing crystallization in a 
\((\text{ZrO}_2)_{0.5}(\text{SiO}_2)_{0.5}\) film after a 900°C RTA, and (c) high-
resolution TEM micrograph showing chemical phase 
separation and crystallization in a 
\((\text{ZrO}_2)_{0.5}(\text{SiO}_2)_{0.5}\) film after a 900°C RTA (from Rayner, et al., figures 1.8b and 
1.8c © 2003 American Vacuum Society). Figures 1.8a and 
1.8b show a sharpening of several features at 900°C, and 
the circled area of figure 1.8c shows streaking. These 
signatures all are indicative of crystallinity. The light and 
dark regions of figure 1.8c show the chemical phase 
separation..........................................................................................27

Figure 1.9: The molecular orbital diagram for oxygen interstitials and 
vacancies in various charge states (from Robertson [23], 
© 2006 IOP Publishing, Ltd.). \(V_0\) represents the doubly-
occupied neutral vacancy with \(A_1\) symmetry. \(V_0^-\) 
represents this vacancy with a trapped electron, pulling 
down an extra state with \(B_1\) symmetry from the conduction 
band. \(V_0^+\) represents the singly-occupied vacancy with \(A_1\) 
symmetry. \(V_0^{2+}\) represents the unoccupied vacancy with \(A_1\) 
symmetry. \(I_0^-\) represents the interstitial with a trapped 
hole. \(I_0\) represents the neutral interstitial, which forms an 
O-O bond and creates a \(\sigma^*\) state in the band gap. Lastly, \(I_0^+\) 
represents the neutral oxygen interstitial which has formed 
an O-O bond and subsequently trapped a hole .........................28

Chapter 2: Thin Film Deposition and Processing

Figure 2.1: A schematic of the RPECVD cluster tool at NCSU used for 
thin film deposition in these experiments (from Hattangady, 
et al. [5], © 1996 American Vacuum Society) .........................44

Figure 2.2: A rendering of the gas injection system for the RPECVD 
used in these studies. Metal precursor gases are brought in 
through the gas dispersal rings, and the rest pass through 
the plasma tube. The sample is held and heated on the 
temperature-controlled stage .........................................................45

Figure 2.3: The use of solid and liquid metal-organic compounds in 
CVD (shown here for a liquid compound). A carrier gas 
enters the bubbler at point (a), is forced through the metal-
organic compound at point (b), picking up gaseous metal-organic molecules through point (c), and exits to the deposition chamber at point (d) ........................................................47

Figure 2.4: The metal-organic delivery subsystem constructed for conducting experiments described in this dissertation. The picture in (a) shows the electronics rack with (listed from bottom to top and left to right): custom valve control switchbox, mass flow controller (MFC) set point controller, upstream pressure gauge, line temperature readout, line temperature zone selector, downstream pressure gauge and controller, pressure gauges power supply, and turbo-molecular pump controller (which never had a home before this). The picture in (b) shows the carrier gas plumbing (top) and (from left to right) the idling pump, variable transformer temperature controls, and constant temperature bath and bubbler (bottom). The figure in (c) is an original design drawing........................................................................................................48

Figure 2.5: A schematic diagram of the germane delivery system (described in the text) for the RPE-MOCVD used in these experiments ........................................................................................................49

Chapter 3: Hafnium-Based High-κ Materials on Germanium Substrates

Figure 3.1: Hole mobility enhancement in (a) bulk p-type Ge devices and (b) in Si devices with Si-capped epitaxially grown p-type Ge channels. These devices use either ZrO$_2$ or HfO$_2$ high-κ gate dielectrics (after Saraswat, et al. [4]).........................83

Figure 3.2: XAS spectra showing the nitrogen K$_1$ edge absorption in plasma-nitrided Ge/HfO$_2$ (top) and Ge/HfSiON (bottom) structures as a function of PDA annealing temperature. (Spectra in each graph are normalized, but not spectra between graphs.) .............................................................................................84

Figure 3.3: The imaginary part of the pseudo-dielectric function of Ge(111) for the various cleaning methods investigated in this study (done in collaboration with Seo [36]). The NH$_4$OH curve corresponds to the NCSU cleaning method; the HF curve corresponds to the Deegan and Hughes method; the H$_2$O$_2$/HF curve corresponds to the Chabal, et
al. method; and the Bare wafer curve corresponds to the as received substrates .................................................................86

Figure 3.4: Real-time measurements of the evolution of the (a) Ge(111) and (b) Ge(100) surfaces during two cycles of the NCSU cleaning method (done in collaboration with Seo [36]). The starting point is after an initial 5 min. H$_2$O bath.................................87

Figure 3.5: Real-time measurements of the surface re-oxidation on NCSU-cleaned Ge(100) and Ge(111) surfaces (done in collaboration with Seo [36]).........................................................................................88

Figure 3.6: AES spectra showing the chemical composition of the native and re-grown Ge(100) overlayers following several of the cleaning methods discussed in the text (done in collaboration with Seo [36]). The wide scan (a) shows the relative concentration of the chemical elements present at the surface, and the narrow scan (b) shows the relative levels of Ge-O and Ge-Ge bonding at the surface .........................89

Figure 3.7: AES spectra showing the chemistry of RPE-MOCVD-deposited thin films of (a) HfO$_2$ and (b) (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$........................................................................................................90

Figure 3.8: MEIS data for HfO$_2$ and (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ high-$\kappa$ materials deposited on plasma-nitrided Ge(100) substrates. Figures on the left are of HfO$_2$, and figures on the right are of (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$. Figures on the bottom correspond to as-deposited films, and figures on the top represent the films after RTA at 800°C. The top of the sample is at 0 nm, with the depth increasing into the sample. The vertical lines are for visual clarity, denoting transition regions. (Special thanks to K.B. Chung.) .........................91

Figure 3.9: Oxygen K$_1$ edge XAS spectra demonstrating the thermal evolution of HfO$_2$ and (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ thin films deposited on plasma-nitrided Ge(100) and Ge(111) surfaces.................93

Figure 3.10: Oxygen K$_1$ edge XAS spectra demonstrating the asymmetry of HfO$_2$ films deposited on Ge(100) and Ge(111) substrates and annealed at 800°C.................................................................94

Figure 3.11: The conduction band defect structure for thin films of (bottom) HfO$_2$ and (top) (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$
annealed at 800°C as deduced by 2nd derivative XAS. The bandgap assignment for HfO2 is taken from Wong and Iwai [13], and all other energies listed in the band diagrams are (529.8 eV → 5.7 eV).

Figure 3.12: Gaussian fitting of the discrete sub-band edge defect states in (a) HfO2 and (b) (HfO2)0.3(SiO2)0.3(Si3N4)0.4 thin films (fitting by K.B. Chung). The area of the Gaussian functions for the normalized spectra are plotted as a function of temperature in (c).

Figure 3.13: The ε2 portion of the dielectric functions for thin films of HfO2 and HfSiON, both as-deposited and annealed at 800°C (data taken and analyzed by Seo [36]).

Figure 3.14: (Left) Depth-resolved cathodoluminescence spectra showing defect level transitions in the bandgap of HfO2 films (CL data is from Brillson, et al. [45], © American Vacuum Society, 2007). (Right) The assigned valence band defect states consistent with the CL and XAS data.

Figure 3.15: XAS spectra for plasma nitried and plasma oxidized Ge surfaces as compared to plasma-nitried and plasma oxidized Si surfaces.

Figure 3.16: Relevant band diagrams for the Ge/HfO2 electrical devices studied in this chapter. Depicted above are (a) flat band conditions for GeO2, Ge3N4, and SiON ITRs, (b) a comparison of substrate injection between GeO2 and SiON ITRs, and (c) a comparison of gate and substrate injection for the GeO2 ITR.

Figure 3.17: Typical C-V hysteresis curves for (bottom) (HfO2)0.3(SiO2)0.3(Si3N4)0.4 and (top) HfO2 gate dielectrics on n-type Ge substrates in these experiments (measurement frequency is 1 MHz). The sweep direction is left-to-right-to left.

Figure 3.18: Typical C-V hysteresis curves for (bottom) (HfO2)0.3(SiO2)0.3(Si3N4)0.4 and (top) HfO2 gate dielectrics on n-type Ge substrates in these experiments (HfSiON was annealed at 650°C, and HfO2 was annealed at 600°C). The sweep direction is left-to-right-to left.
Figure 3.19: Anomalously low leakage current in (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$-based devices on Ge as a result of high series resistance (top). Evidence of the high series resistance is in the C-V curve (bottom). (Figures are from Lucovsky, et al. [12], © 2008 Elsevier B.V. All rights reserved.)

Figure 3.20: The effects of annealing on the I-V characteristics of the (a) HfO$_2$ and (b) (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ devices studied in this chapter. A Si device is shown in each plot as a reference for comparison. The flat band voltage used is only approximate as a result of the large C-V hysteresis.

Chapter 4: Hafnium Based High-κ Materials on Pseudomorphic Germanium Surfaces Grown on Silicon

Figure 4.1: The evolution of the Si/Ge interface as a function of Ge deposition time as measured by AES.

Figure 4.2: The nitrogen K$_1$ edge in Si/Ge/HfO$_2$ samples, both as-deposited and annealed. The interfacial Ge layer was nitrided before HfO$_2$ deposition. Note that a non-trivial amount of nitrogen remains after annealing at 700ºC.

Figure 4.3: XAS spectra showing variations in Si/Ge/high-κ samples due to substrate orientation, high-κ material, and annealing temperature. Figures (a) on the left depict (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ films, and figures (b) on the right depict HfO$_2$ films.

Figure 4.4: 2$^{nd}$ derivative XAS spectra showing the conduction band structures of Si/Ge/high-κ samples and their evolution with post-deposition annealing. The left-hand side depicts (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ films, and the right-hand side shows HfO$_2$ films. Note not only the d-state splitting terms, but the sub-band edge defects as well.

Figure 4.5: The energy band structures of the Si/Ge/high-κ samples.

Figure 4.6: The C-V characteristics for various Si/PM-Ge/HfSiON MOS capacitor structures as a function of interfacial Ge layer thickness and measurement sweep frequency.
Figure 4.7: The C-V hysteresis characteristics for various Si/PM-Ge/HfSiON MOS capacitor structures as a function of interfacial Ge layer thickness. The upper left-hand curve is from a Si/HfSiON device fabricated for reference purposes. The arrows indicate the sweep directions.

Figure 4.8: A summary of the hysteresis, frequency dispersion, fixed charge density, and interface trap density of the various Si/Ge/HfSiON devices studied in this chapter. The lines shown in the figures are simply cubic splines; they are meant only as a guide to the eye.

Figure 4.9: The 1 MHz C-V characteristics for the thinnest HfSiON layers deposited. Note the bump near the accumulation region.

Figure 4.10: A plot of measured EOT versus HfSiON deposition time. Note the non-zero intercept.

Figure 4.11: A summary of the gate leakage characteristics as a function of annealing temperature, RPAN time, and Ge deposition time. The leakage curve obtained from a Si/SiO₂ device with the same EOT (approximately 3 nm) is shown as a reference.

Figure 4.12: Some additional gate leakage I-V curves of interest.

Figure 4.13: Gate leakage characteristics for Si/Ge/HfSiON devices with EOT of approximately 20 Å.

Figure 4.14: Best fits to substrate injection data obtained from two different devices with the Si/Ge/HfSiON device structure. The fits both use a trap-assisted tunneling model, which is a combination of Frenkel-Poole and tunneling leakage mechanisms.

Figure 4.15: Several hypothetical band diagrams depicting some of the potential changes which may occur during gate dielectric deposition: (a) the idealized Si/Ge/high-κ band diagram, (b) intentional plasma nitridation of the interfacial Ge surface, (c) accidental conversion of the interfacial Ge surface into...
GeON, and (d) intentional plasma nitridation of a thin interfacial Ge surface, which unintentionally consumes the entire Ge layer and nitrides the top layer of the Si substrate as well ..................................................................................................................148

Chapter 5: Summary, Future Work, and Concluding Remarks

NONE

Appendix A: The Theory of Point Groups

Figure A.1: A pictorial depiction of (a) a counter-clockwise rotation by 90° around the z-axis, (b) a reflection across the plane normal to the <1, 1, 0> direction, and (c) inversion ..................173

Figure A.2: The symmetry of the ammonia molecule........................................174

Appendix B: Molecular Orbital Theory

Figure B.1: A representative molecular orbital diagram for a rutile transition metal oxide (modeled after P.A. Cox [1]) .........................192

Appendix C: Crystal Field Splitting and Jahn-Teller Distortions

NONE

Appendix D: The Structure and Polymorphisms of Group IV–B Transition Metal Oxides and Related Alloys

Figure D.1: A schematic representation of (a) the local site symmetry and (b) the atomic orbitals in rutile TiO₂. The blue atoms are titanium, and the green atoms are oxygen.................................217

Figure D.2: The oxygen K₁ x-ray absorption edge spectrum for titanium dioxide (data from Lucovsky, et al. [4]) ...............................218
Figure D.3: The removal of d-state degeneracy in an octahedral transition metal oxide complex, such as TiO$_2$ (modeled after Lucovsky, et al. [6]).................................219

Figure D.4: The idealized 7-fold transition metal local bonding environment in the baddeleyite structure. The blue atoms are Zr/Hf, and the green atoms are oxygen..........................220

Figure D.5: A schematic representation of (a) the local site symmetry and (b) the hypothetical atomic orbitals in 8-fold coordinated cubic ZrO$_2$ and HfO$_2$. The blue atoms are Zr/Hf, and the green atoms are oxygen..........................221

Figure D.6: The oxygen K$_1$ x-ray absorption edge spectrum for hafnium dioxide (data from Lucovsky, et al. [5]).............................222

Figure D.7: An approximate molecular orbital diagram for the electronic states of pseudo-cubic hafnia structure having 7-fold coordinated Hf ions........................................................................................................223

Figure D.8: The removal of d-state degeneracy in a pseudo-cubic, monoclinic transition metal oxide complex, such as HfO$_2$...........224

Figure D.9: The local arrangement of atoms in the idealized structure of pseudo-ternary (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$........................................225

Figure D.10: A schematic representation of (a) the local site symmetry and (b) the hypothetical oxygen 2p atomic orbitals in the pseudo-ternary, 4-fold coordinated, tetragonal (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$. The blue atoms are hafnium, and the green atoms are oxygen..........................226

Figure D.11: The oxygen K$_1$ x-ray absorption edge spectrum for pseudo-ternary (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ (data from Lucovsky, et al. [13]).................................................................227

Figure D.12: An approximate molecular orbital diagram for the electronic states of pseudo-ternary, tetragonal (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ structure having 4-fold coordinated Hf ions........................................................................................................228

Figure D.13: The removal of d-state degeneracy in a tetragonal transition metal oxide complex, such as (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ (modeled after Lucovsky, et al. [6]).................................229
Appendix E: A Description of Intermediate Phases and Self-Organization in Non-Crystalline Solids

Figure E.1: A schematic representation of floppy, stressed-rigid, and just-rigid modes in a non-crystalline, covalent network solid (modeled after Boolchand, et al. [5]) .................................................................245

Figure E.2: The glass forming difficulty in Ge_xSe_1-x as a function of germanium composition, x (from Phillips [1], © 1979 North-Holland Publishing Company) .........................................................246

Figure E.3: Site percolation on the square lattice resistor network described in the text (modeled after Zallen [7]). The percolation limit, p_c, is 0.5 and is one of the few analytic solutions known in percolation theory ........................................247

Figure E.4: Rigidity in a bond-diluted, generic two-dimensional network near the percolation limit. Dark regions indicate over-constrained regions, solid lines indicate rigid regions, and open circles indicate floppy regions. (Figure from Jacobs and Thorpe [8], © 1996 The American Physical Society.) .......................................................................................248

Figure E.5: The fraction of floppy modes as a function of average coordination using BCT (solid line) and a computer model for a randomly diluted diamond lattice (dot-dash line). The difference in the onset of rigidity between the two cases indicates an intermediate phase. (From Thorpe, et al. [8], © 2000 Elsevier Science B.V. All rights reserved.) .........................249

Figure E.6: The viscosity (left axis) and reversible heat flow (right axis) measurements indicating the existence of an intermediate phase in the As_xSe_1-x alloy system (after Boolchand, et al. [9]) .............................................................250

Appendix F: Auger Electron Spectroscopy (AES)

Figure F.1: The KVV Auger process responsible for electrons detected in AES experiments discussed in this dissertation .........................257

Figure F.2: A stereotypical experimental set-up for Auger electron detection in AES by a cylindrical mirror analyzer .........................258
Figure F.3: Example AES spectra for two different compositions of the pseudo-ternary compound, HfSiON. The elemental AES peak positions are identified, as are the pseudo-ternary compositions computed from the relative atomic concentrations ................................................................. 259

Figure F.4: The AES spectra resulting from the incremental oxidation of a clean Si surface by remote plasma ................................................................. 260

Figure F.5: The oxidation kinetics of a clean Si surface by remote plasma. This process follows the model for silicon surface oxidation very well, as indicated by the fitting function in the lower right corner [5, 6] ................................................................. 261

Appendix G: X-Ray Absorption Spectroscopy (XAS)

Figure G.1: Representative depictions of the x-ray absorption processes detectable in the TEY: (a) core level excitation into an anti-bonding state, relaxed by Auger electron emission, and (b) core level excitation to the vacuum level (photoelectron emission), relaxed by fluorescent photon emission. Of course, these excitation/relaxation processes do not have to be paired as they are above ................................................................. 269

Figure G.2: A schematic representation of beamline 10-1 at SSRL ............. 270

Figure G.3: Degeneracy removal as a function of electronic state symmetry as it relates to transition metal compounds (modeled after Lucovsky, et al. [10]) ................................................................. 271

Figure G.4: XAS data showing how d-state degeneracy removal is observed spectroscopically (data from Lucovsky, et al. [10]) ............................................................................................................... 272

Figure G.5: Complete d-state degeneracy removal in HfO₂ as seen in the 1st derivative XAS spectrum (data from Lucovsky, et al. [10]) ............................................................................................................... 273

Figure G.6: The evolution of the type I nano-crystalline to type II nano-crystalline transition as observed by XAS in HfO₂ on Ge(100) ............................................................................................................... 274
Figure G.7: The conduction band defect states in a thin film of HfO$_2$ on Ge(100) as detected by XAS (figure from Seo [11]) .................275

Appendix H: Spectroscopic Ellipsometry (SE)

Figure H.1: The experimental set-up for conducting rotating-compensator ellipsometry. The light source is on the left. Variable-frequency incoming light is passed through a polarizer, reflected by the sample, and passed through a rotating compensator and an analyzer. The final intensity of the beam is measured by the detector .................................................285

Figure H.2: A non-homogeneous thin film which can be described by an effective medium theory .......................................................286

Figure H.3: A flow chart representing the method for solving for the dielectric function of a sample measured by SE .........................287

Figure H.4: An example of the absorption constant for a HfO$_2$ sample as measured by SE. The defect states clearly are identified in this measurement (figure taken from Takeuchi, et al. [5], © 2004 American Vacuum Society) .......................................................288

Figure H.5: A representative MO diagram for a transition metal oxide, such as TiO$_2$, with octahedrally coordinated transition metal atoms and oxygen vacancies. The oxygen vacancies are shown to be within the bandgap at the atomic d state energy. These are the states responsible for the sub-band edge defect levels observed in figure H.4 .................................................289

Appendix I: Medium Energy Ion Scattering (MEIS)

Figure I.1: A schematic representation of the basic ion scattering process in MEIS. In this picture, an ion from the incident ion beam impacts an atom near the surface of a thin film ..........294

Figure I.2: A representative atomic channel for an ion used in channeling ..................................................................................295

Figure I.3: The process of scattered ion blocking. The scattered ions which are not de-channeled essentially are blocked by other atoms in the sample as they try to exit it .........................................296
Figure I.4: The experimental set-up for MEIS (top) and ion detection (bottom). (The above figures are from Baumvol [1], © 1999 Elsevier Science B.V. All rights reserved.) ............................297

Figure I.5: An example of depth profiling for a nitried thin film of SiO$_2$ on a crystalline Si substrate. The inset shows the depth profile for the ion energy spectrum shown in the main body of the graph. (The figure is taken from Gusev, et al. [6], © 2001 Elsevier Science B.V. All rights reserved.) .........................................................................................298

Appendix J: Capacitance-Voltage (C-V) Measurements

Figure J.1: Ideal C-V curves for each of the four possible regions (i.e. accumulation, inversion, depletion, and deep depletion). The C-V curves shown here were computed for an ideal $n$-type Si MOS structure with a substrate doping level of $1 \times 10^{17}$/cm$^3$, an area of 10µm x 10µm, and an EOT of 10 nm. The device has aluminum contacts .................................................................309

Figure J.2: The effects of interface and oxide trapped charges on the classical C-V characteristics of a MOS capacitor. The C-V curves shown here were computed for an $n$-type Si MOS structure with a substrate doping level of $1 \times 10^{17}$/cm$^3$, an area of 10µm x 10µm, and an EOT of 10 nm. The device has aluminum contacts .............................................................................310

Figure J.3: A comparison of the most commonly used capacitance modeling programs for calculating C-V curves of MOS devices in the ultrathin regime (from Richter, et al. [9], © 2001 IEEE) ........................................................................................................311

Figure J.4: The effects of series resistance on the capacitance-voltage characteristics of a MOS device (after Yang and Hu [10]) ...........312

Appendix K: Current-Voltage (I-V) Measurements

Figure K.1: Band diagrams depicting six of the most common gate leakage mechanisms. The flat band condition is shown as a reference in (a). The leakage mechanisms are: (b) direct tunneling, (c) Fowler-Nordheim tunneling, (d) thermionic
emission, (e) thermionic-field emission, (f) trap-assisted tunneling, and (g) Poole-Frenkel emission

Figure K.2: Direct tunneling and Fowler-Nordeim tunneling in Si/SiO$_2$ MOS devices with varying oxide thickness (data from Schuegraf and Hu [4]). The change in shape of the I-V characteristics shows that the onset of direct tunneling occurs at roughly $t_{ox} \leq 3.5$ nm.
Chapter 1

Introduction

1.1 – The Semiconductor Roadmap

The International Technology Roadmap for Semiconductors (ITRS) is an international panel, containing experts from Europe, Taiwan, Korea, Japan, and the US, whose mission is to perpetuate the various advancement trends in the semiconductor industry. The ITRS addresses all aspects of the industry, from device design and manufacture, to device reliability and testing, to the development of new materials. Since many of these developments are coupled, the ITRS regularly examines the areas in which progress is being made and those in which more work is needed, constantly creating and revising relevant benchmarks and timelines to ensure that the industry as a whole advances [1].

One of the most famous trends addressed by the ITRS is device scaling, known as Moore’s Law. In 1965, Gordon Moore stated that the number of bits available for memory on integrated circuits (ICs) would double every year, and that trend has continued, remarkably, for decades since [2]. In their present form, the basic idea behind device scaling is to make memory elements smaller by a factor, $\alpha$, and to increase the electric fields within devices by a factor, $\varepsilon$, without sacrificing the quality of device performance metrics [3].
Of particular interest is the relationship between the channel length and gate capacitance in complimentary metal-oxide-semiconductor (CMOS) transistors. The gate length has a scaling factor of $1/\alpha$; whereas, gate capacitance has a scaling factor $\alpha$. Hence, as gate length is reduced to allow for more memory elements on a microchip, the gate capacitance must be increased by a factor of $\alpha$ in order to maintain similar current-voltage (I-V) characteristics [3-5]. Treated as an ideal planar capacitor, the gate of a CMOS device has a capacitance given by the following simple relationship (in SI units):

$$C = \frac{\kappa \varepsilon_0 A}{d}$$  \hspace{1cm} (1.1)

in which $C$ is the gate capacitance, $\varepsilon_0$ is the permittivity in free space, $\kappa$ is the dielectric constant of the gate material, $A$ is the device area, and $d$ is the gate’s physical thickness.

In order to achieve smaller devices having higher gate capacitance, this leaves only two options for gate design: (1) decrease the physical thickness of the gate, or (2) increase the dielectric constant of the gate material. The relevant metric, then, is the equivalent oxide thickness (EOT), defined by the equation:

$$EOT = \left( \frac{\kappa_{SiO_2}}{\kappa} \right) d$$ \hspace{1cm} (1.2)

Figure 1.1 depicts the scaling trend in EOT as a plot of EOT versus time.
For decades, option (1) has been simple and reliable (as indicated in figure 1.1). The gate of choice for silicon CMOS devices has been SiO₂, which leads to many very desirable properties in this material system, such as low stress-induced leakage (SILC), low interface trap density (D_it), low fixed charge density (Qox), small capacitance-voltage (C-V) hysteresis, and low C-V frequency dispersion [2]. As shown by figure 1.1, however, the reduction in the physical thickness of SiO₂ gates has reached its ultimate limits (approximately 0.8-1.2 nm), as dictated by quantum mechanics and chemical bonding [5, 6].

With option (1) no longer available, the only remaining option is to increase the dielectric constant of the gate material. Since this is fundamentally a material-dependent quantity, the semiconductor industry has been forced to explore new materials for use as gates in CMOS devices, which have a higher dielectric constant than SiO₂ (in which \( \kappa = 3.9 \)). Unfortunately, the same physical mechanisms that give such materials a higher dielectric constant often lead to a reduction in device properties, such as D_it, Qox, and C-V hysteresis and frequency dispersion [2-6]. These problems are difficult to overcome, but doing so is vital to the progress of the semiconductor industry. As a result, intensive research into so-called high-\( \kappa \) gate dielectrics is being pursued.

CMOS scaling trends also make an increase in device mobility very desirable. The easiest way in which to illustrate this fact is to examine the drive current associated with a field-effect transistor (FET), which is given by the following expression [7]:

\[
I_D = \frac{W}{2L} \mu C_s V_D^2
\]
In equation 1.3, $I_D$ is the drain current, $W$ is the transistor channel width, $L$ is the channel length, $\mu$ is the channel mobility, $C_{inv}$ is the capacitance per unit area of the gate when the channel is inverted (i.e. when a $p$-type channel is full of electrons or an $n$-type channel is full of holes), and $V_G$, $V_T$, and $V_D$ are the gate, threshold, and drain voltages, respectively. A large $I_D$ is desirable for a variety of reasons, such as improved speed and device performance. This can be achieved by scaling EOT and channel length, as discussed earlier; however, an increase in the channel mobility, $\mu$, will also increase the drain current (as in equation 1.3). Furthermore, as channel thickness is reduced, the mobility of silicon decreases, as shown in figure 1.2 [3]. As mentioned above, the reduction in channel length requires an increase in gate capacitance, yet an increase in gate capacitance leads to an increase in the electric field, which in turn, gives rise to a narrower channel. Hence, in the process of device scaling, there comes a point at which an undesirable reduction in channel mobility occurs. As figure 1.2 illustrates, this happens for an EOT below approximately 10 nm, placing current technology well within this realm.

By increasing the channel mobility in CMOS devices, the potential benefits include not only the offset of this collapse in channel mobility, but reduced power consumption, enhanced device performance, and improved charge transport properties as well [8]. One approach in increasing channel mobility has been strain-engineering of
silicon channels, but using alternative substrates with larger mobilities may provide an even better solution. One such material is germanium.

1.2 – Possible Benefits and Drawbacks of Using Germanium Substrates

The primary benefit in using germanium substrates for microelectronic applications is that it has a higher mobility for both holes and electrons in the bulk than does silicon. Figure 1.3 shows these mobility curves for visual comparison. Note that germanium shows up to a ten-fold increase for electron mobility in places. As a result, germanium is regarded as one of the most promising channel materials for incorporation into future devices [8-12].

Looking at germanium in the context of effective mass theory alone, it appears very promising, indeed. First of all, both holes and electrons have lower effective masses in germanium than in silicon, which enhances the saturation velocity of the carriers, thereby increasing the on current of the device. Germanium also has ideal anisotropy in its effective mass; the effective mass in the $x$ direction is much lower than it is in the $z$ direction. This leads to an increased saturation velocity along the channel for carriers, but it also leads to a reduced inversion layer thickness and a lower effective EOT [8].

Germanium substrates are not without problems, however, many of which are intrinsic and unavoidable. Perhaps the most fundamental of these, and certainly the most relevant to the work contained here, is the lack of a suitable gate dielectric material for germanium substrates. Making an analogy with silicon, the most chemically compatible material should be the natural oxide, GeO$_2$; however, this material is frankly a train
wreck from the point of view of microelectronics, causing immense problems for virtually every phase of device fabrication and ultimately leading to poor device performance.

First of all, the germanium surface is incredibly reactive with oxygen, forming native GeO$_2$ layers almost immediately upon exposure to air (see chapter 3). This makes chemical surface preparation impossible in ambient conditions, and it makes engineering the Ge/GeO$_2$ interface very difficult, to say the least. Thus, native oxide layers pose serious problems even at the onset of device fabrication. In is also worth noting that GeO$_2$ comes in three varieties: (1) a tetragonal rutile form, (2) a hexagonal quartz modification, and (3) a glassy amorphous form [13, 14]. Commonly, however, more than one of these forms can, and generally will, be present concurrently in GeO$_2$ films, leading to mechanical and electrical defects [14]. Moreover, because materials (2) and (3) are water-soluble, conventional wet chemical processing impossible on this material [13].

The thermal properties of the Ge/GeO$_2$ system are a nightmare as well. The germanium surface actually reacts with the GeO$_2$ overlayer, decomposing thermally into gaseous GeO via the following reaction [13,15]:

$$\text{Ge} + \text{GeO}_2 \xrightarrow{\text{heat}} 2\text{GeO}$$

This thermal decomposition occurs at approximately 430ºC, as illustrated dramatically by the x-ray photoemission spectroscopy (XPS) results shown in figure 1.4 [15]. As a result, the thermal budget for devices utilizing a Ge/GeO$_2$ structure severely limits processing
options during fabrication even further. Finally, it should be mentioned that devices fabricated with the Ge/GeO$_2$ gate structure have demonstrated fixed charge density of $5 \times 10^{11}$/cm$^2$ and interface state density of $1 \times 10^{12}$/cm$^2$/eV [14]. These values simply are far too high to be competitive with silicon in any way.

Since GeO$_2$ clearly is doomed to failure as a gate material on germanium substrates, the next most likely candidate for this task is Ge$_3$N$_4$ or a GeON alloy. These materials have demonstrated far more promise as gate dielectrics for germanium substrates than pure GeO$_2$, but they are not without their own set of problems.

First of all, like GeO$_2$, Ge$_3$N$_4$ is thermally unstable and decomposes fully at around 650$^\circ$C. This decomposition is shown by the XPS spectra in figure 1.5 [16]. At around 580$^\circ$C, nitrogen begins to leave the material, and by 640$^\circ$C, almost all of the nitrogen has been removed. Because it is simply an alloy of GeO$_2$ and Ge$_3$N$_4$, GeON is expected to begin decomposing at around 430$^\circ$C and to decompose fully by 650$^\circ$C. In addition to their poor thermal properties, Ge$_3$N$_4$ and GeON have water-soluble forms. Experiments have shown that just like GeO$_2$, Ge$_3$N$_4$ has both water-soluble and water-insoluble forms. Furthermore, GeON, depending on the deposition process, is likely to have mixed solubility properties that favor either GeO$_2$ etching or Ge$_3$N$_4$ etching [17]. Lastly, both Ge$_3$N$_4$ and GeON have been shown to have very poor C-V and I-V properties [17, 18]. Unlike GeO$_2$, however, both Ge$_3$N$_4$ and GeON do have one good property. Namely, they may serve as excellent passivation layers, as both have demonstrated good resistance to subsequent oxidation [17, 18].
It is clear from the above discussion that, unlike their silicon-containing counterparts on silicon substrates, GeO2, GeON, and Ge3N4 all make very poor choices as gate dielectrics on germanium substrates. The one positive than can be said for these materials, however, is that Ge3N4 and GeON are useful as germanium passivation layers. Fortunately, these are not the only choices for gate dielectric materials available. High-κ transmission metal oxides are another avenue well worth exploring, deposited either directly onto a germanium surface or onto a nitrogen-passivated germanium surface. The possibility of using transition metal oxides for Ge-based devices is still an open question, but much is known about these materials on Si substrates. It is important, therefore, to understand the properties of high-κ transition metal oxides on Si surfaces before extending those principles into another new realm.

1.3 –High-κ Transition Metal Oxides as Gate Dielectrics on Si

As briefly mentioned in section 1.1, the scaling of SiO2 has reached its fundamental limits. Further reduction in the thickness of SiO2 gates leads to a host of problems, such as reduced drive current, higher leakage current, and reliability degradation. Furthermore, it is difficult to grow uniform layers at those thicknesses [6,19]. In fact, these ultrathin layers of SiO2 are forced to change their stoichiometry to that of a suboxide, SiOx<2, when they are reduced below 7 Å [7].

The existence of this suboxide region has been verified independently by several groups through the use of surface XPS and spectroscopic ellipsometry (SE) [20-22]. Perhaps the most conclusive of these results came from Himpsel, et al. in 1988 [20],
though Aspnes, et al. demonstrated its existence as early as 1979 [22]. Himpsel’s XPS spectra are shown in figure 1.6 for both Si (100) and Si (111) substrates, and they demonstrate very clearly that there are silicon atoms bonded to 1, 2, and 3 oxygen atoms in this region (less than the 4 in SiO\textsubscript{2}). This suboxide region appears to be fundamental and related to strain at the Si/SiO\textsubscript{2} interface. The XPS spectra change following rapid thermal annealing (providing strain relaxation), but the suboxide signatures remain [21]. These results suggest that this region is intrinsic and unavoidable.

Because further scaling is no longer an option using the conventional materials, much attention has been focused on high-\(\kappa\) dielectrics [2, 5, 23, 24]. As equation 1.1 shows, a larger dielectric constant can provide a gate capacitance equal to that of a SiO\textsubscript{2} layer while having a larger physical thickness. Unfortunately, a larger dielectric constant is not the only requirement for a SiO\textsubscript{2} replacement. Alternative gate materials for CMOS applications must also demonstrate the following properties: (1) thermodynamic stability with the silicon substrate, (2) stability against processing temperatures as high as 1000ºC for 5s (as current processing necessitates), (3) act as an insulator with conduction and valence band offsets of at least 1 eV to prevent carrier injection into the gate, (4) form a high-quality bonding interface with silicon to prevent defects and a reduction in carrier mobility, (5) have few electrically active defect states, preferably below 10\textsuperscript{11}/cm\textsuperscript{2}/eV, and (6) be compatible with a suitable gate electrode material [23, 25]. Though there exists a variety of materials satisfying condition (3), the vast majority of available materials do not satisfy the remaining requirements.
Consider first the issue of thermodynamic stability with the silicon substrate, which has been addressed in several papers [23, 24, 26, 27]. In these, it is shown that silicate and silicide formation is thermodynamically favored at high-κ/Si interfaces during annealing at 1000°C. Stremmer, for example, has discussed and investigated the issue of silicide formation due to reactions between ZrO$_2$ and Si for ZrO$_2$/Si interfaces [24]. There exists a variety of thermodynamically favorable reactions for the formation of a ZrSi$_2$ silicide predicted to occur at 1000°C in these systems, the two most likely reaction pathways being (1) the reaction of ZrO$_2$ with gaseous SiO and (2) the reaction of Si with ZrO$_2$ via oxygen vacancies. Though these reactions are difficult to confirm experimentally, the existence of the ZrSi$_2$ has been confirmed in high resolution transmission electron microscopy (HRTEM) experiments, and a similar HfSi$_2$ silicide has been observed at the HfO$_2$/Si interface. A result from one such experiment, conducted on ZrO$_2$/Si by Stremmer, is shown in figure 1.7, demonstrating quite clearly the presence of ZrSi$_2$ following a 1000°C anneal. The existence of such regions at these interfaces degrades electrical performance substantially, shorting out the field effect as a result of their metallic properties and leading to interface traps [5, 23].

In addition to thermally driven reactions between the gate dielectric and the substrate, the thermodynamic stability of the gate itself is a concern. Consider the ability of high-κ dielectric materials (specifically Zr/Hf transition metal oxides and silicates) to resist crystallization and phase separation. This topic has been studied extensively through the use of x-ray absorption spectroscopy (XAS), extended x-ray absorption fine structure (EXAFS), x-ray diffraction (XRD), Fourier transform infrared spectroscopy
(FTIR), and HRTEM [28-33]. It has been shown by both EXAFS and XAS that thin films of HfO$_2$ deposited at low temperatures (<400ºC) are not non-crystalline as in the case of SiO$_2$ [28, 29]. Instead, these films are nano-crystalline as-deposited, showing evidence of the monoclinic phase typical of crystalline HfO$_2$, and the size of these crystallites increases with the temperature of rapid thermal annealing (RTA) treatments. For very thin films of HfO$_2$ (less than approximately 3-4 nm), however, the growth of crystalline regions is suppressed by size effects. In these cases, the film is simply so thin that it lacks the room required to form the monoclinic unit cells in one direction, inhibiting the overall size of the crystallites. Above this mesoscopic regime, the crystallites are free to grow just as in the bulk material. Hence, while thicker films of HfO$_2$ clearly crystallize, leading to the formation of grain boundaries and defects, there remains the possibility of using very thin nano-crystalline films as gate dielectrics.

Similar crystal growth suppression effects have been observed in both zirconium and hafnium silicates [29-31]. These ([$\text{Zr,Hf}O_2$]$_x$($\text{SiO}_2$)$_{1-x}$ material systems are unstable against chemical phase separation for all values of x, separating into [$\text{Zr,Hf}O_2$] and SiO$_2$ regions during RTA processing at 900ºC. The evolution of these films as a function of temperature is illustrated by the FTIR and XRD spectra shown in figure 1.8a and 1.8b, respectively, which clearly indicate the sharp crystalline transition between 800ºC and 900ºC. The HRTEM micrograph shown in figure 1.8c is representative of the chemical phase separation into SiO$_2$ and ZrO$_2$ regions and the subsequent crystallization of the ZrO$_2$ regions. For SiO$_2$-rich films, the [$\text{Zr,Hf}O_2$] regions are encapsulated as inclusions in a non-crystalline SiO$_2$ matrix, and the size of the inclusions is determined by global
strain minimization requirements. For concentrations of [Zr,Hf]O₂ less than 25%, these considerations limit the crystalline regions to the mesoscopic size regime (< ~3 nm) exactly as in the thin HfO₂ films mentioned above.

Though these transition metal oxides and silicates demonstrate at least some crystalline order in all of their forms following the required processing temperatures, there does exist a special non-crystalline, self-organized phase in a transition metal nitro-silicate, namely ([Ti,Zr,Hf]O₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄ (see appendix D). Films composed of other relative concentrations in this nitro-silicate material system do phase separate and crystallize in the same ways as their silicate counterparts [32, 33]. This particular alloy (to within 5% or so of these relative concentrations), however, does not chemically phase separate or crystallize following anneals at temperatures up to and including 1100°C [31, 33]. These materials are special because of a confluence of two separate phenomena – one which minimizes global strain, and one which forces the local bonding into an optimal configuration (see appendix E). As a result, the films become locked into a non-crystalline state, much like SiO₂. In many ways, (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄ can be thought of SiO₂ with a dielectric constant of 12.

Consider now the condition requiring that high-κ materials form high-quality interfaces with the silicon substrate. Again, this requirement exists in order to ensure that there is negligible carrier mobility degradation in inversion layers due to surface roughness scattering and surface dipole scattering [25]. The lattice constants for HfO₂ and ZrO₂ fall in the range between 5.07 Å and 5.16 Å, as compared to 5.43 Å in crystalline Si. Hence, from a lattice mismatch perspective, one might expect epitaxial
HfO₂ and ZrO₂ to form reasonably high-quality interfaces on Si; however, this is not the case. In general, these transition metal atoms are quite ionic and can have several different bonding coordinations. Both of these characteristics contribute to more dangling bonds and bonding defects, thereby increasing surface roughness scattering. Furthermore, the ionic nature of the transition metal bonds creates a strong surface dipole and exacerbates surface dipole scattering [5, 23].

There are defects in the bonding of bulk transition metal oxides as well. Spectroscopic investigations by XAS, SE, and photoconductivity (PC) experiments have shown that there are intrinsic, electrically active band-edge defects in these materials, attributable to crystal field (CF) splitting and Jahn-Teller (J-T) distortions [34-37]. Furthermore, there exist oxygen vacancies and interstitials in these films, which is largely a result of the ionic nature of their bonds [23]. These discrete defects create trapping states in the band gap as shown in figure 1.9. In addition to creating electrically active traps, oxygen vacancies can also act as precursors for silicide and silicate formation [24]. The O vacancies can be reduced by annealing in hydrogen, although this increases the fixed charge in the films [5]. Annealing in oxygen can also reduce these vacancies, but the degree of crystallization is increases, creating even more shallow traps through grain boundary defects [5]. All of these traps can lead to increased leakage, poor reliability, bias temperature instability (BTI), and charge trapping (hysteresis) [23, 25, 38].

Another requirement for high-κ gate materials is that the interface trap density be low (on the order of 10¹¹/cm²/еВ). There have been several claims for achieving this benchmark with transition metal oxides in direct contact with silicon substrates; however,
the dielectric constant in such films has invariably been much lower than in the bulk oxide, or the EOT has been larger than anticipated [39, 40]. This is due to the oxidation of the substrate and the formation of an SiO$_x$ layer as a result [7, 39]. More accurate measurements of $D_{it}$ have been done on ZrO$_2$ and HfO$_2$, for example, and it was found that these films were at levels of $10^{12}$/cm$^2$/eV [5] – a full order of magnitude larger than required for CMOS applications. These poor results are likely due to poor bonding quality and silicide/silicate formation at the Si/high-$\kappa$ interface [5, 7, 23, 24].

Consider finally the issue of gate electrodes for high-$\kappa$ gate materials. Current silicon processing uses highly doped poly-silicon materials as gates; however, these are fundamentally incompatible with the transition metal oxides [23]. Silicide formation occurs at the interface, and the poly-silicon grains are too large, both of which result in leakage paths. Metal gates are almost certainly required for transition metal oxides if they are to be used at all for CMOS [23]. Metals that place the Fermi level in the middle of the Si bandgap (mid-gap metals) have been explored for use on both $n$- and $p$-type devices to keep the processing simple, but the threshold voltage is simply too high with the Fermi level placed at mid-gap [7]. Using two separate metals has also been explored. Unfortunately, most metals react with the gate material at the required processing temperatures [7]. There are several options, such as Ta and TaN, for $n$-MOS. The options for $p$-MOS are Pt and Au, though both of these are expensive and difficult to process.

To summarize the status of the high-$\kappa$ transition metal oxides and silicates with respect to the criteria mentioned earlier:
(1) **Thermodynamic Stability with the Silicon Substrate** – The high-κ transition metal oxides form silicides and silicates at the silicon/gate interface during the required processing temperatures.

(2) **Stability Against Processing at 1000ºC** – The high-κ transition metal oxides crystallize at this temperature; the high-κ transition metal silicates phase separate and crystallize at this temperature.

(3) **Act as an Insulator with at Least 1 eV Band Offset on Si** – This criterion is no problem for these materials.

(4) **Form a High-Quality Interface with Si** – The ionic nature of the bonds in these materials make them more prone to dangling bonds and defects, increasing surface roughness scattering and surface dipole scattering.

(5) **Exhibit Less Than 10^{11}/cm²/eV Interface Traps** – Silicide formation and bonding defects lead to numbers more like 10^{12}/cm²/eV.

(6) **Be Compatible with a Suitable Gate Electrode** – While there are options, they are few in number, and some are expensive and hard to process.

Based on the above discussion, transition metal oxides and silicates do not seem like viable options at this point in time for silicon CMOS. However, recall the Zr/Hf/Ti nitro-silicate mentioned earlier. In fact, because of its unique self-organized phase, this material is known to satisfy requirements (2) and (3), and it has the potential to satisfy the
remaining four [41]. Hence, the nitro-silicate is still regarded as a viable option for silicon CMOS technology.

1.4 – The Possibility High-κ Transition Metal Oxides on Ge Surfaces

Knowing and understanding the properties of HfO₂ and HfSiON, as well as the properties of their interfaces with silicon, provides a good baseline for exploring their usage on high-mobility germanium substrates. This is one area explored by this dissertation; however, there is another focus as well. In addition to its higher mobility, germanium has another potential advantage over silicon.

As mentioned in section 1.3, there is an inherent suboxide interfacial transition region (ITR) at the Si/SiO₂ interface that is approximately 6-8 Å thick. This is one major problem in fabricating a Si/high-κ interface in which the two constituent materials are in direct contact. Because the SiOₓ layer has a low dielectric constant, the capacitance of this layer contributes significantly to the total gate capacitance in devices striving for minimal EOT.

As discussed in section 1.2, Ge₃N₄ and GeO₂ are thermally unstable. In Ge₃N₄, the nitrogen leaves the film at around 650°C as molecular N₂ [16], and in GeO₂, the oxygen leaves the film as gaseous GeO at around 430°C [15]. Given this information, the prospect of using nitrogen as a passivation layer on Ge substrates was investigated. It was observed that the interfacial nitrogen was liberated from a nitrided Ge/high-κ device structure following an RTA treatment, leaving the Ge and high-κ materials in direct
contact [42]. This observation opens an opportunity for silicon device scaling in which the suboxide layer is eliminated.

By depositing a pseudomorphic (PM) Ge layer onto a silicon substrate, the silicon suboxide ITR will not form, and upon post-deposition annealing (PDA) by RTA, the Ge/high-κ ITR will be eliminated, placing the germanium in direct contact with the gate material. Since the germanium layer should be semiconducting, it should provide no significant contribution to the gate capacitance. Without the suboxide region, the EOT of the device should be lowered by 6-8 Å.

Epitaxial Ge layers have been deposited on Si substrates by a number of techniques, including molecular beam epitaxy (MBE) [43], chemical vapor deposition (CVD) [44, 45], and remote plasma-enhanced chemical vapor deposition (RPECVD) [46], suggesting that a Si/PM-Ge/high-κ device structure is feasible. Moreover, band structure calculations have been carried out which suggest that the valence band offset between PM-Ge and Si (100) is $E_{v,PM-Ge} - E_{v, Si(100)} = 0.74$ eV, leading to a conduction band offset of approximately $E_{c,PM-Ge} - E_{c, Si(100)} = 0.28$ eV [47, 48]. As a result, there should be no energy barrier to holes and only a slight energy barrier to electrons during substrate injection.

1.5 – Overview of the Dissertation

Semiconductor scaling trends have been followed for decades, but they now face increasingly difficult obstacles to overcome if they are to continue. The purpose of this dissertation is to discuss studies of various electrical and spectroscopic
properties of Hf-based high-κ dielectric thin films on germanium and pseudomorphomic germanium surfaces – materials and interfaces that are potentially very important if semiconducto scaling trends are to survive much longer. In particular, the Hf-based high-κ materials referred to in these discussions are HfO₂ and (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄. These investigations are meant to (1) probe the physical/chemical properties of those structures, (2) correlate electrical properties with those physical/chemical properties, and (3) seek potential engineering pathways that may lead to improvement to future devices.

Because the construction and modification of laboratory equipment has been integral to being able to conduct these studies, a discussion of this work is included as Chapter 2. Bulk germanium possesses carrier mobilities higher than silicon, but it lacks a good native insulating material. As a result, Ge/high-κ devices and interfaces are discussed in Chapter 3 in order to vet the possibility of using high-κ gates in future Ge-based devices. PM-Ge surfaces may be able to eliminate the interfacial suboxide region in Si/high-κ stack structures, thereby reducing the EOT in Si-based devices by as much as 6-8 Å. Hence, Si/PM-Ge/high-κ devices and interfaces are studied in Chapter 4 as one possible way to further scale silicon CMOS. Chapter 5 concludes with a summary and indicates the areas in which more study is needed.

The theoretical framework for these investigations includes point group theory, molecular orbital theory, crystal field theory, bond constraint theory (BCT), and percolation theory. Appendix A introduces the theory of point groups, which is followed by a discussion of molecular orbital theory in appendix B and a description of crystal field theory and Jahn-Teller distortions in appendix C. Appendix D discusses several
geometries of the group IV-B transition metal oxides and related alloys, applying various ideas from appendices A-C to them. BCT and percolation theory are discussed in appendix E.

The experimental techniques used in order to investigate these materials and devices are as follows: (1) Auger electron spectroscopy (AES), (2) x-ray absorption spectroscopy (XAS), (3) spectroscopic ellipsometry (SE), (4) medium energy ion scattering (MEIS), (5) capacitance-voltage (C-V) measurements, and (6) current-voltage (I-V) measurements. Appendices F-K, respectively, introduce and explain these techniques.
Figure 1.1: Scaling in the EOT for CMOS devices, both historically and as projected by the ITRS (after Wong and Iwai [5]). Note that alternative high-κ dielectrics are now the only available option for future scaling.
Figure 1.2: Measured channel mobilities in silicon CMOS devices (after Frank, et al. [3] and references therein). Note the decline in mobility for channels with thickness less than approximately 10 nm.
Figure 1.3: A comparison of electron and hole mobilities in both silicon and germanium (data from refs. [9-12]).
Figure 1.4: X-ray photoemission spectra showing the thermal decomposition of GeO$_2$ into gaseous GeO (from Prabhakaran, et al. [15], © 2000 Elsevier Science S.A. All rights reserved.).
**Figure 1.5:** X-ray photoemission spectra showing the thermal decomposition of pure Ge$_3$N$_4$ formed by plasma nitridation of the Ge surface. (from Kutsuki, et al. [16], © 2008 The Japan Society of Applied Physics).
Figure 1.6: Core level XPS spectra showing the structure of the suboxide region at the Si/SiO₂ interface on both Si (100) and Si (111) substrates (from Himpsel, et al., © 1988 The American Physical Society).
Figure 1.7: Cross-section HRTEM micrographs showing the existence of the ZrSi$_2$ silicide phase in ZrO$_2$/Si interfaces after annealing at 1000°C (from Stemmer [24], © 2004 American Vacuum Society).
Figure 1.8: The thermal stability of zirconium silicate films as demonstrated by (a) an FTIR spectrum showing crystallization in a (ZrO$_2$)$_{0.6}$(SiO$_2$)$_{0.4}$ film after a 900°C RTA, (b) an XRD spectrum showing crystallization in a (ZrO$_2$)$_{0.5}$(SiO$_2$)$_{0.5}$ film after a 900°C RTA, and (c) high-resolution TEM micrograph showing chemical phase separation and crystallization in a (ZrO$_2$)$_{0.5}$(SiO$_2$)$_{0.5}$ film after a 900°C RTA (from Rayner, et al., figures 1.8b and 1.8c © 2003 American Vacuum Society). Figures 1.8a and 1.8b show a sharpening of several features at 900°C, and the circled area of figure 1.8c shows streaking. These signatures all are indicative of crystallinity. The light and dark regions of figure 1.8c show the chemical phase separation.
**Figure 1.9:** The molecular orbital diagram for oxygen interstitials and vacancies in various charge states (from Robertson [23], © 2006 IOP Publishing, Ltd.). $V_0$ represents the doubly-occupied neutral vacancy with $A_1$ symmetry. $V_0^-$ represents this vacancy with a trapped electron, pulling down an extra state with $B_1$ symmetry from the conduction band. $V_0^+$ represents the singly-occupied vacancy with $A_1$ symmetry. $V_0^{2+}$ represents the unoccupied vacancy with $A_1$ symmetry. $I_0^-$ represents the interstitial with a trapped hole. $I_0$ represents the neutral interstitial, which forms an O-O bond and creates a $\sigma^*$ state in the band gap. Lastly, $I_0^+$ represents the neutral oxygen interstitial which has formed an O-O bond and subsequently trapped a hole.
References


[41] G. Lucovsky (private communication).


2.1 – RPECVD Process Overview

The thin film deposition tool used in these investigations was originally designed to deposit SiO$_2$- and Si$_3$N$_4$-based by remote plasma-enhanced chemical vapor deposition (RPECVD). In this technique, very low flow rates of reactant gases are used, and the films are grown at low temperature (approx. 300ºC). Chemical reactions occur between plasma-excited radicals much like in a direct plasma-enhanced process, except that the plasma is kept away from the surface of the film. The combination of low flow rates, the remote nature of the plasma and the low deposition temperature leads to very slow, controlled, and reproducible film growth. While as-deposited thin films of SiO$_2$ using this method have very high levels of $D_{it}$ (on the order of $10^{12}$/cm$^2$/eV), a simple PDA using RTA relaxes the material and exhibits $D_{it}$ on the order of $10^{10}$-$10^{11}$/cm$^2$/eV [1-4].

An original schematic of the RPECVD cluster tool at North Carolina State University (NCSU) used in these experiments is shown in figure 2.1. The films are deposited in the RPECVD chamber, and AES characterizations of the sample surfaces are carried out in the analysis chamber. The chamber labeled RTP is not used. Figure 2.2 shows a rendering of the RPECVD gas injection system. Metal precursors are brought in through gas dispersal rings, kept away from the plasma region (downstream introduction), and all other gases are excited by the plasma as they are introduced through
the plasma tube (upstream introduction). The advantage of this design is that the plasma is remote (i.e. away from the sample and certain reactant gases).

The primary differences between a remote plasma processes and a direct one were investigated more than twenty years ago by Lucovsky, Parsons, and Tsu [6, 7]. The results of these investigations are illustrated by figure 2.3. In a direct PECVD process, all of the reactant gases enter the deposition chamber and are excited by the plasma. These excited species then react with the sample surface, which sits within the plasma region. In a remote PECVD process, however, only upstream gases are excited by the plasma; downstream gases are not. This limits the types of excited species present during the reaction. These exited species are then extracted from the plasma region and react with the downstream gases at the sample surface, which is kept away from the chaotic and potentially damaging plasma region.

The various reaction pathways for thin film formation have been discussed in a number of studies [2-6, 8, 9]. For example, RPECVD growth of SiO₂ and Si₃N₄ proceed by the following two-step reactions:

\[
\text{SiO}_2 : \left[ \text{He} + \text{O}_2 \right]^* + \text{SiH}_4 \rightarrow (\text{SiH}_3)_3\text{O} + [\text{H}_2\text{O}, \text{H}_2, \text{etc.}]
\]

\[
\left[ \text{He} + \text{O}_2 \right]^* + (\text{SiH}_3)_3\text{O} \rightarrow \text{SiO}_2 + [\text{H}_2\text{O}, \text{H}_2, \text{etc.}]
\]

\[
\text{Si}_3\text{N}_4 : \left[ \text{He} + \text{N}_2 \right]^* + \text{SiH}_4 \rightarrow (\text{SiH}_3)_3\text{NH} + [\text{NH}_3, \text{H}_2, \text{etc.}]
\]

\[
\left[ \text{He} + \text{N}_2 \right]^* + (\text{SiH}_3)_3\text{NH} \rightarrow \text{Si}_3\text{N}_4 + [\text{NH}_3, \text{H}_2, \text{etc.}]
\]
The growth of \((\text{SiO}_2)_x(\text{Si}_3\text{N}_4)_{1-x}\) is just a combination of these reactions, with \(\text{N}_2\text{O}\) substituting for \(\text{O}_2\) and with different byproducts. Similarly, atoms near the substrate surface in an environment containing only excited nitrogen or oxygen species can be replaced by \(\text{N}\) or \(\text{O}\) atoms, leaving a thin nitrided or oxidized surface.

The adjustable process parameters available in this RPECVD process are the gas flow rates, the RF power, the temperature, and the chamber pressure. The gas flow rates affect the deposition rates and relative compositions of the films, but they do not affect the plasma significantly. The temperature simply needs to be high enough to desorb organic contaminants and chemical byproducts, such as \(\text{H}_2\text{O}\), from the sample surface. The RF power and chamber pressure do affect the plasma properties significantly. RF power affects the energy and types of excited species in the plasma, and the pressure affects the position of the plasma within the chamber. Following device studies, as well as studies investigating the chemical species present in the chamber as a function of position, it was determined by prior students that the optimal process parameters are: (1) a 300°C sample temperature, (2) a 30 W RF plasma, and (3) a pressure between 200-300 mTorr [10-12].

Even under these optimal conditions, films are deposited in a relatively chaotic environment at low temperatures. This combination means that films are initially very disordered, and they are not given sufficient thermal energy for reconstruction and configurational energy minimization. As a result, a thin film of RPECVD-deposited \(\text{SiO}_2\) initially has fairly poor properties [1]. Following a subsequent RTA at 900°C, however, significant rearrangement and relaxation occurs within the film [1, 5]. The nature of the
interfacial stress and its reduction following annealing treatments has been investigated through photoluminescence and second harmonic generation [13], photoreflectance [14], and difference x-ray reflectivity [15]. These studies link improved device properties to a viscoelastic relaxation which accompanies (1) the rearrangement and regrowth of the suboxide ITR discussed in chapter 1 and (2) global strain reduction. Without this rearrangement, these films would not be viable for device applications; therefore, this step is essential in processing RPECVD-deposited thin films.

The only inherent problem in depositing gate oxides onto semiconductor surfaces by RPECVD is the subcutaneous oxide that forms. It has been shown that an oxidized surface layer forms during the RPECVD deposition of SiO₂ not only on Si, but on Ge and compound semiconductors, such as GaAs and CdTe, as well [16, 17]. Since plasma-excited radicals are present at the surface, they will react with that surface; there is no reason for the chemically active species to form only the thin film precursors. Because it forms at the substrate surface, this parasitic surface oxidation has a critical influence on device properties. Since this layer is unavoidable on Si during RPECVD, previous researchers intentionally deposited a surface oxidation layer, engineering it very carefully [12]. In the context of device scaling, however, this layer simply adds to the EOT.

**2.2 – Hafnium-Based High-κ Films by RPECVD**

Unlike silicon, there is no simple hydride precursor for use in CVD processes involving hafnium. As a result, more exotic means for providing hafnium atoms to the reaction are required. As early as the 1960s, a technique called metal-organic chemical
vapor deposition (MOCVD) was being employed in order to deposit epitaxial layers of GaAs [18]. In this method, a carrier gas (such as hydrogen, helium, etc.) flows through a solid or liquid metal-organic chemical precursor, held at a constant temperature and pressure. The finite vapor pressure of the metal-organic compound, determined by the temperature and pressure of its containing vessel (typically referred to as a bubbler), contributes gaseous reactant molecules to the carrier gas stream. These molecules are then carried into the deposition chamber where they can react as in any other method of CVD. The process of ‘bubbling’ a carrier gas through a liquid precursor is illustrated in figure 2.4. MOCVD deposition of group IV-B transition metal oxides was demonstrated as early as 1972 for TiO$_2$, ZrO$_2$, and HfO$_2$ [19], and this is the method chosen for the deposition of the Hf-based high-κ materials in this work.

There are many different hafnium-containing precursors, however, whose properties can vary. It is important to select a precursor which (1) has very high purity, (2) is readily available, (3) is not too expensive, (4) is sufficiently volatile at easily-maintained bubbler pressures and temperatures, (5) leaves low levels of residual species in deposited films, and (6) produces high-quality MOCVD-deposited material. The benefits of using metal-alkoxides for MOCVD were suggested as early as 1992, and researchers since have used them to deposit transition metal oxide films [20]. In particular, the use of hafnium tetra-tert-butoxide (HTB), or Hf(OC(CH$_3$)$_3$)$_4$, as a precursor for HfO$_2$ films has been investigated and found suitable with respect to the six criteria listed above [21, 22]. Thus, the metal-organic precursor used here is hafnium tetra-tert-butoxide.
As mentioned in section 2.1, the RPECVD system used in this work was originally designed for the deposition of SiO$_2$- and Si$_3$N$_4$-based materials. It did not possess any capabilities for metal-organic precursor delivery or incorporation. As a result, the author has undertaken the task of designing, constructing, and integrating into the existing RPECVD system a suitable metal-organic delivery subsystem. Pictures of the end result are shown in figure 2.5. The stand-alone system is designed such that a helium carrier gas first flows through a mass flow controller (MFC), which regulates its flow rate. The carrier gas then passes through either the bubbler (for processing) or a bypass (when idling). A needle valve, which is downstream from the bubbler and connected in a feedback loop to an adjacent pressure gauge, controls the pressure in the gas lines. Finally, the gas passes into either the chamber (for processing) or the idling pump (when idling). Other features include an over-pressure interlock, in addition to convenient back-streaming monitors and safeties, such as an upstream pressure gauge and a back-streaming trap near the bubbler inlet. The carrier lines are wrapped in resistive heating ropes and kept between 50ºC – 100ºC by variable transformers in order to prevent metal-organic condensation inside the lines. The HTB bubbler is submerged in ethylene glycol and kept at 20ºC by a constant temperature bath.

In addition to the construction of this new subsystem, a modification of the deposition chamber itself was required. The chamber had never before been heated on a regular basis. With the metal-organic source addition, it was wrapped in resistive heating ropes and kept between 50ºC – 100ºC by variable transformers in order to prevent metal-organic condensation inside the lines. The HTB bubbler is submerged in ethylene glycol and kept at 20ºC by a constant temperature bath.

* Back-streaming refers to the condition in which a liquid-containing bubbler’s outlet pressure exceeds its inlet pressure, forcing liquid metal-organic material through the inlet and contaminating upstream lines and equipment.
tapes and required constant heating at near 100°C in order to prevent metal-organic condensation. The plasma feedthroughs proved unable to withstand the thermal expansion, falling inside of the chamber. To increase mechanical strength, custom-made, double-gasketed, quick-disconnect seals were designed and constructed for a new plasma tube. These new fittings also included a physical standoff to act as mechanical stress relief.

Upon completion of the various construction and modifications described above, the original RPECVD chamber had been converted into a remote plasma-enhanced metal-organic chemical vapor deposition (RPE-MOCVD) system, now capable of fabricating thin films containing hafnium.

2.3 – Germanium and Germanium-Containing Films by RPECVD

Germanium-containing films can be fabricated via the RPECVD process quite easily through the use of the germanium hydride gas, germane (GeH₄). In fact, a number of researchers have used germane as a germanium precursor in PECVD reactions. Epitaxial thin films of Ge and SiₓGe₁₋ₓ have been deposited onto various surfaces (e.g. Si, Ge, and GaAs) via PECVD [23-25] and RPECVD [26, 27] using germane. Similarly, germane has been used for experiments in which nanocrystalline Ge [28] and Ge nanoclusters [29] were deposited onto silicon substrates.

In analogy with silane (SiH₄), germanium oxide and germanium nitride films are expected to proceed by the following two-step reactions (see section 2.1):
Germanium oxynitride films are simply a combination of these reactions, with $\text{N}_2\text{O}$ substituting for $\text{O}_2$ and with different byproducts. Pure germanium layers are expected to form through reactions of cracked hydride ions, such as $\text{GeH}_2^-$ (generated by interactions with plasma excited He* ions), with the substrate surface.

The use of germane, however, is very dangerous. Not only is it both incredibly flammable and reactive in oxidizing environments, more importantly, it is highly toxic [30, 31]. The source gas utilized in these experiments is a 2% mixture of germane in a helium balance gas, but even this dilute mixture is considered extremely hazardous. The materials safety and data sheet (MSDS) for this mixture of germane states the following [31]:

“Germane is a highly toxic gas. Inhalation of low concentrations can be harmful or fatal. The symptoms produced should be similar to that of arsine which are general malaise, headache, nausea, vomiting, tightness in the chest and pain in the abdomen and loins. Since it has been shown to be a hemolytic agent, there could also be hemolysis (damage to red blood cells) and renal (kidney) damage.”
In accordance with these hazards, the TLV (threshold limit value, or the amount that can be tolerated without an unreasonable risk to one’s health) is only 200 ppb [31]. A full-face respirator in pressure demand mode is required to install/uninstall the cylinder, and continuous gas monitoring of the immediate usage area is required [31].

In order to satisfy these safety requirements, more system upgrades were required. An overview of the germane delivery system constructed for these experiments is shown in figure 2.6. Both the germane bottle and the MFC are kept in exhausted enclosures. Initially, the exhaust in these areas was insufficient for this purpose and had to be increased by widening the ducts, thereby providing more throughput. The RPE-MOCVD system’s valves are interlocked to this exhaust flow. Between the germane bottle and MFC, the gas line is pressurized, and coaxial, double-walled tubing which terminates inside of the exhausted areas is required. Between the MFC and the chamber, however, the line is ordinarily under negative pressure, and any leaks would be inward. Single-walled tubing is acceptable in this region. In order to prevent gas from flowing while the chamber is open, a pressure transducer set at 750 Torr is connected to the valve immediately downstream from the MFC. This ensures that germane cannot flow to the chamber if its pressure is above 750 Torr (i.e. if it is vented). Finally, an MDA Scientific CM4 gas analyzer is used to monitor the following three areas simultaneously: (1) the enclosure in which the germane bottle is kept, (2) the enclosure which houses the MFC, and (3) the room air adjacent to the point at which the germane flows into the deposition chamber. This gas monitor is interlocked to the RPE-MOCVD system, closing the gas
supply valves if germane is detected at levels in excess of 175 ppb. Additionally, alarms signal research and safety personnel (including first-responders, if necessary) to respond appropriately.

2.4 – Thin Film Deposition and Processing: Summary and Conclusions

The basic RPECVD process was described and discussed, including inherent subcutaneous surface oxidation during film deposition and the need for PDA treatments of RPECVD-deposited films using RTA. The incorporation of hafnium into the RPECVD process was then discussed, and technical details of system upgrades completed by the author to do so were described. Finally, safety concerns regarding the use germane as a source gas for Ge in RPECVD processes were discussed, including a description of the modifications carried out by the author to meet the relevant safety requirements.

Following completion of the various modifications discussed in this chapter, the RPE-MOCVD system used for the experiments detailed in this dissertation possessed the ability to incorporate the following materials into thin films: (1) silicon via a silane (SiH₄) source gas, (2) germanium via a germane (GeH₄) source gas, (3) hafnium via an HTB metal-organic precursor, (4) oxygen via O₂ and N₂O source gases, and (5) nitrogen via a molecular N₂ source gas. With these new capabilities, it became possible to study HfO₂ and HfSiON on bulk Ge and PM-Ge/Si surfaces as described in chapter 1.
Figure 2.1: A schematic of the RPECVD cluster tool at NCSU used for thin film deposition in these experiments (from Hattangady, et al. [5], © 1996 American Vacuum Society).
**Figure 2.2:** A rendering of the gas injection system for the RPECVD used in these studies. Metal precursor gases are brought in through the gas dispersal rings, and the rest pass through the plasma tube. The sample is held and heated on the temperature-controlled stage.
**Table 2.1:** The differences between direct and remote plasma CVD processes (after Lucovsky and Tsu [6]). The left side depicts the progression of a direct process; whereas, the right side depicts the steps in a remote process.

<table>
<thead>
<tr>
<th><strong>Direct Plasma-Enhanced CVD</strong></th>
<th><strong>Remote Plasma-Enhanced CVD</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1) Delivery of Reactant Gases:</strong></td>
<td><strong>1) Excitation of 1st Reactant Gases:</strong></td>
</tr>
<tr>
<td>e.g. SiH₄, GeH₄, O₂, N₂, H₂, N₂O, etc.</td>
<td>e.g. O₂, N₂, H₂, He, N₂O, etc.</td>
</tr>
<tr>
<td><strong>2) Excitation of Reactant Gases:</strong></td>
<td><strong>2) Extract Excited Species from the Plasma:</strong></td>
</tr>
<tr>
<td>e.g. SiH₂, GeH₃, NO, O₂⁻ etc.</td>
<td>e.g. NO, He⁺, O₂⁻ etc.</td>
</tr>
<tr>
<td><strong>3) CVD Reaction at Heated Substrate</strong></td>
<td><strong>3) Mixing with 2nd Reactant Gases:</strong></td>
</tr>
<tr>
<td><em>Within the Plasma Region:</em></td>
<td>e.g. O₂⁻ + SiH₄ → H₃Si-O-SiH₃</td>
</tr>
<tr>
<td>e.g. GeH₃ + N₂⁺ → Ge₃N₄</td>
<td></td>
</tr>
<tr>
<td><strong>4) CVD Reaction at Heated Substrate,</strong></td>
<td><strong>Away from the Plasma Region:</strong></td>
</tr>
<tr>
<td><em>Away from the Plasma Region:</em></td>
<td>e.g. H₃Si-O-SiH₃ + O₂⁻ → SiO₂</td>
</tr>
</tbody>
</table>
Figure 2.3: The use of solid and liquid metal-organic compounds in CVD (shown here for a liquid compound). A carrier gas enters the bubbler at point (a), is forced through the metal-organic compound at point (b), picking up gaseous metal-organic molecules through point (c), and exits to the deposition chamber at point (d).
Figure 2.4: The metal-organic delivery subsystem constructed for conducting experiments described in this dissertation. The picture in (a) shows the electronics rack with (listed from bottom to top and left to right): custom valve control switchbox, mass flow controller (MFC) set point controller, upstream pressure gauge, line temperature readout, line temperature zone selector, downstream pressure gauge and controller, pressure gauges power supply, and turbo-molecular pump controller (which never had a home before this). The picture in (b) shows the carrier gas plumbing (top) and (from left to right) the idling pump, variable transformer temperature controls, and constant temperature bath and bubbler (bottom). The figure in (c) is an original design drawing.
Figure 2.5: A schematic diagram of the germane delivery system (described in the text) for the RPE-MOCVD used in these experiments.
References


Chapter 3

Hafnium Based High-κ Materials on Germanium Substrates

3.1 – Restatement of the Motivation and Proposed Solutions

The merits of using germanium substrates have been discussed already in section 1.2. The primary benefit is increased carrier mobility (as evidenced by figure 1.3), which leads to increased drive current and a narrower inversion layer [1]. As discussed in section 1.4, germanium also has shown promise in eliminating the suboxide ITR that hinders silicon CMOS scaling. This is not only a potential benefit for PM-Ge/Si structures, but for Ge substrates as well.

Devices on $p$-type Ge substrates, in addition to devices on Si substrates with Si-capped and epitaxially grown $p$-type Ge channels, already have been fabricated (incorporating ZrO$_2$ and HfO$_2$ high-κ gate dielectrics) [2-4]. The results of these experiments, carried out by Saraswat, et al., are summarized in figure 3.1. The devices on germanium substrates clearly show a marked improvement in hole mobility over even the ideal silicon device. The silicon-based devices with epitaxial Ge channels also outshine their pure Si rival, showing nearly a three-fold increase in hole mobility in one case. Note also the effect of strain on the mobility of these devices. Because of the lattice mismatch between Ge and Si, thinner channel regions are under stress, which leads to further enhancement of the mobility [4].
Given the demonstrably improved carrier mobility in germanium-based devices, the task is to find a compatible gate material and fabrication process; however, as discussed in section 1.2, this is a difficult task. The related oxide, GeO₂, is problematic as a gate material in almost every way imaginable. Nevertheless, the Ge surface oxidizes quickly in air, and in analogy with silicon, it is expected to oxidize subcutaneously at the onset of any RPECVD process involving the deposition of an oxide layer (see section 2.1). As a result, a way to eliminate native surface oxide layers is required, along with a good way to effectively passivate the germanium surface and protect it during gate deposition.

Studies suggest that one effective strategy in improving germanium-based device quality is to replace the GeO₂ layer with a GeON or GeNx layer by using a combination of either thermal or plasma based oxidation and nitridation techniques [5-8]. Diffusion and mixing of Ge atoms with the high-κ film is a problem without a passivation layer; however, nitridation of the surface has been shown to prevent Ge diffusion and to improve device properties noticeably as a result [7, 8]. In section 1.2, the extreme volatility of germanium nitride layers formed on germanium surfaces was discussed, and it was stated that GeNx breaks up completely at temperatures of approximately 650°C (see figure 1.5) [9]. This suggests that not only are GeNx passivation layers relatively stable at temperatures approaching 650°C, but that they may be removed by PDA treatments following a low-temperature deposition process as well. In fact, this phenomenon has been observed in XAS spectra of Ge/GeNx/HfO₂ structures [10-12], and similar results are shown in figure 3.2. In agreement with figure 1.5, figure 3.2 indicates
that nitrogen begins to leave the sample at temperatures as low as 600°C. Unlike in figure 1.5, given the blocking effects of the high-κ overlayer, nitrogen is not completely removed (to background levels) in these samples until a temperature of approximately 750°C – 800°C is reached. There is a sharp transition between 650°C and 700°C in HfO₂, however. This sharp transition is present in HfSiON as well, although it is delayed by 50°C.

Hence, plasma-nitridation of germanium surfaces is potentially very desirable for two reasons: (1) the resulting GeNx layer serves as a diffusion blocking and protective layer during gate deposition [7, 8] and (2) the nitrogen can be removed following gate deposition by a PDA treatment between approximately 650°C – 750°C, leaving the gate and germanium substrate in direct contact [9-12]. The effects of plasma nitridation and PDA temperature, therefore, are topics addressed by studies described later in the chapter.

The gate materials selected for these investigations were HfO₂ and its related nitro-silicate, (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄. Though both materials show promise for a reduction in EOT, HfO₂ is better in this regard (κ_HfO₂ ≈ 6 · κ_SiO₂, and κ_HfSiON ≈ 3 · κ_SiO₂) [13-15]. On the other hand, the nitro-silicate shows a self-organized phase that is more thermodynamically stable and much more resistant to crystallization (see sections 1.3, D.3, and E.2) [16-20]. As discussed at length in section 1.2, the natural oxide (GeO₂) and nitride (Ge₃N₄) simply are not good gate dielectrics for germanium for a variety of reasons. Their properties are investigated, but solely for the purpose of deducing band alignments in this material system.
3.2 – Germanium Substrate Cleaning

Germanium is reactive with oxygen, forming GeO and GeO₂ [21]. This means that germanium surfaces will acquire a native oxide layer in air, which includes both divalent and tetravalent varieties (i.e. GeO and GeO₂). Even a purely divalent oxide layer will likely include a hexagonal quartz modification and a tetragonal rutile form [22], both of which have very different solubility and chemical properties [22, 23]. Moreover, even a purely tetragonal GeO₂ layer will likely include two different phases in the cristobalite crystal configuration [24]. With so much variety, it is unclear how best to remove all of these species from this native germanium oxide layer. Furthermore, it is unclear how to chemically passivate a cleaned germanium surface effectively against subsequent re-oxidation (or even if it is possible).

Silicon’s native oxide, on the other hand, comes in a purely divalent SiO₂ form, which has only four-fold coordinated silicon atoms and is purely amorphous [25]. It can be etched in hydrofluoric acid (HF), leaving a hydrogen-terminated silicon surface that is stable against re-oxidation. In analogy with this process, a method for cleaning germanium using a cyclic H₂O/HF rinse was evaluated using XPS by Deegan and Hughes [26]. They found that the germanium surface is as stable as silicon against re-oxidation, and they assumed that this surface, like silicon, was hydrogen-passivated. A later study confirmed the resistance to re-oxidation but found the assumption of hydrogen passivation to be incorrect.

A similar surface preparation process, developed by Chabal, et al., begins with a water etch, is followed by a wet chemical oxidation (WCO) using H₂O₂, and is finished
with an etch in HF [27]. Their investigations of this chemical treatment employed both infrared absorption spectroscopy (IRAS) and XPS to study the hydrogen-terminated germanium surface. The results confirm that, as expected, HF does etch GeO₂ to a high degree and that the remaining surface is hydrogen-terminated. However, almost all of the hydrogen is removed from the germanium surface within minutes of this procedure, at which time hydrocarbon contamination sets in via bonds to bridging oxygen atoms (i.e. Ge-O-hydrocarbon interfacial bonding). This suggests that the long-term stability against GeO₂ re-growth at the germanium surface following HF treatments is attributable not to hydrogen termination, but rather to atmospheric hydrocarbon contamination.

As a review of the literature shows, the majority of researchers working on Ge-based devices use one of the above methods or a slight variant thereof. The Stanford group, for example, uses the Deegan and Hughes surface treatment method [28]. There is also a third method of interest, which was used by Aspnes and Studna to achieve atomically clean Ge surfaces for measurements of its dielectric function [29]. In the case of Ge, they used a bromine/methanol solution to oxidize the germanium surface, followed by a methanol rinse to reduce organic contamination, and finished with a buffered HF etch. It was this basic philosophy, coupled with that of Chabal, et al., which was adopted for these experiments.

More specifically, the NCSU cleaning method has the following general outline: (1) a primary DI H₂O rinse/etch, (2) a WCO step, (3) an organic contaminant removal step, (4) a chemical etching step, and (5) a final DI H₂O rinse/etch. Step (1) is self-explanatory. Step (2) is used by Chabal, et al. (H₂O₂), as well as Aspnes and Studna.
A 6% solution of H₂O₂ was selected for this step. For step (3), methanol was selected as in the Aspnes and Studna method. Although all three of the referenced methods implement HF to etch the GeO₂, only the hexagonal form of GeO₂ is soluble in this acid. The tetragonal form is not [23]. The complete dissolution of all forms of GeO₂ in strong bases, such as NaOH and NH₄OH, has been reported [23, 30-32]. Hence, a 15% NH₄OH solution was used for step (4), rather than HF. The DI H₂O rinse is simply to clean any chemical residue from the Ge surface.

In examining the relevant literature prior to conducting these experiments, it became apparent that there was no satisfactory quantitative measurement of the efficacy of the cleaning methods or of the subsequent surface evolution immediately following the GeO₂ removal. The prior research used crude thickness estimates based on electron attenuation to study the surfaces on a time scale of minutes, days, and weeks [26, 27]. As a result, real-time, quantitative evaluations of etching steps and of the immediate evolution of germanium surface re-oxidation were conducted by using in-situ SE (see appendix H) and AES (see appendix F). Three cleaning methods were examined in total: (1) the Deegan and Hughes method, (2) the Chabal, et al. method, and (3) the NCSU method. These treatments are described in table 3.1.

One very good way to compare the relative efficacy of these three cleaning methods is to examine the height of the E₂ peak in the imaginary component of their pseudo-dielectric functions [33-35]. For this comparison, the “biggest is best” rule applies since overlayers generally have smaller dielectric functions than the substrate at this wavelength. A thicker overlayer leads to a phase shift which reduces the amount of
light reflected by the substrate, thereby reducing the height of the E₂ peak [34]. It should be emphasized that this reduction in the E₂ peak height can be caused not only by an overlayer of GeO₂, for example, but by adsorbed atmospheric contaminants, such as H₂O and hydrocarbons, and by surface roughness. Accordingly, in computing the overlayer thicknesses for this study, an equivalent GeO₂ thickness is given (similar to the EOT concept), based on the atomically clean reference provided by Aspnes and Studna [29] and using an idealized three-phase optical model (Ge-GeO₂-air).

The E₂ peaks of the various cleaning methods for Ge are shown in figure 3, which were obtained through in situ measurements of Ge (111) carried out in an ambient atmosphere of dry nitrogen. Hence, the contributions to the effective overlayer thickness by surface contamination and re-oxidation are expected to be minimal. Clearly, the NCSU method results in the cleanest surface, corresponding to an effective GeO₂ thickness of only 3 Å. Aspnes and Studna used a pad polishing technique in order to achieve atomically flat surfaces [29]; whereas, this was not done for the current study. The 3 Å difference between these results and those of Aspnes and Studna, then, is likely due to surface roughness; the NH₄OH is expected to remove the oxide layer completely [30-32]. The HF cleaning method of Deegan and Hughes yields a result of 17 Å – only 3 Å better than the as-received Ge wafer’s effective thickness of 20 Å. This is attributed to (1) an increase in surface roughness caused by the HF and (2) the insolubility (in both HF and H₂O) of the tetragonal GeO₂ phases present in the native oxide layer [22, 23]. The method of Chabal, et al. actually adds to the thickness of the as-received wafer. This is expected based on reports that H₂O₂-grown GeO₂ films are primarily of the insoluble
tetragonal variety [24]. As a result, the oxidation step in his method is not undone by 
H2O or HF treatments, and a net increase in GeO2 occurs.

Having chosen the method providing the cleanest (and/or flattest) surface, another 
experiment was carried out in order to examine the real-time evolution of the Ge surface 
during that cleaning procedure. The results are shown in figure 3.4, and the thickness 
conversion scale on the left axis was computed using the same method as described 
above. The starting point is after the initial 5-minute de-ionized water bath. Note that the 
minimum thickness achieved for Ge (111) was 2 Å; whereas, the minimum thickness 
achieved for Ge (100) was 6 Å. Note also that the rate of thickness increase (i.e. the 
slope of the lines) is greater for Ge (100) than it is for Ge (111). Both of these traits are 
attributed to a difference in surface reactivity, which is enhanced in Ge (100) by its two 
dangling bonds versus the one dangling bond of the Ge (111) surface. Note also that the 
initial H2O bath reduces the Ge(100) surface to a thickness of 12 Å, indicating that much 
of its native oxide layer is comprised of the insoluble tetragonal form of GeO2. To the 
contrary, the initial H2O bath on Ge (111) reduces the apparent thickness to 3 Å, 
indicating that its native oxide layer is comprised almost completely of the soluble 
hexagonal GeO2 phase. The effect of the H2O2 step is clear for both surfaces – it grows a 
chemical oxide very quickly, as expected. The methanol does little to affect the overlayer 
thickness either way, and is likely to be unnecessary in this procedure. On both 
orientations, the NH4OH etches the oxide layer better than the initial H2O step, again as 
expected. The final water rinse rids the surface of NH4OH residue, reducing the
thickness slightly again. After cycling through this process twice, a minimum thickness is reached, and further cycling does nothing to reduce the overlayer thickness further.

The next phase of this study was to examine in real time the overlayer re-growth of the NCSU-cleaned Ge (100) and Ge (111) surfaces. The results are shown in figure 3.5. Note the incredibly rapid re-growth process, leaving both surfaces saturated within minutes of surface treatment. The Ge (100) surface saturates at 17 Å, barely better than the as received wafer and on par with the Deegan and Hughes method. The Ge (111) surface saturates at 12 Å, which is a significant improvement over the as-received wafer, but much larger than the 2 Å minimum achieved while etching. This re-grown layer is not necessarily GeO₂, however.

To investigate the composition of the re-grown overlayer, AES experiments were conducted. The results are shown in figure 3.6. The wide scan in figure 3.6a shows the relative atomic concentrations on the surface, and the narrow scan in figure 3.6b shows the relative amounts of Ge-O and Ge-Ge bonding at the surface. Note first the large levels of both oxygen and carbon on the as-received wafer (i.e. the bare wafer in figure 3.6a). These are presumed to be due to the native oxide layer and to hydrocarbon adsorbates, respectively. As expected from the discussions above, the HF cleaning method of Deegan and Hughes and the NCSU cleaning method show comparable amounts of oxygen. The Deegan and Hughes method, however, shows slightly less oxygen, indicating that a relatively larger part of its 17 Å effective overlayer thickness is due to surface roughness when compared to the NCSU method. The NCSU method shows much less carbon, which suggests that the NH₄OH is more effective at removing
and/or preventing hydrocarbon contamination at the Ge surface. When compared to the as-received wafer, figure 3.6b indicates that both cleaning methods considered reduce the amount of Ge-O bonding at the surface relative to the amount of Ge-Ge bonding. Moreover, they do so to comparable levels. As a final note, figure 3.6a shows that a short plasma oxidation step removes residual carbon from the Ge surface. Because a thin (approx. 2-3 nm) GeO$_2$ layer will decompose thermally at temperatures of around 450ºC [37] (see section 1.2), an oxidation step followed by a 450ºC annealing step is seen as one possible pathway for producing atomically clean Ge surfaces for high-κ gate dielectric deposition.

In summary, the cleaning method of Chabal, et al. was shown to actually increase the amount of GeO$_2$ at the Ge surface, and is therefore unsuitable for wet cleaning of Ge surfaces. The NCSU cleaning method was shown to produce the most atomically clean Ge surfaces during etch treatments; however, overlayer re-growth is so quick that for all intents and purposes, the method of Deegan and Hughes produces surfaces with equivalent apparent thickness. On the other hand, these two methods do show differences. The NCSU cleaning method reduces carbon contamination much more effectively than the Deegan and Hughes method, and it appears to produce a smoother surface as well. For these two reasons, the NCSU method appears to have a slight advantage and was employed in the high-κ studies discussed later in this chapter.
3.3 – Sample Preparation

Before delving into the properties of the germanium-based structures studied in this work, it is necessary first to describe exactly how they were made. This will include descriptions of (1) wet sample preparation, (2) the remote plasma-assisted nitridation (RPAN) process, (3) thin film deposition, and (4) PDA treatments by RTA.

First, consider the wet sample preparation. These experiments employed the NCSU germanium surface cleaning method, as described in detail in section 3.2 and summarized in table 3.1. First, the substrates were placed in de-ionized water for 5 minutes. Next, they were rinsed for 5 seconds in a 6% aqueous solution of H$_2$O$_2$, followed by a 15 second rinse in methanol, followed by a 15 second rinse in a 15% aqueous solution of NH$_4$OH, followed by a 15 second rinse in de-ionized water. This rinsing process was repeated twice, and it was done simply by squirting the samples using reagent bottles. Finally, the substrates were submerged into a de-ionized water bath and transported underwater to the RPE-MOCVD load lock chamber. There, they were blown dry with nitrogen, mounted onto the sample holder, and introduced as quickly as possible into the vacuum chamber (typically on the order of 1-2 minutes).

Once inside the RPE-MOCVD chamber, the germanium substrates were heated at 300°C for 15 minutes. During this step, the samples were kept at a pressure of 300 mTorr in an inert atmosphere of research-grade helium which was flowing at 200 sccm. This step was done in order to outgas any volatile materials from the sample surface and to reach the steady-state processing temperature.
The samples then underwent a remote plasma-assisted nitridation (RPAN) treatment for 90 seconds in a 30W RF plasma at a frequency of 13.56 MHz. During this step, the sample temperature was 300ºC, and the chamber pressure was 300 mTorr. The flow rates of the gases were 200 sccm of research grade (99.9999% purity) helium and 60 sccm of research grade nitrogen, both of which were used in an upstream process (see section 2.1). Given these process conditions, the mechanism for the germanium nitride layer formation is most likely excited N$_2^*$ species rather than excited N$^+$ species [38, 39]. As discussed in section 3.1, nitridation of the germanium surface has several potential benefits. First of all, it acts as a passivation layer during the gate deposition, preventing surface oxidation and germanium diffusion [9-12, 38]. Due to its volatility, the GeN$_x$ layer can also be removed following gate deposition by RTA at 650ºC-750ºC, leaving the gate and substrate in direct contact [10-12].

Following the RPAN process, the gate was deposited by RPE-MOCVD (as discussed at length in chapter 2). Two gate materials were examined: HfO$_2$ and the self-organized pseudo-ternary, (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$. Both materials were deposited at 300ºC, at a pressure of 300 mTorr, and with a 30W, 13.56 MHz plasma excitation. All gases used were of research grade. The HTB bubbler was maintained at 20ºC and at 30 mTorr for both material depositions as well, using a He carrier gas. In the case of HfO$_2$, the gas flow rates were 200 sccm of He (upstream), 20 sccm of O$_2$ (upstream), and 10 sccm of HTB carrier He (downstream). In the case of (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$, the gas flow rates were 200 sccm of He (upstream), 6 sccm of 1% N$_2$O in a N$_2$ balance gas (upstream), 74 sccm of N$_2$ (upstream), 20 sccm of 2% SiH$_4$ in a He balance gas.
(downstream), and 11 sccm of HTB carrier He (downstream). The chemistry of the resulting thin films was measured by taking AES spectra of thick (approx. 10 nm) films, and the deposition rates were estimated by taking AES spectra of thin (less than approx. 3 nm) films (see appendix F). Representative spectra showing the chemistry of these materials are shown in figure 3, indicating the relative amounts of atomic species present in the films.

As an aside, one noticeable difference between these spectra (aside from the compositional differences) is the presence of carbon in the \((\text{HfO}_2)_{0.3}(\text{SiO}_2)_{0.3}(\text{Si}_3\text{N}_4)_{0.4}\). In HfO\(_2\) deposition, there is an excess of oxygen, which likely consumes the organic HTB ligands by forming CO, OH, and other similar species (note the removal of surface carbon by O\(_2\) in figure 3.6, for example). In \((\text{HfO}_2)_{0.3}(\text{SiO}_2)_{0.3}(\text{Si}_3\text{N}_4)_{0.4}\), however, there is very little oxygen present. Because it is so reactive, the oxygen levels are kept extremely low just in order to achieve this film composition. This leaves less oxygen available during this deposition for reacting with the organic HTB ligands, resulting in detectable levels of carbon.

Once the gate dielectric depositions were complete, the samples were annealed for 1 minute durations at various temperatures by an RTA process. In the following discussions, when a sample’s temperature is referred to, it is this PDA temperature that is being referenced – not the actual growth temperature. Again, this step was carried out in order to relax strain in the material (see section 2.1) and to eliminate the interfacial nitride region (see section 3.1).
For spectroscopic measurements, the sample preparation ended here; however, electrical devices required further processing. Following the PDA, those samples were subjected to a forming gas anneal at 450°C for 30 minutes. Aluminum gate electrodes were then deposited by resistive heating evaporation through a shadow mask, and backside contacts were made using a GaIn (76% Ga and 24% In) eutectic paint.

3.4 – Spectroscopic Results for Ge/High-κ Stack Structures

Three different spectroscopic techniques were employed to examine Ge/high-κ stack structures: MEIS, XAS, and SE. MEIS (see appendix I) experiments were conducted in order to study the chemical make-up of the interface, showing evidence for diffusion and mixing. XAS (see appendix G) experiments were done to probe changes in the conduction band of the high-κ materials, revealing information about their bonding configurations, crystallinity, and defect levels. Similarly, SE (see appendix H) measurements were made in order to investigate the defect levels in the high κ films by looking at their conduction band structure.

The MEIS results are shown in figure 3.8. First, examine the spectra of the films as-deposited. The samples show three basic regions (moving right to left): (1) a GeON region, (2) a region transition region, and (3) the HfSiON film. The HfO₂ film also shows a surface with slight oxygen enhancement. The important things to notice about these spectra are that (1) the interfacial nitrogen region acts as a very good barrier against hafnium and silicon diffusion into the substrate, (2) that the nitrided region is converted into an oxynitride region, presumably at the onset of deposition, and (3) that there is a
transition region of approximately 6-10 Å. Next, examine the MEIS spectra of the samples annealed at 800ºC. The first thing to note is that the interfacial nitrogen is gone in both cases, as predicted (see section 3.1). Both also have a GeOₓ region that contains a small amount of hafnium and a hafnium germinate/germano-silicate region. On the other hand, the HfO₂ film contains germanium throughout not only these transition regions, but throughout the entire film. The HfSiON film does not. Lastly, the HfO₂ shows a slight enhancement of germanium at the very surface of the film. These results for HfO₂ are all consistent with prior work done by angle-resolved XPS, MEIS, and HR-TEM [40-43].

The structure of these films is summarized in table 3.2, and the main conclusions that can be taken away from these MEIS results are as follows:

- The interfacial nitrogen layer acts as a very good barrier against diffusion during deposition on Ge substrates.
- The interfacial nitride layer is converted to an interfacial oxynitride layer during deposition.
- Annealing at 800ºC (and presumably at any temperature sufficient to remove the interfacial nitrogen from the film) converts the interfacial oxynitride layer into an interfacial oxide layer.
- Annealing at 800ºC distributes Ge throughout HfO₂ films, and it creates a uniformly graded chemical transition region from the GeOₓ layer to the Ge-doped HfO₂ film.
• Aside from the interfacial nitride layer, HfSiON films on Ge are very stable chemically both during deposition and annealing.

Having examined the chemical compositions of these films, consider now how this affects their bonding and electronic structure by analyzing their XAS spectra. Figure 3.9 shows the thermal evolution of crystallinity in thin films of both HfO$_2$ and (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ deposited on plasma-nitrided Ge (100) and Ge (111) substrates. Before discussing these spectra, however, a couple of quick definitions are required (after Lucovsky and Phillips [19]). Let type I nano-crystalline films refer to films which have very small grain sizes, no coherent grain boundaries, but rather a disordered interfacial transition region and band tail defects. Let type II nano-crystalline films refer to films that demonstrate relatively larger grain sizes, coherent grain boundaries, single crystal formation, and discrete defects.

Consider first the HfO$_2$ spectra shown in figure 3.9. Through processing temperatures up to 650°C, these films show a characteristic type I nano-crystalline structure [10, 12, 19]. At annealing temperatures in excess of 650°C, the structure of the HfO$_2$ films changes noticeably. The $e_g$ features split and sharpen, and the $a_{1g}$ and $t_{1u}$ features appear. These changes also accompany precisely the removal of interfacial nitrogen shown in figure 3.2, and the films shift to type II nano-crystallinity [10, 12, 19]. With the onset of single crystal formation in these HfO$_2$ films, the $e_g$ and $t_{2g}$ states are split by Jahn-Teller distortions and by the crystal field of their new pseudo-cubic $C_{1}$
symmetry. This is consistent with their seven-fold coordinated hafnium bonding arrangements (see appendices C & D). Note also the difference between Ge(111) and Ge(100) substrates with respect to the HfO$_2$ films. As shown in figure 3.10, there is a clear difference in the ratio of the split $e_g$ states. This indicates that the substrate geometry is affecting the films during annealing, bolstering the assertion of type II nanocrystallinity in which larger grains are present, single crystal formation occurs, and grain boundaries are very likely to be present. Moreover, oxygen vacancies cluster around such grain boundaries, creating electrically-active traps and leakage paths [10, 12].

Consider next the (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ films shown in figure 3.9. As-deposited, these structures demonstrate the typical type I nanocrystallinity described above. Upon annealing at 600ºC, both the $e$ and $t_2$ symmetry states sharpen slightly, but the films still exhibit type I nanocrystallinity. Even for annealing temperatures up to 800ºC these films do not show any further structural changes, maintaining their type I nanocrystallinity. Furthermore, the Jahn-Teller distortions are completely suppressed. Hence, the HfSiON films exhibit not only better chemical stability than their HfO$_2$ counterparts (as seen in the MEIS spectra), but better bonding and structural stability as well. On the other hand, (following annealing) the absorption edge in HfO$_2$ is much sharper than in HfSiON, which is likely due to band tail defects in the HfSiON film.

Fortunately, in addition to the information provided about the conduction band states, these XAS spectra can provide insight into such defect levels. Consider first figure 3.11, which shows the 2$^\text{nd}$ derivative absorption spectra for thin films of HfO$_2$ and (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$. The bandgap for HfO$_2$ is 5.7 eV [13]; therefore, since the
absorption edge (529.8 eV) in the HfO$_2$ spectrum corresponds to absorption by states at
the conduction band edge, 529.8 eV is taken to be 5.7 eV above the valence band edge
(which are filled states and do not absorb in XAS). Using this as a reference point, the
conduction band edge structure can be completely determined by examining the 2$^{nd}$
derivative absorption spectra. They show quite clearly the relative energy differences
between not only the d, s, and p states, but also two observable sub-band edge defect
states (observable in figure 3.11). Moreover, the first feature in the 2$^{nd}$ derivative is sharp
in HfO$_2$ but more smeared out in HfSiON, indicating larger amounts of band tail defects
in HfSiON films (as expected from the discussion of type I and type II nano-crystallinity
above).

Additionally, the 2$^{nd}$ derivative XAS spectra provide a measurement of the
HfSiON bandgap, based on the shift in the absorption band edge position relative to
HfO$_2$. The O 2p non-bonding states, which do not hybridize, dominate the valence band
position; they are simply atomic p states and should be at the same energy in both HfO$_2$
and HfSiON. The conduction band edge is dominated by Hf d states, which are hybrid
orbitals and are expected to change energy between HfO$_2$ and HfSiON (see appendix D
for details). Hence, the change in the conduction band edge position is the change in the
bandgap. Since the absorption edge shifts from 529.8 eV in HfO$_2$ to 530.0 eV in
HfSiON, the HfSiON band gap is taken to be 5.9 eV. All of this information is
summarized by the band diagrams on the right side of figure 3.11.

Aside from crystal field and Jahn-Teller splitting effects in the HfO$_2$ film, these
spectra provide a nearly identical structure for the conduction band edges, especially the
sub-band defect states. These discrete sub-band edge defect states are at 1.0 eV and 2.5 eV below the conduction band edge in both films and are attributable to Hf$^{3+}$ states associated with oxygen di-vacancies [10, 15]. The d states just above the conduction band edges also show striking similarities. Though the $e_g$ states are split in HfO$_2$, their centroid is 1.4 eV above the band edge, which is the precise position of the $e$ states in HfSiON relative to its band edge. Similarly, the centroid of the $t_{2g}$ states in HfO$_2$ is at the same position as the $t_2$ states in HfSiON – approximately 5.3 eV above their respective conduction band edges.

The 2$^{nd}$ derivative absorption is not the only means for defect state analysis, however. Gaussian fitting of the raw spectrum also provides insight into these states. Figure 3.12 shows such fitting for HfO$_2$ and HfSiON films on Ge. The full spectra were fit with 10 Gaussian functions each (2 defect states + 5 non-degenerate d states + 1 s state + 1 triply degenerate p state + 1 background peak), but only the limited sub-band regions are shown. The position of these two peaks relative to the absorption band edge agrees very well with the 2$^{nd}$ derivative absorption analysis. The areas under the curves are proportional to the number of states, and figure 3.12c shows how the areas under the curves changes as a function of temperature. Note that not only is the number of defects lower in HfSiON films by at least ½ of an order of magnitude for both defect states, but the change in the number of defect states upon annealing is smaller by roughly 50% as well. Furthermore, the critical transition temperature for an increase in defects shown in figure 3.12c is delayed by 50ºC, shifting from 650ºC in HfO$_2$ to 700ºC in HfSiON. This agrees with the nitrogen K$_1$ spectra shown in figure 3.2. For HfO$_2$/HfSiON/Ge stacked
structures, there is an observable increase in defect states which arises from the HfO\textsubscript{2} overlayer, but there are still less defects than in pure HfO\textsubscript{2} films. The 50\degree C delay in critical transition temperature observed in pure HfSiON films is observed in these stacked structures as well. Given the results from the MEIS studies, at least part of these two defect states is likely to be related to Ge diffusion. The same two defects are observed in Si/HfO\textsubscript{2} structures as well, and they are attributed to Hf\textsuperscript{3+} states arising from oxygen di-vacancies \cite{44}. Since these di-vacancies occur primarily at grain boundaries, and since Ge diffusion likely occurs along grain boundaries as well, the Hf\textsuperscript{3+} di-vacancy defects are almost certainly modified by Ge diffusion.

In addition to XAS, SE can provide insight into the conduction band structure of these materials (see appendix H). Figure 3.13 shows the $\varepsilon_2$ portion of the dielectric function for (HfO\textsubscript{2})\textsubscript{0.3}(SiO\textsubscript{2})\textsubscript{0.3}(Si\textsubscript{3}N\textsubscript{4})\textsubscript{0.4} and HfO\textsubscript{2} films, both as-deposited and annealed at 800\degree C. These spectra were obtained by analyzing SE data with a four-phase model consisting of a Ge substrate, a GeON layer, a high-\kappa layer, and an air ambient. The best fit was determined by minimizing the strain-related artifacts. Just as in the XAS spectra, the dielectric function reveals two sub-band edge defect states. This analysis gives their positions to be at approximately 1.3 eV and 2.3 eV below the conduction band edge, which is slightly different than, yet still largely in agreement with, the analysis done on the XAS spectra and with theoretical calculations \cite{15}. This difference may be attributable to excitations from valence band defect states. Such transitions can be observed in SE, but they are not possible in XAS. The relative change in the quantity of defect states upon annealing at 800\degree C also qualitatively agrees with the XAS analysis.
The crystallization of HfO₂ is accompanied by a major increase in sub-band defects related to oxygen di-vacancies at grain boundaries. The HfSiON film, however, shows very little change in the strength of absorption in these defect states.

The valence band defects have been studied by others, including Brillson, et al. by depth-resolved cathodoluminescence spectroscopy (DRCLS) [45]. Their data is shown in figure 3.14, along with the assigned transitions that are consistent with the XAS data presented above. By placing the valence band edge defect state energies at 0.8 eV and 1.4 eV, the transition energies consistent with the XAS data are 1.8 eV, 2.4 eV, 3.3 eV, 3.9 eV, and 4.3 eV. The DRCLS spectra show transitions for 2.0 eV, 2.5 eV, 3.5 eV, 3.9 eV, and 4.3 eV, in very good agreement with this assignment. Surface-sensitive XPS measurements have been done as well [44], and those results are also in good agreement with this assignment for the valence band edge defect states.

3.5 – Electrical Results for Ge/High-κ Stack Structures

Having discussed the physical and chemical properties of these materials in detail, a discussion of electrical devices is warranted in order to correlate spectroscopic observations with device properties. Simple MOS capacitor structures were fabricated as described in section 3.3, and both C-V and I-V characterizations were done in order to study some of the electrical performance characteristics of Ge/high-κ gate stacks.

Before discussing the electrical characteristics, however, consider the band alignment in these stacks. XAS spectra were collected for oxidized and nitrided Ge surfaces, and the results are shown in figure 3.15. The relative shifts in the bandgap are
-3.4 eV for GeO$_2$ and -0.9 eV for Ge$_3$N$_4$, relative to their silicon counterparts. As discussed in the previous section with respect to the bandgap of HfSiON relative to HfO$_2$, since the O 2p non-bonding states dominate the top of the valence band, this shift in conduction band edge position translates into a shift in bandgap. Since the bandgaps for SiO$_2$ and Si$_3$N$_4$ are 8.9 eV and 5.3 eV, respectively, the bandgaps for GeO$_2$ and Ge$_3$N$_4$ are taken to be 5.5 eV and 4.4 eV, respectively. This value for GeO$_2$ is in good agreement with a prior estimate as well [46]. The valence band offsets of these materials relative to Ge have been measured by other researchers via ultraviolet XPS experiments. Their findings show that the valence band offsets are 3.3 eV in GeO$_2$ [47] and 2.2 eV in Ge$_3$N$_4$ [48]. Given that the bandgap of Ge is approximately 0.7 eV, the conduction band offsets between Ge and GeO$_2$ and Ge$_3$N$_4$ both are 1.5 eV. The band offsets in HfO$_2$ relative to both Si and Ge have also been measured, and they were found to be approximately 1.8 eV [47, 48] and 2.0 eV [48], respectively. These values are all in good agreement with other work as well [49-52].

Using the above information, instructive band diagrams relevant to the devices discussed in this chapter were constructed and are given in figure 3.16. Note that for a GeO$_2$ ITR, there is an intrinsic well between GeO$_2$ and HfO$_2$ that exists for substrate injection in an $n$-type Ge device. For a $p$-type device, this barrier does not exist. This may, at least in part, explain the much higher reported success rate with respect to $p$-type versus $n$-type devices [3-5, 13-15, 28, 41, 42, 48].

Consider now the C-V characteristics of these structures (see appendix J). Traces showing their C-V hysteresis properties at 1 MHz are plotted in figure 3.17 as a function
of annealing temperature, and traces showing their C-V hysteresis properties as a function of measurement frequency are shown in figure 3.18. In general, the C-V hysteresis curves show a lot of both mobile and fixed charge trapping. Fixed charge traps are those which capture electrons permanently, adding charge to the dielectric. The net effect is a shift in the flat band voltage of the device. In order to achieve a desired capacitance, more gate voltage is required in order to cancel out the additional field generated by these fixed charges. Mobile charge trapping, on the other hand, is visible in the amount of hysteresis present. While sweeping from depletion to accumulation, electrons “fall” into these traps. While sweeping back from accumulation to depletion, these same traps are emptied; however, there is a barrier to overcome in “lifting” these electrons out of the traps. Just as in fixed charge traps, the mobile traps acquire extra negative charge when occupied, thereby shifting the flat band voltage to the left. Hence, the amount of hysteresis is related to both the depth of these mobile traps and to the amount of traps present. Ideally, the hysteresis should be zero, and the flat band voltage should be equal to the metal-semiconductor work function difference (which is usually designed to be at or close to zero). Lastly, the interface trap density can be seen qualitatively by how abrupt the depletion to inversion transition is. The sharper this transition, the fewer interface states are present in the device.

The discussion of the C-V properties that follows is largely qualitative, rather than quantitative, as the author is reluctant to attach numbers primarily for the two following reasons: (1) the large hysteresis makes accurate flat band and surface potential determination very difficult (since it is a function of trap occupation) and (2) C-V curves
were unobtainable for frequencies less than 100 kHz. Because of (1), calculations of fixed charge are very difficult, and because of both (1) and (2), calculations of interface trap densities are inaccurate at best. Moreover, the flat band voltage depends on both fixed and interface charge. This means that the fixed charge calculation is coupled to the interface state calculation and is, therefore, also unreliable. Because the devices are so far from ideal, a qualitative discussion should suffice for the identification of trends.

In figure 3.17, it is apparent that HfO$_2$ devices have a lot of mobile traps and that annealing does very little to mitigate these effects. On the other hand, HfSiON devices show a significant reduction in hysteresis at an annealing temperature of 650ºC. The HfO$_2$ devices, however, possess fewer fixed traps than their HfSiON counterparts. Annealing does not affect the fixed charge significantly in either type of device. Based on the spectroscopic evidence in section 3.4, these fixed charges are likely related to oxygen di-vacancies.

Given the spectroscopic data in section 3.4, it is known that there are certain chemical and structural changes which occur in these devices upon annealing. The EOT (which is related to the accumulation capacitance) steadily decreases as a function of annealing temperature in HfO$_2$, most likely because of germanium diffusion and hafnium germanate formation throughout the gate. The EOT for HfSiON devices, on the other hand, is the same at 650ºC as it is in the as-deposited film, and it increases at 700ºC. This enhancement actually may be due to a physical reduction in size. At 700ºC, nitrogen leaves the film, but Ge diffusion does not occur, and the gate material retains its
properties. (Note that these devices came from one HfO₂ deposition and one HfSiON deposition, which are known to be physically uniform.)

On a final note with regard to figure 3.17, annealing also reduces the number of interface traps observed in both HfO₂ and HfSiON devices. This enhancement is likely due to chemical and structural re-organizations at the interface which reduce strain and the number of dangling bonds (see section 2.1). Based on very crude Terman model estimates, the author estimates that the interface trap densities in the annealed HfO₂ and HfSiON devices are on the order of 1-3 x 10¹³/cm²/eV and 5-8 x 10¹²/cm²/eV, respectively. Similar numbers in the mid-10¹²/cm² to low-10¹³/cm² are estimated for the fixed oxide charge density.

Figure 3.18 shows the frequency dispersion for optimally annealed HfO₂ and HfSiON devices. The large density of trap states near the depletion region at 100 kHz is obvious in both devices. At higher frequencies, these traps are too slow to respond to the changing fields; however, at lower frequencies, this is no longer a problem, and the frequency-dependent traps become apparent. In fact, this is likely the reason for which sub-100 kHz measurements were not attainable. There were simply too many trapping states available at lower frequencies, and the internal charges caused the field effect to break down. These frequency-dependent traps have been observed by others as well [41, 48].

Consider now the I-V and gate leakage characteristics of these HfO₂- and HfSiON-based devices (see appendix K). Before delving into the results, however, a discussion about the effects of series resistance is warranted. When poor electrical
contact is made to the semiconductor substrate, the resistance of the device increases, and the amount of current flowing through it for a given gate voltage decreases. This is termed series resistance due to its role in the equivalent circuit of the MOS capacitor. If one neglects a sizeable series resistance, assuming that the total resistance of the device is due only to the impedance of the gate, then the surface potential is computed incorrectly, and the gate leakage is under-reported. To think of this another way, much of the voltage applied to the device is dropped across the semiconductor contact due to its large resistance; therefore, the electric fields within the MOS device are much less than expected. This leads directly to a lower internal gate bias, and in the end, to less measured leakage current at a given voltage. Basically, it is Ohm’s Law in effect.

Though these effects are difficult to observe in the I-V measurement itself, they can be seen in the C-V curves. The capacitance in accumulation tends to be reduced, producing a stereotypical bend downward in the C-V curve [53]. Figure 3.19 shows an extreme example of such a C-V curve. Not only does the capacitance bend over prematurely, indicating a very high series resistance, but it collapses. This indicates that the gate has also shorted out. In the corresponding I-V curve, however, an anomalously low leakage is reported. A shorted gate should lead to infinite, not small, leakage current. Hence, it is likely that the device shown in figure 3.19 is dominated by series resistance due to the \textit{n}-Ge/Al contact scheme. As a result, the GaIn contact scheme mentioned in section 3.3 was adopted for the devices studied in this chapter.

Figure 3.20 shows the I-V characteristics for those devices, which are very similar between the two materials. Annealing does reduce the leakage current by 1 – 2 orders of
magnitude in substrate injection, though it does little to affect the leakage in gate injection. The characteristic shape of the leakage current also changes after annealing. Though Si/SiO₂ devices typically deal only with direct and Fowler-Nordheim tunneling processes [54], high-κ devices are dominated typically by other leakage mechanisms as well, including conduction along grain boundaries and trap-assisted (Poole-Frenkel) tunneling [13-15]. Furthermore, they often have thin interfacial oxides/nitrides between the substrate and high-κ which can affect the shape of the leakage current [14]. The change in the characteristic shape of the leakage current upon annealing, then, is most likely due to at least one of the three following causes: (1) conduction along grain boundaries, (2) a change in the interfacial layer, and (3) trap-assisted tunneling. Traps are pervasive throughout both films at all annealing conditions, as evidenced by the C-V curves. The trap density does lessen noticeably in the HfSiON devices upon annealing but not significantly in HfO₂ devices. Hence, reason (3) cannot account for their mutual and strikingly similar changes in leakage current shapes. Similarly, reason (1) is ruled out as the major driving force behind the qualitative change in leakage current. XAS data showed that HfSiON films do not crystallize to a high degree as observed in HfO₂ films; therefore, conduction along the grain boundaries cannot account for the mutual change in leakage current shape. Reason (2), however, can. Upon annealing at 800°C, the MEIS data clearly showed a conversion from a GeON interfacial layer to a GeOₓ layer in both materials. Though the conversion is unlikely to be complete at the 600°C and 650°C annealing temperatures shown here, the amount of nitrogen in each interfacial region is reduced (as witnessed by figure 3.2). Of course, leakage mechanisms (1) and (3)
undoubtedly do play a significant role in these devices, but it is suspected that mechanism (2) is behind the qualitative change in the shape of the gate leakage described above. These results are in good agreement with measurements and observations made on similar device structures by others [13-15, 55].

3.6 – Summary of Results for High-κ Thin Films on Ge Substrates

In this chapter, two hafnium-based high-κ materials were discussed with regard to their incorporation onto germanium substrates. The motivation for these investigations was discussed in terms of Ge carrier mobility and its lack of an acceptable native insulator, and processing details, such as Ge surface passivation and PRE-MOCVD deposition of high-κ materials, followed. Real-time SE studies of Ge substrate cleaning methods and the subsequent oxidation in air were also discussed. Finally, this chapter gave spectroscopic results from MEIS, XAS, and SE investigations, as well as electrical results from C-V and I-V measurements of $n$-type Ge-based MOS capacitors.

The real-time SE studies found that while the most commonly used technique is not necessarily the best at removing the Ge native oxide, oxidation kinetics are a much larger factor in determining the condition of Ge surfaces prior to gate deposition. Ultimately, it seems that Ge(100) surfaces acquire a layer prior to gate deposition that is optically equivalent to about 17 Å of GeO$_2$, and Ge(111) surfaces acquire an analogous layer optically equivalent to roughly 12 Å of GeO$_2$.

MEIS studies revealed that the in-situ nitried Ge surface serves as a very good diffusion blocking layer in Ge/high-κ structures, but it seems to be somewhat less
effective against oxidation. During deposition, the interfacial GeNₙ layer converts to a GeON layer, which ultimately becomes a GeOₓ layer following thermal processing in at 800°C. This is true in both HfO₂ and (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄ structures. MEIS also revealed that (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄ is more stable chemically than HfO₂, staying resistant against Ge diffusion through a PDA process at 800°C for 1 minute. HfO₂, on the other hand, becomes impregnated with Ge throughout the material following the same process.

XAS experiments provided further insight into the removal of nitrogen from Ge/high-κ interfacial layers, as well as insight into the thermal evolution of crystallinity and conduction band edge defect states within these structures. The level of interfacial nitrogen undergoes a very big reduction between annealing temperatures of 650°C and 700°C in HfO₂ and between 700°C and 750°C in HfSiON. In addition, HfO₂ switches from type I crystallinity to type II crystallinity at this transition temperature; whereas, HfSiON seemingly maintains its type I crystalline structure following thermal treatments at temperatures up to and including 800°C (the maximum considered in these experiments). Two discrete defect levels were identified at approximately 1.0 eV and 2.5 eV below the conduction band edge and were attributed to oxygen di-vacancies (which likely are modified and exacerbated by diffused Ge atoms within the films). SE experiments conducted generally substantiated this information. Moreover, the number of observable defects was higher in HfO₂ than in HfSiON, and both materials showed an increase in defect levels coinciding with interfacial nitrogen removal (i.e. 650°C in HfO₂ and 700°C in HfSiON). This increase was much larger in HfO₂ than in HfSiON.
Lastly, electrical characterizations of Ge/high-κ MOS capacitors were carried out in order to correlate electrical properties with spectroscopic ones. Band alignment was investigated with the aid of XAS data, and several band diagrams were presented for Ge/high-κ MOS device structures. As expected from the spectroscopic data, very high levels of interface, fixed, and mobile charge traps were observed in both HfO₂ and HfSiON gates on Ge. Not only are the physical defects suspected to contribute to this phenomenon, but the band diagrams show an intrinsic potential well for electrons that develops in n-type Ge/high-κ MOS devices, which is one possible explanation for the relatively poor electrical performance of n-type versus p-type substrates observed by others. Gate leakage current was also investigated and found to be quite high in both materials. Though HfSiON was found to have superior properties relative to HfO₂ by virtually all material and device metrics considered here (except for κ, of course), Ge/HfSiON MOS devices are still far from competitive with current Si-based technologies at this time. Still, HfSiON does provide reason to suspect improvements with better processing capabilities. HfO₂, on the other hand, seems to have too many intrinsic deficits ever to be useful in such applications.
Figure 3.1: Hole mobility enhancement in (a) bulk $p$-type Ge devices and (b) in Si devices with Si-capped epitaxially grown $p$-type Ge channels. These devices use either ZrO$_2$ or HfO$_2$ high-$\kappa$ gate dielectrics (after Saraswat, et al. [4]).
Figure 3.2: XAS spectra showing the nitrogen K$_1$ edge absorption in plasma-nitrided Ge/HfO$_2$ (top) and Ge/HfSiON (bottom) structures as a function of PDA annealing temperature. (Spectra in each graph are normalized, but not spectra between graphs.)
Table 3.1: The three germanium chemical surface treatments investigated using real-time SE and AES.

<table>
<thead>
<tr>
<th></th>
<th>(1) 10 sec. etch in 10% HF</th>
<th>(2) 20 sec. etch in DI H₂O</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deegan and Hughes</strong></td>
<td></td>
<td>* Repeat six times</td>
</tr>
<tr>
<td><strong>Chabal, et al.</strong></td>
<td>(1) 2 min. DI H₂O etch</td>
<td>(2) 10 sec. dip in 30% H₂O₂</td>
</tr>
<tr>
<td></td>
<td>(3) 10 min. etch in 10% HF</td>
<td>(4) DI H₂O rinse</td>
</tr>
<tr>
<td><strong>NCSU</strong></td>
<td>(1) 5 min. DI H₂O bath</td>
<td>(2) 6% H₂O₂ rinse</td>
</tr>
<tr>
<td></td>
<td>(3) CH₃OH rinse</td>
<td>(4) 15% NH₄OH rinse</td>
</tr>
<tr>
<td></td>
<td></td>
<td>* Repeat two times</td>
</tr>
</tbody>
</table>
Figure 3.3: The imaginary part of the pseudo-dielectric function of Ge(111) for the various cleaning methods investigated in this study (done in collaboration with Seo [36]). The NH$_4$OH curve corresponds to the NCSU cleaning method; the HF curve corresponds to the Deegan and Hughes method; the H$_2$O$_2$/HF curve corresponds to the Chabal, et al. method; and the Bare wafer curve corresponds to the as received substrates.
Figure 3.4: Real-time measurements of the evolution of the (a) Ge(111) and (b) Ge(100) surfaces during two cycles of the NCSU cleaning method (done in collaboration with Seo [36]). The starting point is after an initial 5 min. H\textsubscript{2}O bath.
Figure 3.5: Real-time measurements of the surface re-oxidation on NCSU-cleaned Ge(100) and Ge(111) surfaces (done in collaboration with Seo [36]).
Figure 3.6: AES spectra showing the chemical composition of the native and re-grown Ge(100) overlayers following several of the cleaning methods discussed in the text (done in collaboration with Seo [36]). The wide scan (a) shows the relative concentration of the chemical elements present at the surface, and the narrow scan (b) shows the relative levels of Ge-O and Ge-Ge bonding at the surface.
Figure 3.7: AES spectra showing the chemistry of RPE-MOCVD-deposited thin films of (a) HfO₂ and (b) (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄.
Figure 3.8: MEIS data for HfO$_2$ and (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ high-κ materials deposited on plasma-nitrided Ge(100) substrates. Figures on the left are of HfO$_2$, and figures on the right are of (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$. Figures on the bottom correspond to as-deposited films, and figures on the top represent the films after RTA at 800ºC. The top of the sample is at 0 nm, with the depth increasing into the sample. The vertical lines are for visual clarity, denoting transition regions. (Special thanks to K.B. Chung.)
Table 3.2: A summary of the approximate chemical make-up of the films shown in figure 3.8.

<table>
<thead>
<tr>
<th>Film</th>
<th>Layer 1</th>
<th>Layer 2</th>
<th>Layer 3</th>
<th>Layer 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>HfO₂ as-dep.</td>
<td>Ge(100) Substrate</td>
<td>12 Å GeON</td>
<td>11 Å HfGeOₓ</td>
<td>29 Å HfO₂</td>
</tr>
<tr>
<td>HfO₂ at 800°C</td>
<td>Ge(100) Substrate</td>
<td>21 Å GeOₓ</td>
<td>→ HfO₂: Ge</td>
<td>23 Å HfO₂: Ge</td>
</tr>
<tr>
<td>HfSiON as-dep.</td>
<td>Ge(100) Substrate</td>
<td>9 Å GeON</td>
<td>9 Å HfGeSiON</td>
<td>27 Å HfSiON</td>
</tr>
<tr>
<td>HfSiON at 800°C</td>
<td>Ge(100) Substrate</td>
<td>6 Å GeOₓ</td>
<td>6 Å HfGeSiON</td>
<td>26 Å HfSiON</td>
</tr>
</tbody>
</table>
Figure 3.9: Oxygen K$_1$ edge XAS spectra demonstrating the thermal evolution of HfO$_2$ and (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ thin films deposited on plasma-nitried Ge(100) and Ge(111) surfaces.
Figure 3.10: Oxygen K_{1} edge XAS spectra demonstrating the asymmetry of HfO_{2} films deposited on Ge(100) and Ge(111) substrates and annealed at 800°C.
Figure 3.11: The conduction band defect structure for thin films of (bottom) HfO₂ and (top) (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄ annealed at 800°C as deduced by ²nd derivative XAS. The bandgap assignment for HfO₂ is taken from Wong and Iwai [13], and all other energies listed in the band diagrams are relative to the conduction band edge in HfO₂ (529.8 eV → 5.7 eV).
Figure 3.12: Gaussian fitting of the discrete sub-band edge defect states in (a) HfO$_2$ and (b) (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ thin films (fitting by K.B. Chung). The area of the Gaussian functions for the normalized spectra are plotted as a function of temperature in (c).
Figure 3.13: The $\varepsilon_2$ portion of the dielectric functions for thin films of HfO$_2$ and HfSiON, both as-deposited and annealed at 800°C (data taken and analyzed by Seo [36]).
Figure 3.14: (Left) Depth-resolved cathodoluminescence spectra showing defect level transitions in the bandgap of HfO$_2$ films (CL data is from Brillson, et al. [45], © American Vacuum Society, 2007). (Right) The assigned valence band defect states consistent with the CL and XAS data.
Figure 3.15: XAS spectra for plasma nitrided and plasma oxidized Ge surfaces as compared to plasma-nitrided and plasma oxidized Si surfaces.
Figure 3.16: Relevant band diagrams for the Ge/HfO$_2$ electrical devices studied in this chapter. Depicted above are (a) flat band conditions for GeO$_2$, Ge$_3$N$_4$, and SiON ITRs, (b) a comparison of substrate injection between GeO$_2$ and SiON ITRs, and (c) a comparison of gate and substrate injection for the GeO$_2$ ITR.
Figure 3.17: Typical C-V hysteresis curves for (bottom) \((\text{HfO}_2)_{0.3}(\text{SiO}_2)_{0.3}(\text{Si}_3\text{N}_4)_{0.4}\) and (top) HfO\(_2\) gate dielectrics on \(n\)-type Ge substrates in these experiments (measurement frequency is 1 MHz). The sweep direction is left-to-right-to left.
Figure 3.18: Typical C-V hysteresis curves for (bottom) \((\text{HfO}_2)_{0.3}(\text{SiO}_2)_{0.3}(\text{Si}_3\text{N}_4)_{0.4}\) and (top) \(\text{HfO}_2\) gate dielectrics on \(n\)-type Ge substrates in these experiments (HfSiON was annealed at 650\(^\circ\)C, and \(\text{HfO}_2\) was annealed at 600\(^\circ\)C). The sweep direction is left-to-right-to left.
Figure 3.19: Anomalously low leakage current in (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$-based devices on Ge as a result of high series resistance (top). Evidence of the high series resistance is in the C-V curve (bottom). (Figures are from Lucovsky, et al. [12], © 2008 Elsevier B.V. All rights reserved.)
Figure 3.20: The effects of annealing on the I-V characteristics of the (a) HfO₂ and (b) (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄ devices studied in this chapter. A Si device is shown in each plot as a reference for comparison. The flat band voltage used is only approximate as a result of the large C-V hysteresis.
References


Chapter 4

Hafnium Based High-κ Materials on Pseudomorphic Germanium Surfaces Grown on Silicon

4.1 – Brief Restatement of the Motivation

As mentioned in section 1.4, the potential benefit in using ultrathin pseudomorphic layers of Ge on Si substrates is to avoid the intrinsic sub-oxide ITR formation that occurs at Si/high-κ interfaces. A thin layer of semiconducting Ge on the Si substrate should not add to the EOT of the device, and the thermal instability of GeO₂ and Ge₃N₄ discussed in prior chapters provides a way in which the Ge/high-κ ITR may be removed (e.g. by a PDA process). Given the thermal instability of Ge, itself, observed during several of the experiments discussed in chapter 3, the probability of success for this application is expected to be low; nevertheless, the results from these experiments may also provide additional details about the physics and chemistry governing the Ge/high-κ material system as well.

4.2 – Sample Preparation

In order to prepare these samples for deposition, silicon wafers were first cleaned in Mallenkrodt-Baker’s JTB-100 solution for 10 minutes before being blown dry with dry nitrogen and placed immediately into a furnace. A sacrificial SiO₂ layer was then grown at 900°C in an oxygen environment for 20 minutes, followed by a 30 minute anneal in
nitrogen. Just before gate deposition, the SiO$_2$ layer was etched in a 1% HF solution until
the substrate became hydrophobic, plus an extra 20% of the time required for this to
occur (see refs. [1, 2], for example). The H-passivated samples were then blown dry and
placed immediately into the load lock. The load lock chamber was pumped to a pressure
of approximately $1 \times 10^{-7}$ Torr, and the samples were transferred into the RPE-MOCVD
chamber. Once in the PRE-MOCVD chamber, the PM-Ge layer was deposited.

Numerous researchers have reported the low-temperature growth of PM-Ge and
related Ge layers by remote and direct plasma methods [3-9]. Because of the inherent
4.4% lattice mismatch between crystalline Si and Ge layers, there is a so-called critical
thickness for such layers, beyond which the Ge is no longer pseudomorphic (crystalline).
Beyond this, threading and misfit dislocations form, the Ge layer develops grain
boundaries, and the layer becomes polycrystalline [10]. Theoretical and experimental
evidence place this critical thickness at approximately 14 Å for PM-Ge layers deposited
on silicon substrates [10, 11].

All gases used in these experiments were of research grade (99.9999% purity).
For the experiments discussed in this chapter, the PM-Ge layers were deposited at a
pressure of 200 mTorr and at a temperature of 300°C. The Ge source was a 2% mixture
of GeH$_4$ in a helium balance (see section 2.3), which was supplied at 10 sccm for a
downstream process (see section 2.1). Helium was supplied at 200 sccm in an upstream
process as the source of plasma-excited species. The RF plasma excitation power was
30W, and the excitation frequency was 13.56 MHz. Because of the surface reactivity of
Ge layers in air (see section 3.2), accurate *ex-situ* measurement of the thickness of such a
thin layer was prohibitively difficult to make. The best thickness estimate available for such layers was AES measurement (see appendix F). The height of the Si-Si substrate bonding peak was monitored as a function of deposition time until it disappeared, as shown in figure 4.1. Experimental deposition times were made roughly in powers of 2 relative to this time, ranging from 7 seconds to 520 seconds (7, 15, 30, 65, 130, 260, and 520 seconds).

Following the Ge deposition, some of the samples were plasma-nitrided for 90 seconds, some were plasma-nitrided for 600 seconds, and some samples were left un-nitrided. This nitridation process was motivated by the preliminary results on Ge substrates, which indicated that this layer protects and passivates Ge surfaces during gate deposition (see section 1.2). Furthermore, the thermal decomposition of GeNₓ layers during PDA treatment (see chapter 3) suggested the possibility of achieving direct contact between Ge and high-κ layers during post-deposition processing. Other researchers have reported, in agreement with the results presented in chapter 3, that Ge surface nitridation also serves to reduce interface trap density [13], prevent accumulation capacitance degradation [13], and prevent Ge migration [14]. The actual plasma nitridation process took place at a pressure of 300 mTorr and at a temperature of 300ºC. Helium was supplied at 200 sccm in an upstream process, and nitrogen was supplied at 60 sccm in an upstream process as well. The RF plasma excitation power was 30W, and the excitation frequency was 13.56 MHz.

Finally, the gate material was deposited. For spectroscopic studies, both HfO₂ and (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄ were chosen, and the electrical devices employed only
(HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$. All gate depositions took place at a pressure of 300 mTorr and at a temperature of 300ºC. The RF plasma excitation power was 30W, and the excitation frequency was 13.56 MHz. Helium was supplied in an upstream process at 200 sccm. For HfO$_2$ films, oxygen was supplied at 20 sccm for an upstream process, and HTB carrier gas (see section 2.2) was supplied at 10 sccm for a downstream process. For (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ films, HTB carrier gas was supplied at 11 sccm for a downstream process, nitrogen was supplied at 74 sccm for an upstream process, a 2% SiH$_4$ mixture in a helium balance was supplied at 20 sccm for a downstream process, and a 1% N$_2$O mixture in a nitrogen balance was supplied at 6 sccm for an upstream process. The growth rates for both films were estimated by AES to be approximately 10 Å/min, and films were generally deposited for between one and four minutes.

Following gate deposition, samples were subjected to a PDA treatment by RTA for one minute at various processing temperatures, ranging from 600ºC to 800ºC (see section 2.1). Again, the benefits of a post-deposition anneal are to remove the interfacial nitride layer [15], to allow mechanical strain relaxation [16, 17], and to allow chemical self-organization [18-20].

The sample preparation for spectroscopic samples ended here; however, the MOS capacitors were subjected to further processing in order to form the electrodes for electrical contact. First, these samples were placed into a furnace and annealed in forming gas at 450ºC for 30 minutes. They were then placed into a resistively heated metal evaporation chamber where their front sides were coated with 3000 Å of aluminum as a gate electrode. Next, the samples were subjected to a 5 minute dehydration bake on
a 115°C hot plate and spin-coated with photoresist. The photoresist was then soft-baked at 115°C for 1 minute. Standard MOS capacitor gate electrodes were patterned by a photolithography mask while the samples were exposed to a 275 W mercury vapor lamp for 20 seconds. The exposed regions of photoresist were removed from the samples by placing them into a developer solution for one minute while constantly agitated. Once blown dry, the remaining photoresist was hard-baked at 115°C for five minutes. In order to remove the exposed aluminum material, the samples were placed into a commercial aluminum etching solution (containing primarily nitric acid) at 40°C and agitated periodically until the aluminum was no longer visible, plus an extra 20% of the time required to achieve this effect. After being blown dry, the samples were then placed into a commercial photoresist stripping solution for 10 minutes to remove the remaining photoresist. After being blown dry again, the gate electrodes were completely formed. The final step was to form the backside (substrate) electrode. The samples were placed again into the resistively heated metal evaporation chamber where 2000 Å of aluminum were deposited. Complete with contacts to the Si substrate and the high-κ gate, the samples were then full-fledged MOS capacitor devices and ready for testing.

4.3 – Spectroscopic Properties of Si/PM-Ge/High-κ Structures

Like the Ge/high-κ samples discussed in chapter 3, XAS spectra were measured for several Si/PM-Ge/high-κ structures at beamline 10-1 of SSRL. The samples investigated by these means consisted either of HfO₂ or of HfSiON overlayers, were either as-deposited at 300°C or annealed at 700°C, and were deposited on either Si(100)
or on Si(111) substrates. All of the samples had an intermediate Ge layer which was deposited for 260 seconds. As electrical results later indicated, this deposition was too long to achieve pseudomorphic growth conditions. Hence, it is suspected that this Ge layer was polycrystalline, with numerous grain boundaries. The XAS results are interpreted in this context.

First, examine the nitrogen K edge in the Si/Ge/HfO$_2$ samples depicted in figure 4.2. Not only does a non-trivial amount of nitrogen remain in the sample following a 700ºC anneal, but the as-deposited spectrum shows significant enhancement in the N $\pi-\pi^*$ peak over the corresponding spectrum shown for Ge/HfO$_2$ in figure 3.2. This feature is associated with molecular nitrogen, and it is hypothesized that grain boundaries in the interfacial Ge layer provide the extra regions in which such molecules can be trapped.

Next, examine the raw XAS spectra for the samples shown in figure 4.3. This data illustrates differences between various structures with respect to high-κ material, substrate orientation, and annealing temperature. Figure 3.10 shows that an anisotropy exists between HfO$_2$ deposited on single-crystal Ge(100) and Ge(111); however, this does not exist in the present system. This is one indication that the interfacial Ge layer has exceeded the critical thickness and is not pseudomorphic. If it was, this layer show a corresponding anisotropy as a result of the underlying Si(100) or Si(111) symmetry.

Just as noticeable is the fact that the HfO$_2$ samples show a trend with annealing temperature that is opposite to the one demonstrated by figure 3.9. Figure 4.4 shows that the Si/Ge/HfO$_2$ samples actually self-organize, reducing their crystal field and Jahn-Teller splitting terms rather than increasing them as in the Ge/HfO$_2$ samples studied in
chapter 3. In particular, the \( e_g \) states of the as-deposited HfO\(_2\) film clearly are split by approximately 1.2 eV; however, the \( e_g \) states in the HfO\(_2\) film annealed at 700\(^\circ\)C are degenerate. Contrast this with the samples shown in figure 3.9. The as-deposited film does not show any obvious \( e_g \) state splitting, but as the annealing temperature is increased, these states show clear separation. In fact, the annealed HfO\(_2\) film depicted in figure 4.4 resembles the HfSiON films depicted in figures 3.11 and 4.4. Because the interfacial Ge likely possesses grain boundaries in the Si/Ge/HfO\(_2\) sample, this provides an easy pathway for the migration of atoms between the HfO\(_2\) and the underlying Ge layer. As a result, these two layers are suspected to mix chemically during annealing, forming a HfGeO\(_x\) alloy mixture and possibly retaining some of the nitrogen which is still present following annealing (see figure 4.2).

Also note the defect structures present in figure 4.5. The HfSiON films have on discrete defect structure at approximately 3.3 eV below the conduction band edge and a broader defect feature centered at roughly 1.3 eV below the conduction band edge. The HfO\(_2\) films, on the other hand, show two discrete conduction band edge defect features – one at approximately 3.4 eV below the conduction band edge and one at approximately 1.1 eV below the conduction band edge. These energy band diagrams are very similar to the ones shown in figure 3.11 and are presumed to have the same origin – namely, oxygen di-vacancies.

Ultimately, the Si/Ge/high-\( \kappa \) samples have XAS signatures very similar to the Ge/high-\( \kappa \) samples. The main differences observed can be attributed to the presence of grain boundaries in the Ge interfacial layers of the former structures. In actual
pseudomorphic interfacial Ge layers, the XAS signatures are anticipated to be nearly identical to the ones discussed in chapter 3; however, more studies are required to confirm this assertion.

### 4.4 – C-V Properties of Si/PM-Ge/High-κ Structures

Similar to the structures discussed in chapter 3, MOS capacitors were fabricated from Si/PM-Ge/HfSiON structures, and both C-V and I-V measurements were made in order to investigate their electrical properties. By studying the frequency dispersion, flat band voltage, and hysteresis on these devices, some insight can be gathered into the relationship between the spectroscopic and chemical structure investigated in the previous section and the electrical properties of related devices. The C-V curves for various samples are shown in figure 4.6. This figure shows how frequency dispersion and flat band voltage are affected by the thickness of the interfacial Ge layer. As expected, an increase in interfacial Ge leads to degradation of device performance; however, there are some interesting trends to notice.

First, note that the 1 MHz C-V curve for each of the three layers having interfacial Ge are roughly equivalent in terms of shape and flat band voltage. These uniform shifts in flat band voltage relative to the Si control sample indicate that there exist permanently occupied traps, even at very fast sweep rates. Moreover, these permanently occupied sources of positive fixed charge are due to the presence of Ge, and since this trait does not scale with Ge thickness, they are most likely related to the Si/Ge and/or the Ge/HfSiON interface(s). Based on AES data (not shown), it is known that the HF-last Si
surface is not chemically pure; there are small traces of carbon and oxygen present on the surface which are expected to degrade the Si/Ge interface. The Ge/HfSiON interface is formed in situ in a chemically controlled environment; therefore, it is expected to be of much higher quality. Of course, the impurities at the Si/Ge interface may lead to reduced quality in the Ge layer, itself, which can then affect the quality of the Ge/HfSiON interface as a result. Hence, these charges may be located at the Si/Ge interface, in the Ge layer, and/or at the Ge/HfSiON interface, but despite their actual location, it is the Si/Ge interface quality that is suspected to be the root cause. In the Si/HfSiON device, carbon impurities are removed during the remote plasma-assisted oxidation (RPAO) step, and the oxygen is actually beneficial at the Si surface. The Si/Ge/HfSiON devices, however, cannot afford the luxury of this processing step. If the Si/Ge interface quality is responsible for the shift in flat band voltage, in situ treatments of the silicon substrate, such as high-temperature heating or low-pressure hydrogen plasma cleaning, may reduce the amount of this type of fixed charge.

Figure 4.6 also shows the frequency dispersion in these devices. Clearly, the frequency dispersion increases with Ge layer thickness; however, the difference between the 30 second and 130 second depositions is very drastic. Not only does the flat band voltage shift dramatically as a function of frequency in the device with the thicker Ge layer, but the depletion/accumulation transition sharpens significantly. This indicates that, in addition to the permanently filled interface traps discussed above, the overlayers provide positively charged, relatively slowly responding traps that scale with Ge layer thickness. Therefore, trapping sources within the Ge and/or HfSiON layers must be
present. Moreover, the sharpening which accompanies this flat band voltage shift suggests that, at low frequencies, more carriers are generated in the channel region of the device structure. While it is difficult to pinpoint the exact causes for this, it is presumed to be related to one or more of the two following factors: (1) grain boundary formation in the Ge layer as a result of the pseudomorphic to non-pseudomorphic transition and/or (2) the relative thickness of the Ge layer as compared to the thickness of the channel.

As for reason (1), not only do grain boundaries provide numerous dangling bonds as positively charged trap sources, they also provide a mechanism for increased atomic migration between layers during PDA treatment (see section 4.3). These altered chemistries and any resultant defects, such as di-vacancies, are likely to have an effect on this as well. Moreover, the Ge layer is likely to have inherent bonding defects and impurities in the bulk, including hydrogen-passivated bonds. The time constant for trapping and de-trapping in these processes may be such that they are not noticeable at high frequencies but become much more apparent as the frequency is reduced.

In regard to reason (2), it is important to note that the Ge layer is undoped. This means that as more of the channel is confined to the Ge layer, the more important carrier recombination and generation rates in this layer become. Because the Si layer is doped, and the Ge layer is not, for all intents and purposes carriers are generated within the Si region. This means that carriers must traverse the undoped Ge region in order to form at least part of the channel. For very thin Ge layers, this problem is minor. For example, the Ge layer deposited for only 7 seconds is similar to the Si reference device in terms of
frequency dispersion. In these devices, Si still comprises the majority of the channel region, which is why the effect is only minor. For very thick Ge layers, however, it is a significant effect because the channel is now confined primarily within the Ge, not the Si. At low frequency, many more carriers can traverse the Ge region in the channel, increasing the effective doping level and leading to a sharper C-V characteristic. Of course more carriers lead to more filled traps, and an increase in the amount of positive fixed charge is observed. At low frequency, these effects seem to dominate, and at high frequencies, interface state trapping appears to dominate. Regardless of the precise mechanism responsible for the increased frequency dispersion, it ultimately degrades device performance and should be minimized.

Next, consider the C-V hysteresis in these device structures. The relevant characteristics, which were obtained at 1 MHz, are shown in figure 4.7. The amount of hysteresis at 1 MHz in each of the three Si/Ge/HfSiON devices is comparable to the hysteresis in the Si/HfSiON control sample. In fact, the sample with the Ge layer which was deposited for 130 seconds actually demonstrates less hysteresis than the control. As with the frequency dispersion curves, all three of the Si/Ge/HfSiON devices show basically the same flat band voltage shift relative to the Si/HfSiON device (approximately 0.5 – 0.7 V). As discussed above, this positive fixed charge is suspected to be derived from impurities at the Si/Ge interface. The difference is that the amount of positive charge traps which are filled is reduced as the sweep direction is reversed. It is these mobile traps which are related to the hysteresis.
Again, the hysteresis in all of the samples is comparable; therefore, it is likely that the origin of the hysteresis is the same in all of the samples. In addition, the hysteresis relates to the response of the carriers in the channel to changes in the applied field. Hence, the hysteresis is suspected to be a result of the semiconductor/HfSiON interface where the channel resides. As for the sample depicted in figure 4.7d, the hysteresis actually is reduced. This reduction should not be interpreted as a reduction in trapping, rather as a change in the response time of the traps. Basically, these traps are no longer mobile, but fixed. Since the interfacial Ge layer in this device is not expected to be pseudomorphic as in the other two, the relative quality of the Ge/HfSiON interface is likely to be much worse (i.e. more dangling bonds, grains, etc.). As a result, the traps may be deeper and slower to respond to changes in the applied field. Put simply, the C-V curve in figure 4.7d may appear to be better than the others when taken at face value, but it actually indicates that the device properties are, on the whole, worse.

Numerical estimates for the hysteresis, fixed charge density, interface trap density, and frequency dispersion are summarized in figure 4.8. The hysteresis and frequency dispersion were measured in terms of their changes in flat band voltage with sweep direction and frequency, respectively. The fixed charge and interface trap densities were measured by modeling the C-V curves (see appendix J), and the interface charge was also measured by using the high-low frequency method. The resulting interface trap densities from these two methods were averaged. Three basic insights can be gleaned from these figures:
(1) The Effect of Interfacial Ge Thickness: Essentially, all of the properties measured get worse monotonically as a function of interfacial Ge thickness, with the exception of the hysteresis. The hysteresis shows a plateau, which is expected to be related to the pseudomorphic to non-pseudomorphic transition as discussed above.

(2) The Effect of Interfacial Ge Nitridation: Nitriding the Ge layer reduces device quality in these structures. The hysteresis and frequency dispersion are increased, and the interface trap density is relatively unchanged. The fixed charge density is reduced, but this is merely at the expense of an increased hysteresis. As explained above, this is simply a shift from mobile charge to fixed charge, which yields a net reduction in device quality in this case.

(3) Annealing Temperature: These figures also show that the optimal annealing temperature is approximately 700°C. The best combination of these properties is attained at that processing temperature.

In addition to the relatively thick devices discussed above (which have EOT in the 30 Å range), an attempt was made at fabricating devices with very low EOT. In fact, deposition times were cut by as much as 75% in the thinnest films. The C-V curves measured on those devices are shown in figure 4.9. First, note that the flat band voltage shift is approximately 0.3 V less than for the devices with thicker HfSiON films. This suggests that at least some of the positive fixed charge discussed earlier resides within the
HfSiON film. The root cause, however, still is suspected to be the contamination of HF-last Si surfaces in these experiments.

Also note that the EOT is only 20 Å. As discussed in appendix J, this thickness is likely to be an overestimate, but it is the accumulation capacitance that is important. From this perspective, the classical EOT is the relevant number, which is 20 Å in this case. Based on the estimated physical thickness of the HfSiON layer, the EOT should have been approximately \( \frac{1}{4} \) of this. Figure 4.10 shows the measured EOT of this device structure as a function of deposition time. This graph demonstrates quite clearly that there is an additional source of capacitance beneath the HfSiON overlayer. Moreover, this layer contributes approximately 14 Å to the total EOT, which by itself is larger than the desired goal of less than 10 Å. This layer likely depends on oxygen and/or nitrogen from the HfSiON overlayer; therefore, the linear extrapolation near the y-axis is probably an oversimplification. Nevertheless, the MEIS data suggests that there is almost certainly an interfacial HfGeO\( _x \) region (which is more like low-\( \kappa \) GeO\( _2 \) near the interface), and this is presumed to be the cause of the additional capacitance. The kink seen near the accumulation region in figure 4.10 is also expected to be due to the parasitic interfacial region. This shoulder softens with annealing temperature as the interfaces in the device structure migrate and chemically re-order themselves.

Based on the above discussions, it seems as though the interfacial HfGeO\( _x \) layer plays the same role in Ge/high-\( \kappa \) based devices as the interfacial SiO\( _x \) region plays in Si/SiO\( _2 \) based devices, and it may be unavoidable.
4.5 – Gate Leakage Properties of Si/PM-Ge/High-κ Structures

The gate leakage characteristics of Si/PM-Ge/HfSiON device structures were also investigated during these studies. A summary of the results are shown in figure 4.11 along with the gate leakage obtained from a Si/SiO₂ reference device having the same EOT (~ 3nm).

The first thing to notice about these I-V curves is that for many of the devices, the minimum leakage current does not occur at 0 V gate bias. The only way for this to occur is for surface charges to be present at one or more of the internal interfaces. Given the direction of this shift, the internal electric field must point in the direction from the gate electrode to the substrate electrode. It should be noted that this effect does not occur in Si/HfSiON structures (not shown) and is, therefore, related to the presence of the Ge layer. Moreover, the effect appears to be diminished by annealing and, in general, more prominent in the samples with thinner Ge layers. As already discussed, the thinner Ge layers are expected to be pseudomorphic, providing a structurally sound template onto which the HfSiON gate can be deposited. The thicker Ge layers are cracked, providing easy pathways for atomic migration along grain boundaries. Hence, the thinner Ge layers provide a more abrupt Ge/HfSiON interface. As the MEIS data shows, annealing also encourages atomic migration and chemical re-organization at this interface. Annealing, therefore, also softens the abruptness of the Ge/HfSiON interface. As a result, samples with thin Ge layers and which are annealed at lower temperatures are expected to have much more well-defined Ge/HfSiON interfaces than do samples with thicker Ge layers and annealed at higher temperatures. The Ge/HfSiON interface, therefore, is presumed to
be one location of the surface charges responsible for the non-zero minima in figure 4.11. This means that the Ge layer, the HfSiON layer, or both are responsible for at least part of the internal field. Because it is reasonable to expect that the Si/Ge interface undergoes similar transitions, it is not possible to rule out that interface as an additional location for surface charges.

Note that annealing at 700°C does not get rid of this effect in the sample with the thinnest nitrided Ge layer. In fact, at 750°C (shown in figure 4.12) the effect is still present, though it is reduced to approximately 0.1 V. As the MEIS data in chapter 3 showed, this is to be expected. The nitridation prevents diffusion at the Ge/HfSiON interface, and the HfSiON overlayer delays interfacial nitrogen removal until a temperature of approximately 750°C. Because the films with nitrided Ge layers deposited for 65 seconds or longer do lose this feature upon annealing at 700°C, it is likely that grains are present in those samples. If this was not the case, the interfaces would not lose their integrity. Hence, it is likely that films with Ge layers deposited longer than 30 seconds are no longer pseudomorphic (which is to be expected due to the presumed 30 – 40 Å/min Ge deposition rate). Moreover, this effect is not completely removed even at 800°C in the samples with un-nitrided Ge layers deposited for 30 seconds or less (see figure 4.12). These samples are still shifted by approximately 0.05 V. By this point, the Ge/HfSiON interface is no longer abrupt. Following anneals at 750°C or more, non-zero leakage minima persist only in samples with Ge layers deposited for 30 seconds or less (all shown in figure 4.12). This suggests that this phenomenon is related to the
pseudomorphicity of the Ge layer and that at least some surface charge is likely to be present at Si/Ge interfaces.

The gate leakage characteristics for devices with thinner HfSiON layers are shown in figure 4.13. These are the devices discussed in the previous section whose EOTs are 20 Å and whose C-V characteristics are depicted in figure 4.9. The results are quite reasonable. The most notable difference between these characteristics and those in figures 4.11 and 4.12 is that the I-V minima for all annealing temperatures in these devices are zero. Given that the HfSiON layers in these devices are thin enough (approximately < 4 nm) to self-organize and ward off mechanical defects [21, 22], the number of charged defect states is expected to be much lower than in the thicker films. This is because not only is there less material (and hence likely fewer traps) present, but also because the material is better structurally (i.e. less dangling bonds, etc.).

Intuitively, these fixed charge traps are almost certainly the reason for which surface charges form at the Ge/HfSiON interfaces in the devices with thicker gates. MOS device modeling (see appendix J) shows that such charged traps lead to non-zero surface potentials at zero gate bias. These internal fields, which were alluded to earlier, are due to a reconfiguration of free carriers at the internal semiconductor/gate interface, which itself is a response to microscopic fields caused by things like fixed charge, polarization, etc. Of course, it should be stressed that these effects are transient in the presence of gate leakage through an external circuit; current cannot flow forever in a device simply by connecting the gate and substrate electrodes. On the other hand, the amount of current that is flowing through the gate during these measurements is very
small, and there is an associated discharging rate. By significantly reducing the I-V measurement sweep rate (depending on the RC time constant of the device), then the minima for all of the devices would be zero.

The last thing to note about figures 4.11 and 4.13 is the asymmetry between gate injection and substrate injection conditions. As demonstrated by the SiO$_2$ sample in figure 4.11, the capacitor’s field effect limits leakage in the gate injection mode. This is to be expected. Note, however, that only the samples with Ge layers which were deposited for 30 seconds or less show this same effect to any significant extent. Again, this demonstrates a relationship between the pseudomorphicity of the Ge layer and device quality.

As discussed in appendix K, gate leakage can be modeled by several different methods, depending on the origin of the leakage mechanism. Figure 4.14 shows the best fits to experimental data obtained from both 10 Å EOT and 40 Å EOT devices with the Si/Ge/HfSiON structure. The fitting was done under substrate injection conditions, and a trap-assisted tunneling (TAT) model yielded the best overall match to the experimental curves. As a result, the dominant leakage mechanism in these devices is suspected to by trap-assisted tunneling. These results are in line with what others have observed for high-$\kappa$ gate dielectric stacks as well [23-25].

Overall, the origins of the electrical properties of these devices are very difficult to pin down. Ultimately, the idealized Si/Ge/HfSiON heterostructure model is not valid because of the physical and chemical changes observed within related samples. Several basic band diagrams are shown in figure 4.15 to illustrate some of the possible changes.
that may occur during deposition alone. Of course, these structures are idealized as well. Upon annealing, the idea of abrupt interfaces with the high-κ film is no longer valid. Atoms migrate and reconstruct, smearing out the interfacial transition region and forming a hafnium germanate region. In some ways, this is beneficial. The inherent quantum wells that would form for electrons in these structures are not as well-defined. For example, the kink observed in figure 4.9 may be due to this effect. Given the MEIS results, a picture similar to figure 4.15c is likely similar to the actual structure of these devices. The kink is possibly due to the filling and emptying of the quantum well that would form under a moderate, positive gate bias when applied to the structure shown in this figure. Upon annealing, interfaces are blunted, the well is no longer well-defined (excuse the pun), and the kinking effect is lessened. On the other hand, these chemically mixed regions make it difficult to achieve the end goal of a very low EOT and, as discussed above, can degrade a number of other device quality metrics.

**4.6 – Summary of Results for Si/PM-Ge/High-κ Structures**

In this chapter, Si/PM-Ge/high-κ structures were studied in the hope that these devices could provide a pathway to a low EOT, avoiding the intrinsic low-κ sub-oxide region that occurs in Si/SiO₂ technology. Such hopes were inspired by preliminary results on Ge/high-κ samples showing that GeNx and GeOx layers are volatile during PDA treatments. Moreover, results suggested that Ge nitridation may provide a suitable passivation layer during gate deposition. Any such oxidation and/or nitridation, intentional or unintentional, which occurred during deposition, could later be removed
during PDA treatments. These observations provided a potential pathway to reduce the device EOT. Specifically, a thin, pseudomorphic, sacrificial layer of Ge could be deposited at the surface of a Si substrate in order to prevent surface oxidation during gate deposition. During PDA treatments, the nitrogen and oxygen at the Ge/high-κ surface would be driven off, leaving the two materials in direct contact with no parasitic low-κ regions.

Various physical, chemical, and electrical properties of the Si/PM-Ge/high-κ stacked structure were investigated by spectroscopic means, as well as by fabricating and testing MOS capacitors. XAS experiments were carried out in order to study the electronic structure of the conduction bands in these samples, yielding information relevant to their bonding as physical structure, as well as their sub-band edge defect levels. C-V experiments investigated the nature of charge traps in these structures and how they affect device properties. Finally, I-V measurements were carried out in order to examine the gate leakage characteristics of these MOS device structures.

XAS results, which are summarized in figures 4.2 – 4.5, depict respectively the following: interfacial nitrogen removal upon annealing, a lack of anisotropy between Si(100) and Si(111) substrates, the thermal evolution of d-state splitting terms and sub-band edge defect levels in the high-κ films, and the conduction band edge energy diagram for the high-κ films. As expected, nitrogen removal in these films is analogous to the Ge/high-κ structures discussed in chapter 3. Unlike those samples, however, the Si/Ge/high-κ samples studied here did not demonstrate any anisotropy between Si(100) and Si(111) substrates. It is thought that this is due to the interfacial Ge layer not being
pseudomorphic, as it is very likely that these layers exceeded the critical thickness for Si/Ge heterostructures. For pseudomorphic layers, it is thought that the same anisotropy observed in the Ge/high-\(\kappa\) samples will be present; however, more studies are required to verify this. The conduction band edge structure of these high-\(\kappa\) films is very similar to those observed in chapter 3 as well. Essentially the Si substrate is hidden from these materials and has little effect on their electronic structure. The crystalline quality of the interfacial Ge layer, however, does affect the conduction band edge structure in HfO\textsubscript{2} films. Rather than phase separating and crystallizing as they are annealed like the samples studied in chapter 3, these HfO\textsubscript{2} films actually reduce their d-state splitting terms when heated. This is almost certainly due to alloying with the underlying Ge layer, which is in excess of its critical thickness in this structure, providing convenient reaction pathways for this to occur. As in the case of anisotropy, PM-Ge layers are expected to resemble their single-crystal Ge substrate cousins with respect to d-state splitting, though more experiments are required to verify this. Finally, two sub-band edge defect levels are again observed, approximately at the same positions as in the samples from chapter 3. The main difference is this case is that the D\textsubscript{1}(CB) state is not as well-defined, especially in the HfSiON films. This defect level is more like a set of diffuse, band-tail defects. The non-pseudomorphicity of the interfacial Ge layer is again blamed for this. The bonding and chemistry is not as well-defined; therefore, neither is the position of this defect state. Nevertheless, the origin of the defects is still suspected to be related to oxygen di-vacancies, some of which may be modified by the presence of Ge at those sites.
C-V studies of Si/Ge/HfSiON MOS stacks revealed a wide variety of information regarding traps in these structures. The results are summarized in figures 4.6 – 4.10, and they respectively depict C-V frequency dispersion, high-frequency C-V hysteresis, the effects due to varying process conditions on various device metrics, and an indication of a parasitic low-$\kappa$ interfacial region. Four main factors (which are not necessarily independent) are expected to contribute to trapping in these devices: (1) chemical impurities at the HF-last silicon surface at the time of Si/Ge interface formation, (2) the relative thickness of the interfacial Ge layer compared to the channel thickness, (3) the appearance of grain boundaries in the interfacial Ge layer, and (4) the appearance of grain boundaries in the HfSiON layer. As discussed in the text, these factors all work together either as sources, or as precursors to sources, of fixed, mobile, and interface trapped charges. Factor (1) may be minimized by \textit{in situ} cleaning by heating or hydrogen plasma, factor (2) may be mitigated through the use of very thin PM-Ge layers (i.e. layers deposited for less than 30 seconds), factor (3) may be controlled by implementing both of the previously mentioned steps, and the effects of factor (4) may be reduced by using very thin (approximately less than 3 nm) HfSiON layers. The density of interface states in these MOS structures was on the order of mid-$10^{12}/\text{cm}^2/\text{eV}$ to low-$10^{13}/\text{cm}^2/\text{eV}$, and the density of fixed charge traps was on the order of low-$10^{13}/\text{cm}^3$ to mid-$10^{13}/\text{cm}^3$. Both of these numbers are far too large for CMOS applications. These studies also suggested that a parasitic low-$\kappa$ interfacial region forms at the Ge/HfSiON layer. Based on MEIS results, it is likely that this is a HfGeO$_x$ region, which forms during gate deposition and is
Ge-rich near the interface. Unfortunately for technology, this region may be analogous to the intrinsic SiOₓ layer at Si/SiO₂ interfaces.

Lastly, I-V measurements of gate leakage were made on Si/Ge/high-κ MOS capacitor structures. These results are summarized in figures 4.11 – 4.14, and they all depict gate leakage characteristics for various devices. The most interesting thing about these samples was the presence of non-zero minima in the gate leakage for some of the samples. This was attributed to the reorganization of surface charges in response to the presence of various fixed charge sources within the structure. In samples with more abrupt interfaces and better quality gates, these charges tended to be held in place capacitively. Because these samples had lower leakage current, the RC time constant did not allow for the rapid discharging of the internal fields. Again, at lower measurement sweep rates this effect must be reduced. In order to conserve energy, either no leakage current can flow, or the fields must be discharged. If current leaks (as is the case here), the only question is how long is takes for the device to discharge. Nevertheless, the presence of this effect does indicate the presence of transient internal fields in at least some of the devices studied, and it is linked, in part, to the thickness of the HfSiON gate. Based on comparison to various theoretical models, the gate leakage mechanism is these devices was attributed to trap-assisted tunneling. Despite the poor C-V characteristics of these devices, their gate leakage properties are quite respectable (at least in the devices with PM-Ge interfacial layers). In some cases the gate leakage is even comparable to, or better than, the gate leakage through ultrathin SiO₂-based devices with comparable EOT.
Figure 4.1: The evolution of the Si/Ge interface as a function of Ge deposition time as measured by AES.
Figure 4.2: The nitrogen K₁ edge in Si/Ge/HfO₂ samples, both as-deposited and annealed. The interfacial Ge layer was nitrided before HfO₂ deposition. Note that a non-trivial amount of nitrogen remains after annealing at 700°C.
Figure 4.3: XAS spectra showing variations in Si/Ge/high-κ samples due to substrate orientation, high-κ material, and annealing temperature. Figures (a) on the left depict (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ films, and figures (b) on the right depict HfO$_2$ films.
Figure 4.4: 2nd derivative XAS spectra showing the conduction band structures of Si/Ge/high-κ samples and their evolution with post-deposition annealing. The left-hand side depicts (HfO₂)₀.₃(SiO₂)₀.₃(Si₃N₄)₀.₄ films, and the right-hand side shows HfO₂ films. Note not only the d-state splitting terms, but the sub-band edge defects as well.
Figure 4.5: The energy band structures of the Si/Ge/high-κ samples.
Figure 4.6: The C-V characteristics for various Si/PM-Ge/HfSiON MOS capacitor structures as a function of interfacial Ge layer thickness and measurement sweep frequency. The upper left-hand curves are from a Si/HfSiON device fabricated for reference purposes.
Figure 4.7: The C-V hysteresis characteristics for various Si/PM-Ge/HfSiON MOS capacitor structures as a function of interfacial Ge layer thickness. The upper left-hand curve is from a Si/HfSiON device fabricated for reference purposes. The arrows indicate the sweep directions.
Figure 4.8: A summary of the hysteresis, frequency dispersion, fixed charge density, and interface trap density of the various Si/Ge/HfSiON devices studied in this chapter. The lines shown in the figures are simply cubic splines; they are meant only as a guide to the eye.
Figure 4.9: The 1 MHz C-V characteristics for the thinnest HfSiON layers deposited.

Note the bump near the accumulation region.
Figure 4.10: A plot of measured EOT versus HfSiON deposition time. Note the non-zero intercept.
Figure 4.11: A summary of the gate leakage characteristics as a function of annealing temperature, RPAN time, and Ge deposition time. The leakage curve obtained from a Si/SiO₂ device with the same EOT (approximately 3 nm) is shown as a reference.
Figure 4.12: Some additional gate leakage I-V curves of interest.
Figure 4.13: Gate leakage characteristics for Si/Ge/HfSiON devices with EOT of approximately 20 Å.
Figure 4.14: Best fits to substrate injection data obtained from two different devices with the Si/Ge/HfSiON device structure. The fits both use a trap-assisted tunneling model, which is a combination of Frenkel-Poole and tunneling leakage mechanisms.
Figure 4.15: Several hypothetical band diagrams depicting some of the potential changes which may occur during gate dielectric deposition: (a) the idealized Si/Ge/high-\(\kappa\) band diagram, (b) intentional plasma nitridation of the interfacial Ge surface, (c) accidental conversion of the interfacial Ge surface into GeON, and (d) intentional plasma nitridation of a thin interfacial Ge surface, which unintentionally consumes the entire Ge layer and nitrides the top layer of the Si substrate as well.
References


Chapter 5

Summary, Future Work, and Concluding Remarks

5.1 – Summary of Experiments and Results

The incorporation of hafnium-based high-$\kappa$ gate dielectrics onto Ge surfaces was studied both spectroscopically and electrically. As discussed at length in chapter 1, these structures are important to the continuation of Moore’s Law and scaling in CMOS technology. Bulk Ge substrates provide higher carrier mobility than do Si substrates, and PM-Ge surfaces were suspected to provide one potential pathway towards avoiding an inherent limitation in Si CMOS scaling. The two gate materials employed in these studies were HfO$_2$ and (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$, and the germanium surfaces used were Ge(111), Ge(100), PM-Ge on Si(100), and PM-Ge on Si(111). Chemical distributions were studied by MEIS; nano-crystallinity and electronic states were probed via XAS measurements; SE measurements provided insight into conduction band defect states; and electrical properties were investigated by measuring C-V and I-V characteristics of MOS structures.

XAS spectra revealed that, as with silicon substrates, HfO$_2$ is very sensitive to crystallization and phase separation upon PDA treatments; whereas, the HfSiON alloy is much more resistant to such effects [1-3]. These spectra also showed the existence of two distinct defect levels below the conduction band edge, which are associated with oxygen di-vacancies [1]. The presence of these defect states was also confirmed by SE
data. MEIS studies showed the formation of both Ge-rich and Hf-rich HfGeO_x regions in these structures, which are expected to contribute to and modify these defects as well. Interfacial nitridation was found to suppress the formation of this region to a degree, though not completely, and HfSiON was much more robust against these atomic migrations than were the HfO_2 films. The interfacial nitridation, while a reasonably good passivating layer during low-temperature deposition processes, is thermally volatile during PDA treatments and was shown by XAS do leave the interfacial region upon heating. In HfO_2, the critical temperature for interfacial nitrogen removal was found to be somewhere around 650°C-700°C, and in HfSiON, it was approximately 700°C-750°C.

While the I-V measurements of MOS capacitors revealed reasonably good results for gate leakage in these devices, the C-V characteristics showed very high degrees of both fixed charge traps and interface states in these materials. In fact, these defects are present to a degree that is 1-2 orders of magnitude too high for CMOS applications. Moreover, the results indicated that the transitional HfGeO_x region adds a parasitic capacitance, putting basic limitations on EOT scaling in this system. This region is essentially analogous to the SiO_x region in Si/SiO_2 MOS devices. Similar phenomena regarding these various physical and electrical properties in the Ge/high-κ material system have been observed by others as well [4-15].

5.2 – Recommendations for Future Work

There are still many unknowns with regard to this materials system into which further studies may provide insight. TEM studies, for example can provide much-needed
information about the quality and structure of the interfacial Ge layer in Si/Ge/high-κ stacks. Such investigations could reveal, for example, if the Ge layer is pseudomorphic and, if not, what the grains look like. For these stacks to work as desired, they must be single-crystal, mimicking the underlying Si substrate properties. Even the layers that are less than the critical thickness may not be pseudomorphic due to its initial nucleation. The Si surface impurities mentioned in chapter 4 may lead to island growth, rather than the uniform, 3-D growth that is desired. Furthermore, the RPE-MOCVD approach may itself present problems in the growth of these materials [16-18]. TEM can also provide additional information about alloying and the suspected HfGeO_x layer, such as how abrupt and well-defined the interfacial transitions are.

As already mentioned, the surface nucleation of the Ge deposition needs better understanding, in particular with respect to the quality of the underlying Si substrate. Others have shown, for example, that a H_2 plasma cleaning step and a thermal bake-out of the Si substrate prior to deposition can improve the Si/Ge interfacial quality [16]. Similarly, it has been shown that this particular deposition may be more sensitive to the growth temperature than other films, such as SiO_2 or the high-κ materials [17, 18]. More experiments focusing on growth conditions, especially when coupled with TEM studies, can provide valuable insight into this issue.

As for bulk Ge surfaces, more investigations into chemical etching and surface passivation by real-time SE may be valuable. The results discussed in chapter 3 show that hydrogen-passivated Ge surfaces are not very well passivated at all. The results of Chabal, et al. [19] suggest that perhaps hydrocarbons provide the real passivation
following HF etching observed by Deegan and Hughes, who presumed a hydrogen-passivated surface [20]. It is possible that hydrocarbon treatments following etching may provide more effective passivation. Since most of these compounds are volatile, they may be easily removed by an in situ, pre-deposition bake-out of the Ge substrate. Sulfur passivation may also be an avenue which provides better chemical passivation than hydrogen. The studies discussed in chapter 3 provide an excellent way by which to evaluate such cleaning procedures. Not only can one monitor their effectiveness at removing native oxide, but the subsequent surface oxidation kinetics can be evaluated in real time. It may be the case that in situ thermal baking of the Ge surface is the best cleaning method available, and implementation of this technique is currently underway in the RPE-MOCVD system used in these studies.

More studies into the nature of the HfGeOx region would be beneficial as well. As stated earlier, TEM can provide one manner by which the physical size and location of this region can be investigated. XAS studies into intentionally deposited hafnium germanate and hafnium nitro-germanate films can provide valuable information about the electronic structure, crystallinity, and thermal stability of these regions as well. In fact, these materials also may provide a chemically self-organized phase analogous to the one observed in (HfO2)0.3(SiO2)0.5(Si3N4)0.2. Since the HfGeOx region seems inherent, it may be less detrimental in a gate oxide which is more chemically similar. Perhaps (HfO2)0.3(GeO2)0.3(Ge3N4)0.4 may provide a better choice as a gate material for Ge substrates. Some of these experiments are forthcoming.
Lastly, investigations of Ge/PM-Si/high-κ structures may be useful. Si surfaces have shown much more compatibility with the high-κ gate materials investigated in this work than have Ge surfaces. This may be the better tactic in trying to develop Ge CMOS technologies, and these experiments are slated to be conducted in the near future.

5.3 – Concluding Remarks

In summary, there is a lot of work remaining do be done if Ge surfaces are to be employed in a viable CMOS technology. While the gate leakage properties in Ge/high-κ MOS capacitors are acceptable, these structures at present still demonstrate far too much trapping and poor overall C-V characteristics. Unlike Si, Ge does not seem to possess much potential for chemical surface passivation, and its native oxide does not have many, if any, of the necessary traits that are desirable in a gate material. Moreover, the high-κ alternatives explored in these studies do not appear to be compatible chemically with Ge surfaces. They tend to alloy with the underlying Ge surface, forming an intrinsic HfGeOx region which not only limits gate scaling possibilities, but which seems also to degrade important device performance metrics. Simply put, this work provides more evidence to the growing body of results which indicate severe difficulties in developing Ge-based CMOS technology.

One of the most interesting possibilities which remains to be explored towards this end is the (HfO2)x(GeO2)y(Ge3N4)1-x-y material system, which may form more effective gates than the pure HfO2 and hafnium nitro-silicate studied here. The RPE-MOCVD system constructed by the author for these studies has demonstrated the
capabilities necessary to fabricate such films, and future studies into this material system are scheduled to be carried out soon by other researchers with the aid of this research tool. Lastly, studies soon will be carried out by others in order to investigate Ge/PM-Si/high-κ structures as yet another pathway towards a viable Ge-based CMOS technology.
References


Appendix
Appendix A

The Theory of Point Groups

A.1 – Definition of a Group

The mathematical definition of a group, \( G \), is a non-empty set of elements, together with a binary operation (which will be referred to as multiplication and denoted by writing elements of \( G \) next to one another), that has the following properties [1]:

1. **Associativity** – \((XY)Z = X(YZ)\) for every \( X, Y, Z \) in \( G \).
2. **Identity** – There is an element, \( E \), in \( G \) such that \(XE = EX = X\) for all \( X \) in \( G \).
3. **Inverse** – For every element, \( X \), in \( G \), there exists an element, \( Y \), in \( G \) such that \( XY = YX = E \).

A group is said to be *closed* under its binary operation, meaning that \( XY \) is in \( G \) for each \( X, Y \) in \( G \). The *order* of a group is simply the number of elements comprising it.

This definition is incredibly simple, yet the concept of a group is so versatile that it has had far ranging consequences in almost all branches of the physical sciences from quantum field theory to structural biology. Simply put, group theory is the theory of symmetry, and nature, it seems, is governed by symmetry.

For the purpose of studying local bonding environments in molecules, the relevant groups are referred to as *point groups*, and they are comprised of several basic symmetry operations. Much of the discussion contained in the remainder of appendix A follows the treatments of R. McWeeny [2] and D.M. Bishop [3].
A.2 – Symmetry Operations

There are three fundamental point group symmetry operations, known as rotation, reflection, and inversion, and they are denoted by C, σ, and i, respectively. Figure A.1 depicts these operations pictorially. There is also a fourth symmetry element, referred to as an improper rotation, that consists of a rotation, followed by a reflection across the plane normal to the rotation axis, and it is denoted by S. Lastly, there must be an identity element, which is denoted by E and does nothing at all.

The operations, C, S, and σ are usually given subscripts and primes to distinguish them. Rotations are always counter-clockwise, and C_n refers to a rotation by an angle of $360^\circ/n$ about the symmetry axis. Similarly, S_n refers to an improper rotation by an angle of $360^\circ/n$ about the symmetry axis. Reflections are written as $\sigma_h$ if they are perpendicular to the principle axis (the C_n axis with the largest n), written as $\sigma_v$ if they contain the principle axis, and written as $\sigma_d$ if they contain the principle axis and bisect two C_2 axes. When more than one symmetry element exists with the same label, they are distinguished by adding primes.

A.3 – Point Groups

In section A.2, the symmetry element of translation, which refers to the movement of the origin, was noticeably omitted. This is what distinguishes the concept of point groups from that of space groups. Space groups are concerned with the overall symmetry of a crystal, for example; whereas, point groups care only about the local
symmetry environment of the basis atoms. This local environment is called site symmetry.

There are various ways of denoting the different point groups, but the most commonly used in chemistry is the so-called Schöenflies notation. In this scheme, $C_n$ refers to cyclic groups (i.e. groups with one n-fold principle axis), $C_{nv}$ refers to cyclic groups with $n$ $\sigma_v$ planes, $C_{nh}$ refers to cyclic groups with a $\sigma_h$ plane, $D_n$ refers to dihedral groups (i.e. $C_n$ groups with $n$ $C_2$ axes perpendicular to the $C_n$ axis), $D_{nh}$ refers to dihedral groups with a $\sigma_h$ plane, $D_{nd}$ refers to dihedral groups with $n$ $\sigma_d$ planes, $T$ refers to the proper rotational symmetry of a regular tetrahedron, $T_d$ refers to the full symmetry of a regular tetrahedron, $T_h$ refers to the full symmetry of a di-dodecahedron, $O$ refers to the proper rotational symmetry of a cube, $O_h$ refers to the full symmetry of a cube, $I$ refers to the proper rotational symmetry of a regular icosahedron, $I_h$ refers to the full symmetry of a regular icosahedron, and $K_h$ refers to the full symmetry of a sphere.

As an example of site symmetry and point groups, consider the ammonia molecule, NH$_3$, shown in figure A.2. The symmetry elements are the identity (E), rotation by 120º about the z-axis ($C_3$), rotation by 240º about the z-axis ($C_3$), and reflection across the three equivalent $\sigma_v$ planes ($\sigma_v$, $\sigma_v'$, and $\sigma_v''$). There is one $C_3$ principle axis, and there are 3 $\sigma_v$ planes; therefore, ammonia belongs to the group, $C_{3v}$.

These elements can be placed into a so-called multiplication table to depict how they combine. Such a table is for $C_{3v}$ is shown in table A.1. An operation from the far left column is applied first, followed by an operation from the top row, and the result is placed at the row/column intersection. It does not matter what symbols are placed into
this table; any set of elements that produces a multiplication table with this same structure belongs to $C_{3v}$. In a very real way, a group’s multiplication table defines it.

**A.4 – Representations of Point Groups**

Since any set of elements that satisfy a group’s multiplication table can be used to define the group, there are clearly very many ways to represent any group. Consider again $C_{3v}$. One representation of this group obviously is the set of atoms in the ammonia molecule, but another equally valid representation is the set of numbers, $\{1,1,1,-1,-1,-1\}$, under the operation of arithmetic multiplication. One could go on and on and discover an infinite number of ways to represent $C_{3v}$, using a variety of objects, arranged in a number of ways. Three of these representations, however, are special. In the same way that the Cartesian unit vectors can be combined to describe any point in 3-dimensional space, these representations can be combined to describe any other representation of $C_{3v}$. Such representations are called *irreducible representations*, and they are the group-theoretical analogs of basis vectors in linear algebra or atoms in chemistry. They are the most basic ways in which to represent a point group.

The most mathematically concrete way to describe irreducible representations is to use square matrices, $D^\mu$, each with dimension $n_\mu$. It can be shown that for a group of order $g$ that the following relation holds [2, 3]:

$$\sum_{\mu=1}^{N} n_\mu^2 = g \quad (A.1)$$
where $N$ is the total number of irreducible representations. For example, $C_{3v}$ has order 6; therefore, there are only two ways to satisfy equation A.1 for $C_{3v}$. There can be 6 one-dimensional irreducible representations, or 2 one-dimensional and 1 two-dimensional irreducible representations.

In order to understand irreducible representations more fully, a two definitions are first in order. The first definition is that of a character. When an element of a group is represented by a matrix, its character is defined as the trace of the matrix (i.e. the sum of its diagonal entries). The second definition needed is that of a class. A class, $C$, of a group, $G$, is a set of all distinct elements $Y = RXR^{-1}$ produced for a single element, $X$, where $R$ runs over the entire group. When this is done for each member, $X$, in $G$, the group is partitioned into its distinct classes. Using table A.1, it is easy to see that the partitioning for $C_{3v}$ takes place as follows: $\{E\}$, $\{C_3, C_3^2\}$, and $\{\sigma_v, \sigma_v', \sigma_v''\}$.

With these definitions, several important results about the irreducible representations of a group can be stated. First, it can be shown that the characters of each of the elements belonging to a class, $C$, are equal. Moreover, it can be shown that the number of irreducible representations of a group, $G$, is simply the number of distinct classes belonging to $G$. This means that $C_{3v}$ must have 3 irreducible representations, and by equation A.1, two of these are one-dimensional, and one of them is two-dimensional. Finally, as its name implies, perhaps the most important result about the irreducible representations of a group, $G$, is the Great Orthogonality Theorem:
\[ \sum_R D^g(R)D^v(R^{-1}) = \frac{G}{n} \delta_{\mu\nu} \delta_{ik} \delta_{jm} \]  
(A.2)

where \( R \) runs over all \( g \) members of \( G [2, 3] \).

The irreducible representations for \( C_{3v} \) are shown in table A.2. Note that all of the rows and columns in table A.2 are orthogonal, which is no accident; equation A.2 says that the must be. The convention for labeling the irreducible representations is to use \( A \) and \( B \) to indicate one-dimensional representations, \( E \) to indicate two dimensional representations, and \( T \) to indicate three-dimensional representations. In some cases, the subscripts, \( g \) and \( u \), are used to distinguish between even and odd symmetry in the representations, and primes are used to distinguish between multiple representations having the same symmetry.

**A.5 – The Reduction of Reducible Representations**

The most basic task in the theory of point groups is to factor a particular representation – that is, to express it in terms of its irreducible representations. The Great Orthogonality Theorem provides a method to do just that.

Consider the characters, \( \chi_\mu(R) = Tr(D^\mu(R)) \), of the representation defined by the matrices, \( D^\mu(R) \), in the context of equation A.2. Setting \( j = i \) and \( k = m \) in equation A.2, the result is:
\[
\sum_{i} D^\mu_{ii}(R) D^\nu_{kk}(R^{-1}) = \frac{g}{n_\mu} \delta^\mu_\nu \delta_{ik} \delta_{ik} = \frac{g}{n_\mu} \delta^\mu_\nu \delta_{ik}
\]

Summing over \( i \) and \( j \), the expression becomes:

\[
\sum_{R} \left[ \sum_{i=1}^{n_\mu} D^\mu_{ii}(R) \sum_{k=1}^{n_\nu} D^\nu_{kk}(R^{-1}) \right] = \sum_{R} \chi^\mu_{\mu}(R) \chi^\nu_{\nu}(R^{-1}) = \frac{g}{n_\mu} \delta^\mu_\nu \sum_{i=1}^{n_\mu} \delta_{ik} = \frac{g}{n_\mu} \delta^\mu_\nu \sum_{i=1}^{n_\mu} \delta_{ii} = \frac{g}{n_\mu} \delta^\mu_\nu n_{\mu}
\]

\[\Rightarrow \sum_{R} \chi^\mu_{\mu}(R) \chi^\nu_{\nu}(R^{-1}) = g \delta^\mu_\nu \]

Each \( D^\mu \) is related to a unitary matrix by a similarity transformation [3], and the trace (i.e. character) of a matrix is unchanged by such transformations. The consequence of this is that, without loss of generality:

\[
\sum_{R} \chi^\mu_{\mu}(R) \chi^\nu_{\mu}(R) = g \delta^\mu_\nu
\]

(A.3)

The characters of members of the same class are equal, so equation A.3 becomes:

\[
\sum_{i=1}^{k} h_i \chi^\mu_{\mu}(C_i) \chi^\nu_{\mu}(C_i) = g \delta^\mu_\nu
\]

(A.4)

where \( k \) is the number of classes in \( G \), and there are \( h_i \) members in class \( C_i \).
Now consider a reducible representation of a group, $G$, where the members, $R$, of $G$ are described by the matrices, $D(R)$. This representation must be able to be expressed in terms of the irreducible representations of $G$. This is done using the direct sum operator as follows:

$$D(R) = a_1D_1(R) \oplus a_2D_2(R) \oplus \cdots \oplus a_kD_k(R) = \begin{pmatrix}
a_1[D_1(R)] & [0] & \cdots & [0] \\
[0] & a_2[D_2(R)] & \cdots & [0] \\
\vdots & \vdots & \ddots & \vdots \\
[0] & [0] & \cdots & a_k[D_k(R)]
\end{pmatrix}$$

where $k$ is the number of classes in $G$, and the $a_i$'s are numerical coefficients representing how many times irreducible representation $i$ occurs in the reducible representation. All reducible representations of $G$ are able to be put into block diagonal form by a unitary transformation, so the above holds without loss of generality. Taking the trace of the above equation, the result is given simply by:

$$\chi(R) = \sum_{v=1}^{k} a_v \chi_v(R)$$

Multiplying the above equation by $\chi_\mu^*$, summing over the symmetry elements, $R$, and applying equation A.3, the result is:

$$\sum_R \chi(R)\chi_\mu^*(R) = \sum_{v=1}^{k} a_v \sum_R \chi_v(R)\chi_\mu^*(R) = \sum_{v=1}^{k} a_v \sum_R \chi_v(R)\chi_\mu^*(R) = g \sum_{v=1}^{k} a_v \delta_{v\mu} = ga_\mu$$

$$\Rightarrow a_\mu = \frac{1}{g} \sum_R \chi(R)\chi_\mu^*(R)$$

With the use of equation A.4, this becomes:

$$a_\mu = \frac{1}{g} \sum_{i=1}^{k} h_i \chi(C_i)\chi_\mu^*(C_i) \quad (A.5)$$
Of course, equation A.5 provides more than just a simple way to factor a reducible representation. It also provides a criterion for reducibility. To see this, simply assume that the reducible representation is irreducible by setting $a_\mu = 1$ and $\chi = \chi_\mu$. This gives the criterion for irreducibility, which states that a representation is irreducible if and only if the following holds:

$$g = \sum_{i=1}^{k} h_i \chi_\mu (C_i) \chi_\mu^* (C_i)$$  \hspace{1cm} (A.6)

As an example of the factorization of a reducible representation, consider again the ammonia molecule. The character table for the group, $C_{3v}$, is shown as table A.3. All that is needed to factor the representation for the ammonia molecule are the characters of the molecule itself. Do determine these, only three matrices are required – one for any member of each class. This is easiest to do in the permutation representation. In this description, only the rearrangement of the atoms is considered. For example, under $\sigma_v$ atoms at the sites, $H_1$ and $H_2$, are exchanged. In matrix form, the three representatives of the classes are described by the following permutation matrices:

$$\begin{align*}
E : \begin{pmatrix} N' \\ H_1' \\ H_2' \\ H_3' \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} N \\ H_1 \\ H_2 \\ H_3 \end{pmatrix} & C_3 : \begin{pmatrix} N' \\ H_1' \\ H_2' \\ H_3' \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} N \\ H_1 \\ H_2 \\ H_3 \end{pmatrix} & \sigma_v : \begin{pmatrix} N' \\ H_1' \\ H_2' \\ H_3' \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} N \\ H_1 \\ H_2 \\ H_3 \end{pmatrix}
\end{align*}$$

and the traces immediately give the characters for $E$, $C_3$, and $\sigma_v$ as 4, 1, and 2, respectively. Note that the characters of these symmetry elements are equal to the number of atoms remaining unaltered by them. With that in mind, the matrices are not even required. A picture and some simple mental computation is all that is required.
Using equation A.5 for each of the three irreducible representations listed in table A.3, the reduction of the ammonia molecule proceeds as:

\[ A_1 : \frac{1}{g} \sum_{i=1}^{k} h_i \chi(C_i) \chi_{\mu}^*(C_i) = \frac{1}{6} (1 \cdot 4 \cdot 1^* + 2 \cdot 1 \cdot 1^* + 3 \cdot 2 \cdot 1^*) = \frac{12}{6} = 2 \]

\[ A_2 : \frac{1}{g} \sum_{i=1}^{k} h_i \chi(C_i) \chi_{\mu}^*(C_i) = \frac{1}{6} (1 \cdot 4 \cdot 1^* + 2 \cdot 1 \cdot 1^* + 3 \cdot 2 \cdot 1^*) = \frac{0}{6} = 0 \]

\[ E : \frac{1}{g} \sum_{i=1}^{k} h_i \chi(C_i) \chi_{\mu}^*(C_i) = \frac{1}{6} (1 \cdot 4 \cdot 2^* + 2 \cdot 1 \cdot 1^* + 3 \cdot 2 \cdot 0^*) = \frac{6}{6} = 1 \]

\[ \Rightarrow \Gamma_{\text{NH}_3} = 2 \Gamma_{A_1} \oplus \Gamma_E \]

where \( \Gamma_{\mu} \) has been used to signify representation \( \mu \).

**A.6 – The Applications of Point Group Theory**

Ultimately, it is a relatively simple task to express the symmetry of a physical system in terms of its fundamental, irreducible symmetries. One needs only to determine the symmetry group and the characters of the particular system in question, and the reduction is carried out through the application of equation A.5. With the factorization in hand, one can then determine a variety of physical and spectroscopic properties, such as normal modes and their optical activity, molecular orbitals and their degeneracies, and allowed and forbidden electronic transitions between states, to name a few. While the uses of these basic ideas are far too extensive to be discussed at length here, several
applications relevant to transition metal oxides are presented in appendix B and appendix J.
Figure A.1: A pictorial depiction of (a) a counter-clockwise rotation by 90° around the z-axis, (b) a reflection across the plane normal to the <1, 1, 0> direction, and (c) inversion.
Figure A.2: The symmetry of the ammonia molecule.
Table A.1: The multiplication table for the group $C_{3v}$. The operation in the left column is applied first, followed by the operation in the top row.

<table>
<thead>
<tr>
<th>$C_{3v}$</th>
<th>E</th>
<th>$C_3$</th>
<th>$C_3^2$</th>
<th>$\sigma_v$</th>
<th>$\sigma_v'$</th>
<th>$\sigma_v''$</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>E</td>
<td>$C_3$</td>
<td>$C_3^2$</td>
<td>$\sigma_v$</td>
<td>$\sigma_v'$</td>
<td>$\sigma_v''$</td>
</tr>
<tr>
<td>$C_3$</td>
<td>$C_3$</td>
<td>$C_3^2$</td>
<td>E</td>
<td>$\sigma_v''$</td>
<td>$\sigma_v$</td>
<td>$\sigma_v'$</td>
</tr>
<tr>
<td>$C_3^2$</td>
<td>$C_3^2$</td>
<td>E</td>
<td>$C_3$</td>
<td>$\sigma_v'$</td>
<td>$\sigma_v''$</td>
<td>$\sigma_v$</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>$\sigma_v$</td>
<td>$\sigma_v'$</td>
<td>$\sigma_v''$</td>
<td>E</td>
<td>$C_3$</td>
<td>$C_3^2$</td>
</tr>
<tr>
<td>$\sigma_v'$</td>
<td>$\sigma_v'$</td>
<td>$\sigma_v''$</td>
<td>$\sigma_v$</td>
<td>$C_3^2$</td>
<td>E</td>
<td>$C_3$</td>
</tr>
<tr>
<td>$\sigma_v''$</td>
<td>$\sigma_v''$</td>
<td>$\sigma_v$</td>
<td>$\sigma_v'$</td>
<td>$C_3$</td>
<td>$C_3^2$</td>
<td>E</td>
</tr>
</tbody>
</table>
Table A.2: The three irreducible representations for $C_{3v}$: $A_1$, $A_2$, and $E$. $A_1$ and $A_2$ are both one-dimensional, and $E$ is two-dimensional (i.e. a $2 \times 2$ matrix). Note that all rows and columns of this table are orthogonal as they must be, due to the Great Orthogonality Theorem. (This data can be found in most texts on point groups, e.g. R. McWeeny [2].)

<table>
<thead>
<tr>
<th>$C_{3v}$</th>
<th>E</th>
<th>$C_3$</th>
<th>$C_3^2$</th>
<th>$\sigma_v$</th>
<th>$\sigma_v'$</th>
<th>$\sigma_v''$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$E_{11}$</td>
<td>1</td>
<td>-½</td>
<td>-½</td>
<td>-½</td>
<td>1</td>
<td>-½</td>
</tr>
<tr>
<td>$E_{12}$</td>
<td>0</td>
<td>-½√3</td>
<td>√3</td>
<td>-½√3</td>
<td>0</td>
<td>√3</td>
</tr>
<tr>
<td>$E_{21}$</td>
<td>0</td>
<td>√3</td>
<td>-½√3</td>
<td>-½√3</td>
<td>0</td>
<td>√3</td>
</tr>
<tr>
<td>$E_{22}$</td>
<td>1</td>
<td>-½</td>
<td>-½</td>
<td>½</td>
<td>-1</td>
<td>½</td>
</tr>
</tbody>
</table>
Table A.3: The character table for the group, $C_{3v}$. Note that table A.3 is simply a compact way of writing the traces of the three representations in table A.2. (This data can be found in most texts on point groups, e.g. R. McWeeny [2] or D.M. Bishop [3].)

<table>
<thead>
<tr>
<th>$C_{3v}$</th>
<th>E</th>
<th>$2C_3$</th>
<th>$3\sigma_v$</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$z$; $x^2+y^2$; $z^2$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>---------------------------</td>
</tr>
<tr>
<td>$E$</td>
<td>2</td>
<td>-1</td>
<td>0</td>
<td>$(x,y)$; $(x^2-y^2,xy)$; $(xz,yz)$</td>
</tr>
</tbody>
</table>
References


Appendix B

Molecular Orbital Theory

B.1 – Molecular Orbitals

Molecular orbital (MO) theory is an independent-electron theory which includes the electron-electron repulsions within a molecule in an average way [1], rather than by using the brute force required by a first-principles approach. The easiest, most obvious, and most common way of doing this is by making linear combinations of atomic orbitals (LCAOs). In the LCAO method, the MOs are just linear combinations of the s, p, d, and f wave functions belonging to each of the bond-forming atoms in a molecule or cluster. Each electron is thought of as being owned by a particular atom and is then shared with other atoms when the molecule is formed. The degree to which this sharing occurs is described by bond ionicity [2]. The resulting MOs can be thought of as hybrid orbitals, retaining characteristics of their constituents while taking on new ones by way of mixing together. In general, some MOs point in the bond direction (referred to as $\sigma$ orbitals); others are perpendicular to the bond direction (called $\pi$ orbitals). Additionally, some orbitals describe average electron-electron interactions favorable to forming bonds (bonding orbitals), and some describe average electron-electron interactions that discourage the formation of bonds (non-bonding orbitals, indicated usually with an asterisk).
While inter-atom electron-electron interactions for the molecule are treated in an average way by the LCAO theory, these interactions within each atom are treated as a complicated, *ab-initio* many-body problem. The AOs are found, usually numerically, by various techniques, such as Slater determinants, the Hartree-Fock approximation, configuration-interaction, many-body perturbation theory, and many-body Green’s functions. Most of these techniques are sophisticated enough to treat an entire molecule or molecular cluster as a many-body system, though almost always, various approximations are required even in determining the AOs of an isolated atom [3, 4]. This process is not trivial, and it generally requires complicated computation methods to do properly, even for individual AOs. To quantify the nature of the bonding within a molecule, this step is a necessity; however, to understand the qualitative nature of the bonds, it is not.

**B.2 – Atomic Orbitals and Their Transformation Properties**

Since MO theory threats atoms individually, the total electronic wave function can be written as a product of one-electron wave functions [5]. Moreover, in this approximation, each of the atomic orbitals (AO) can be written as the product of radial and angular functions. This is because the electronic potentials are spherically symmetric; the radial functions will vary depending on the particular atom, but the angular components of the wave functions will not. In fact, all of these angular wave functions satisfy the same angular equation that arises in the solution of the hydrogen atom. The result is that the angular dependences of the AOs are described simply by the
spherical harmonics. This means that the familiar notion of s, p, d, and f orbitals in the hydrogen atom are still valid for all of the AOs used in MO theory under the LCAO approximation.

Appendix A discusses at length how matrices can be used to represent symmetry groups, but functions also possess symmetry, and they can be used to represent point groups as well. This situation is analogous to the one in quantum mechanics in which Schrödinger’s wave functions and Heisenberg’s matrices are both equivalent descriptions of the same phenomena. To understand how functions can represent point group symmetry, consider the form of the s, p, and d orbitals. Table B.1 shows the dependence of these orbitals in terms of both spherical and Cartesian coordinate systems. When a symmetry operator belonging to a molecule’s point group is applied to one of these orbitals, the atomic coordinates are changed, and the AOs, tied to these coordinates, transform as well [6].

Consider how the operator, $C_3$ (counter-clockwise rotation by 120° about the z-axis), from the group, $C_{3v}$, affects the functions, $d_{xz}$ and $d_{yz}$. The local coordinates of the atoms change as follows:

$$
\begin{bmatrix}
    x' \\
    y' \\
    z'
\end{bmatrix}
= C_3
\begin{bmatrix}
    x \\
    y \\
    z
\end{bmatrix}
= \begin{bmatrix}
    -1/2 & \sqrt{3}/2 & 0 \\
    -\sqrt{3}/2 & -1/2 & 0 \\
    0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    z
\end{bmatrix}
\Rightarrow
\begin{align*}
x' &= -\frac{1}{2}x + \frac{\sqrt{3}}{2}y \\
y' &= -\frac{\sqrt{3}}{2}x - \frac{1}{2}y \\
z' &= z
\end{align*}
$$

From this, it follows that under $C_3$: 

181
Similarly, under the other five symmetry operators of $C_{3v}$, these two AOs transform into linear combinations of one another. Thus, $d_{zx}$ and $d_{yz}$ form a two-dimensional representation of $C_{3v}$. In fact, they serve as a basis for the irreducible representation, $E$. In general, the AOs form bases for many of the irreducible representations in the various point groups. This fact leads to a surprisingly simple, yet elegant, way to qualitatively describe molecular bonding.

**B.3 – Bond Hybridization**

Once a representation of the molecular bonds in a system has been factored as described in appendix A, the irreducible representations in that factorization actually represent the symmetry of the various MOs in that molecular system. The basis functions of these irreducible representations are then given by the various AOs. In other words, once the molecular bond representation has been factored, one can determine which MOs are present, what their symmetries are, and which AOs hybridize to form them simply by looking at the appropriate character table.

As an explicit example, consider the $\sigma$ bonds in ammonia. Since these bonds point from the central nitrogen atom to the hydrogen atoms, the $\sigma$ molecular bond representation is identical to the symmetry of the ammonia molecule itself. It is a fact
that in this picture, the factorization of the representation takes place as \( \Gamma_a = 2\Gamma_{A_1} + \Gamma_E \) (see appendix A). Referring to the far right column in table A.3, it is apparent that \( A_1 \) can be represented by \( s, p_z, \) and \( d_{x^2-y^2}, \) and \( E \) can be represented by \((p_x, p_y), (d_{x^2-y^2}, d_{xy}), \) and \((d_{2z^2}, d_{yz})\). Choosing to ignore the \( d \) orbitals in this example (since they are unpopulated in the ground state), \( A_1 \) is represented by a linear combination of \( s \) and \( p_z, \) and \( E \) is represented by \((p_x, p_y)\). Already, a qualitative picture has emerged regarding the symmetry of the \( \sigma \)-type MOs (two \( A_1 \)’s and an \( E \)), and which AOs interact to form them (\( s \) and \( p_z \) in the \( A_1 \) states, and \( p_x \) and \( p_y \) in the \( E \) state).

A schematic representation of a more complicated example of MO bond hybridization is given in figure B.1, which is the so-called molecular orbital diagram for a stereotypical rutile transition metal oxide. The AOs of the metal atom are shown on the left, the AOs of the oxygen atoms are shown on the right, and the MOs are shown in the center. The various lines between the AOs and the MOs show which AOs hybridize to form the MOs. While the ordering of the MOs in this diagram cannot be ascertained by symmetry alone, everything else in figure B.1 can.

**B.4 – Allowed and Forbidden Transitions**

Before considering the details of electronic transitions, the definition of a new operation is in order. The *direct product*, \( \Gamma_\mu \otimes \Gamma_\nu \), of two representations, \( \Gamma_\mu \) and \( \Gamma_\nu \), forms a new representation whose basis is created by taking all possible products between every basis function of \( \Gamma_\mu \) and every basis function of \( \Gamma_\nu \). As an example,
consider the spaces, $S_1$ and $S_2$, spanned by the spins of two electrons, and let $|\uparrow\rangle$ and $|\downarrow\rangle$ represent spin up and spin down states, respectively. As basic quantum mechanics indicates, these two functions form a basis for each independent electron [7]. The direct product of these two spaces is then spanned by linear combinations of $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, $|\downarrow\uparrow\rangle$, and $|\downarrow\downarrow\rangle$, where the first arrow indicates the state of the $S_1$ electron, and the second arrow represents the state of the $S_2$ electron. If the basis is defined in terms of matrices, the direct product of two matrices is given by the following, shown here in block-diagonal form:

$$A \otimes B = \begin{pmatrix}
A_{11}B & A_{12}B & \cdots & A_{1m}B \\
\vdots & \vdots & \ddots & \vdots \\
A_{m1}B & A_{m2}B & \cdots & A_{mm}B
\end{pmatrix}$$

In general, the direct product of two representations, $\Gamma_\mu$ and $\Gamma_\nu$, will not be irreducible, even if $\Gamma_\mu$ and $\Gamma_\nu$ are themselves both irreducible. Hence, the general rule is that $\Gamma_\mu \otimes \Gamma_\nu = \sum_\lambda a_\lambda \Gamma_\lambda$. From the matrix definition of the direct product above, it is clear that the characters of the new representation are given simply by the product of the characters of $\Gamma_\mu$ and $\Gamma_\nu$. With this definition, it is now possible to discuss electronic transitions.
In the first-order approximation, the dominant transitions between electronic states in an atom when it absorbs electromagnetic radiation are the dipole-induced transitions [7]. When “forbidden” transitions are mentioned, the reference is to only dipole-induced transitions. Magnetic dipole, electric quadrupole, and higher order multipole terms, as well as multiple photon processes, all play a role in electronic transitions, but they are usually not dominant.

Consider a plane wave, polarized in the z direction. In this case, the transition probability for a dipole-induced electronic transition is given by \( \langle \psi_f^\mu | z^+ | \psi_i^\nu \rangle \), where \( \psi_f^\mu \) and \( \psi_i^\nu \) are the final and initial states, respectively, belonging to representations \( \Gamma_\mu \) and \( \Gamma_\nu \), respectively [6, 7]. The operator, \( z \), belongs to the representation, \( \Gamma_\lambda \). It can be shown that if \( \Gamma_\lambda \otimes \Gamma_\nu \) does not contain \( \Gamma_\mu \) at least once in its factorization, then \( \langle \psi_f^\mu | z^+ | \psi_i^\nu \rangle \) vanishes identically [5, 6]. This statement is true not just for the operator, \( z \), but for any operator, in general, belonging to some \( \Gamma_\lambda \).

As the simplest example, again consider the \( \sigma \)-MO states in an ammonia molecule, which is now excited by a plane wave that is polarized along the z-axis. There are three possible types of transitions: \( A_1 \rightarrow A_1 \), \( A_1 \rightarrow E \), and \( E \rightarrow A_1 \). In \( C_3v \), the \( z \) operator belongs to \( A_1 \), as shown in table A.3. It is necessary, then, only to factor two direct products: \( \Gamma_{A_1} \otimes \Gamma_{A_1} \) and \( \Gamma_{A_1} \otimes \Gamma_E \). Since \( \Gamma_{A_1} \) has characters \{1, 1, 1\}, it is essentially like an identity, leaving all other representations unchanged in direct products. Hence, \( \Gamma_{A_1} \otimes \Gamma_{A_1} = \Gamma_{A_1} \), and \( \Gamma_{A_1} \otimes \Gamma_E = \Gamma_E \). This means that only \( A_1 \rightarrow A_1 \) or \( E \rightarrow E \) type transitions are possible, but since there is only one E state, the \( E \rightarrow E \) possibility is
ruled out. Thus, $A_1 \leftrightarrow A_1$ transitions are allowed, but $A_1 \leftrightarrow E$ transitions are forbidden in the $\sigma$-MO states of ammonia under the dipole approximation.

The techniques presented here apply not only to electronic transitions, but to any quantum mechanical transitions, governed by any operator. The spectroscopic applications are profound; by using symmetry arguments alone, one can deduce the relationship between various spectra and molecular symmetry. Fourier transform infrared (FTIR) spectroscopy, Raman spectroscopy, and x-ray absorption spectroscopy (XAS) are just a few examples of techniques where local bonding symmetry plays a vital role.

**B.5 – Symmetry Adapted Linear Combinations of Atomic Orbitals**

The final aspect of MO theory that garners mention here is the method by which one can derive the specific linear combinations of the AOs in a molecule to produce its MOs. The Great Orthogonality Theorem (see appendix A) provides an elegant technique for doing just that. Using the orthogonality relationships, a projection operator can be defined by the following [6]:

$$\rho_{\mu}^{\nu} = \frac{g_{\mu}}{g} \sum_{R} \bar{D}_{\mu}^{\nu}(R)R$$  \hspace{1cm} (B.1)

where $g_{\mu}$ is the dimension of representation $\Gamma_{\mu}$, $g$ is the order of the group, the $R$’s are the symmetry operations of the group, $\bar{D}_{\mu}^{\nu}(R)$ is the contragradient matrix representation of $R$ in $\Gamma_{\mu}$ (i.e. the transposed inverses of the matrix representation of $R$ in $\Gamma_{\mu}$), and $\rho_{\mu}^{\nu}$
produces the component of a vector or function along the $i^{th}$ basis vector of $\Gamma_\mu$ when applied. In other words, if one wishes to find the projection of a vector or function, $f$, in $\Gamma_\mu$, it can be found by the following relationship:

$$\rho_\mu^i f = \left[ \frac{g_\mu}{g} \sum_R \bar{D}_\mu^i (R) R \right] f = f_\mu^i \quad \text{(B.2)}$$

As an example, consider yet again the $\sigma$ bonds of an ammonia molecule. Denote the bonds between the nitrogen atom and the three hydrogen atoms, $H_1$, $H_2$, and $H_3$ (as labeled in figure A.2), as $\psi_1$, $\psi_2$, and $\psi_3$, respectively. Upon applying the six symmetry functions discussed in appendix A to each of these wave functions, the result is:

$$E \psi_1 = \psi_1 \quad C_3 \psi_1 = \psi_2 \quad C_3^2 \psi_1 = \psi_3 \quad \sigma_v \psi_1 = \psi_2 \quad \sigma_h \psi_1 = \psi_3 \quad \sigma_h'' \psi_1 = \psi_1$$
$$E \psi_2 = \psi_2 \quad C_3 \psi_2 = \psi_3 \quad C_3^2 \psi_2 = \psi_1 \quad \sigma_v \psi_2 = \psi_1 \quad \sigma_h \psi_2 = \psi_2 \quad \sigma_h'' \psi_2 = \psi_3$$
$$E \psi_3 = \psi_3 \quad C_3 \psi_3 = \psi_1 \quad C_3^2 \psi_3 = \psi_2 \quad \sigma_v \psi_3 = \psi_3 \quad \sigma_h \psi_3 = \psi_1 \quad \sigma_h'' \psi_3 = \psi_2$$

and the $\bar{D}_\mu^i (R)$’s are given by (see table A.2):

$$\bar{D}_h^i (E) = 1 \quad \bar{D}_h^i (C_3) = 1 \quad \bar{D}_h^i (C_3^2) = 1$$
$$\bar{D}_s^i (E) = 1 \quad \bar{D}_s^i (C_3) = 1 \quad \bar{D}_s^i (C_3^2) = 1$$
$$\bar{D}^e (E) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \bar{D}^e (C_3) = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix} \quad \bar{D}^e (C_3^2) = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$$
\[
\begin{align*}
D^A_1(\sigma_v) &= 1 & D^A_1(\sigma_v') &= 1 & D^A_1(\sigma_v'') &= 1 \\
D^A_2(\sigma_v) &= -1 & D^A_2(\sigma_v') &= -1 & D^A_2(\sigma_v'') &= -1 \\
D^E_1(\sigma_v) &= \left( -\frac{1}{2}, -\sqrt{3}/2 \right) & D^E_1(\sigma_v') &= \left( 0, 0 \right) & D^E_1(\sigma_v'') &= \left( -\sqrt{3}/2, \sqrt{3}/2 \right)
\end{align*}
\]

In section B.3, a basis was defined for this system after factoring the representation of the \( \sigma \) bonds in the ammonia molecule. Specifically, the \( A_l \) basis was chosen to be \( \phi = \lambda s + \eta p_z \) (in which the coefficients have no bearing on the symmetry), and the \( E \) basis was chosen to be \( (p_x, p_z) \). Upon applying equation B.2 to \( \psi_1 \) and normalizing the projections, the results are:

\[
\begin{align*}
\rho^{A_1}_{11}\psi_1 &= \left[ \frac{g_{\mu}}{g} \sum_R D_{11}^A(R)R \right] \psi_1 = \frac{1}{6} \left[ \psi_1 + \psi_2 + \psi_3 + \psi_2 + \psi_3 + \psi_1 \right] \Rightarrow \frac{1}{\sqrt{3}} \left[ \psi_1 + \psi_2 + \psi_3 \right] \\
\rho^{A_2}_{11}\psi_1 &= \left[ \frac{g_{\mu}}{g} \sum_R D_{11}^A(R)R \right] \psi_1 = \frac{1}{6} \left[ \psi_1 + \psi_2 + \psi_3 - \psi_2 - \psi_3 - \psi_1 \right] \Rightarrow 0 \\
\rho^{E}_{11}\psi_1 &= \left[ \frac{g_{\mu}}{g} \sum_R D_{11}^E(R)R \right] \psi_1 = \frac{1}{3} \left[ \psi_1 - \frac{1}{2} \psi_2 - \frac{1}{2} \psi_3 - \frac{1}{2} \psi_2 + \psi_3 - \frac{1}{2} \psi_1 \right] \Rightarrow \frac{1}{\sqrt{6}} \left[ 2\psi_2 - \psi_1 - \psi_3 \right] \\
\rho^{E}_{22}\psi_1 &= \left[ \frac{g_{\mu}}{g} \sum_R D_{22}^E(R)R \right] \psi_1 = \frac{1}{3} \left[ \psi_1 - \frac{1}{2} \psi_2 - \frac{1}{2} \psi_3 + \frac{1}{2} \psi_2 - \psi_3 + \frac{1}{2} \psi_1 \right] \Rightarrow \frac{1}{\sqrt{2}} \left[ \psi_1 - \psi_3 \right]
\end{align*}
\]
Since three linearly independent functions have already been projected from $\psi_1$, there is no need to project the other two wave functions; they would simply reproduce the same results, yielding no new information.

Recall that $\phi$ provides a basis for $A_1$ and that $(p_x, p_y)$ provides a basis for $E$. Combining the above result with this fact implies that:

$$\phi = \frac{1}{\sqrt{3}} [\psi_1 + \psi_2 + \psi_3] \quad p_x = \frac{1}{\sqrt{6}} [2\psi_2 - \psi_1 - \psi_3] \quad p_y = \frac{1}{\sqrt{2}} [\psi_1 - \psi_3]$$

or in matrix form (note the unitary relationship):

$$\begin{pmatrix} \phi \\ p_x \\ p_y \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} \Rightarrow \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 2/\sqrt{6} & 0 \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \phi \\ p_x \\ p_y \end{pmatrix}$$

which finally yields the so-called symmetry adapted linear combinations of the atomic orbitals (SALCs) for the ammonia molecule:

$$\psi_1 = \frac{1}{\sqrt{3}} \phi - \frac{1}{\sqrt{6}} p_x + \frac{1}{\sqrt{2}} p_y; \quad \psi_2 = \frac{1}{\sqrt{3}} \phi + \frac{2}{\sqrt{6}} p_x; \quad \psi_3 = \frac{1}{\sqrt{3}} \phi - \frac{1}{\sqrt{6}} p_x - \frac{1}{\sqrt{2}} p_y$$

Though simple, this procedure is rather time-consuming relative to the other group-theoretical techniques described in the rest of appendix A and appendix B;
however, it demonstrates very clearly the relationship between the MOs and the AOs. From simple symmetry arguments alone, it has been demonstrated that one can obtain which AOs hybridize, in what ratios they do so, and which transitions are allowed between these hybrid MOs. The only important information that cannot be obtained from basic symmetry arguments is the exact ordering of the MOs. With the aid of a good computational chemist, however, MO diagrams, such as the one depicted in figure B.1, can be produced.
**Table B.1:** The symmetry of the atomic orbitals commonly used in molecular orbital theory.

<table>
<thead>
<tr>
<th>Atomic Orbital</th>
<th>Spherical</th>
<th>Cartesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>no $\theta,\varphi$ variation</td>
<td>$x^2+y^2+z^2$</td>
</tr>
<tr>
<td>$p_x$</td>
<td>$\sin(\theta)\cos(\varphi)$</td>
<td>$x$</td>
</tr>
<tr>
<td>$p_y$</td>
<td>$\sin(\theta)\sin(\varphi)$</td>
<td>$y$</td>
</tr>
<tr>
<td>$p_z$</td>
<td>$\cos(\theta)$</td>
<td>$z$</td>
</tr>
<tr>
<td>$d_{z^2}$</td>
<td>$3\cos^2(\theta) - 1$</td>
<td>$2z^2-x^2-y^2$</td>
</tr>
<tr>
<td>$d_{x^2-y^2}$</td>
<td>$\sin^2(\theta)\cos(2\varphi)$</td>
<td>$x^2-y^2$</td>
</tr>
<tr>
<td>$d_{xy}$</td>
<td>$\sin^2(\theta)\sin(2\varphi)$</td>
<td>$xy$</td>
</tr>
<tr>
<td>$d_{yz}$</td>
<td>$\sin(\theta)\cos(\theta)\sin(\varphi)$</td>
<td>$yz$</td>
</tr>
<tr>
<td>$d_{zx}$</td>
<td>$\sin(\theta)\cos(\theta)\cos(\varphi)$</td>
<td>$zx$</td>
</tr>
</tbody>
</table>
Figure B.1: A representative molecular orbital diagram for a rutile transition metal oxide (modeled after P.A. Cox [1]).
References


Appendix C

Crystal Field Splitting and Jahn-Teller Distortions

C.1 – Crystal Field Splitting

In an isolated environment with spherical symmetry, the orbitals of an atom have the familiar $2l+1$ degeneracy. Specifically, the s states are non-degenerate, the p states are triply degenerate, the d states have 5-fold degeneracy, the f states are 7-fold degenerate, etc. In the context of transition metal oxide complexes, the d and f states are by far the most interesting, and it is where many of their unique electrical and magnetic properties originate [1, 2].

Crystal field theory attempts to explain what happens to these degenerate atomic states when they are bonded to neighboring atoms, commonly referred to as ligands. Crystal field theory can be thought of as a rather crude approximation to molecular orbital (MO) theory (see appendix B), which considers only the electrostatic interaction between the transition metal ions and their ligands. It was pioneered by Hans Bethe in 1929 and provided a relatively simple way to study some very complicated systems. The results of crystal field theory calculations are far less reliable than those now provided by MO theory, however, since it treats the metal-ligand interaction from a purely ionic, electrostatic point of view [1].

Despite this, the effect of crystal field symmetry on the removal of degeneracy in the d (and f) states of transition metal ions is still very relevant. The easiest way to
understand this effect is in the context of point group theory, and since this type of
degeneracy removal is a direct result of the metal-ligand bond interaction, it still is
commonly referred to as crystal field splitting. While there certainly are other
mechanisms by which the degeneracy of the d states is lifted (see section C.2, for
example), the crystal field splitting is usually the most dominant [2]. Using point group
theory and molecular orbital theory (see appendices A and B), it is remarkably simple to
explain crystal field splitting as a necessary consequence of moving the transition metal
ion from a completely symmetric spherical environment into one with a lower symmetry.

As described in appendix B, the d state functions provide bases for various
irreducible representations of the common chemical point groups. Consider the cubic
group, O_h, for example. Table C.1 shows the character table for O_h, along with basis
functions for several of its irreducible representations. Referring to the Cartesian
symmetry of the d states as shown in table B.1, it is clear that the set of orbitals, (d_{xz},
d_{yz}, d_{x^2-y^2}), forms a basis for E_g, and the set, (xy, zx, yz), forms a basis for T_{2g} in the point
group, O_h. Hence, in going from the spherical environment of the free ion to an O_h
bonding environment, the crystal field will split the 5-fold degeneracy of the free ion into
two states – a doubly degenerate state with E_g symmetry, and a triply degenerate state
with T_{2g} symmetry. Furthermore, to a good approximation, the states are split such that
the center of the splitting is equal to the energy of the original 5-fold degenerate states [1,
2]. This means that in O_h environments, 2/5 of the total crystal field splitting energy is
assigned to e_g, and 3/5 is assigned to t_{2g}. (It is customary to write the states in lower case
and the irreducible representations in capital letters).
Much as is the case in molecular orbital theory, the relative ordering and magnitude of the crystal field splitting cannot be ascertained, in general, without detailed quantum mechanical calculations. In some cases, however, the relative ordering of the split states can be determined from symmetry arguments. As an interesting example, consider two transition metals, both belonging to the group, $O_h$. One is has a six-fold coordinated, octahedral bonding environment, and the other has an 8-fold coordinated, cubic one.

As an approximation, consider the metal ion’s interaction with its ligands electrostatically as the repulsion between negative point charges or point dipoles [1]. In the octahedral bonding environment, the $d_{x^2-y^2}$ and $d_{z^2}$ states point in the directions of the ligands; whereas, the $xy$, $zx$, and $yz$ states point to the space between the ligands. As a result, the electrostatic repulsion of the $d_{x^2-y^2}$ and $d_{z^2}$ states is greater than in the $xy$, $zx$, and $yz$ states, and the $t_{2g}$ states are more stable. In the 8-fold, cubic bonding environment, however, the $xy$, $xz$, and $yz$ states point more towards the ligands than do the $d_{x^2-y^2}$ and $d_{z^2}$ states. In this case, the $e_g$ states are more stable. Though both of these molecular clusters belong to the same point group, they provide very different representations for that group, and in this example, the relative order of the states split by the crystal field is reversed as a result. In the 6-fold coordinated octahedron, the $t_{2g}$ states are lower in energy, and in the 8-fold coordinated cube, the $e_g$ states are lower.

Again, the only way to quantify these details properly is with explicit quantum mechanical cluster calculations, but the relative order of these split states has been
verified quite generally [1, 2]. For example, the calculations summarized in figure B.1 demonstrate the crystal field splitting in the tetrahedral $O_h$ environment quite well.

As a final point of interest with regard to crystal field splitting, transition metal oxides often form 4-fold coordinated, tetrahedral bonding arrangements as well, belonging to $T_d$. In these systems, the crystal field splitting lowers the $e$ states and raises the $t_2$ states [1].

**C.2 – Jahn-Teller Distortions**

In 1937, H.A. Jahn and E. Teller published a paper in which they proved a striking result about the stability of molecular configurations, now referred to as the Jahn-Teller theorem. In that paper, they showed that a non-linear, polyatomic molecule cannot simultaneously possess both a stable equilibrium and orbital degeneracy (i.e. degeneracy not arising from spin or the two-fold Kramers degeneracy). In addition, this instability is not significant when the degeneracy affects only electronic states which do not play a significant role in bonding [3].

This phenomenon arises, ironically, because of the existence of symmetry in a molecular system. Symmetry gives rise to orbital degeneracy, but it also defines normal modes – the ways in which a molecule can bend, stretch, vibrate, etc. It is these specific normal mode distortions that, in turn, provide the mechanism by which the orbital degeneracy of a molecule is removed – namely, the electronic states couple with the normal mode motions. By moving through some of these normal modes, a molecule can become slightly distorted. Since it is symmetry that gives rise to orbital degeneracy in
the first place, such distortions remove the degeneracy of its electronic states by destroying that symmetry [3-7]. Some of these distorted configurations are lower in energy than the more symmetric arrangement, and as a result, molecules tend to give up their degeneracy by distorting into one of those lower-energy states [6].

Because of the mechanism behind Jahn-Teller distortions, they come in two varieties – static and dynamic. As their names imply, the static Jahn-Teller effect refers to the permanent distortion of a molecule; whereas, the dynamic Jahn-Teller effect refers to the active motions of a molecule through its normal modes. Typically, the static distortion arises when one of the distorted configurations sits at a noticeably lower energy than the others. Because there is a minimum in the energy for one preferred configuration, the molecule simply distorts itself into that configuration. The dynamic distortion, on the other hand, occurs when multiple distorted configurations have similar energies. Because there is no well-defined minimum in this situation, the molecule resonates back and forth continually between several low-energy, distorted states [6, 7].

The simplest realization of the dynamic Jahn-Teller effect is when doubly-degenerate electronic states couple to doubly-degenerate vibrational states. Clearly then, there must be at least one or more doubly-degenerate electronic states in a system for the dynamic Jahn-Teller effect to be observed. While the effects of such distortions are weak (or even absent) in many systems, they can have significant consequences in others, such as the transition metal oxides. In these systems, the Jahn-Teller effect is observed frequently, and it often affects the magnetic, electronic, and optical properties in a non-trivial way [6-11]. Hence, not only is symmetry of profound importance in the electronic
structure of transition metal oxides (e.g. crystal field splitting and molecular orbitals), but the destruction of symmetry is as well.
Table C.1: The character table for the cubic point group, \( O_h \).

<table>
<thead>
<tr>
<th>( \text{O}_h )</th>
<th>E</th>
<th>8( C_3 )</th>
<th>3( C_2 )</th>
<th>6( C_2 ) (^-)</th>
<th>i</th>
<th>8( S_6 )</th>
<th>3( \sigma_h )</th>
<th>6( C_4 )</th>
<th>6( \sigma_d )</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_{1g} )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>( x^2+y^2+z^2 )</td>
</tr>
<tr>
<td>( A_{2g} )</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>( E_g )</td>
<td>2</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>( 2z^2-x^2-y^2, x^2-y^2 )</td>
</tr>
<tr>
<td>( T_{1g} )</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-----</td>
</tr>
<tr>
<td>( T_{2g} )</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>( xy,xz,yz )</td>
</tr>
<tr>
<td>( A_{1u} )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-----</td>
</tr>
<tr>
<td>( A_{2u} )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-----</td>
</tr>
<tr>
<td>( E_u )</td>
<td>2</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>-2</td>
<td>1</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>-----</td>
</tr>
<tr>
<td>( T_{1u} )</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>-3</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>( x,y,z )</td>
</tr>
<tr>
<td>( T_{2u} )</td>
<td>3</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>-3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-----</td>
</tr>
</tbody>
</table>
References


Appendix D

The Structure and Polymorphisms of Group IV–B Transition Metal Oxides and Related Alloys

D.1 – The Symmetry of Rutile TiO$_2$

By far, the most common crystalline form of titanium dioxide is the rutile structure [1, 2]. This structure is an, octahedral complex whose local symmetry is depicted in figure D.1a. Given this site symmetry, it is clear that TiO$_2$ in the rutile structure belongs to the point group, $O_h$ (see appendices A, B, and C); however, it should be noted that even its other crystalline and non-crystalline forms still tend to favor 6-fold coordinated titanium ions and to have related local bonding environments [2].

In order to examine the symmetry of the electronic states in rutile TiO$_2$, it is helpful to factor the representation of its atomic orbitals (AOs), labeled here as $\Gamma_{\text{rutile}}$. To do this, one must first select the appropriate atomic orbitals to consider in such a representation. A top-down approach is perhaps the easiest way to proceed.

Note first that no symmetry element of $O_h$ can transform a titanium AO into an oxygen AO, and vice versa. Hence the space spanned by $\Gamma_{\text{rutile}}$ can already be reduced into the sum of a titanium AO space, $\Gamma_{\text{Ti}}$, and the space spanned by the oxygen AOs, $\Gamma_{\text{O}}$.

In mathematical terms:

$$\Gamma_{\text{rutile}} = \Gamma_{\text{Ti}} \bigoplus \Gamma_{\text{O}}$$  \hspace{1cm} (D.1)
The electrons available for bonding in the titanium atom are the 3d, 4s, and 4p orbitals. In building the local TiO₆ cluster, the starting point is simply a Ti atom in empty space, so it is easiest just to choose the s, p, and d atomic orbitals of the titanium atom as the basis for $\Gamma_{\text{Ti}}$. Consulting the character tables listed in Bishop [3], it is readily apparent that the 3d states belong to the $E_g$ and $T_{2g}$ representations, the 4p states belong to the $T_{1u}$ representation, and the 4s state belongs to the $A_{1g}$ representation. Hence,

$$\Gamma_{\text{Ti}} = \Gamma_{A_{1g}} \oplus \Gamma_{E_g} \oplus \Gamma_{T_{1u}} \oplus \Gamma_{T_{2g}}$$ \hspace{1cm} (D.2)

The oxygen AOs require a bit more finesse in their selection, though it is clear that the 2p valence electron states be involved in some symmetry-adapted way. Thus far, there exists an isolated titanium atom, and now six oxygen atoms will be brought in around that central atom in an octahedral configuration. By definition, the $\sigma$ orbitals will point from the oxygen atoms to the titanium atom, as shown in figure D.1b. Mutually perpendicular to each other and to each of the six $\sigma$ bonds will be a pair of $\pi$ bonds, also shown in figure D.1b. Note that no symmetry operator can transform a $\sigma$ bond into a $\pi$ bond, and vice versa. As a result,

$$\Gamma_{\text{O}} = \Gamma_{\sigma} \oplus \Gamma_{\pi}$$ \hspace{1cm} (D.3)
The next task is to factor the $\sigma$ bonds. To do this, it is necessary to determine the characters of the $\sigma$ bond representation. Following the methods described in appendix A, it is simple to show that these characters are:

\[
\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\text{Character} & 8C_3 & 6C_2 & 6C_4 & 3C_2 & i & 6S_4 & 8S_6 & 3\sigma_h & 6\sigma_d \\
\hline
\text{Value} & 2 & 4 & 0 & 0 & 0 & 2 & 2 & 0 & 0 & 6 & 6 & 3 & 8 & 63 & 6 & 6 & 8 & 64 & 24 & 23 \\
\hline
\end{array}
\]

From here, the factorization is straight-forward through the application of equation A.5. The result is:

\[\Gamma_\sigma = \Gamma_{A_g} \oplus \Gamma_{E_g} \oplus \Gamma_{T_u}\] (D.4)

Lastly, it is necessary to factor $\Gamma_\pi$. The method is exactly the same as the one used to factor $\Gamma_\sigma$. The characters for the $\pi$ bond representation are as follows:

\[
\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\text{Character} & 8C_3 & 6C_2 & 6C_4 & 3C_2 & i & 6S_4 & 8S_6 & 3\sigma_h & 6\sigma_d \\
\hline
\text{Value} & 12 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
\end{array}
\]

After applying equation A.5, the result is:

\[\Gamma_\pi = \Gamma_{T_{1g}} \oplus \Gamma_{T_{2g}} \oplus \Gamma_{T_{1u}} \oplus \Gamma_{T_{2u}}\] (D.5)
Putting together equations D.1 – D.5, the factorization for the rutile representation of $O_h$ is:

$$\Gamma_{\text{rutile}} = 2\Gamma_{A_{1g}} \oplus 2\Gamma_{E_g} \oplus \Gamma_{T_{1g}} \oplus 2\Gamma_{T_{2g}} \oplus 3\Gamma_{T_{1u}} \oplus \Gamma_{T_{2u}}$$  \hspace{1cm} (D.6)

Using equations D.1 – D.6, the states will, therefore, be as follows: $2a_{1g}(\sigma)$, $2e_{g}(\sigma)$, $t_{1g}(O_{\sigma})$, $2t_{2g}(\pi)$, $3t_{1u}(\sigma, \pi)$, and $t_{2u}(O_{\pi})$. The parenthetical terms represent the symmetry, whereby $\sigma$ is an O – Ti $\sigma$ state, $\pi$ is an O – Ti $\pi$ state, and $O_{\pi}$ is a state arising solely from the O 2p $\pi$ orbitals. Since $t_{1g}$ and $t_{2u}$ do not hybridize with any of the Ti states, there is a set of $t_{1g} \oplus t_{2u}$ states at the same energy as the O 2p atomic states. Most of the remaining orbitals come in pairs, which are likely to be bonding/anti-bonding pairs. Hence it is expected that $a_{1g}(\sigma)$, $e_{g}(\sigma)$, $t_{2g}(\pi)$, and $t_{1u}(\sigma, \pi)$ will be at or below the Ti and O atomic orbital energies, and $a_{1g}(\sigma^*)$, $e_{g}(\sigma^*)$, $t_{2g}(\pi^*)$, and $t_{1u}(\sigma^*, \pi^*)$ will be at or above the Ti and O atomic orbital energies. The last $t_{1u}$ could be either bonding or anti-bonding in nature, but it will have both $\sigma$ and $\pi$ character.

In addition to the above information, it is known from crystal field theory that the Ti 3d states will split such that the $e_{g}$ states lie above the $t_{2g}$ states in the conduction band (see appendix C), and this will be mirrored in the valence band. The lowest valence band state is usually symmetric, so $a_{1g}$ is likely to be below $e_{g}$ and $t_{2g}$ there, and this is likely to be mirrored (i.e. $a_{1g}$ is above $e_{g}$ and $t_{2g}$) in the conduction band. Ignoring the $t_{1u}$ states, then, one would expect that the following ordering applies:
The only missing link is the ordering of the mysterious, mixed-character $3t_{1u}(\sigma, \pi)$ states, at least one of which is expected to be of an anti-bonding nature, and at least one of which is expected to have a bonding character. Comparing these results to the calculations represented by figure B.1 reveals how well one can deduce the electronic structure of a system through symmetry alone.

An x-ray absorption (XAS) spectrum measuring the oxygen K\textsubscript{1} edge in TiO\textsubscript{2} is shown in figure D.2, revealing the anti-bonding states in the conduction band [4, 5]. The two lowest-energy peaks are characteristic of the crystal field-split d states in transition metal oxides, with the larger peak representing the $t_{2g}$ states and the smaller representing the $e_g$ states. Based on the symmetry arguments above, the two broader peaks are the $a_{1g}$ and $t_{1u}$ anti-bonding states, with the peak having the larger area being assigned to $t_{1u}$ (since it has more states and should, therefore, absorb more radiation). Thus, the energy of the electronic states in the conduction band of TiO\textsubscript{2} are ordered such that $t_{2g}(\pi^*) < e_g(\sigma^*) < a_{1g}(\sigma^*) < t_{1u}(\sigma^*, \pi^*)$, in agreement with the symmetry-based predictions made above. Since there are no other peaks observed, it is likely that the last $t_{1u}$ state is in the valence band.

Having analyzed its electronic states and crystal field splitting, a final note on Jahn-Teller distortions in TiO\textsubscript{2} is warranted (see appendix C). Figure D.3 shows the degeneracy removal of the d states in an octahedral complex (with the splitting greatly exaggerated). Beginning with the free titanium ion, the d-states are 5-fold degenerate in a

\[ a_{1g}(\sigma) < e_g(\sigma) < t_{2g}(\pi) < t_{1g} \oplus t_{2u} < t_{2g}(\pi^*) < e_g(\sigma^*) < a_{1g}(\sigma^*) \]  

\(\text{(D.7)}\)
spherical \((K_h)\) symmetry. When the oxygen ligands are brought in, the \(O_h\) crystal field splits the d states into \(t_{2g}\) and \(e_g\) states as shown above. The first-order Jahn-Teller distortion is a stretching along one of the principle axes [6], which reduces the cubic \(O_h\) symmetry to the dihedral symmetry of \(D_{4h}\). This reduction in symmetry splits the \(e_g\) states into \(b_{1g}\) and \(a_{1g}\) states, and it splits the \(t_{2g}\) states into \(e_g\) and \(b_{2g}\) states. Finally, distortions of the axes in the plane normal to the principle axis of \(D_{4h}\) remove the \(e_g\) degeneracy, leaving five, non-degenerate states.

**D.2 – The Symmetry of Baddeleyite \(\text{ZrO}_2\) and \(\text{HfO}_2\)**

By far, the most common structure for \(\text{ZrO}_2\) and \(\text{HfO}_2\) crystals is the monoclinic structure known as baddeleyite, though cubic and tetragonal forms exist as well [7-10]. In this structure, the Zr and Hf transition metal ions have a 7-fold coordination as shown in figure D.4. This structure clearly has \(C_1\) symmetry (i.e. no symmetry at all), but it does have a pseudo-cubic nature. It seems as though hafnia and zirconia want to be cubic, but they cannot.

In order to understand this seemingly unusual bonding configuration, it is instructive to examine first the symmetry of the hypothetical molecular orbitals in the cubic structure of figures D.5a and D.5b. As in octahedral \(\text{TiO}_2\), the point symmetry group is \(O_h\), though this representation, \(\Gamma_{\text{cubic}}\), is a bit different. Again, no symmetry operation can transform a Hf/Zr state (considered Hf for the rest of this section, for brevity) into an O state, so the factorization begins with the following:
As with TiO$_2$, the first step is to consider an isolated Hf ion, along with its s, p, and d orbitals (5s, 5p, and 4d in Zr – 6s, 6p, and 5d in Hf). Just as for titanium in TiO$_2$, the result is:

\[
\Gamma_{cubic} = \Gamma_{Hf} \oplus \Gamma_O
\]  

(D.8)

\[
\Gamma_{Hf} = \Gamma_{4g} \oplus \Gamma_{5g} \oplus \Gamma_{4u} \oplus \Gamma_{5g}
\]  

(D.10)

Once again, as in equation D.3, the oxygen 2p states can be factored into \(\sigma\) and \(\pi\) states since no symmetry operator in O$_h$ can transform a \(\sigma\) orbital into a \(\pi\) orbital, and vice versa.

\[
\Gamma_O = \Gamma_{\sigma} \oplus \Gamma_{\pi}
\]  

(D.11)

Hence, the critical task is to factor \(\Gamma_{\sigma}\) and \(\Gamma_{\pi}\). The characters for \(\Gamma_{\sigma}\) are given by:

\begin{array}{cccccccccccc}
E & 8C_3 & 6C_2 & 6C_4 & 3C_2 & i & 6S_4 & 8S_6 & 3\sigma_h & 6\sigma_d \\
--- & --- & --- & --- & --- & --- & --- & --- & --- & --- & --- & --- \\
8 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \\
\end{array}

which, upon the application of equation A.5, yields the following factorization:
\[ \Gamma_\sigma = \Gamma_{A_g} \oplus \Gamma_{T_{2g}} \oplus \Gamma_{A_{2u}} \oplus \Gamma_{T_{1u}} \]  
(D.12)

The characters for \( \Gamma_\pi \) are given by the following:

<table>
<thead>
<tr>
<th>( E )</th>
<th>8( C_3 )</th>
<th>6( C_2 )</th>
<th>6( C_4 )</th>
<th>3( C_2 )</th>
<th>( i )</th>
<th>6( S_4 )</th>
<th>8( S_6 )</th>
<th>3( \sigma_h )</th>
<th>6( \sigma_d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

which, upon the application of equation A.5, yields the following factorization:

\[ \Gamma_\pi = \Gamma_{E_g} \oplus \Gamma_{T_{2g}} \oplus \Gamma_{T_{1u}} \oplus \Gamma_{T_{2u}} \oplus \Gamma_{E_u} \oplus \Gamma_{T_{1g}} \]  
(D.13)

The problem with \( \Gamma_\pi \) is that the available atomic states provide a basis only for \( \Gamma_{E_g} \) and \( \Gamma_{T_{2g}} \) (by using the Hf 5d states). To get the symmetry for the \( \Gamma_{T_{1u}} \) and \( \Gamma_{T_{2u}} \) states, the \( 4f_y(x^2-y^2), 4f_y(3z^2-r^2), 4f_z(r^2-x^2), 4f_z(r^2-y^2), \) and \( 4f_{yz} \) atomic orbitals of Hf are required. In general, however, \( f \) states do not hybridize efficiently (since they are too localized, and since there is so much electron-electron repulsion due to their low symmetry), and they do not play an important role in bonding. Moreover, there are no atomic orbitals that provide a basis for the \( \Gamma_{E_u} \) or \( \Gamma_{T_{1g}} \) states. The only orbitals that make sense, then, would be the Hf 5d states that span \( \Gamma_{E_g} \) and \( \Gamma_{T_{2g}} \). Hence, \( \pi \) orbitals are not
expected to play a big role in the bonding for this local cubic structure, but the \( \Gamma_{\nu g} \) and \( \Gamma_{T_{zg}} \) states will remain as O 2p non-bonding orbitals.

Returning to the factorization for \( \Gamma_{\sigma} \), the appropriate bases for \( A_{1g} \), \( T_{2g} \), and \( T_{1u} \) are \( s \), \((d_{zx}, d_{yz}, d_{xy})\), and \((p_x, p_y, p_z)\), respectively. To get the appropriate symmetry for the \( \Gamma_{A_{2u}} \) states, however, the \( f_{xyz} \) atomic orbital is required. Again, the \( f \) states should not play a significant role in bonding, so one of the \( \sigma \) orbitals must be removed. This leaves only seven \( \sigma \) orbitals for bonding, so the cube distorts slightly to accommodate the electron-electron repulsion between these seven \( \sigma \) bonds. This is the pseudo-cubic (i.e. nearly-cubic) monoclinic arrangement, shown in figure D.4. The \( \Gamma_{A_{2u}} \) state will be a non-bonding O 2p state.

The hybridized molecular orbitals are expected to be cubic-like, retaining some of their cubic character, but they will have their degeneracy removed. From the above arguments, the remaining molecular orbitals should provide the following factorization for \( \Gamma_{\text{cubic}} \):

\[
\Gamma_{\text{cubic}} = 2\Gamma_{A_{1g}} \oplus 2\Gamma_{E_{g}} \oplus 2\Gamma_{T_{1a}} \oplus 3\Gamma_{T_{2g}} \oplus \Gamma_{A_{2u}}
\]

(D.14)

with the seven \( \sigma \) bonds being represented (approximately) by

\[
\Gamma_{\sigma} = \Gamma_{A_{1g}} \oplus \Gamma_{T_{1a}} \oplus \Gamma_{T_{1a}}
\]

(D.15)
The s state spans $A_{1g}$, $(p_x, p_y, p_z)$ spans $T_{1u}$, and $(d_{zx}, d_{yz}, d_{xy})$ spans $T_{2g}$. Hence, these bonds can be described as $sd^{3}p^{3}$ hybrid orbitals.

Figure D.6 shows a representative XAS spectrum for HfO$_2$, which illuminates the nature of the conduction band states in this material [4-6, 11, 12]. The $e_{g}$, $t_{2g}$, $a_{1g}$, and $t_{1u}$ states are all detected, and they demonstrate an ordering such that $e_{g} < t_{2g} < a_{1g} < t_{1u}$. This is expected to be mirrored in the valence band, and since the $e_{g}$ states do not hybridize, it is expected to be at roughly the same energy as the atomic d states. The O 2p non-bonding orbitals are expected to retain their atomic character. Based on this information, a qualitative molecular diagram is proposed in figure D.7, in which the degeneracy removal has been downplayed. While this treatment certainly is not exact, it does provide a qualitative picture for the natural way in which the unusual 7-fold hafnium coordination arises.

Based on this argument, it seems as though a 7-fold coordinated, pseudo-cubic structure is as close as a molecule can get to being cubic when bonding covalently. Obviously, molecules can be cubic, and there are many examples. They can have 8-fold (e.g. Ca$^{2+}$ in CaF$_2$) and even 12-fold (e.g. Sr$^{2+}$ in SrTiO$_3$) coordinated atoms as well. In these cases, however, it is well-known that there is little overlap between atomic states; the nature of the bonds is highly ionic (i.e. electrostatic), and any hybridization between states is minimal.

Again, as a final note on these materials, a discussion of the d-state degeneracy removal is warranted. As the molecular orbital treatment above shows, the $e_{g}$ and $t_{2g}$ states in the conduction band are split by the crystal field as depicted in figure D.8.
however, the opposite ordering of that splitting relative to the crystal field splitting of these states in the octahedral TiO$_2$ molecular cluster. As figure D.6 shows, the additional degeneracy removal of the $e_g$ states is apparent, and it is attributed to the lower symmetry of the monoclinic structure relative to that of the idealized cubic structure, plus Jahn-Teller distortions.

**D.3 – The Symmetry of ([Hf, Ti, Zr]O$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ Compounds**

Having considered the symmetries of the natural structures of TiO$_2$, ZrO$_2$, and HfO$_2$, a discussion of symmetry in the self-organized phase of group IV-B transition metal nitro-silicates is merited. (For brevity, the transition metal will be assumed to be hafnium for the remainder of the discussion.) These materials are pseudo-ternary alloys of HfO$_2$, SiO$_2$, and Si$_3$N$_4$ with concentrations close to 30%, 30%, and 40%, respectively. As FTIR and XAS studies show, they demonstrate a 4-fold coordination of the transition metal ions with a local arrangement as depicted in figure D.9 [13-15]. As a result, the treatment of the local symmetry in these compounds will discuss the tetrahedral HfO$_4$ molecular cluster.

The point symmetry group of this cluster is $T_d$, which is shown in figure D.10a along with the hypothetical, symmetrized O 2p atomic orbitals in figure D.10b. As in sections D.1 and D.2, this representation can be factored initially as:

$$\Gamma_{\text{tetragonal}} = \Gamma_{\text{Hf}} \oplus \Gamma_{O}$$

(D.16)
since none of the Hf states can be transformed into O states by any of the symmetry operations of \(T_d\), and vice versa. The 6s, 6p, and 5d atomic states of the Hf atom are considered first, and referring once again to the character tables in Bishop [3], the factorization of \(\Gamma_{Hf}\) can be written as:

\[
\Gamma_{Hf} = \Gamma_{A_1} \oplus \Gamma_{E} \oplus 2\Gamma_{T_2} \tag{D.17}
\]

with the s state spanning \(A_1\), the p states spanning \(T_2\), and the d states spanning \(T_2\) and \(E\).

The symmetrized oxygen 2p orbitals can be written again, as in sections D.1 and D.2, such that:

\[
\Gamma_{O} = \Gamma_{\sigma} \oplus \Gamma_{\pi} \tag{D.18}
\]

since none of the \(\sigma\) orbitals can be transformed into \(\pi\) orbitals, and vice versa.

As before, the task then becomes to factor \(\Gamma_{\sigma}\) and \(\Gamma_{\pi}\) into their irreducible representations. The characters for \(\Gamma_{\sigma}\) are:

\[
\begin{array}{c|c|c|c|c|c}
E & 8C_3 & 3C_2 & 6S_4 & 6\sigma_d \\
\hline
4 & 1 & 0 & 0 & 2
\end{array}
\]

which leads to a factorization of:
\[ \Gamma_\sigma = \Gamma_{\sigma_1} \oplus \Gamma_{\tau_2} \]  

(D.19)

The characters of \( \Gamma_\pi \) are determined to be:

<table>
<thead>
<tr>
<th></th>
<th>8C_3</th>
<th>3C_2</th>
<th>6S_4</th>
<th>6\sigma_d</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

which leads to a factorization of:

\[ \Gamma_\pi = \Gamma_E \oplus \Gamma_{\tau_1} \oplus \Gamma_{\tau_2} \]  

(D.20)

As in the hypothetical situation treated in section D.2, the \( \Gamma_\pi \) states pose a problem. The atomic states in the HfO_4 molecule cannot provide a basis for \( \Gamma_1 \), so the \( \pi \) bonds are not expected to play a role in bonding for this structure (although, again, some of the \( \pi \) states do exist). The \( \sigma \) bonds, however, are acceptable in this representation since the s states form a basis for \( A_1 \), and the p states form a basis for \( T_2 \). This is the famous sp\(^3\) bonding for tetrahedral molecules. The \( \Gamma_E \) and \( \Gamma_{\tau_2} \) states will remain as non-bonding O 2p states. The total factorization for \( \Gamma_{\text{tetrahedral}} \) is then:

\[ \Gamma_{\text{tetragonal}} = 2\Gamma_{\sigma_1} \oplus 2\Gamma_E \oplus 4\Gamma_{\tau_2} \]  

(D.21)

215
Figure D.11 shows the XAS spectrum of the O K edge for pseudo-ternary \((\text{HfO}_2)_{0.3}(\text{SiO}_2)_{0.3}(\text{Si}_3\text{N}_4)_{0.4}\), revealing its conduction band structure. The states present are \(e\), \(t_2\), and \(a_1\), ordered such that their energies are \(e < t_2 < a_1\). Since the \(e\) states belong solely to the Hf ion’s d states, they should be at approximately the same energy as the atomic d states. The valence band is expected mirrored the conduction band structure, with the remaining states ordered such that their energies are \(a_1 < t_2 < t_2\). The \(\Gamma_E\) and \(\Gamma_T\) non-bonding states are expected to retain their atomic character. The proposed molecular orbital diagram is shown in figure D.12. As in HfO\(_2\), this is again not expected to be exact, but it is expected to capture the qualitative electronic structure of \((\text{HfO}_2)_{0.3}(\text{SiO}_2)_{0.3}(\text{Si}_3\text{N}_4)_{0.4}\).

Once again, this section concludes with a discussion of d-state degeneracy removal. Tetrahedral systems experience crystal field splitting similar to the splitting observed in cubic complexes, whereby the energy of the \(t_2\) states is raised relative to the \(e\) states. This mechanism is shown in figure D.13. The 5-fold degenerate d states of the free Hf ion are spilt into \(e\) and \(t_2\) states by the crystal field, lowering the symmetry from the spherically symmetric \(K_h\) to the tetrahedral symmetry of \(T_d\). Jahn-Teller distortions (exaggerated in figure D.13) then remove the degeneracy further by lowering the symmetry from the tetragonal symmetry of \(T_d\) to the completely asymmetric point group, \(C_1\). This process completely removes the Hf d-state degeneracy.
Figure D.1: A schematic representation of (a) the local site symmetry and (b) the atomic orbitals in rutile TiO$_2$. The blue atoms are titanium, and the green atoms are oxygen.
Figure D.2: The oxygen $K_\alpha$ x-ray absorption edge spectrum for titanium dioxide (data from Lucovsky, et al. [4]).
Figure D.3: The removal of d-state degeneracy in an octahedral transition metal oxide complex, such as TiO₂ (modeled after Lucovsky, et al. [6]).
Figure D.4: The idealized 7-fold transition metal local bonding environment in the baddeleyite structure. The blue atoms are Zr/Hf, and the green atoms are oxygen.
Figure D.5: A schematic representation of (a) the local site symmetry and (b) the hypothetical atomic orbitals in 8-fold coordinated cubic ZrO$_2$ and HfO$_2$. The blue atoms are Zr/Hf, and the green atoms are oxygen.
Figure D.6: The oxygen K\textsubscript{1} x-ray absorption edge spectrum for hafnium dioxide (data from Lucovsky, et al. [5]).
Figure D.7: An approximate molecular orbital diagram for the electronic states of pseudo-cubic hafnia structure having 7-fold coordinated Hf ions.
**Figure D.8:** The removal of d-state degeneracy in a pseudo-cubic, monoclinic transition metal oxide complex, such as HfO$_2$.  

---

**d-State Degeneracy Removal in a Pseudo-Cubic Monoclinic System**

- $t_{2g}$
- $e_g$

Crystal Field + Jahn-Teller Splitting

Symmetry transitions:

- $K_n$ Symmetry $ightarrow$ ($O_n$ Symmetry) $ightarrow$ $C_1$ Symmetry
**Figure D.9:** The local arrangement of atoms in the idealized structure of pseudo-ternary (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$. 
Figure D.10: A schematic representation of (a) the local site symmetry and (b) the hypothetical oxygen 2p atomic orbitals in the pseudo-ternary, 4-fold coordinated, tetragonal (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$. The blue atoms are hafnium, and the green atoms are oxygen.
Figure D.11: The oxygen $K_\alpha$ x-ray absorption edge spectrum for pseudo-ternary \((\text{HfO}_2)_{0.3} \text{(SiO}_2)_{0.3} \text{(Si}_3\text{N}_4)_{0.4}\) (data from Lucovsky, et al. [13]).
Figure D.12: An approximate molecular orbital diagram for the electronic states of pseudo-ternary, tetragonal $(\text{HfO}_2)_{0.3}(\text{SiO}_2)_{0.3}(\text{Si}_3\text{N}_4)_{0.4}$ structure having 4-fold coordinated Hf ions.
Figure D.13: The removal of d-state degeneracy in a tetragonal transition metal oxide complex, such as (HfO$_2$)$_{0.3}$(SiO$_2$)$_{0.3}$(Si$_3$N$_4$)$_{0.4}$ (modeled after Lucovsky, et al. [6]).
References


Appendix E
A Description of Intermediate Phases and Self-Organization in Non-Crystalline Solids

E.1 – Short Range Order: Bond Constraint Theory

Because non-crystalline solids lack any regular, periodic order, there exists no complete, rigorous mathematical theory of their structural or electrical properties. The Fourier-based methods employed to investigate such properties in many other solid-state systems simply cannot be applied to a non-periodic lattice. Despite their macroscopic disorder, however, non-crystalline solids do possess some degree of local uniformity. The similarities in bond lengths, bond angles, and coordination numbers of atoms in these materials constitute a short-range, topological order.

As a result, a topological description of non-crystalline solids, based on ideas of network connectivity and structural rigidity, was put forth as a way to describe the average properties of such materials. These ideas were first proposed by J.C. Phillips and modified by G.H. Döhler et al. nearly 30 years ago [1, 2]. Since then, this so-called bond constraint theory (BCT) has been studied extensively, both experimentally and computationally [3-6], and it has provided a great deal of insight into the nature of covalent, non-crystalline material systems.

In BCT, constituent atoms are first assumed to follow the famous 8-N rule for covalent bonding (i.e. that their coordination is equal to 8 minus the number of valence
electrons in their outer shell). Next, two types of valence force fields (VFFs) are assumed to be present: (1) “strong” bond stretching forces (denoted as $\alpha$) that depend on the distance between the atoms ($i$) and their nearest neighbors ($j$) as $|\vec{R}_{ij}|^2$ and (2) “weak” bond bending forces (denoted as $\beta$) that depend on the atoms ($i$), their nearest neighbors ($j$), and their second-nearest neighbors ($k$) as $\vec{R}_{ij} \cdot \vec{R}_{jk}$. Lastly, the Van der Waals attractions and electrostatic repulsions between non-bonding, lone-pair electrons in constituent atoms are considered.

To illustrate the ideas simply, the case of a non-crystalline network solid in the form of $A_xB_{1-x}$ is treated. The average coordination, $\langle m \rangle$, in such a material is defined by the following equation:

$$\langle m \rangle = x m_A + (1 - x) m_B \quad (E.1)$$

where $m_A$ and $m_B$ are the coordinations of A and B, respectively. Furthermore, it is clear that $N_B = 8 - N_A$ as required by the 8-N rule. Note that this definition permits non-integral coordination to exist in the representative pseudo-atoms defined by equation E.1.

The first task is to count the number of forces, or constraints, that are present in the system. Consider first the $\alpha$ forces which depend on $|\vec{R}_{ij}|^2$. For each atom, $i$, the index, $j$, runs from 1 to $m_i$, the number of nearest neighbors of atom $i$. These constraints can be considered ideally as rods or stiff springs connecting atoms $i$ and $j$ such that the
stretching constraint is shared equally between the two. Hence, the number of $\alpha$ constraints is given by the following:

$$n_{\alpha} = x\left(\frac{m_A}{2}\right) + (1-x)\left(\frac{m_B}{2}\right)$$

$$= \frac{m}{2} \quad \text{(E.2)}$$

In counting the $\beta$ forces, two cases must be considered since in $N_d$ dimensions only $N_d$ angles are linearly independent. If $\langle m \rangle \leq N_d-1$, the indices in the $\vec{R}_{ij} \cdot \vec{R}_{jk}$ dependence run over $j = 1$ to $j = m_i$ and $k = 1$ to $k = m_j-1$ (to avoid double-counting). Additionally, since bonds are shared between atoms, only $1/2$ of each constraint belongs to a particular atom. Hence, the number of $\beta$ constraints is given by the following for $\langle m \rangle \leq N_d - 1$:

$$n_{\beta} = x\left(\frac{m_A \left[xm_A + (1-x)m_B\right]}{2} - \left[xm_A + (1-x)m_B\right]\right)$$

$$+ (1-x)\left(\frac{m_B \left[xm_A + (1-x)m_B\right]}{2} - \left[xm_A + (1-x)m_B\right]\right)$$

$$= x\left(\frac{\langle m \rangle m_A - \langle m \rangle}{2}\right) + (1-x)\left(\frac{\langle m \rangle m_B - \langle m \rangle}{2}\right)$$
\[ \langle m \rangle (\langle m \rangle - 1) \leq N_d - 1 \]  \quad m \leq N_d - 1 (E.3)

If \( \langle m \rangle > N_d - 1 \), however, \( \langle m \rangle - N_d \) of the \( \beta \) constraints are linearly dependent; only \( N_d \) are linearly independent. Considering only these constraints, equation E.3 is modified to read [2]:

\[ n_\beta = \frac{N_d - 1}{2} (2\langle m \rangle - N_d) \quad \langle m \rangle > N_d - 1 (E.4) \]

Putting together equations E.2 – E.4, the constraint counting is given by:

\[ N_{CO} = \begin{cases} 
\frac{\langle m \rangle^2}{2} & \langle m \rangle \leq N_d - 1 \\
\frac{\langle m \rangle}{2} + \frac{N_d - 1}{2} (2\langle m \rangle - N_d) & \langle m \rangle > N_d - 1 
\end{cases} (E.5) \]

Using Phillips’s condition that \( N_{CO} = N_d \) for ideal networks:

\[ \langle m \rangle = \begin{cases} 
\sqrt{2N_d} & N_d \geq \langle m \rangle + 1 \\
\frac{N_d (N_d + 1)}{2N_d - 1} & N_d < \langle m \rangle + 1 
\end{cases} (E.6) \]
which gives in three dimensions the condition for an ideal non-crystalline network solid as:

\[ \langle m \rangle = 2.4 \quad \text{(E.7)} \]

If any of the constraints are too weak to be counted, such as the flexible Si-O-Si bridging bonds in SiO₂, there are additional internal degrees of freedom (or equivalently, broken constraints). In order to account for the extra internal degrees of freedom, \( N_d \) must be changed in the following way in the statement of the ideal network condition:

\[
xN_{CO}^A + (1-x)N_{CO}^B = x\left(N_d^A + N_i^A\right) + (1-x)\left(N_d^B + N_i^B\right)
\]

where \( N_i \) represents the internal degrees of freedom, and the superscripts denote the particular species. In general, for more than two atoms (i.e. a compound with chemical formula \( A_{n_a} B_{n_b} C_{n_c} \cdots \)), the equation is modified to read:

\[
\sum_{\lambda} x_{\lambda} N_{\lambda\lambda CO} = \sum_{\lambda} x_{\lambda} \left(N_{\lambda\lambda d}^A + N_{\lambda\lambda i}^A\right)
\]

\[
\Rightarrow \frac{\sum_{\lambda} n_{\lambda} N_{\lambda CO}^\lambda}{\sum_{\mu} n_{\mu}} = \frac{\sum_{\lambda} n_{\lambda} \left(N_{\lambda\lambda d}^A + N_{\lambda\lambda i}^A\right)}{\sum_{\mu} n_{\mu}}
\]

236
\[ \sum_{\lambda} n_{\lambda} N_{\lambda}^{CO} = \sum_{\lambda} n_{\lambda} \left( N_{\lambda}^{d} + N_{\lambda}^{i} \right) \]  

(E.9)

where \( x_{\lambda} \) is the percentage of atomic species \( \lambda \) in the compound, \( n_{\lambda} \) is the stoichiometric coefficient of atomic species \( \lambda \) in the chemical formula, and the various \( N \)'s are written in analogy with earlier notations. Considering only 3-dimensional networks with no singly coordinated atomic species, with the help of equation E.5, equation E.9 becomes:

\[ \sum_{\lambda} n_{\lambda} \left( \frac{5m_{\lambda}}{2} - 3 \right) = \sum_{\lambda} n_{\lambda} \left( 3 + N_{\lambda}^{i} \right) \]  

(E.10)

in which \( m_{\lambda} \) is the coordination of atomic species \( \lambda \).

Thorpe, et al. [3] showed that it is beneficial to define this in terms of the fraction of zero-frequency modes. In analogy, it is useful to define the same quantity (slightly modified):

\[
\begin{align*}
\frac{\sum_{\lambda} n_{\lambda} \left( 3 + N_{\lambda}^{i} \right) - \sum_{\lambda} n_{\lambda} \left( \frac{5m_{\lambda}}{2} - 3 \right)}{\sum_{\lambda} n_{\lambda} \left( 3 + N_{\lambda}^{i} \right)} \\
\end{align*}
\]

\[ f = \frac{\sum_{\lambda} n_{\lambda} \left( 3 + N_{\lambda}^{i} \right) - \sum_{\lambda} n_{\lambda} \left( \frac{5m_{\lambda}}{2} - 3 \right)}{\sum_{\lambda} n_{\lambda} \left( 3 + N_{\lambda}^{i} \right)} \]  

(E.11)
which reduces to Thorpe’s expression when internal degrees of freedom are ignored or are zero. In light of equation E.11, the following definition for a critical average coordination becomes useful (derived by setting the numerator in equation E.11 equal to zero):

\[ \langle m \rangle_c = \frac{12}{5} + \frac{2}{5} \sum x_i N_{i_x} \]  

(E.12)

Three cases are considered:

- \( \langle m \rangle > \langle m \rangle_c \)
- \( \langle m \rangle < \langle m \rangle_c \)
- \( \langle m \rangle = \langle m \rangle_c \)

Because \( f \) must be zero or larger, it is necessary that \( f = 0 \) for \( \langle m \rangle > \langle m \rangle_c \) (since otherwise it would be negative). Since there are no zero-frequency modes for this condition, these networks are over-constrained, and this case is referred to as stressed-rigid. Similarly, for \( \langle m \rangle < \langle m \rangle_c \), \( f > 0 \), and there do exist zero-frequency modes. These networks are under-constrained and referred to as floppy. For the special case in which \( \langle m \rangle = \langle m \rangle_c \), the network is ideally constrained; it is isostatically rigid since \( f = 0 \), but it is not over-constrained by additional connectivity [5]. These networks are called just-rigid.
Using stiff springs to depict the bonds, figure E.1 shows these three cases schematically for a square network composed of four atoms.

In its original context, Phillips’s BCT was concerned with trying to describe glass formation in Ge\textsubscript{x}Se\textsubscript{1-x} and As\textsubscript{x}Se\textsubscript{1-x} alloys [1]. By examining the quenching rate in these materials, he noticed that for certain compositions glasses were easily formed, yet for others glass formation was difficult. Some of his results are shown in figure E.2. Note the minimum in the formation difficulty around 14%. Without considering Van der Waals forces or internal degrees of freedom, BCT as presented above predicts a minimum for Ge\textsubscript{x}Se\textsubscript{1-x} with x = 0.2 (i.e. GeSe\textsubscript{4}). While this is in the range of the glass forming minimum, it is not exact. When Van der Waals forces resulting from non-bonding Se electrons are included, along with the internal degrees of freedom of Se, BCT predicts that x = 1/7 ≈ 0.14 – right at the minimum shown in figure E.2 [2].

Success by other researchers has confirmed BCT to be a very good description of non-crystalline networks as well (e.g. refs. [4-6]). The way in which internal constraints are counted tends to be a bit less rigorous than is intellectually or philosophically desirable, but BCT captures the essence of local order in these material systems well, and (to this author’s knowledge) there has yet to be a competitive alternative put forth.

The most important result of BCT is that ideal, non-crystalline networks are of the just-rigid variety. Too few restrictions means that the network is floppy and has too much freedom to rearrange and crystallize; too many restrictions means that the network is under stress and will try to crystallize exothermically to relieve it. With a just-rigid
configuration, a network is rigid, but stress free, avoiding both of these crystallization mechanisms [5].

**E.2 – Long Range Order: Intermediate Phases and Percolation Theory**

The short range order of non-crystalline materials was discussed in the previous section in the context of bond constraint theory (BCT), which describes these systems in terms of mechanical constraints and rigidity. Because BCT is a statistical, mean-field theory, the local rigidity is likely to differ from the macroscopic, mean-field average at various locations throughout a material. A discussion of the long range order, therefore, must address these variations, effectively describing how local rigidity relates to global rigidity. The correct ideas for such a treatment come from percolation theory.

Since it is difficult to provide a satisfactory linguistic definition of percolation theory, the concept will be demonstrated through two examples from Zallen’s book on amorphous solids [7]. The first is a two-dimensional example of *bond* percolation on a square lattice. Consider a network of resistors, connected in a chain-link array, with each resistor attached to two 4-fold coordinated nodes. At opposing ends of this network are two conducting plates to which the terminal resistors are attached and across which a voltage applied. The current flowing through this system is measured as the resistors are removed *at random* and plotted as a function of the fraction, \( p \), of resistors remaining. Lastly, assume that the system is very large, essentially in the thermodynamic limit. The result is shown in figure E.3. Note that the current is reduced linearly until \( p \) reaches a
value of approximately 0.5, at which time the current falls to zero. This cut-off value is called the percolation limit, and it is given the symbol $p_c$.

As a second example, consider a three-dimensional example of *site* percolation. In this experiment, a large (in diameter and height) glass cylinder is filled with small beads. The bottom of the cylinder is coated with a conductor and attached to one terminal of a battery. Each of the beads is composed of either SiO$_2$ or Al, and the fraction of Al beads as compared to the total number of beads is $p$. This mixture of beads is shaken up and randomized, then poured into the cylinder where it is lightly shaken to allow the beads to settle into a random close-packed (rcp) structure. A conducting foil is then placed tightly against the beads at the top of the cylinder and attached to the other terminal of the battery. The current flowing through the system is then measured, and the experiment is repeated for other values of $p$. The behavior of this system is similar that shown in the resistor network of figure E.3, but this experiment has a different geometry. As a result, $p_c$ should be different. In fact, it is found that $p_c = 0.64$ [7].

If this experiment was repeated using a diamond lattice, rather than an rcp structure, the result would be $p_c = 0.34$. For a face-centered cubic (fcc) structure, $p_c$ would be 0.74. Despite these variations, there is an invariant (dimensionally-dependent) description of this critical phenomenon. Letting $v$ be the filling factor of the lattice (i.e. the ratio of space occupied by the beads relative to the total space available to them), the quantity, $vp_c$ provides this invariant. For example, despite the fact that the diamond structure has a $p_c$ of 0.34, and the face-centered-cubic (fcc) structure has a $p_c$ of 0.74, the percolation density, $vp_c$, is equal to 0.146 for the diamond structure and equal to 0.147 for
the fcc structure. Averaging over the various crystalline structures in three dimensions, as well as the rcp structure, the average percolation density for site percolation in three dimensions is found to be $0.16 \pm 0.02$ [7]. For bond percolation, the invariant quantity is $z p_c$, where $z$ is the coordination of the sites. The average $z p_c$ for bond percolation in three dimensions is found to be $1.5 \pm 0.1$ [7].

These two experiments capture the nature of percolation theory quite well. In essence, percolation theory is concerned with how much of a random, local characteristic needs to be present in a system before becoming collective and being reflected in the macroscopic properties of the system.

Rigidity percolation in BCT is exactly this same phenomenon. Namely, a material goes from floppy to rigid as constraints are added at random (i.e. as average coordination and internal degrees of freedom are increased). As constraints are added, eventually local clusters of rigidity emerge. As more and more of these small, local clusters form, they begin to merge, thereby forming larger clusters. Eventually, the material becomes one large cluster at the percolation limit, and the system becomes completely rigid. This process was investigated in two dimensions on a bond-diluted random network by Jacobs and Thorpe, and their results are shown in figure E.4 to illustrate this phenomenon [8].

Recall from section E.1 that (in the absence of internal degrees of freedom) the floppy – rigid transition takes place at an average coordination of 2.4. Equations E.11 and E.12 indicate that the fraction, $f$, of zero-frequency modes at this critical coordination goes abruptly from zero to $f > 0$. As computer models by Thorpe, et al. suggest, this
transition in reality is not abrupt, but continuous [3]. His plot of $f$ versus average coordination is shown in figure E.5. Note the indication of an intermediate phase. This predicted behavior was confirmed experimentally by Boolchand, et al. by using differential scanning calorimetry (DSC) and Raman spectroscopy to study a variety of chalcogenide glasses [9]. His DSC results for $As_xSe_{1-x}$ are shown in figure E.6, along with viscosity measurements made much earlier by Nemilov, et al. [10]. All of this data verifies the existence Thorpe’s window for the formation of intermediate phases. It is the formation of these intermediate phases that are suspected to be the cause of spontaneous self-organization in a variety of materials, from glasses to proteins [5].

The beauty of the ([Ti, Zr, Hf]O$_2$)$_x$(SiO$_2$)$_y$(Si$_3$N$_4$)$_{1-x-y}$ material systems is that there exists an intermediate phase close to the composition $x = y = 0.3$, although the percolation phenomenon in this case is more complicated. Rather than a single rigidity percolation as in $As_xSe_{1-x}$, for example, there is a confluent double percolation involving not only short-range bonding constraints, but long-range strain as well [11]. The first issue involves the local order of the cluster shown in figure D.9 and is a site percolation phenomenon. Using BCT (and including the broken constraints properly), the number of constraints per atom in this cluster can be shown to be 3.08 – very close to the ideal value of 3 at the percolation limit for rigidity (as required by BCT) [11]. The second issue deals with the way in which these local clusters embed themselves into the larger Si$_3$N$_4$ matrix in order to relieve macroscopic strain. This is a site percolation phenomenon, and the percolation pathway is provided by considering the ratio of [Ti, Zr, Hf]/Si on 4-fold coordinated sites. In the self-organized phase, this ratio is $1/6 = 16.7\%$, which is right at
the percolation density ($v_p$) for site percolation [7]. Any materials in this system that satisfy only one of the two confluent percolation paths will chemically phase separate into [Ti, Zr, Hf]O$_2$ and SiO$_2$. That there exists a chemically ordered phase in these materials which occurs concurrently with not just one, but with two percolation phenomena, is nothing short of astonishing.
Figure E.1: A schematic representation of floppy, stressed-rigid, and just-rigid modes in a non-crystalline, covalent network solid (modeled after Boolchand, et al. [5]).
Figure E.2: The glass forming difficulty in Ge$_x$Se$_{1-x}$ as a function of germanium composition, x (from Phillips [1], © 1979 North-Holland Publishing Company).
Figure E.3: Site percolation on the square lattice resistor network described in the text (modeled after Zallen [7]). The percolation limit, $p_c$, is 0.5 and is one of the few analytic solutions known in percolation theory.
Figure E.4: Rigidity in a bond-diluted, generic two-dimensional network near the percolation limit. Dark regions indicate over-constrained regions, solid lines indicate rigid regions, and open circles indicate floppy regions. (Figure from Jacobs and Thorpe [8], © 1996 The American Physical Society.)
Figure E.5: The fraction of floppy modes as a function of average coordination using BCT (solid line) and a computer model for a randomly diluted diamond lattice (dot-dash line). The difference in the onset of rigidity between the two cases indicates an intermediate phase. (From Thorpe, et al. [8], © 2000 Elsevier Science B.V. All rights reserved.)
Figure E.6: The viscosity (left axis) and reversible heat flow (right axis) measurements indicating the existence of an intermediate phase in the As$_x$Se$_{1-x}$ alloy system (after Boolchand, et al. [9]).
References


Appendix F

Auger Electron Spectroscopy (AES)

F.1 – The Auger Process, Auger Electron Detection, and AES Spectra*

The Auger process (shown in figure F.1) is a two-electron process which can provide chemical information about a surface or thin film by what is referred to as Auger electron spectroscopy (AES). In this class of experiments, a core electron is excited first into the vacuum level by an electron beam (generally at 3 keV, 5 keV, or 10 keV). Next, an electron from an outer shell of one of the atoms in the film relaxes into the vacant atomic core state, and another electron is emitted to vacuum in order to balance the energy given off by this electronic relaxation. This is why the process is referred to as a two-electron process. The electrons emitted to vacuum during the electronic relaxation are known as Auger electrons, and their kinetic energies are characteristic of the chemistry from which they originate.

The nomenclature associated with AES uses three letters to characterize the core level electron, the relaxation electron, and the Auger electron, respectively. The letters K, L, M, … are used to designate principle quantum numbers of 1, 2, 3, …, respectively. Similarly, V is used to designate an electron from the valence band of a solid. Subscripts are used to designate non-zero orbital quantum numbers of the electrons involved, if applicable. If the spin-orbit splitting is too small to detect, more than one subscript may

*Most of the information in this section can be found in any relevant text. For example, see Woodruff and Delchar [1] or the Handbook of Auger Electron Spectroscopy [2].
be used to designate that both types of electron are involved in the Auger process. For example, the Hf and Si transitions discussed in this dissertation are KVV processes. In this process, a core s electron is liberated from an atom into the vacuum level, an electron from the valence band relaxes into the newly-created core hole, and another electron from the valence band is emitted into vacuum with enough kinetic energy left over to conserve overall energy in this process.

The Auger electrons are typically generated and detected via a concentric cylindrical or spherical geometry. For experiments in this work, a concentric cylindrical mirror analyzer (CMA) and electron gun set-up was used, as depicted in figure F.2. The incident electron beam is generated along the axis of the cylinder and directed at the sample target. Some of the emitted Auger electrons are ejected in the direction of the CMA entrance and enter the cylinder. Once inside the cylinder, a control voltage is applied to the walls of the CMA by the energy analyzer in order to curve the paths of the Auger electrons via a Lorentz force. Since the curvature of an electron’s path depends on its kinetic energy, tuning the CMA control voltage will allow only electrons with the desired energy into the analyzer. The detector in the Auger energy analyzer is simply a photomultiplier tube. A complete AES spectrum is generated by sweeping the CMA control voltage over the desired range, and it is comprised of the Auger electron kinetic energy versus the number of electrons detected by the CMA at that energy.

Many, and in fact most, Auger electrons are scattered as they make their way out of the sample. This process reduces their kinetic energy by varying degrees and erases from them the characteristic signature of their parent atoms. Hence, raw AES spectra are
generally quite smooth. Some Auger electrons are not scattered, however, and retain their characteristic energy. As a result, there are small Auger-related features riding on top of a large, smooth background. Differentiation is typically done in order to enhance the magnitude of these features and to quantify their exact positions. Such 1st derivative AES curves are what others generally refer to as the AES spectra (leaving out mention of the implied derivative analysis).

F.2 – Elemental Analysis by AES*

The kinetic energy of Auger electrons is governed by atomic (and in some cases molecular orbital) energy levels; therefore, AES provides a manner in which to analyze not only the types, but also the relative amounts, of atoms present in a thin film. The types of atoms present in a sample are easily identifiable based on their peak positions and shapes when compared to reference spectra. Because of differences in transition probabilities and ionization cross sections, however, there are relative differences in how efficiently various peaks are detected in AES. These effects are taken into account by the so-called atomic sensitivity factors, $S_x$, for a given electronic transition, $x$. If $I_x$ refers to a particular AES peak intensity, then the relative amount, $P_x$, of the atomic species responsible for that transition and present in the sample is given by the following (as a percentage):

* Most of the information in this section can be found in any relevant text. For example, see the Handbook of Auger Electron Spectroscopy [2].
As an example, figure F.3 shows the AES spectra for two HfSiON thin films of different relative atomic concentrations. The pseudo-ternary compositions listed in the figure were calculated using equation F.1.

\[
P_x = \frac{I_x}{S_x} \times 100
\]

(F.1)

F.3 – Thickness Estimations by AES

A crude estimate for the thickness of thin films can also be attained through the use of AES. When the electron beam interacts with the sample, the incident electrons are scattered by atoms inside of the film. This spreads them out in the plane and slows them down as they penetrate into the film, producing a teardrop shaped interaction volume [2]. The characteristic length of this interaction is called the penetration depth. Similarly, outgoing Auger electrons will be scattered and slowed by atoms within the sample. The characteristic length an electron can travel without being scattered is referred to as the inelastic mean free path, or imfp, and only the Auger electrons which do not scatter at all will escape with their parent atom’s signature, thereby contributing to the AES spectrum. Given the physics of the Auger process, incident electrons by definition must have a higher energy than the outgoing Auger electrons. As a result, the deepest Auger electrons detectable in AES are determined not by the penetration depth of the electron beam, but by the inelastic mean free path of the Auger electrons [2, 3].
Experiments have shown that the thickness of a thin film (approximately less than 120 Å) can be estimated to within roughly 30% by the following formula [3]:

\[
d = 0.77 \lambda \ln \left( 1 + \alpha \frac{I_{\text{film}}}{I_{\text{substrate}}} \right)
\]  

(F.2)

in which \(d\) is the approximate film thickness, \(\lambda\) is the imfp, \(I_{\text{film}}\) is the height of the peak corresponding to the film, \(I_{\text{substrate}}\) is the height of the substrate peak, and \(\alpha\) is the ratio of the peak height for a pure substrate to the peak height of a pure overlayer material. The quantity, \(\alpha\), can be measured by taking two simple AES spectra: one of a cleaned substrate, and one of a sample onto which the appropriate overlayer has been deposited for an extended period of time (i.e. until the AES spectrum no longer changes). Their ratio gives \(\alpha\). The United States National Institute of Standards and Technology (NIST) provides a database for calculating the imfp, \(\lambda\) [4]. With \(\lambda\) and \(\alpha\) determined for a given substrate and overlayer chemistry, the thickness estimate becomes routine and can be determined simply by analyzing the AES spectrum of the sample being investigated.

As an example, consider the oxidation of a silicon surface by remote plasma. By incrementally oxidizing a clean Si surface and measuring the resulting AES spectra, the SiO\(_2\) overlayer can be plotted as a function of oxidation time. The AES spectra are shown sequentially in figure F.4, and the resulting thicknesses as analyzed by equation F.2 are plotted in figure F.5. The growth kinetics provided by this analysis follow the model exactly as expected for Si surface oxidation [5, 6].
**Figure F.1:** The KVV Auger process responsible for electrons detected in AES experiments discussed in this dissertation.
**Figure F.2:** A stereotypical experimental set-up for Auger electron detection in AES by a cylindrical mirror analyzer.
Figure F.3: Example AES spectra for two different compositions of the pseudo-ternary compound, HfSiON. The elemental AES peak positions are identified, as are the pseudo-ternary compositions computed from the relative atomic concentrations.
Figure F.4: The AES spectra resulting from the incremental oxidation of a clean Si surface by remote plasma.
Figure F.5: The oxidation kinetics of a clean Si surface by remote plasma. This process follows the model for silicon surface oxidation very well, as indicated by the fitting function in the lower right corner [5, 6].
References


Appendix G

X-Ray Absorption Spectroscopy (XAS)

G.1 – The X-Ray Absorption Process and Experiment

The x-ray absorption process proceeds in two main phases – excitation and relaxation. Depending on the type of excitation and the kind of relaxation measured, spectra may vary slightly in the information they contain.

The x-ray absorption process begins when an x-ray interacts with the sample, exciting a core electron from one of its constituent atoms in one of two primary ways. If the incident x-ray energy is near the absorption threshold for a transition, the core electron will be excited into a non-bonding or anti-bonding state at or just above the Fermi level. On the other hand, if the incident x-ray energy is much larger than that, the core electron will be excited into the vacuum level. Furthermore, the core hole can be filled by any radiative or non-radiative transition the transfers an electron from a state with lower binding energy into the core electronic state. As a result, there are many modes of measurement in x-ray absorption, typically coming from one of two main classes: fluorescent detection and electron yield.

The XAS experiments discussed in this dissertation use total electron yield (TEY) detection, which measures the total number of electrons liberated from atoms within the

* Most of the information in this section can be found in any relevant text. For example, see X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS, and XANES [1], NEXAFS Spectroscopy [2], or J.G. Chen [3].
interaction volume of the sample. This includes photoelectrons and Auger electrons which are detected as current flowing through the sample. Two processes representative of the XAS studies carried out for this dissertation are depicted in figure G.1. Figure G.1a shows the excitation of a core electron into an anti-bonding state just above the Fermi level, which is relaxed by the emission of an Auger electron. The hole left in the valence band is filled by the flow of current into the sample. Figure G.1b shows the excitation of a photoelectron from a core electronic state, which is relaxed by the radiative emission of a fluorescent photon. Again, the hole left in the valence band is filled by the flow of current into the sample. In fact, TEY is quantified in just that way – by measuring sample current.

Of course, detection is the easy part of XAS; obtaining a variable-energy, monochromatic source of x-rays is a bit trickier. In this case they were provided by the Stanford Synchrotron Radiation Laboratories (SSRL), located at Stanford University in Menlo Park, CA. SSRL is an electron storage ring that has been added at the end of a large linear particle accelerator (the Stanford Linear Accelerator Center, or SLAC), and in it, electrons zoom around at near the speed of light. At various points along the ring, some of these electrons are siphoned off down linear tubes (beamlines) along which wigglers and/or undulators (i.e. fancy magnets) bend their trajectories dramatically, thereby producing x-rays through bremsstrahlung. These x-rays are generated over a wide range of energies, but they are passed through a monochromater in order to select just one.
The x-ray generation process described above is more or less the same at any synchrotron; however, the instrumentation at the end of each beamline can vary widely, depending on the particular experiment. A schematic of beamline 10-1 at SSRL is given in figure G.2, and it depicts the experimental set-up for the XAS experiments conducted in this work. Polychromatic x-rays from the synchrotron are focused first through an entrance slit to create a narrow beam. They then pass through a monochromater, which selects the incident energy, and through an exit slit, further narrowing the beam. A final focusing mirror ensures that the beam is aligned with the sample, and a very fine mesh grid (approx. 100 wires/cm) is placed in the path of the beam in order to monitor the incident beam’s intensity. Because the incident beam flux can vary over time, the sample current is divided by the total current in the grid in order to filter out such unwanted variations. By sweeping the monochromater over the desired energy range, the XAS spectrum is generated.

G.2 – XAS Spectra

Having described what XAS measures and how these measurements are made, a discussion about what types of information are contained in the spectra is needed. A treatment of all such information fills books (e.g. refs. [1, 2]), but the information relevant to transition metal oxides can be found in review papers, such as refs. [3-5]. In particular, this work is concerned with band edge defect states [6-8, 11], nanocrystallinity [9], and degeneracy removal via crystal field splitting and/or Jahn-Teller distortions [10].
The reason that these effects can be observed is rooted in the fact that the x-ray energies can be finely tuned to particular absorption processes. In covalent (or at least partially covalent) solids, the states lying just above the Fermi level are empty hybrid molecular orbital states. In transition metal oxides, these are typically mixtures of transition metal d states and oxygen 2p states. Hence, by honing in on absorptions from oxygen core electrons into empty molecular orbital states, the influences of the neighboring transition metal atoms via molecular orbital hybridization are observed. In fact, the oxygen K₁ near-edge spectra for transition metal oxides are dominated by d-like states, which never would be observed in atomic or molecular oxygen. In a sense, inter-atomic effects are observed through these intra-atomic transitions because of molecular orbital hybridization.

Often, the simplest observations are related to molecular orbital degeneracy removal. Figure G.3 depicts the evolution of the symmetry of electronic d states in transition metal compounds during their degeneracy removal (see appendices A-D). Figure G.4 shows two spectra which demonstrate how this degeneracy removal is observed spectroscopically. Because HfO₂ and TiO₂ have different local bonding geometries, the crystal field splitting is actually reversed in the two materials, but both clearly show the two d-state absorption peaks (e₉ and t₂g). HfO₂ splits according to the O₁₁₁₁ symmetry on the left-hand side of figure G.3, and TiO₂ splits according to the O₁₁₁₁ symmetry on the right-hand side. Moreover, the slightly distorted pseudo-cubic bonding environment of HfO₂ (see appendix D) is manifested clearly in a further splitting of the e₉g peaks in accordance with the C₁₁₁₁ geometry on the left-hand side of figure G.3. Further
splitting of the $t_{2g}$ peaks is also present in HfO$_2$, though it is difficult to see in the raw spectrum. Figure G.5 shows clearly the complete d-state degeneracy removal in HfO$_2$. The five d-state features, split by a combination of the crystal field and Jahn-Teller distortions, show up quite distinctly in this derivative spectrum.

The effects of nano-crystallization are also readily detected via XAS [9], as depicted in figure G.6. Between annealing temperatures of 650ºC and 700ºC, the d-state features in the spectra sharpen and split, and even the s and p states are greatly enhanced. This is the onset of the transition between type I and type II nano-crystallization (see section 3.4). These effects are observed for exactly the same reasons as the crystal field splitting. When crystallization occurs, the symmetry of the transition metal oxide changes as larger crystallites and grain boundaries form. The results are the enhanced crystal field and Jahn-Teller splitting effects observed in figure G.6.

Though often quite difficult to do so, sub-band defect features can be observed in XAS spectra; however, its sensitivity to these states is low, and corroboration by other methods is generally necessary for a definitive analysis [6-8, 11]. Figure G.7 shows the defect states detected in a thin film of HfO$_2$, which later were confirmed by SE (see appendix H). These small shoulders located in the pre-edge region of the XAS spectrum are discernable, but the absorption there is very low relative to the rest of the spectrum. Hence, these features often may be buried in other signals, though this is not always the case. For example, quantitative measurements have been made on HfO$_2$ films, yielding both relative defect state energy levels and relative defect state densities [11]. Band tail, rather than discrete sub-band, defects also can be detected (not shown here). In those
cases, the absorption edge of one sample may be much more gradual than in a related sample, but discrete peaks would not be observed. Of course, quantification of such an effect is very difficult, indeed; relative qualifications are much more useful in those cases.
Figure G.1: Representative depictions of the x-ray absorption processes detectable in the TEY: (a) core level excitation into an anti-bonding state, relaxed by Auger electron emission, and (b) core level excitation to the vacuum level (photoelectron emission), relaxed by fluorescent photon emission. Of course, these excitation/relaxation processes do not have to be paired as they are above.
**Figure G.2:** A schematic representation of beamline 10-1 at SSRL.
Figure G.3: Degeneracy removal as a function of electronic state symmetry as it relates to transition metal compounds (modeled after Lucovsky, et al. [10]).
Figure G.4: XAS data showing how d-state degeneracy removal is observed spectroscopically (data from Lucovsky, et al. [10]).
Figure G.5: Complete d-state degeneracy removal in HfO$_2$ as seen in the 1$^{st}$ derivative XAS spectrum (data from Lucovsky, et al. [10]).
**Figure G.6:** The evolution of the type I nano-crystalline to type II nano-crystalline transition as observed by XAS in HfO$_2$ on Ge(100).
Figure G.7: The conduction band defect states in a thin film of HfO$_2$ on Ge(100) as detected by XAS (figure from Seo [11]).
References


Appendix H

Spectroscopic Ellipsometry (SE)

H.1 – The Theory and Experimental Set-Up for SE*

Spectroscopic ellipsometry (SE) is a non-destructive optical measurement technique which relies on changes in the state of polarization of an incident beam of light when it interacts with a thin film. The fundamentals regarding the polarization of light are discussed in great detail in any text on optics, such as Born and Wolf [2] or Hecht [3], and will not be treated in depth here. Essentially, the polarization of light can be broken into two types: s- and p-polarizations. An s-polarized beam has its electric field oscillating in the plane defined by the surface of the film, and p-polarized light has its electric field oscillating in the plane of incidence, which is perpendicular to the plane defined by the film’s surface. In general, an incident beam will be elliptically polarized, having both s- and p-polarized components.

An electromagnetic wave traveling in the \( \hat{z} \) direction can be described by the following:

\[
\begin{bmatrix}
E_x \\
E_y
\end{bmatrix}
= 
\begin{bmatrix}
E_{x0} e^{i(kz-\alpha_x)} \\
E_{y0} e^{i(kz-\alpha_y)}
\end{bmatrix}
= 
\begin{bmatrix}
E_{x0} e^{i\delta_x} \\
E_{y0} e^{i\delta_y}
\end{bmatrix} e^{i(kz-\alpha)}
\]  
(H.1)

* The information contained in this section can be found in any textbook on spectroscopic ellipsometry. This treatment uses information which comes from Fujiwara’s Spectroscopic Ellipsometry: Principles and Applications [1].
In general, such a wave will be elliptically polarized, and this ellipse of polarization can be described in a multitude of coordinate systems. In the \((\psi, \Delta)\) description, the coordinates are defined as follows:

\[
\begin{align*}
E_{x0} &= \sin \psi \\
E_{y0} &= \cos \psi \\
\Delta &= \delta_x - \delta_y
\end{align*}
\]  

(H.2)

Since the phase angles are relative quantities, the electric field intensities can be written as follows without loss of generality:

\[
\begin{align*}
E_{x0} &= \sin \psi e^{i(\delta_x - \delta_y)} \\
E_{y0} &= \cos \psi
\end{align*}
\]  

(H.3)

The ratio of p- to s-polarized light is defined as \(\rho\). In a right-handed coordinate system, the p-polarized field corresponds to \(E_x\), and the s-polarized field corresponds to \(E_y\). From equation H.1, this ratio is given by:

\[
\rho = \frac{E_x}{E_y} = \frac{E_{x0}e^{i(\delta_x - \delta_y)}}{E_{y0}}
\]

\[
\Rightarrow \rho = \tan \psi e^{i\Delta}
\]  

(H.4)
There are several different experimental set-ups for doing SE, but the method used for measurements conducted in these studies was so-called rotating-compensator ellipsometry (RCE), which was done in reflection. The experimental set-up is depicted in figure H.1. In this scenario, the polarization of the light is modified as it travels through this system in the following way:

\[
\begin{bmatrix}
E_x \\
E_y
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\cos \phi & \sin \phi \\
-\sin \phi & \cos \phi
\end{bmatrix}
\begin{bmatrix}
\cos \gamma & -\sin \gamma \\
\sin \gamma & \cos \gamma
\end{bmatrix}
\begin{bmatrix}
1 & 0 \\
0 & e^{-i\delta}
\end{bmatrix}
\begin{bmatrix}
\cos \gamma & \sin \gamma \\
-\sin \gamma & \cos \gamma
\end{bmatrix}
\times
\begin{bmatrix}
\sin \psi & e^{i\Delta} \\
0 & \cos \psi
\end{bmatrix}
\begin{bmatrix}
\cos \xi & -\sin \xi \\
\sin \xi & \cos \xi
\end{bmatrix}
\begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
E_x \\
E_y
\end{bmatrix}
\]

(H.5)

in which \(\delta\) is the phase shift of the slow axis due to the compensator, \(\xi\) is the polarization angle due to the initial polarizer, \(\gamma\) is the angle of the rotating compensator, and \(\phi\) is the angle of the analyzer. These are the Jones matrices which describe the polarization effects due to these optical elements [1, 2]. In reality, the compensator is rotating with an angular speed of \(\omega\) such that \(\gamma = \omega t\). Carrying out this matrix multiplication, the following result is obtained for the intensity of the detected light [4]:

\[
I = I_0 \left( \alpha_0 + \alpha_{2c} \cos 2\omega t + \alpha_{2s} \cos 2\omega t + \alpha_{4c} \cos 4\omega t + \alpha_{4s} \cos 4\omega t \right)
\]

(H.6)

in which \(I_0\alpha_0\) is the average incident light irradiance, and the remaining variables are defined as follows:
\[ \alpha_0 = \frac{1}{2} (1 + \cos \delta) \left( \cos 2\varphi \cos 2\xi - \cos 2\xi \cos 2\varphi + \sin 2\varphi \sin 2\xi \sin 2\psi \cos \Delta \right) - \cos 2\varphi \cos 2\psi + 1 \]

\[ \alpha_{2c} = -\sin 2\varphi \sin 2\xi \sin \delta \sin 2\psi \sin \Delta \]

\[ \alpha_{2s} = \sin 2\varphi \cos 2\xi \sin \delta \sin 2\psi \sin \Delta \]

\[ \alpha_{4c} = \frac{1}{2} (1 - \cos \delta) \left( \cos 2\varphi \cos 2\xi - \cos 2\xi \cos 2\varphi - \sin 2\varphi \sin 2\xi \sin 2\psi \cos \Delta \right) \]

\[ \alpha_{4s} = \frac{1}{2} (1 - \cos \delta) \left( \cos 2\varphi \sin 2\xi - \sin 2\xi \cos 2\varphi \sin 2\psi \cos \Delta \right) \]

(H.7)

By solving for \( \psi \) and \( \Delta \), the final result becomes [4]:

\[ \tan 2\psi = -\frac{\left[ \left( \alpha_{2c}^2 + \alpha_{2s}^2 \right) \left( 1 - \cos \delta \right)^2 + 4 \left( \alpha_{4s} \cos 2\xi - \alpha_{4c} \sin 2\xi \right) \right]^{\frac{1}{2}}}{2 \left( \alpha_{4c} \cos 2\xi + \alpha_{4s} \sin 2\xi \right)} \]  

(H.8)

\[ \tan \Delta = \left( \frac{1 - \cos \delta}{2 \sin \delta} \right) \frac{\alpha_{2s} \sin 2\xi - \alpha_{2c} \cos 2\xi}{\alpha_{4c} \sin 2\xi - \alpha_{4s} \cos 2\xi} \]  

(H.9)

In RCE, the \( \alpha_n \) are determined from Fourier analysis of the measured light’s intensity in relation to equations H.6 – H.7. From these quantities, the polarization parameters, \( \psi \) and

281
$\Delta$, are calculated from equations H.8 – H.9. The quantity, $\rho$, then can be determined from equation H.4.

These quantities do nothing to describe the film without the aid of a suitable optical model. The simplest model is a two-phase model, consisting of a sample/air interface. Analyses made using this model yield so-called pseudo-dielectric functions, $\langle \varepsilon \rangle$, given by the following:

$$
\langle \varepsilon \rangle = \sin^2 \theta \left[1 + \tan^2 \theta \left(\frac{1-\rho}{1+\rho}\right)^2\right]
$$

(H.12)

In this equation, $\theta_i$, is the angle of incidence of the light beam on the sample. More complicated 3- and 4-phase models can be derived for particular samples with a known structure, yielding overlayer thicknesses and real dielectric functions.

So-called effective medium theories can be used to describe non-homogeneous films, which contain inclusions. A schematic representation of such a scenario is depicted in figure H.2. The most useful of these models when the volume of the inclusions is large in comparison to the encapsulating film is the Bruggeman model, which is described by the following equation [5]:

$$
\sum_k f_k \frac{\varepsilon_k - \varepsilon}{\varepsilon_k + 2\varepsilon} = 0
$$

(H.13)
In equation H.13, $f_k$ represents the volume fraction of inclusion $k$, and $\varepsilon_k$ is the dielectric constant of inclusion $k$. By solving this equation, $\varepsilon$ for the sample can be determined. In particular, this technique is very useful in determining the composition of alloys.

Despite the model used to describe a sample, the basic process for arriving at the best solution is the same. A flow chart representing this process is given in figure H.3. First, an optical model is defined for a given sample, and a dielectric function is derived from this model. After making the SE measurement, $\psi$ and $\Delta$ are extracted as described above, and the error between the theoretical model and the measurement is computed. The extraction of $\psi$ and $\Delta$ is repeated until the $\psi$ and $\Delta$ which minimize this error are attained. Once this is done, the physical quantities which describe the sample are extracted and judged by the experimenter. If these quantities are reasonable, the analysis is done; if not, a new optical model must be derived that may describe the sample better, and the entire process is repeated until a good fit is obtained. Clearly, SE is quite model-dependent; however, the measurement itself is incredibly sensitive, and a good analysis can provide very accurate information about a wide variety of samples.

H.2 – SE as it Relates to Transition Metal Oxides and Related Alloys

Like XAS spectra, the dielectric functions of transition metal oxides can provide valuable information with regard to the conduction band and, in particular, with regard to sub-band edge defect levels. Because the transitions involved in SE are valence band to conduction band transitions, rather than core level to conduction band transitions as in XAS, the transition probabilities and matrix elements are different. As a result, SE is
much more sensitive to absorption by defect states, which shows up in the imaginary part of the dielectric function (typically denoted as $\varepsilon_2$).

An example of the absorption constant as determined by SE for a HfO$_2$ sample is given in figure H.4. The defect states clearly are detected in this measurement below the conduction band absorption edge. In this case, the defects are due to oxygen vacancies. When oxygen atoms are removed from a transition metal atom, the d states which normally hybridize with oxygen 2p states revert to their atomic d-like character and energy. This is shown schematically by the MO diagram in figure H.5, which shows how oxygen vacancies are responsible for the sub-band edge defect levels.

Again, SE is very sensitive to absorption by such defect states, and it can provide much information about such states. Not only this, but by studying absorption at the conduction band edge, much can be gleaned with respect to the nature of conduction band edges in transition metal oxides and their related alloys as well [6, 7]. Finally, because of the nature of SE measurements, physical quantities such as the dielectric constant, absorption coefficient, index of refraction, alloy composition, surface roughness, and physical thickness of samples can be determined by this technique [1].
Figure H.1: The experimental set-up for conducting rotating-compensator ellipsometry. The light source is on the left. Variable-frequency incoming light is passed through a polarizer, reflected by the sample, and passed through a rotating compensator and an analyzer. The final intensity of the beam is measured by the detector.
Figure H.2: A non-homogeneous thin film which can be described by an effective medium theory.
Figure H.3: A flow chart representing the method for solving for the dielectric function of a sample measured by SE.
Figure H.4: An example of the absorption constant for a HfO₂ sample as measured by SE. The defect states clearly are identified in this measurement (figure taken from Takeuci, et al. [5], © 2004 American Vacuum Society).
Figure H.5: A representative MO diagram for a transition metal oxide, such as TiO$_2$, with octahedrally coordinated transition metal atoms and oxygen vacancies. The oxygen vacancies are shown to be within the bandgap at the atomic d state energy. These are the states responsible for the sub-band edge defect levels observed in figure H.4.
References


Appendix I

Medium Energy Ion Scattering (MEIS)

I.1 – The Theory and Experimental Set-Up for MEIS

Medium energy ion scattering (MEIS) is a chemically sensitive technique which can be used to profile the composition of very thin films as a function of depth into the sample. In this experiment, a collimated beam of light ions (typically H⁺ or He⁺) with a well-defined energy (usually on the order of 100 keV) is directed at a sample, and the energies of the backscattered ions are measured. By using sophisticated theories of energy loss and atomic scattering, the distribution of energies in this spectrum can be used to identify not only which isotopes are present near the sample’s surface, but to measure their concentrations as well [1].

Before discussing depth profiling by MEIS, a treatment of the basic scattering processes is required. The basic ion scattering process is depicted in figure I.1. An ion of mass, \( M_1 \), and energy, \( E_0 \), from the ion beam approaches the sample surface at an angle, \( \phi \), and collides with an atom of mass, \( M_2 \), near the surface of a sample, knocking it free from the lattice (i.e. sputtering it). The sputtered atom takes off with energy, \( E_2 \), at an angle, \( \theta_2 \), relative to the incident ion’s trajectory. The scattered ion changes its trajectory such that it travels at an angle, \( \theta_1 \), relative to its initial path with energy, \( E_1 \). The scattered ion is then detected by an ion energy analyzer. The ratio of the ion’s energy
after an elastic scattering event relative to its energy initially is given by the following [2]:

$$\frac{E_1}{E_0} = \frac{4M_1M_2}{(M_1 + M_2)^2} \cos^2 \theta_1$$  \hspace{1cm} (I.1)

While actual ion scattering into and out of the film is exceedingly more difficult than this simple picture, usually involving multiple scattering events, this equation does illustrate the basic relationship between the scattered ion and the atom(s) which scatter it. Since the mass of the scattering atom is unique to each isotope, the identity of the scattering atom can be identified by its effect on the ion energy. Depth profiling is carried out via the complex modeling of energy loss through various scattering processes within the sample.

For thin films, the substrate will also influence the MEIS spectrum, often burying important information beneath its signal. Two techniques are generally used in combination in order to minimize the effects of such a crystalline substrate. These are called channeling and blocking. Channeling [1, 3] employs substrate orientation-specific incident ion beam angles in order to direct the ions down atomic “channels.” The ions will still interact with the atoms comprising the overlayer, but the ions which penetrate beyond that will travel very far into the sample. By the time they scatter, only certain back-scattered ions can exit the sample – namely the ones that happen to travel down another atomic channel (called de-channeling) [3]. Figure I.2 shows what the substrate
may look like to a channeled ion. Because de-channeling can occur, another technique, called blocking, is used [1, 3]. In this process, the detector is placed at an angle relative to the sample such that the de-channeled ions are obscured. This process is depicted in figure I.3. When the detector is placed at the appropriate angle, most of the ions scattered within the substrate are blocked by other atoms in the sample and are not detected [1, 3]. When channeling and blocking are used together, channeling forces ions to travel deep into the substrate, and blocking prevents the scattered ions from reaching the detector.

This set-up is fairly basic conceptually; however, the models required to analyze the back-scattered ions are anything but simple (e.g. see Harrison [4] and Gusev, et al. [5]), and the research equipment required is very expensive. For example, one experimental set-up is shown in figure I.4. Among the required elements are a linear accelerator, ion pumps and high vacuum equipment, beam-guiding magnets and deflection plates, and a position-sensitive microchannel plate analyzer. Despite these setbacks, MEIS has proven to be a valuable technique for high-resolution (approximately 3 Å) depth profiling in thin films [1, 5, 6]. An example of such a depth profile is shown in figure I.5. In this figure, a very thin layer of SiO₂ deposited on top of a nitrided crystalline Si substrate was profiled by MEIS. The resolution and sensitivity clearly are quite high, which is part of the reason for which this technique is far superior to others, such as secondary ion mass spectroscopy (SIMS), for profiling very thin films.
**Figure I.1:** A schematic representation of the basic ion scattering process in MEIS. In this picture, an ion from the incident ion beam impacts an atom near the surface of a thin film.
Figure I.2: A representative atomic channel for an ion used in channeling.
Figure I.3: The process of scattered ion blocking. The scattered ions which are not de-channeled essentially are blocked by other atoms in the sample as they try to exit it.
Figure I.4: The experimental set-up for MEIS (top) and ion detection (bottom). (The above figures are from Baumvol [1], © 1999 Elsevier Science B.V. All rights reserved.)
Figure I.5: An example of depth profiling for a nitrided thin film of SiO$_2$ on a crystalline Si substrate. The inset shows the depth profile for the ion energy spectrum shown in the main body of the graph. (The figure is taken from Gusev, et al. [6], © 2001 Elsevier Science B.V. All rights reserved.)
References


Appendix J

Capacitance-Voltage (C-V) Measurements

J.1 – The Theory of Capacitance in MOS Structures

Capacitance-voltage (C-V) measurements are useful in examining many properties of MOS devices, such as fixed charge trap density, mobile charge trap density, interface charge trap density, EOT, carrier density, and flat band voltage. Typically, the error between experimental data and a model with these or other related parameters is minimized in order to determine the relevant values. Schröeder [1], Nicollian and Brews [2], and Sze [3] all have written in-depth descriptions of MOS devices if more detail than provided here is desired. For simplicity, the following discussions here will refer to $n$-type MOS capacitors since all of the devices studied in this work were $n$-type. The extension to $p$-type devices can be made by simple analogy.

There are four basic regions possible in a C-V curve: (1) accumulation, (2) inversion, (3) depletion, and (4) deep depletion. The C-V characteristics for each of these are shown in figure J.1, and each one will be discussed, in turn.

In accumulation, a positive bias is applied across the gate, and the majority carriers (i.e. electrons) are attracted towards the semiconductor/gate interface. This situation is analogous to a parallel plate capacitor and (assuming the gate behaves properly) results in a high capacitance. In an ideal device, the majority carriers respond
almost immediately over the entire 0-1 MHz frequency range; therefore, accumulation ideally results despite the measurement frequency.

Inversion gets its name because this situation is exactly opposite to the scenario discussed for accumulation. A negative gate bias is applied to the gate, and minority carriers (i.e. holes) are attracted towards the semiconductor/gate interface. Unlike the majority carriers, however, minority carriers do not necessarily respond instantly over the 0 -1 MHz frequency range. If the holes can follow the measurement frequency, they will begin to pile up at the semiconductor/gate interface. This case is exactly like the parallel plate capacitor arising in accumulation, only with the bias reversed. Since the bias does not affect the capacitance in that situation, the inversion capacitance ideally is equivalent to the accumulation capacitance (again, assuming the gate works properly).

If the holes cannot follow the measurement frequency under a negative gate bias, the depletion case arises. The electrons are repelled from the semiconductor/gate interface by the negative gate bias, and the holes do not reach the interface because they cannot respond quickly enough to the measurement frequency. With no charge present in this region, a depletion layer forms, and the capacitance is very low. The width of this depletion layer saturates, however, and is determined by the response time of the holes. In some cases, the measurement frequency is too fast even for the generation of holes to occur, and the deep depletion scenario results. In this case, the depletion layer thickness never saturates and continues to widen as the gate bias is made more negative. This drives the capacitance increasingly lower.
The ideal low-frequency (i.e. inversion to accumulation) capacitance for the semiconductor can be described by the following equation \[1\]:

\[
C_{s,df} = \frac{\text{sgn}(\phi_s)}{2L_{Di}} \frac{\kappa_s \varepsilon_0}{n_i} \left[ \frac{n_D}{n_i} \left( 1 - e^{-\frac{q\phi_s}{kT}} \right) + \frac{n_i}{n_D} \left( e^{\frac{q\phi_s}{kT}} - 1 \right) \right]
\]

(J.1)

In equation J.1, \(\phi_s\) is the surface potential, \(\kappa_s\) is the semiconductor dielectric constant, \(\varepsilon_0\) is the dielectric permeability of free space, \(n_i\) is the intrinsic semiconductor carrier density, \(n_D\) is the donor density, \(k\) is Boltzmann’s constant, \(T\) is the ambient temperature, and \(L_{Di}\) is the intrinsic Debye length, given by \[1\]:

\[
L_{Di} = \sqrt{\frac{\kappa_s \varepsilon_0 kT}{2q^2 n_i}}
\]

(J.2)

in which \(q\) is the electronic charge. The function, \(F(\phi_s, n_D)\), is given by \[1\]:

\[
F(\phi_s, n_D) = \frac{n_D}{n_i} \left( e^{-\frac{q\phi_s}{kT}} + \frac{q\phi_s}{kT} - 1 \right) + \frac{n_i}{n_D} \left( e^{\frac{q\phi_s}{kT}} - \frac{q\phi_s}{kT} - 1 \right)
\]

(J.3)
The ideal high-frequency (i.e. depletion to accumulation) capacitance for the semiconductor can be described by the following equation [1]:

\[
C_{S,\delta f} = \text{sgn}(\phi_e) \frac{\kappa_s \epsilon_0}{2L_{Dd}} \left[ \frac{n_D}{n_i} \left( 1 - e^{\frac{q\phi}{kT}} \right) + \frac{n_i}{n_D (1 + \delta)} \left( e^{\frac{q\phi}{kT}} - 1 \right) \right] \quad \text{(J.4)}
\]

in which \( \delta \) is given by the following [1]:

\[
\delta = \left( \frac{e^{\frac{q\phi_e}{kT}} - \frac{q\phi_e}{kT}}{1 - e^{\frac{q\phi_e}{kT}}} \right) \left( \frac{1 - e^{\frac{q\phi}{kT}}}{e^{\frac{q\phi}{kT}} - 1} \right) \left( \frac{1 - e^{\frac{q\phi}{kT}}}{e^{\frac{q\phi}{kT}} - 1} \right) \quad \text{(J.5)}
\]

The ideal deep depletion to accumulation capacitance for the semiconductor can be described by the following equation [1]:

\[
\int_{\delta}^{\infty} \frac{e^{\frac{q\phi_e}{kT}} \left( 1 - e^{\frac{q\phi_e}{kT}} \right)}{2[F(\phi_e, n_D)]^2} \, d\phi
\]
\[
C_{\text{s,dd}} = \left\{ \begin{array}{ll}
\text{sgn}(\phi_s) \frac{\kappa_s \varepsilon_0}{2L_{Dk}} \left[ \frac{n_D}{n_i} \left( 1 - e^{-\frac{q\phi_s}{kT}} \right) + \frac{n_i}{n_D} \left( \frac{q\phi_s}{e^{\frac{q\phi_s}{kT}} - 1} \right) \right] F(\phi_s, n_D) ; & \phi_s \leq \frac{2kT}{q} \ln \left( \frac{n_D}{n_i} \right) \\
\frac{C_{\text{ox}}}{\sqrt{1 + 2(V_G - V_{FB})/V_0}} ; & \phi_s > \frac{2kT}{q} \ln \left( \frac{n_D}{n_i} \right) 
\end{array} \right.
\]

(J.6)

in which \( V_g \) is the gate voltage, \( V_{FB} \) is the flat band voltage, \( C_{\text{ox}} \) is the oxide capacitance, and \( V_0 \) is given by the following [1]:

\[
V_0 = \frac{q\kappa_s \varepsilon_0 N_D}{C_{\text{ox}}^2} 
\]

(J.7)

The oxide capacitance is given by [1]:

\[
C_{\text{ox}} = \frac{\kappa_{\text{ox}} \varepsilon_0 A}{t_{\text{ox}}} 
\]

(J.8)

in which \( t_{\text{ox}} \) is the oxide thickness, \( \kappa_{\text{ox}} \) is the oxide material’s dielectric constant, and \( A \) is the MOS device area. The gate voltage is related to all of this through the following relationship [1]:

304
\[ V_G = V_{FB} + \phi_s + \text{sgn}(\phi_s) \frac{kT\kappa_s}{q\kappa_{ox}L_{Di}} F(\phi_s, n_D) \quad (J.9) \]

Assuming that the oxide charges are distributed uniformly, the flat band voltage in equations J.6 and J.9 is given by the following expression [1]:

\[ V_{FB} = \phi_M - \phi_s - \frac{Q_{ox}}{C_{ox}} - \frac{Q_{it}(\phi_s)}{C_{ox}} \quad (J.10) \]

in which \( \phi_M \) is the contact metal’s work function, \( Q_{ox} \) is charge trapped in the oxide, and \( Q_{it} \) is the interface charge (which is a function of the surface potential). \( Q_{it} \) is given by [2]:

\[ Q_{it} = q \int_{E_{c},q} d\phi \left( f_0(\phi - \phi_s - \phi_B) D_{it}^a(\phi) + \left[ 1 - f_0(\phi - \phi_s - \phi_B) D_{it}^p(\phi) \right] \right) \quad (J.11) \]

in which \( E_c \) is the conduction band energy, \( E_v \) is the valence band energy, \( \phi_B \) is the gate barrier height, and the \( D_{it} \) terms are the acceptor and donor interface charge trap densities, and \( f_0 \) is the Fermi function given by [1]:

\[ f_0(\phi) = \frac{1}{n_i \frac{q}{N_D} e^{\frac{\phi}{kT}} + 1} \quad (J.12) \]
The interface capacitance is given by the following expression [2]:

\[
C_{it} = -q \int_{E_{it}}^{E_{it}} d\phi \phi \left( \phi - \phi_s - \phi_b \right) \left[ D_0' + D_{it}' \left( \phi \right) \right] \quad (J.13)
\]

Last, but not least, the total MOS capacitance is given by [2]:

\[
C = \frac{C_{ox} \left( C_s + C_{it} \right)}{C_{ox} + C_s + C_{it}} \quad (J.14)
\]

The MOS capacitance versus gate voltage can then be computed iteratively by using equations J.1 – J.14. Using this method, computer algorithms were written by the author in order to calculate the classical C-V characteristics of ideal MOS devices with interface and oxide charges. Figure J.1, for example, was computed in this manner. Figure J.2 also was computed by this program, and it illustrates the effects of oxide and interface trapped charges. Oxide trapped charges simply translate the curve along the \(V_G\) axis; whereas interface trapped charges both translate and stretch the curve along the \(V_G\) axis. Figure J.2 also shows how a curve with a very large number of both fixed and trapped charge densities can combine to yield a flat band voltage similar to the ideal case. The “stretchout,” however, is a tell-tale sign of interface charge, and fixed charge clearly is present to offset some of those effects.
There are also many other programs available [4-8]. Hauser’s CVC program [4], for example, showed very good agreement with the results obtained from the author’s code. Inevitably, different techniques are employed by all of these programs, and some differences are unavoidable, especially in the ultrathin regime. The pros and cons of these C-V modeling methods are discussed by Richter, et al. [9], and a comparison is shown in figure J.3. It should be noted that the author’s code, for instance, is classical and does not account for quantum mechanical effects. As a result, it overestimates the EOT for very thin films (see figure J.3). Moreover, it does not take into account series resistance effects. Again, this leads to additional EOT overestimation [10], as shown in figure J.4. The EOT typically is determined from the 1 MHz C-V curve, which clearly is reduced in the high-frequency regime when appreciable series resistance is present. Hence, the EOT values computed by the author in chapter 4 are likely to be larger than in reality; however, the accumulation capacitance is the important value from a device perspective, regardless of the true gate thickness or EOT.

In terms of experimental C-V curves, there are a number of ways to determine interface trapped charge densities, \( D_{it} \) [1-3]. One is through modeling, such as with the programs described above (Terman method) [1, 2]. Another is by using the experimental conductance values, but the analysis is complicated, and the amount of data required is very large [1, 2]. The easiest and perhaps most common, is the high-low frequency method, which requires a high-frequency (approx. 1 MHz) C-V curve and a low-frequency (approx. 0.1 – 1 kHz) C-V curve. The interface trapped charge density is then given by the following expression [1]:
in which $C_{lf}$ and $C_{hf}$ are the low- and high-frequency C-V curves, respectively.

As a final note, it should be pointed out that the modeling described above is insufficient for examining mobile charge traps. Mobile charge traps show up in the C-V curve as hysteresis. As the gate voltage is swept in one direction, charges are trapped, translating the C-V curve in one direction. As the gate bias is swept back towards the initial value, the charges are de-trapped, but there is an extra activation energy required to do so. This translates the C-V curve back in the opposite direction. Hence, neither sweep direction represents the true (i.e. ideal) C-V characteristic for the device. The measurement of these parameters is difficult because the amount of hysteresis is increased by not only the number of traps, but by the de-trapping energy as well. In the end, it is hysteresis that diminishes device performance; therefore, examining the hysteresis, regardless of its root causes, still is important.

\[ D_n = \frac{C_{ox}}{q} \left( \frac{C_{lf}/C_{ox}}{1 - C_{lf}/C_{ox}} - \frac{C_{hf}/C_{ox}}{1 - C_{hf}/C_{ox}} \right) \] (J.15)
Figure J.1: Ideal C-V curves for each of the four possible regions (i.e. accumulation, inversion, depletion, and deep depletion). The C-V curves shown here were computed for an ideal $n$-type Si MOS structure with a substrate doping level of $1 \times 10^{17}$/cm$^3$, an area of $10 \mu$m x $10 \mu$m, and an EOT of 10 nm. The device has aluminum contacts.
Figure J.2: The effects of interface and oxide trapped charges on the classical C-V characteristics of a MOS capacitor. The C-V curves shown here were computed for an $n$-type Si MOS structure with a substrate doping level of $1\times10^{17}/\text{cm}^3$, an area of $10\mu\text{m} \times 10\mu\text{m}$, and an EOT of $10$ nm. The device has aluminum contacts.
Figure J.3: A comparison of the most commonly used capacitance modeling programs for calculating C-V curves of MOS devices in the ultrathin regime (from Richter, et al. [9], © 2001 IEEE).
Figure J.4: The effects of series resistance on the capacitance-voltage characteristics of a MOS device (after Yang and Hu [10]).
References


Appendix K

Current-Voltage (I-V) Measurements

K.1 – The Theory of Leakage Current in MOS Structures

Current-voltage measurements are vital to the understanding of gate integrity in MOS devices. The main function of the gate in MOS structures is to control carrier distributions within the semiconductor capacitively. If sufficient, however, the controlling gate bias actually can force charge to move through the gate, causing a drop in potential across the oxide and reducing its capacity to confine and modulate charge distributions within the semiconductor. Such current is called leakage current, and it may be caused by a variety of mechanisms depending on the gate material, thickness, and bias condition. A MOS device which leaks not only operates with reduced performance, but it causes resistive heating and, ultimately, leads to the failure of the device or circuit.

The six most common leakage mechanisms are shown in figure K.1 in terms of their corresponding band diagrams [1-3]. In the Si/SiO₂ material system, the two most common of these are direct tunneling and Fowler-Nordheim tunneling [2]. On the other hand, it is expected that Poole-Frenkel emission and trap-assisted tunneling play a much larger role in leakage through high-κ gate materials [3]. Thermionic emission and thermionic-field emission tend to play a role when devices are operated at elevated temperatures [1, 2]. By examining the I-V curves and comparing them to theoretical models for the various leakage mechanisms, the integrity of the insulator can be better
understood. As an example, figure K.2 shows the leakage current in several Si/SiO₂ MOS devices, which is viewed in terms of direct tunneling and Fowler-Nordheim tunneling. Note that the change in shape between the 5.6 nm and 3.5 nm curves indicates the onset of direct tunneling somewhere around $t_{ox} \leq 3.5$ nm.

Specifically, the various leakage mechanisms can be described using analytical expressions. Direct tunneling is described by the following equation [2]:

$$J_{direct} = \frac{q^2 (m_e/m_{ox})}{8\pi \hbar \Phi_B} E_{ox}^2 \exp \left( \frac{-8\pi \sqrt{2m_{ox}\Phi_B^3}}{3\hbar E_{ox}} \left( 1 - \left( 1 - \frac{qE_{ox}}{\Phi_B t_{ox}} \right)^{3/2} \right) \right)$$

$$\sim A(V_G - \Delta V)^2 \exp \left( \frac{-B \left( 1 - \left( 1 - C(V_G - \Delta V) \right)^{3/2} \right)}{(V_G - \Delta V)} \right)$$

(K.1)

in which $E_{ox}$ is the oxide electric field, $t_{ox}$ is the oxide thickness, $m_e$ is the electron mass, $m_{ox}$ is the effective electron mass in the insulator, $\hbar$ is Planck’s constant, $\Phi_B$ is the potential barrier between the gate and the semiconductor, $q$ is the electronic charge, $V_G$ is the gate bias, and $A$, $B$, $C$, and $\Delta V$ are constants. Similarly, Fowler-Nordheim tunneling is given by the following [2]:
Thermionic emission is given by the following equation [1]:

\[
J_{F-N} = \frac{q^3 (m_e/m_\alpha) E_{ox}^2}{8\pi \hbar \Phi_B} \exp \left( \frac{-8\pi \sqrt{2m_\alpha \Phi_B^3}}{3q\hbar E_{ox}} \right) \sim A(V_G - \Delta V)^2 \exp \left( -\frac{B}{(V_G - \Delta V)} \right)
\]

(K.2)

in which \( k \) is Boltzmann’s constant, \( T \) is the temperature, \( \varepsilon_0 \) is the dielectric permittivity of free space, and \( \kappa_{ox} \) is the dielectric constant of the insulator. Because thermionic-field emission is just a combination of thermionic emission and Fowler-Nordheim tunneling, it is represented by the sum of equations K.2 and K.3. Frenkel-Poole emission is governed by the following equation [1]:

\[
J_{\text{Thermionic}} = AT^2 \exp \left( -\frac{q}{kT} \left( \Phi_B - \frac{qE_{ox}}{4\pi \kappa_{ox} \varepsilon_0} \right) \right) \sim AT^2 \exp \left( -\frac{B \sqrt{V_G - \Delta V}}{T} - C \right)
\]

(K.3)

Since trap-assisted tunneling is a combination of Frenkel-Poole emission and direct tunneling, it will likely be governed by a combination of equations K.1 and K.4.
Again, by analyzing the I-V curves obtained from a particular MOS device in the context of equations K.1 – K.4, one can understand the dominant leakage mechanisms present. In the end, a small leakage current during the application of a large gate bias is desirable in a device, regardless of the particular leakage mechanism. Understanding the mechanism and its exact physical origin, however, can provide insight into how (or even if) the integrity of the gate material can be improved.
Figure K.1: Band diagrams depicting six of the most common gate leakage mechanisms. The flat band condition is shown as a reference in (a). The leakage mechanisms are: (b) direct tunneling, (c) Fowler-Nordheim tunneling, (d) thermionic emission, (e) thermionic-field emission, (f) trap-assisted tunneling, and (g) Poole-Frenkel emission.
Figure K.2: Direct tunneling and Fowler-Nordeim tunneling in Si/SiO₂ MOS devices with varying oxide thickness (data from Schuegraf and Hu [4]). The change in shape of the I-V characteristics shows that the onset of direct tunneling occurs at roughly $t_{ox} \leq 3.5$ nm.
References


