INVESTIGATION AND SUPPRESSION OF SEMICONDUCTOR–OXIDE RELATED DEFECT STATES
FROM SURFACE SCIENCE TO DEVICE TESTS

Jaakko Mäkelä
Abstract

Jaakko Mäkelä

Investigation and Suppression of Semiconductor–Oxide Related Defect States: From Surface Science to Device Tests

Many present challenges in semiconductor technology are related to utilizing solid structures with atomic scale dimensions and materials with higher charge carrier mobility and/or other readily controllable properties. These include many surface-related problems because the ratio of surface parts of devices to the whole material volume increases all the time in practical device structures. One of the major problems has been oxidation of semiconductor surfaces during the manufacturing of devices.

This PhD work deals with the surface and oxide interface properties of different III–V semiconductors induced by the oxidation, the study of which is imperative in realizing devices with desired characteristics. The general goal has been in finding answers to these problematic issues on atomic scale, and whether they can be resolved with simple parameter control of a thermal oxidation treatment. Much of the work leans on a previous novel finding of crystalline oxide phases on indium-containing III–V semiconductor (100) surfaces. Various aspects of applicability of such a structure in real semiconductor devices are considered in this work.

Common denominator in all of the experiments and studies is that the initial investigations were carried out in very controlled environment in ultrahigh-vacuum: detailed basics and initial characterizations were carried out with high resolution and precision surface science methods. In particular, this work has resulted in novel crystalline oxide phases observed on GaSb(100) and InSb(111)B semiconductor surfaces. They have been extensively discussed from an applied point of view as well as their fundamental characteristics, relating to their already previously studied counterpart, InSb(100). Furthermore, beneficial passivating characteristics of a stabilizing crystalline InO$_x$ interfacial layer beneath an Al$_2$O$_3$ and reasons behind such behavior are demonstrated for InGaAs IR detector device structure.
INVESTIGATION AND SUPPRESSION OF SEMICONDUCTOR–OXIDE RELATED DEFECT STATES

From Surface Science to Device Tests

Jaakko Mäkelä
University of Turku
Faculty of Science and Engineering
Department of Physics and Astronomy
Materials Research Laboratory

Supervised by
Professor Kalevi Kokko
Department of Physics and Astronomy
University of Turku
Turku, Finland

Docent Pekka Laukkanen
Department of Physics and Astronomy
University of Turku
Turku, Finland

Dr. Mikhail Kuzmin
Ioffe Physical-Technical Institute
St. Petersburg, Russian Federation

Docent Marko Punkkinen
Department of Physics and Astronomy
University of Turku
Turku, Finland

Reviewed by
Professor Anders Mikkelsen
Department of Physics,
Synchrotron Radiation Research
Lund University
Lund, Sweden

Professor Tapio Rantala
Department of Physics
Tampere University of Technology
Tampere, Finland

Opponent
Professor Mats Göthelid
Materials and Nano Physics
KTH Royal Institute of Technology
Stockholm, Sweden

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-7544-0 (PRINT)
ISBN 978-951-29-7545-7 (PDF)
ISSN 0082-7002 (Print)
ISSN 2343-3175 (Online)

Grano Oy - Turku, Finland 2019
“One shouldn’t work on semiconductors, that is a filthy mess; who knows whether any semiconductors exist.”

“God made the bulk; surfaces were invented by the devil.”

— Wolfgang Pauli
Abstract

Jaakko Mäkelä

Investigation and Suppression of Semiconductor–Oxide Related Defect States: From Surface Science to Device Tests

Many present challenges in semiconductor technology are related to utilizing solid structures with atomic scale dimensions and materials with higher charge carrier mobility and/or other readily controllable properties. These include many surface-related problems because the ratio of surface parts of devices to the whole material volume increases all the time in practical device structures. One of the major problems has been oxidation of semiconductor surfaces during the manufacturing of devices.

This PhD work deals with the surface and oxide interface properties of different III–V semiconductors induced by the oxidation, the study of which is imperative in realizing devices with desired characteristics. The general goal has been in finding answers to these problematic issues on atomic scale, and whether they can be resolved with simple parameter control of a thermal oxidation treatment. Much of the work leans on a previous novel finding of crystalline oxide phases on indium-containing III–V semiconductor (100) surfaces. Various aspects of applicability of such a structure in real semiconductor devices are considered in this work.

Common denominator in all of the experiments and studies is that the initial investigations were carried out in very controlled environment in ultrahigh-vacuum: detailed basics and initial characterizations were carried out with high resolution and precision surface science methods. In particular, this work has resulted in novel crystalline oxide phases observed on GaSb(100) and InSb(111)B semiconductor surfaces. They have been extensively discussed from an applied point of view as well as their fundamental characteristics, relating to their already previously studied counterpart, InSb(100). Furthermore, beneficial passivating characteristics of a stabilizing crystalline InO$_x$ interfacial layer beneath an Al$_2$O$_3$ and reasons behind such behavior are demonstrated for InGaAs IR detector device structure.
This thesis provides background of semiconductors, their surfaces, interfaces, and semiconductor technology as appropriate, research methods utilized, and briefly summarizes the findings of the work.
Tiivistelmä

Jaakko Mäkelä

Puolijohde-oksidi-liitoksen kidevikatilojen tutkiminen ja vähentäminen: pintatieteestä laitetesteihin


Tämä väitöskirjatyö käsittelee hapetuksen aiheuttamia pintoihin ja raja-pintaominaisuksia III–V-puolijohdeilla, joiden tutkiminen on välttämätöntä haluttujen ominaisuuksien toteuttamiseksi. Tavoitteena on ollut löytää vastauksia hapetuksen aiheuttamiin ongelmiiin atomitasolla ja se, voidaanko niihin löytää ratkaisuja korotetun lämpötilan hapetuksen yksinkertaisilla parametrisäädöillä. Suuri osa työstä nojaa aiempaan uudenlaiseen löytöön kiteisistä oksidifaaseista indiumia sisältävillä (100) III–V-puolijohdepinnoilla. Tässä työssä on käsitelty useita näkökulmia kyseisenlaisen rakenteen sovellettavuudesta todellisissa puolijohdelaitteissa.

Työn kaikkien tutkimusten yhteisenä tekijänä on ollut erittäin hyvin kontrolloitu koeypäräistä ultrasuuryhjäössä: yksityiskohtaiset perusteet ja lähtökohtaiset ominaisuuudet selvitettiin korkean resolution ja tarkkuuden pintatieteestä menetelmin. Erityisesti tämä työ on tuottanut uudet havainnot kiteisistä oksidipinnoista GaSb(100) ja InSb(111)B-puolijohdepinnoilla. Niitä on käsitelty laajasti sovelletustaka näkökulmasta, ja toisaalta perustavanlautuisista lähtökohdista verraten niiden jo aiemmin tutkittuun InSb(100)-vastineeseen. Lisäksi stabiloinvan, kiteisen InOₓ-päälläkerroksen hyödyllisiä passivoivia ominaisuuksia Al₂O₃-kerroksen alla ja syitä niiden takana on osoittettu InGaAs infrapunadetektori-laiterakenteelle.
Tämä väitöskirja esittelee työn kannalta oleellista taustaa puolijohteis-
ta, niiden pinnoista ja rajapinnoista, puolijohdeteknologiasta, sekä hyödyn-
netystä tutkimusmenetelmistä, ja lyhyesti kokoaa yhteen työssä tehtyjen
tutkimusten tulokset.
Acknowledgements

I gratefully acknowledge the doctoral programme in Physical and Chemical Sciences (PCS) and University of Turku Graduate School (UTUGS) for funding that enabled working as a full time PhD student. Financial support from the Academy of Finland (project no. 296469) that enabled carrying out the work in the last few months until graduation is also gratefully acknowledged. Personnel at MAX-lab, Lund University, Sweden and BESSY II, Helmholtz-Zentrum, Berlin, Germany are acknowledged for enabling our high quality synchrotron radiation photoelectron spectroscopy results at the beamtimes in these facilities.

I am most thankful for the warm and easy working environment of the Materials Physics group at Materials Research lab at the Department of Physics and Astronomy. Here, it has never been difficult to bring forth my opinions or visions, which is the very first prerequisite for any innovative research and also the seed for fruitful discussions, correcting misunderstandings, and learning something new. Doc. Pekka Laukkanen has been my main supervisor and beside me the biggest contributor to this work. His optimism and enthusiasm are very inspiring, and make us a good team of opposites due to realism and sceptism of yours truly. Prof. Kalevi Kokko has been the director of this work as the head of the group and nowadays the Department. It is always a pleasure to discuss with him about anything and he always finds time for students even despite the increasingly busy schedules. Some of the most profound discussions have been with my second supervisor from Ioffe Physical-Technical Institute in Russia, Doc. Mikhail Kuzmin, most skilled and analytical surface science specialist. Doc. Marko Punkkinen has also supervised the thesis and especially lately we have had good talks about interfaces between experimental and computational research and between oxides and semiconductors also.

Most thanks for the warm atmosphere go to the past and present personnel of our group with whom I have had the pleasure to work with, who have been friendly and provided peer-support during this work: Dr. Muhammad Yasir, M.Sc. Marjukka Tuominen, Dr. Johnny Dahl, M.Sc. Juha-Pekka Lehtio, M.Sc. Zahra Jahanshah Rad, Dr. Henrik Levämäki, Dr. Jouko Lång, Dr. Matti Ropo, M.Sc. Veikko Tuominen, M.Sc. Antti Lahti, and
B.Sc. Esa Vuorinen. Thanks also to Dr. Sari Granroth from the upstairs division of our laboratory, for many good discussions, and to the head of Materials Research laboratory, Prof. Edwin Kukk.

I also wish to thank the colleagues during my time working at Tampere University of Technology. The semiconductor research at Optoelectronics Research Centre is top of the line, certainly owing much to the nicest work atmosphere one can find. I would like to specifically thank Prof. Mircea Guina and Dr. Jukka Viheriälä for having me as a part of the semiconductor processing team, M.Sc. Antti Aho, Mervi Koskinen and Mariia Bister for their guidance with processing, nanoimprint lithography, lasers, etc., and Dr. Ville Polojarvi for leading me into many secrets of processing and solar cells. My time at the Surface Science laboratory brought me into this field, for which I feel very grateful to Prof. Mika Valden and other personnel at the laboratory. It was easy to start building my expertise there together with M.Sc. Markku Hannula.

It was always easy to uphold one’s mental health because of the best times, dinners, beers, and events with friends from Kyläilykanava. Biggest thanks go to my family and parents, who have always supported me and seen me as more than myself, giving me something to aim for. My grandfather, Veli Tiainen, deserves big thanks for guiding me into mathematics and natural sciences.

Still, all of this would be pointless without my beloved wife, Noora, and our dear children, Touko and Heino, who bring happiness to my life every single day. Thank you.

Turku, January 21, 2019

Jaakko Mäkelä
List of Publications


The author is principal author in publications I-IV that comprise this work. Comments on the author’s contribution to each publication:

I  Measurements carried out jointly. Responsible for data analysis and interpretation.

II  Planning and carrying out measurements jointly. Responsible for data analysis and interpretation.

III  Planning and carrying out measurements jointly. In charge of device design and processing. Responsible for analysis of the results.

IV  Executed designing, integration and building of a UHV sample preparation chamber to an existing STM system. In charge of planning and carrying out measurements and analysis.
The following publications were contributed to alongside PhD work but are not included in the thesis:


# Contents

Abstract v

Tiivistelmä vii

Acknowledgements ix

List of Publications xi

Glossaries xix

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>xix</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Symbols</td>
<td>xxii</td>
</tr>
<tr>
<td>List of Physical Constants</td>
<td>xxiv</td>
</tr>
</tbody>
</table>

1 Introduction 1

2 Semiconductors, Surfaces and Interfaces 5

| 2.1 Fundamental Properties of Semiconductors | 5 |
| 2.1.1 Bonding and Energy Levels | 5 |
| 2.1.2 Characteristic Semiconductor Properties | 10 |
| 2.2 Semiconductor Surfaces | 14 |
| 2.2.1 Atomic Structure | 14 |
| 2.2.2 Surface and Interface States | 16 |
| 2.3 Oxide–Semiconductor Stack | 19 |
| 2.3.1 Thin Oxide Films | 20 |
| 2.3.2 Oxide–Semiconductor Interface | 21 |
| 2.3.3 Formation of Crystalline Oxides on III–V’s | 24 |
| Thermal Oxidation of III–V in O₂ Ambient | 24 |
| III–V Crystalline Oxide Properties | 26 |

3 Experimental Methods 27

| 3.1 Ultrahigh Vacuum | 27 |
| 3.1.1 Vacuum Ranges and Exposure Magnitudes | 27 |
| 3.1.2 Pumping | 28 |
| 3.1.3 Pressure Measurement | 29 |
3.2 Sample Treatment Facilities ......................................... 29
  3.2.1 Ion Bombardment, Sputtering ................................. 29
  3.2.2 Sample Annealing ............................................. 30
  3.2.3 Deposition ................................................... 31
  3.2.4 Gas Exposure .................................................. 31
  3.2.5 Atomic Layer Deposition ..................................... 33
  3.2.6 Surface Science Setup ....................................... 36
3.3 Photoelectron Spectroscopy ..................................... 38
  3.3.1 Laboratory Sources ......................................... 43
  3.3.2 Synchrotron Sources ........................................ 44
  3.3.3 PES Data Analysis .......................................... 46
3.4 Low Energy Electron Diffraction ................................ 47
3.5 Scanning Tunneling Microscopy .................................. 50
  3.5.1 Scanning Tunneling Spectroscopy ............................ 52
  3.5.2 Low Temperature STM ....................................... 55
3.6 Photoluminescence ................................................ 56
4  Semiconductor Devices .............................................. 59
  4.1 pn-Junction Diodes ............................................. 59
  4.2 Photodiode Devices ............................................. 60
  4.3 Transistor Devices ............................................... 64
    4.3.1 Metal Oxide Semiconductor Field Effect Transistor .... 64
    4.3.2 High Electron Mobility Transistor ......................... 68
  4.4 Device Fabrication .............................................. 70
    4.4.1 Molecular Beam Epitaxy ................................... 71
    4.4.2 Device Processing ......................................... 71
5  Summary and Conclusions .......................................... 75
  5.1 Thermally Assisted Oxidation of GaSb(100) .................... 75
  5.2 Crystalline Oxide on InGaAs(100) Photodetector .............. 77
  5.3 Crystalline Oxide on InSb(111)B ................................ 78
  5.4 Conclusions .................................................. 80

Bibliography .......................................................... 83

Original Research Articles ........................................... 105
## Glossaries

### Abbreviations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-DEG</td>
<td>two-dimensional electron gas</td>
</tr>
<tr>
<td>ALD</td>
<td>atomic layer deposition</td>
</tr>
<tr>
<td>ALE</td>
<td>atomic layer epitaxy</td>
</tr>
<tr>
<td>AR</td>
<td>anti-reflection</td>
</tr>
<tr>
<td>ARPES</td>
<td>angle-resolved photoelectron spectroscopy</td>
</tr>
<tr>
<td>B</td>
<td>bulk</td>
</tr>
<tr>
<td>BSF</td>
<td>back-surface field</td>
</tr>
<tr>
<td>BTBT</td>
<td>band-to-band tunneling</td>
</tr>
<tr>
<td>CB</td>
<td>conduction band</td>
</tr>
<tr>
<td>CBM</td>
<td>conduction band minimum</td>
</tr>
<tr>
<td>CITS</td>
<td>current imaging tunneling spectroscopy</td>
</tr>
<tr>
<td>CLS</td>
<td>core-level shift</td>
</tr>
<tr>
<td>CMOS</td>
<td>complementary MOS</td>
</tr>
<tr>
<td>D</td>
<td>drain</td>
</tr>
<tr>
<td>DFT</td>
<td>density-functional theory</td>
</tr>
<tr>
<td>DIBL</td>
<td>drain-induced barrier lowering</td>
</tr>
<tr>
<td>DIGS</td>
<td>disorder-induced gap state</td>
</tr>
<tr>
<td>DOS</td>
<td>density of states</td>
</tr>
<tr>
<td>e-beam</td>
<td>electron beam</td>
</tr>
<tr>
<td>EAL</td>
<td>effective attenuation length</td>
</tr>
<tr>
<td>FCC</td>
<td>face-centered cubic</td>
</tr>
<tr>
<td>FET</td>
<td>field effect transistor</td>
</tr>
<tr>
<td>FinFET</td>
<td>fin-type field effect transistor</td>
</tr>
<tr>
<td>Notation</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>FLP</td>
<td>Fermi-level pinning</td>
</tr>
<tr>
<td>FWHM</td>
<td>full width at half-maximum</td>
</tr>
<tr>
<td>G</td>
<td>gate</td>
</tr>
<tr>
<td>GAA</td>
<td>gate-all-around</td>
</tr>
<tr>
<td>GAA-FET</td>
<td>gate-all-around field effect transistor</td>
</tr>
<tr>
<td>GGO</td>
<td>Ga$_2$O$_3$ (Gd$_2$O$_3$)</td>
</tr>
<tr>
<td>HAXPES</td>
<td>hard x-ray photoelectron spectroscopy</td>
</tr>
<tr>
<td>HEMT</td>
<td>high electron mobility transistor</td>
</tr>
<tr>
<td>HR</td>
<td>high reflection</td>
</tr>
<tr>
<td>HSA</td>
<td>hemispherical analyzer</td>
</tr>
<tr>
<td>HV</td>
<td>high-vacuum</td>
</tr>
<tr>
<td>IBA</td>
<td>ion bombardment and annealing</td>
</tr>
<tr>
<td>IC</td>
<td>integrated circuit</td>
</tr>
<tr>
<td>ICP</td>
<td>inductively coupled plasma</td>
</tr>
<tr>
<td>IMFP</td>
<td>inelastic mean free path</td>
</tr>
<tr>
<td>IR</td>
<td>infrared</td>
</tr>
<tr>
<td>LDOS</td>
<td>local density of states</td>
</tr>
<tr>
<td>LED</td>
<td>light emitting diode</td>
</tr>
<tr>
<td>LEED</td>
<td>low energy electron diffraction</td>
</tr>
<tr>
<td>LHe</td>
<td>liquid helium</td>
</tr>
<tr>
<td>LIDAR</td>
<td>light detection and ranging</td>
</tr>
<tr>
<td>LN2</td>
<td>liquid nitrogen</td>
</tr>
<tr>
<td>LT-STM</td>
<td>low temperature scanning tunneling microscopy</td>
</tr>
<tr>
<td>MBE</td>
<td>molecular beam epitaxy</td>
</tr>
<tr>
<td>mHEMT</td>
<td>metamorphic HEMT</td>
</tr>
<tr>
<td>MIGS</td>
<td>metal-induced gap state</td>
</tr>
<tr>
<td>ML</td>
<td>monolayer</td>
</tr>
<tr>
<td>MOS</td>
<td>metal oxide semiconductor</td>
</tr>
<tr>
<td>MOSFET</td>
<td>metal oxide semiconductor field effect transistor</td>
</tr>
<tr>
<td>NEP</td>
<td>noise equivalent power</td>
</tr>
<tr>
<td>PBN</td>
<td>pyrolytic boron nitride</td>
</tr>
<tr>
<td>PDA</td>
<td>post-deposition annealing</td>
</tr>
<tr>
<td>Notation</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>PECVD</td>
<td>plasma-enhanced chemical vapor deposition</td>
</tr>
<tr>
<td>PEEM</td>
<td>photoemission electron microscopy</td>
</tr>
<tr>
<td>PES</td>
<td>photoelectron spectroscopy</td>
</tr>
<tr>
<td>pHEMT</td>
<td>pseudomorphic HEMT</td>
</tr>
<tr>
<td>PIC</td>
<td>photonic integrated circuit</td>
</tr>
<tr>
<td>PL</td>
<td>photoluminescence</td>
</tr>
<tr>
<td>PVD</td>
<td>physical vapor deposition</td>
</tr>
<tr>
<td>QD</td>
<td>quantum dot</td>
</tr>
<tr>
<td>QW</td>
<td>quantum well</td>
</tr>
<tr>
<td>RF</td>
<td>radio frequency</td>
</tr>
<tr>
<td>RHEED</td>
<td>reflected high energy electron diffraction</td>
</tr>
<tr>
<td>RIE</td>
<td>reactive ion etching</td>
</tr>
<tr>
<td>RT</td>
<td>room temperature</td>
</tr>
<tr>
<td>RTA</td>
<td>rapid thermal annealing</td>
</tr>
<tr>
<td>S</td>
<td>source</td>
</tr>
<tr>
<td>SCE</td>
<td>short-channel effect</td>
</tr>
<tr>
<td>SCLS</td>
<td>surface core-level shift</td>
</tr>
<tr>
<td>SIBA</td>
<td>simultaneous ion bombardment and annealing</td>
</tr>
<tr>
<td>SIMS</td>
<td>secondary ion mass spectroscopy</td>
</tr>
<tr>
<td>SOI</td>
<td>silicon on insulator</td>
</tr>
<tr>
<td>SOS</td>
<td>spin-orbit splitting</td>
</tr>
<tr>
<td>SPEM</td>
<td>scanning photoelectron emission microscopy</td>
</tr>
<tr>
<td>SPM</td>
<td>scanning probe microscopy</td>
</tr>
<tr>
<td>SRH</td>
<td>Shockley-Read-Hall</td>
</tr>
<tr>
<td>SR-PES</td>
<td>synchrotron radiation photoelectron spectroscopy</td>
</tr>
<tr>
<td>SS</td>
<td>subthreshold swing</td>
</tr>
<tr>
<td>STM</td>
<td>scanning tunneling microscopy</td>
</tr>
<tr>
<td>STS</td>
<td>scanning tunneling spectroscopy</td>
</tr>
<tr>
<td>TDMAH</td>
<td>tetrakis(dimethylamido)hafnium(IV) $\text{[Hf(N(CH}_3\text{)_2}]_4}$</td>
</tr>
<tr>
<td>TDMAZ</td>
<td>tetrakis(dimethylamido)zirconium(IV) $\text{[Zr(N(CH}_3\text{)_2}]_4}$</td>
</tr>
<tr>
<td>TFET</td>
<td>tunnel field effect transistor</td>
</tr>
<tr>
<td>TIBB</td>
<td>tip-induced band bending</td>
</tr>
<tr>
<td>TMA</td>
<td>trimethylaluminum $\text{[Al(CH}_3\text{)]_3}$</td>
</tr>
<tr>
<td>TRPL</td>
<td>time-resolved photoluminescence</td>
</tr>
</tbody>
</table>
### Notation Description

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDM</td>
<td>unified defect model</td>
</tr>
<tr>
<td>UHV</td>
<td>ultrahigh-vacuum</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>VB</td>
<td>valence band</td>
</tr>
<tr>
<td>VBM</td>
<td>valence band maximum</td>
</tr>
<tr>
<td>XPS</td>
<td>x-ray photoelectron spectroscopy</td>
</tr>
</tbody>
</table>

### List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>area</td>
</tr>
<tr>
<td>$BE$</td>
<td>binding energy</td>
</tr>
<tr>
<td>$C$</td>
<td>capacitance</td>
</tr>
<tr>
<td>$D_{it}$</td>
<td>interface defect state density</td>
</tr>
<tr>
<td>$D_n, D_p$</td>
<td>diffusion constant (diffusivity) for electrons or holes</td>
</tr>
<tr>
<td>$E$</td>
<td>energy</td>
</tr>
<tr>
<td>$E_{F,s}$</td>
<td>Fermi-level of energy of the sample</td>
</tr>
<tr>
<td>$E_{F,t}$</td>
<td>Fermi-level of energy of the STM tip</td>
</tr>
<tr>
<td>$E_F$</td>
<td>Fermi-level of energy</td>
</tr>
<tr>
<td>$E_g$</td>
<td>energy band gap</td>
</tr>
<tr>
<td>$E_{hk}, E_{in}$</td>
<td>energy of $hk$-diffracted/incident electron beam</td>
</tr>
<tr>
<td>$I$</td>
<td>signal intensity</td>
</tr>
<tr>
<td>$I_{DS}$</td>
<td>source-drain current in a transistor</td>
</tr>
<tr>
<td>$I_{OFF}, I_{ON}$</td>
<td>off-state/on-state current in a transistor</td>
</tr>
<tr>
<td>$I_t$</td>
<td>tunneling current</td>
</tr>
<tr>
<td>$J_0$</td>
<td>saturation current density of a reverse biased diode</td>
</tr>
<tr>
<td>$J_d$</td>
<td>current density of a diode</td>
</tr>
<tr>
<td>$J_{ph}$</td>
<td>photoexcited charge carrier induced current density</td>
</tr>
<tr>
<td>$J_{sc}$</td>
<td>short circuit current density of a photodiode</td>
</tr>
<tr>
<td>$J_{tot}$</td>
<td>total current density</td>
</tr>
<tr>
<td>$KE$</td>
<td>kinetic energy</td>
</tr>
<tr>
<td>$L_{dn}, L_{dp}$</td>
<td>diffusion length for electrons or holes</td>
</tr>
<tr>
<td>$N_d, N_a$</td>
<td>donor or acceptor dopant impurity concentration</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>$N_{sd}, N_{sa}$</td>
<td>surface donor or acceptor traps per unit area</td>
</tr>
<tr>
<td>$N_{ts}$</td>
<td>surface trap states per unit area</td>
</tr>
<tr>
<td>$P$</td>
<td>power</td>
</tr>
<tr>
<td>$S$</td>
<td>sticking probability</td>
</tr>
<tr>
<td>$T$</td>
<td>temperature</td>
</tr>
<tr>
<td>$T(E, eV)$</td>
<td>transmission probability</td>
</tr>
<tr>
<td>$V$</td>
<td>potential difference or voltage</td>
</tr>
<tr>
<td>$V_D$</td>
<td>transistor drain voltage</td>
</tr>
<tr>
<td>$V_G$</td>
<td>transistor gate voltage</td>
</tr>
<tr>
<td>$V_{bi}$</td>
<td>built-in voltage of a diode</td>
</tr>
<tr>
<td>$V_{ext}$</td>
<td>external bias voltage</td>
</tr>
<tr>
<td>$V_{gap}$</td>
<td>voltage between STM tip and sample</td>
</tr>
<tr>
<td>$V_{oc}$</td>
<td>open circuit voltage of a photodiode</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>absorption coefficient</td>
</tr>
<tr>
<td>$G_{hk}$</td>
<td>translation vector defined by indices $h$ and $k$</td>
</tr>
<tr>
<td>$S$</td>
<td>Poynting vector</td>
</tr>
<tr>
<td>$a_1^<em>, a_2^</em>$</td>
<td>primitive vectors of 2-d reciprocal lattice</td>
</tr>
<tr>
<td>$k_{hk}, k_{in}$</td>
<td>wavevector of $hk$-diffracted/incident electron beam</td>
</tr>
<tr>
<td>$\delta$</td>
<td>penetration depth</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>dielectric constant</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>wavelength</td>
</tr>
<tr>
<td>$\lambda_{al}$</td>
<td>attenuation length of electrons</td>
</tr>
<tr>
<td>$\mathcal{E}$</td>
<td>electric field strength</td>
</tr>
<tr>
<td>$\Delta E$</td>
<td>FWHM of spectral line broadening</td>
</tr>
<tr>
<td>$\Delta G$</td>
<td>Gibbs free energy</td>
</tr>
<tr>
<td>$\mu_n, \mu_p$</td>
<td>mobility of electrons or holes in a material</td>
</tr>
<tr>
<td>$\nu$</td>
<td>electromagnetic frequency of a photon</td>
</tr>
<tr>
<td>$\phi_0$</td>
<td>charge capture potential</td>
</tr>
<tr>
<td>$\phi_{sp}$</td>
<td>spectrometer work function</td>
</tr>
<tr>
<td>$\phi_s$</td>
<td>surface potential</td>
</tr>
<tr>
<td>$\phi_{tip}, \phi_{sub}$</td>
<td>STM tip / substrate work function</td>
</tr>
<tr>
<td>$\rho_{sub}$</td>
<td>LDOS of the substrate</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>conductivity</td>
</tr>
<tr>
<td>$\tau_n, \tau_p$</td>
<td>excited electron or hole lifetime</td>
</tr>
<tr>
<td>$\tau_r$</td>
<td>charge carrier relaxation time</td>
</tr>
<tr>
<td>$\theta$</td>
<td>electron emission angle with respect to the surface normal</td>
</tr>
<tr>
<td>$\varrho$</td>
<td>space charge</td>
</tr>
<tr>
<td>$a$</td>
<td>lattice constant</td>
</tr>
</tbody>
</table>
Symbol | Description |
--- | --- |
$b$ | small offset in STS data handling |
$c_n, c_p$ | electron and hole capture coefficients |
$d$ | film thickness |
$f_T$ | cut-off frequency of transistor current gain |
$f_{\text{max}}$ | maximum oscillation frequency for a unity power gain of a transistor |
$g_m$ | transconductance (mutual conductance) |
$h\nu$ | photon energy |
$k$ | wavenumber at a particular wavelength $\lambda$ |
$k_d$ | wavefunction decay constant into the vacuum |
$m_{c_n}^*, m_{c_p}^*$ | electron or hole conductivity effective mass |
$m_{n}^*, m_{p}^*$ | electron or hole density of states effective mass |
$n_0$ | electron concentration at equilibrium conditions |
$n_i$ | intrinsic carrier concentration |
$p_0$ | hole concentration at equilibrium conditions |
$v_s$ | surface recombination velocity |
$z$ | distance from the surface |
$z_{\text{dep}}$ | width of charge carrier depletion layer |

List of Physical Constants

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hbar$</td>
<td>reduced Planck’s constant: $\hbar/2\pi$</td>
<td>$1.054 571 80 \times 10^{-34}$ J s</td>
</tr>
<tr>
<td>$\varepsilon_0$</td>
<td>vacuum permittivity</td>
<td>$8.854 187 82 \times 10^{-12}$ F m$^{-1}$</td>
</tr>
<tr>
<td>$c$</td>
<td>speed of light</td>
<td>$2.997 924 58 \times 10^{8}$ m s$^{-1}$</td>
</tr>
<tr>
<td>$e$</td>
<td>elementary charge</td>
<td>$1.602 176 62 \times 10^{-19}$ C</td>
</tr>
<tr>
<td>$h$</td>
<td>Planck’s constant</td>
<td>$6.626 069 57 \times 10^{-34}$ J s</td>
</tr>
<tr>
<td>$k_B$</td>
<td>Boltzmann constant</td>
<td>$1.380 648 52 \times 10^{-23}$ J K$^{-1}$</td>
</tr>
<tr>
<td>$m_0$</td>
<td>mass of an electron</td>
<td>$9.109 383 56 \times 10^{-31}$ kg</td>
</tr>
</tbody>
</table>
1 Introduction

Semiconductor materials are in vital role in the information and mobility era of today. The advances in modern technology are most significantly driven by improvements in field effect transistor (FET) technology, implemented as smallest functional units of microprocessors. The improvements have been enabled by strive of chip technology companies to execute the famous Moore’s law from 1965 [1], revised into its current form in 1975 [2]: it is possible to double the number of transistors in a given area of an integrated circuit about every two years. However, it is impossible to carry on with this indefinitely since chip technologies are already reaching sub-10 nm features and would need to reach atomic dimensions in a decade or so. In fact, even if the device physics would uphold the scaling, the economic aspect is suspected to bring barrier to scaling in just a few years [3]. Future challenges concerning consumer electronics lie in seamless integration of different technological and materials systems. Thus, the industrial challenges are largely issues of the material technology in many cases.

Remarkable parallel improvements in metal oxide semiconductor field effect transistor (MOSFET) structures have been implemented already in the last ten to fifteen years, most notably the use of atomic layer deposition (ALD) to have unsurpassed controllability of thickness of pinhole-free, very uniform oxide materials, high-κ oxides to enable higher capacitance for a given film thickness, i.e., more charge carriers with a given voltage [4], and shift from planar MOSFET structure to fin-type field effect transistors (FinFETs) and to gate-all-around (GAA) FETs for enhanced electrostatic control [5]. Many of the steps taken have already required a deep understanding of atomic processes on surfaces and interfaces. It is worth noting here that the Millennium technology prize (Technology Academy Finland) was given in 2018 to Tuomo Suntola for the development of ALD.

To keep up with the required pace of improvements, other alternative approaches are also needed, a potential one being the use of III–V compound semiconductor materials [6] instead of Si that is mainly used today in the MOSFETs of chip technology. The main benefit of III–V’s in FET technology is their higher electron and hole mobilities than in Si. This brings several benefits, e.g., shorter delays, higher output currents, or more significantly
as transistor density increases, similar output with significantly lower voltages and thus, lower power consumption [7]. III–V’s are used today mostly in a variety of optoelectronics such as light emitting diodes (LEDs), concentrated photovoltaic cells, lasers, and also in some specific high electron mobility transistor (HEMT) structures, etc., but not significantly in microprocessor FETs largely due to their weaker oxide interface characteristics and the massive momentum of Si technology market. However, the shift from SiO$_2$/Si system to ALD-deposited high-$\kappa$ oxide/semiconductor system is already bridging the gap between benefits of using Si as compared to other semiconductor materials in MOSFETs [8].

This being said, it is still important to note the significance of present and emerging optoelectronics technologies and their growth enabled by III–V material structures. There is a demand of e.g. efficient photodetectors and laser diodes for automotive light detection and ranging (LIDAR) systems, large bandwidth transistors for wireless communication, and optical amplifiers for photonic integrated circuits (PICs), to mention a few. Ample margin for device development for these and other emerging technologies exists, especially concerning passivation of surfaces, that would result in higher efficiencies and/or signal to noise ratios. From the aforementioned, specifically photodetector arrays can benefit from more efficient surface passivation in many ways, enabling smaller unit components in arrays with better performance, and ultimately better resolution and detection efficiency.

The more complex (i.e., advanced) the patterning on a semiconductor wafer becomes, the more facets and different surfaces are exposed to ambient during the device processes. For example, if we consider the surface of a material to consist of two outermost atomic layers, a 5 nm GAA FET already contains almost as many surface as bulk atoms. It is therefore obvious that surface properties and passivation are in a crucial role to continue such developments in technology that we have seen in the past, especially for FETs but to implement benefits for also other technologies. Indeed, current trends in semiconductor industry are moving towards utilizing ultra-thin, low-dimensional device structures (the surface of which consists of several different crystal planes), especially nanowires [9–14] that offer improvements and find even broader range of applications than the traditional bulk device structures. In fact, tunnel field effect transistors (TFETs) that avoid the current thermal limitations in switching speed [15], have been shown with III–V nanowire components [14, 16, 17]. Another particular, interesting example is utilization of nanostructured InSb networks as a template material for quantum computers [18–20]. It is therefore obvious that investigation of III–V surfaces and interfaces is of paramount importance for future technological breakthroughs.
In this work, we have investigated III–V materials and their oxidation, implemented recently discovered crystalline oxide films [21], shown to result in major reductions in interface defect density [22], into practical III–V material surfaces and interfaces. Our research facilities include highly surface sensitive ultrahigh-vacuum (UHV) methods such as photoelectron spectroscopy (PES), low energy electron diffraction (LEED), and scanning tunneling microscopy (STM). Using these complementary methods, we have been able to investigate atomic processes that occur on the surface. Subsequently depositing protective metals or ALD films in situ, we have been able to probe the interface properties with ex situ methods such as separate synchrotron radiation photoelectron spectroscopy (SR-PES) and photoluminescence (PL). Using this procedure, we have also proceeded into device tests using practical III–V device structures grown with molecular beam epitaxy (MBE), indicating advances also in e.g. infrared (IR) detector devices. The last part of this work is giving promising results for also non-planar device concepts consisting of several crystal plane surfaces, a key point for future technologies, by proof-of-concept crystalline oxidation of (111) surfaces.

Chapter 2 gives introduction into fundamental properties of semiconductors, surfaces and interfaces, and interplay between them. In particular, semiconductor–oxide junctions are discussed. In chapter 3, the experimental methods used in this work are introduced, as well as their basic working principles and theory, along with the benefits and information each of the methods offers. Chapter 4 focuses on the applied part; exemplary semiconductor devices that facilitate functional oxide-interfaces or contain such interfaces by necessity are presented along with the background on building blocks of such structures. Meaning of passivation on interfaces for these device concepts is discussed. Chapter 5 summarizes the results of the work and presents concluding remarks. Papers I–IV present the research that comprises this work.
2 Semiconductors, Surfaces and Interfaces

Semiconductors have a variety of distinct properties that make them beneficial for various applications in photonics and electronics. Some of the beneficial aspects of a particular semiconductor material can be hindered by degradation resulting from ambient conditions or disruption of symmetry at the surface and hence difference in electrical properties. Many of the detrimental properties are related to effects on the surfaces and interfaces, a key point regarding the work carried out in this dissertation, and a key challenge regarding the optimal utilization of these materials. To provide a sufficient background on the subject, this chapter first presents some basics of semiconductors in general.

2.1 Fundamental Properties of Semiconductors

Semiconducting material is traditionally loosely defined as a solid material that has an energy gap between the valence band (VB) and conduction band (CB) of electrons, that is smaller than for insulators, but non-zero as opposed to metals. This property implies that charge carriers can be excited fairly easily from and to conduction band, in which they can transport charge in the material, i.e., conduct electricity. They can also recombine, producing photons with the wavelength corresponding to the difference between the energy levels, or bands in this case. By tailoring and engineering the band structure, several desirable functionalities can be obtained.

2.1.1 Bonding and Energy Levels

In this, and the following sections, properties characterizing semiconductors are introduced. In the notations and subscripts used, $n$ refers to conduction electrons, $p$ to valence band holes, $s$ to surface, $d$ to donor character, $a$ to acceptor character, and $i$ to an undoped (intrinsic) semiconductor, where Fermi-level $E_F$ is approximately at the middle of the energy band gap $E_g$. 
Chapter 2. Semiconductors, Surfaces and Interfaces

The energy bands are characteristic to solid-state materials; when two atoms are brought into the vicinity of one another, their wave functions start to significantly overlap. This leads to the superposition of the individual wave functions, which in turn can result in a lower energy (bonding) state, which valence electrons from both of the atoms tend to occupy, and a higher energy (antibonding) state, which is left empty. If the system is brought into lower total energy in this way, a stable compound will form spontaneously. Similar effect can occur if several electron pairs are involved and any number of atoms are added periodically, and thus, solid state material is formed. When the number of atoms is on macroscopic scale ($\sim1 \times 10^{23}$ cm$^{-3}$), there is a similar number of energy states to occupy, and these are the quasi-continuous (valence and conduction) energy bands of the material. Depending on the location of the electron energy levels, which allow the first excited states of a solid (with respect to the ground state), the material can be categorized as metal, insulator, or semiconductor, as defined above. A simplistic scheme of band formation is illustrated in Fig. 2.1 [23].

![Figure 2.1: Energy band splitting, from atoms to solid](image)

Semiconductors are traditionally divided into sets consisting of elemental groups that build the structure, the most common ones being group IV semiconductors (Si, Ge, Si$_x$Ge$_{1-x}$) and group III–V compound semiconductors (such as GaAs, Ga$_x$In$_{1-x}$As, GaSb, GaN, InP, InSb) [23, 24]. These materials form covalent bonds to fulfill the octet rule. This is easily understood by considering bulk Si; a single Si atom has 4 valence electrons: $3s^23p^2$, that hybridize into $sp^3$ orbital. The hybrid orbitals are tetrahedrally configured, so that a single Si atom finds four Si atoms as nearest neighbors.
due to symmetry of bonding energy states. Repeating such configuration for the adjacent atoms results in a diamond structure \( \{(0,0,0),(1/4,1/4,1/4)\} \) basis face-centered cubic (FCC) lattice, depicted in Fig. 2.2 (a). The energy required to lift one electron from 3s to 3p orbital in order to form the hybrid orbital is much less than the lowering in the total energy of the system. Binding energy for a Si–Si bond is 1.8 eV per atom, making the solid diamond lattice structure (d) a very stable phase for bulk Si [23]. Similar configurations are observed for many III–V materials, but for III–V’s, the structure consists of group III atom FCC lattice with group V atoms at \( (1/4,1/4,1/4) \) basis. This is known as zincblende (zb) structure (Fig. 2.2 (b)). Another frequently observed type for III–V’s is the würtzite (w) structure, in which the hexagonal closely packed planes are separated by a different distance from what is found in cubic structure (Fig. 2.2 (c)). Group IV elemental semiconductors have a perfectly covalent bonding, whereas III–V’s have a mixed covalent/ionic, or polarized bonding due to the difference of electronegativity of the different atoms. However, their ionic nature is significantly less than e.g. II–VI or I–VII compounds [23]. Semiconducting nature is generally found in covalently bonded solids as the charge carriers are more weakly bound than in strongly ionic materials (correspondingly, commonly found in insulators), but stronger than in metals.

\[ \text{Figure 2.2: Diamond (a), zincblende (b), and würtzite (c) structures} \]

The exact description of a materials properties can be done utilizing quantum mechanics, treating electrons as waves rather than particles. The energetics are solved from the Schrödinger equation, a task which is greatly simplified by the periodicity of the lattice and the corresponding potential that the electrons experience, each wave function being characterized by a wavevector \( \mathbf{k} \). It follows from such preconditions and construction of a reciprocal lattice in \( \mathbf{k} \)-space that the general solutions for one-electron time-independent Schrödinger equations are spatially periodic (Bloch–Floquet theorem), and that in a similar manner, the solutions are periodic also in the \( \mathbf{k} \)-space. Thus, it is sufficient to inspect a primitive unit cell in both real and reciprocal space, to completely characterize the properties of a material.
Qualitative description of semiconductor band structure \((i.e., E - k)\) relation of valence electrons can be constructed by nearly free electron approximation using by adding a small perturbation (attractive potential near the nuclei) to the Hamiltonian of the Schrödinger equation. The result shows that for weakly bound electrons, a band gap \(E_g\), or a discontinuity in energy at the Brillouin zone boundary and \(\Gamma\) point \((i.e. \ k = \pm \pi/a\) where \(a\) is the lattice constant, or \(k = 0\), in one-dimensional case) opens up. \(E_g\) is observed also in disordered solids if the electrons are strongly bound, \(i.e.,\) bonds are strongly ionic, such as in SiO\(_2\). More sophisticated yet specific approaches, such as using pseudopotentials or orthogonalized plane wave method are required for quantitative analysis. The most widely utilized concept in computational studies is the density-functional theory (DFT), in which the many-particle system is solved with high accuracy using approximate functionals of the particle density. Validity of such analyses are still always guided by experimental results, such as optical absorption, PL or PES/angle-resolved photoelectron spectroscopy (ARPES) studies [25, 26]. Fig. 2.3 (a) shows the first Brillouin zone of FCC lattice and (b) the band structure and projected density of states (DOS) for GaAs.

Band structure of a semiconductor is instrumental for solving any of its properties, such as DOS, directness of \(E_g\) (whether conduction band minimum (CBM) is on the same or different \(k\)-value as valence band maximum (VBM)), and effective masses \(m_n^*, m_p^*\) that are used to compensate for the different response to external electric field \(E\) in a crystal as compared to free particles: \(-eE = m_n^* \frac{d\psi}{dt}\), and acquired from the local relationship between energy and momentum \((E - k)\) relation approximated as parabolic in the \(\Gamma\)-valley:

\[
E(k) = E_0 + \frac{\hbar^2 |k|^2}{m_n^*}. \tag{2.1}
\]

The different symbols in x-axis of Fig. 2.3 (b) correspond to the symmetry directions seen in (a).

In an intrinsic, pure semiconductor material all of the valence electrons occupy bonding orbitals at VB at 0 K temperature \((i.e.,\) all the levels below the band gap). When temperature is increased, thermal excitations start to occur and some of the electrons occupy also CB states. For example, an intrinsic semiconductor which is completely insulating at 0 K can conduct electricity in room temperature (RT) because of non-zero statistical occupation in the conduction band. The extent of this effect is dependent on temperature, band gap of the material, and the positioning of Fermi-level \(E_F\), \(i.e.,\) the level up to which the energy levels are statistically more probably occupied than not. Generally, the energy level occupation distribution
2.1. Fundamental Properties of Semiconductors

for fermions such as electrons from Fermi-Dirac statistics is characterized by the Fermi-distribution \( f_e(E) \), which is valid for also characterizing occupation in VB (holes) and CB (electrons) for highly doped, degenerate (metal-like occupation of CB states for electrons) semiconductors:

\[
f_e(E) = \frac{1}{\exp\left(\frac{E - E_F}{k_B T}\right) + 1} \tag{2.2}
\]

Here, \( k_B \) represents the Boltzmann constant and \( T \) the temperature. For non-degenerate semiconductors (low doping and thus, discrete dopant levels and low CB occupation of electrons), only the tail of this distribution occupies VB or CB, and is simplified to Boltzmann-distribution when \( E - E_F \gg k_B T \):

\[
f_e(E) = \exp\left(-\frac{E - E_F}{k_B T}\right) \tag{2.3}
\]

In a statistical sense, as seen from Eq. (2.2), \( E_F \) is the level in which the probability of finding an electron is 0.5, the above and below of which the probability is characterized by \( f_e \). Equilibrium conduction electron density \( n_0 \) in a semiconductor can be calculated as a product of the distribution \( f_e(E) \) and CB DOS and integrating over the CB range.

By introducing impurities into the material, one can bring extra electrons or holes that are utilized as mobile charge carriers in the structure. In practice, \( e.g. \) using donor atoms (\( n \)-type doping) such as Si in GaAs melt, the

Figure 2.3: (a) First Brillouin zone of FCC (that constructs the zb structure of GaAs) lattice in \( k \)-space. (b) Calculated GaAs band structure and DOS from empirical nonlocal pseudopotential scheme, with experimental PES spectrum marked with dashed lines, after [26, 27]
Si atoms substitute some of the Ga atoms in solid GaAs, but since there is an extra electron, it is donated into the structure and is fairly mobile. In energy space, the effect of \( n \)-type impurity atom is seen as a state above the intrinsic Fermi-level \( (E_i) \), from which the electron is easily excited into the CB, and thus, \( n \)-type conductivity is introduced. With sufficient \( n \)-doping, the \( E_F \) is raised close to the CB due to significant proportion of electrons occupying high energy levels, and thus, the material, rich in electrons, exhibits much higher conductivity in a given temperature. Similar but opposite effect is realized with acceptor atoms such as Zn or Be in GaAs, which replace Ga atoms with a charge deficiencies, \( i.e. \), holes. Si works as an amphoteric dopant, so that if it substitutes As, the Si atom is an acceptor. A hole propagates in the structure by drawing an electron from an adjacent bond; the \( p \)-type conduction, or movement of holes, is actually net movement of adjacent electrons in the opposite direction. Analogously, in an energy diagram, holes are introduced to the VB through the excitation of electrons to low energy states induced by acceptors below \( E_i \), which effectively lowers the \( E_F \). Efficient doping requires low ionization energies of the dopant atoms in the semiconductor matrix, with respect to the CBM or VBM, and also appropriate doping levels to avoid formation of compensating defects \([23]\).

### 2.1.2 Characteristic Semiconductor Properties

The key properties that dictate electrical characteristics of a semiconductor are majority carrier density, \( n_0 \) (electrons) or \( p_0 \) (holes), and charge carrier mobility, \( \mu_n \) or \( \mu_p \). \( \mu_n \) (or \( \mu_p \) similarly with corresponding values for holes) is given by:

\[
\mu_n = e\langle \tau_r \rangle / m^*_{cn}.
\]

\( m^*_{cn} \) is the conductivity electron effective mass and \( \langle \tau_r \rangle \) the mean relaxation time of charge carriers. Most III–V’s, such as GaAs and InP, have spherical constant energy-surface (parabolic band), and the DOS effective electron mass, \( m^*_{n} \) is an isotropic scalar quantity and thus, equivalent to \( m^*_{cn} \).

Variation in carrier density and mobility are often observed in performance in the electrical conductivity, \( \sigma \), of a material:

\[
\sigma = \sigma_n + \sigma_p = en_0\mu_n + ep_0\mu_p,
\]

which is straightforwardly increased by doping the material and thus increasing \( n_0 \) or \( p_0 \). Too heavy doping will however make the electronic structure degenerate and change fundamental properties of the material. Furthermore, ionized impurity scattering becomes important in \( e.g. \) Si and GaAs for doping concentrations of \( \geq 1 \times 10^{17} \text{cm}^{-3} \), and as dopants impurities
are introduced into the structure, the crystal quality is affected, which can degrade some properties to an extent. Increasing doping level reduces \( \langle \tau_r \rangle \), which is seen as reduced mobility. \cite{25}

To maintain a high mobility and thus fast switching in transistors, undoped regions of semiconductor material are utilized, \textit{e.g.}, in HEMT structures, discussed briefly in section 4.3.2. Mobility in several III–V semiconductors is much higher than in Si due to their band structure, which results in a much lower \( m^*_n \), but also affects the scattering mechanisms involved \cite{25}. Much faster and higher frequency operation of electronic and photonic devices are thus potentially achieved with III–V semiconductors. Some of the important materials parameters that are readily quantifiable for intrinsic semiconductors are listed in Table 2.1 for some selected materials \cite{23, 26, 28}. It is noted that in addition to direct band gaps and high mobilities, many of the ternary and quaternary III–V compounds can be grown with crystal growth methods such as MBE lattice matched on binary substrates such as GaAs by tuning their composition. These properties make III–V’s highly versatile materials and enable \textit{e.g.} growth of quantum well structures and high efficiency multi-junction solar cells.

Other highly important factors in semiconductors concerning especially photonic devices include their absorption and recombination characteristics. These properties can also be investigated and utilized to probe the material in more general, using \textit{e.g.} PL method described in section 3.6. An electron can be excited to a higher energy state when it absorbs a photon, leaving behind a hole in its origin. Fundamentally, direct transition from VB to CB is the most likely event that occurs when the photon energy \( h\nu \) exceeds the direct \( E_g \) of the material. In this regime, a significant absorption coefficient \( \alpha \) is observed. \( \alpha \) dictates attenuation of the Poynting vector \( S \), \textit{i.e.}, the flow of energy (that is proportional to photon flux):

\[
S(z) = S_0 \exp (-\alpha z).
\] (2.6)

However, the absorption (and thus, also attenuation) are dependent on the band structure; transition probability is proportional to the transition states available, \textit{i.e.}, DOS at VBM and CBM and time-dependent perturbation matrix element. The matrix element is independent of wave vector for direct allowed state transitions, and near the \( \Gamma \) point, it is valid to assume the vicinity of VBM and CBM parabolic. From these assumptions one obtains:

\[
\alpha = K \sqrt{h\nu - E_g},
\] (2.7)

where \( K \) is a constant. Typical values of \( \alpha \) for III–V semiconductors for visible light are in the range of \( 1 \times 10^3 \text{ cm}^{-1} \) to \( 1 \times 10^5 \text{ cm}^{-1} \), with higher
Table 2.1: Some of the discussed physical properties of common (undoped) semiconductor materials at room temperature. \( n_i \) denotes intrinsic carrier concentration. Representative values are shown from vast literature of experimental results [23, 26, 28]. For materials with an indirect band gap, the lowest energy VB \( \rightarrow \) CB as well as direct \( \Gamma \) transitions are shown.

<table>
<thead>
<tr>
<th>Material</th>
<th>Lattice</th>
<th>( E_g ) (eV)</th>
<th>( a ) (Å)</th>
<th>( n_i ) (cm(^{-3}))</th>
<th>( \mu_n ) (cm(^2) V(^{-1}) s(^{-1}))</th>
<th>( \mu_p ) (cm(^2) V(^{-1}) s(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs</td>
<td>zb</td>
<td>1.42 (( \Gamma ))</td>
<td>5.65</td>
<td>( 2.1 \times 10^6 )</td>
<td>7000 – 10000</td>
<td>250 – 400</td>
</tr>
<tr>
<td>GaSb</td>
<td>zb</td>
<td>0.70 (( \Gamma ))</td>
<td>6.10</td>
<td>( 1.5 \times 10^{12} )</td>
<td>2600 – 7700</td>
<td>250 – 1000</td>
</tr>
<tr>
<td>( \text{Ga}<em>{0.47}\text{In}</em>{0.53}\text{As} )</td>
<td>zb</td>
<td>0.73 (( \Gamma ))</td>
<td>5.87</td>
<td>( 6.3 \times 10^{11} )</td>
<td>12000</td>
<td>300</td>
</tr>
<tr>
<td>( \text{Ga}<em>{0.52}\text{In}</em>{0.48}\text{P} )</td>
<td>zb</td>
<td>1.96 (( \Gamma ))</td>
<td>5.65</td>
<td></td>
<td>3000</td>
<td>45</td>
</tr>
<tr>
<td>AlAs</td>
<td>zb</td>
<td>2.16 (( \Gamma ) ( \rightarrow ) X)</td>
<td>5.66</td>
<td></td>
<td>1000</td>
<td>80</td>
</tr>
<tr>
<td>InP</td>
<td>zb</td>
<td>1.35 (( \Gamma ))</td>
<td>5.87</td>
<td>( 3.3 \times 10^7 )</td>
<td>4000 – 6000</td>
<td>150 – 600</td>
</tr>
<tr>
<td>InAs</td>
<td>zb</td>
<td>0.36 (( \Gamma ))</td>
<td>6.06</td>
<td>( 1.3 \times 10^{15} )</td>
<td>20000 – 40000</td>
<td>100 – 500</td>
</tr>
<tr>
<td>InSb</td>
<td>zb</td>
<td>0.18 (( \Gamma ))</td>
<td>6.48</td>
<td>( 2.0 \times 10^{16} )</td>
<td>70000 – 100000</td>
<td>850 – 1700</td>
</tr>
<tr>
<td>( \alpha)-GaN</td>
<td>w</td>
<td>3.43 (( \Gamma ))</td>
<td>3.19 (a)</td>
<td></td>
<td>150 – 1500</td>
<td>&lt; 350</td>
</tr>
<tr>
<td>( \beta)-GaN</td>
<td>zb</td>
<td>3.22 (( \Gamma ))</td>
<td>4.51</td>
<td></td>
<td>50 – 1000</td>
<td>50 – 350</td>
</tr>
<tr>
<td>AlN</td>
<td>w</td>
<td>6.2 (( \Gamma ))</td>
<td>3.11 (a)</td>
<td></td>
<td>(14)</td>
<td></td>
</tr>
<tr>
<td>Si</td>
<td>d</td>
<td>1.11 (( \Gamma ) ( \rightarrow ) ( \Delta ))</td>
<td>5.43</td>
<td>( 1.0 \times 10^{10} )</td>
<td>1350 – 1500</td>
<td>370 – 500</td>
</tr>
<tr>
<td>Ge</td>
<td>d</td>
<td>0.66 (( \Gamma ) ( \rightarrow ) ( \text{L} ))</td>
<td>5.66</td>
<td>( 2.3 \times 10^{13} )</td>
<td>3900</td>
<td>1800 – 1900</td>
</tr>
</tbody>
</table>
2.1. Fundamental Properties of Semiconductors

$h\nu$, or lower $\lambda_0$, photons being absorbed closer to the surface [25, 26, 29]. Penetration depth, $\delta$, is an illustrative quantity for comparing the magnitude of absorption [29]. It is defined as the depth, by which 63\% of the photon flux has been absorbed. That is, the intensity has attenuated to $e^{-1}$ (37\%):

\[
S(\delta) = S_0 \exp (-\alpha \delta) = S_0 \exp (-1) \quad (2.8)
\]

\[
\Leftrightarrow \delta = 1/\alpha. \quad (2.9)
\]

After excitation, the charge carriers can drift in the material due to an electric field $E$ or randomly diffuse without external force. Eventually, if the charge carriers are not collected to an external electrode, the carrier will recombine radiatively or non-radiatively, depending on the energy states involved. Radiative recombination commonly occurs in high-quality direct $E_g$ material where only small amount or no gap states (states that lie within $E_g$) are found, and this produces light with wavelength that corresponds to the $E_g$. If trap recombination centers (due to deep-level impurities, or, vacancy, interstitial, or antisite defects) exist in the material, the charge carrier can recombine non-radiatively through the corresponding, commonly introduced gap states. The recombination through these centers facilitates phonon vibrations. The process and recombination rates are described in the Shockley-Read-Hall (SRH) model [30, 31], from which it is observed that the recombination rates are dictated by charge carrier densities and minority carrier lifetimes $\tau_n$ and $\tau_p$. Lifetimes are inversely proportional to their capture coefficients ($c_n$, $c_p$), and trap state densities $N_{ts}$, and so higher $N_{ts}$ is always observed as reduced $\tau$ [25].

Using diffusivity $D_n$ for electrons according to Einstein relations:

\[
D_n = \frac{k_B T}{e} \mu_n, \quad (2.10)
\]

one obtains the mean length that an electron propagates before recombination, diffusion length $L_{dn}$:

\[
L_{dn} = \sqrt{D_n \tau_n}. \quad (2.11)
\]

The same equations are valid for holes, correspondingly. Typical carrier lifetimes for high crystal quality GaAs based materials can be found in the range of $5 \times 10^{-9}$ s and $3 \times 10^{-6}$ s for electrons and holes, respectively (low injection levels, low doping). Using mobilities listed in Table 2.1, corresponding $L_d$ are in the range of 10 $\mu$m and 50 $\mu$m in RT. Diffusion lengths of around 10 $\mu$m are commonly observed for high quality, pure III–V crystals [25, 26, 28].
Chapter 2. Semiconductors, Surfaces and Interfaces

2.2 Semiconductor Surfaces

Bulk properties of a semiconductor dealt with so far are necessary for understanding the properties of semiconductor surfaces, and principles behind the methods that are used for investigating them. The surface part of semiconductor material is however an important focus area here and exhibits a range of important phenomena to study by itself, in order to understand further interactions with the environment.

2.2.1 Atomic Structure

In order to create a surface from bulk, bonds along a particular lattice plane (commonly (100), (110) or (111)) need to be broken. The covalent bonds, that build the semiconductor lattice, towards this plane are left partially unfilled and “dangling”. This is energetically unfavorable, and the bonds are highly reactive. The surface energy is lowered tremendously as the dangling bonds find their nearest surface neighbor atoms to bond with. A bulk truncated (100) surface of a diamond structure contains two dangling bonds per surface atom, which tend to form dimer bonds between two of these atoms to lower the surface energy. In such a configuration, each of the atoms still contains one dangling bond, which might be filled with electrons or empty, depending on the valence of the atom. From this process it follows that for semiconductors, the surface unit cell is most often larger than the (1×1) truncated bulk unit cell. Famous examples are e.g. Si(100)(2×1) observed as a result from fluctuating buckled dimers of c(4×2) and p(2×2) \[32\], and Si(111)(7×7), illustrated in Fig. 2.4 (a) and (b), respectively. Also (ζ) atomic models for Ga-rich conditions on GaAs(100) are shown in (c) (by Lee et al. \[33\]) and (d) (by Kumpf et al. \[34\]). The particular unit cells describing the resulting lattice are called surface reconstructions, which are obtained when lowering of energy on the surface through rebonding or rehybridization overcompensates the increase of energy through the strain that accompanies the reconfiguration in surface layers. \[35, 36\]

A few principles that deal with physics and intrinsic features of semiconductor crystals can be distinguished to identify and predict the nature of surface reconstructions. \[35\]

1. Surface reconstruction tends to form in such a fashion that saturates the dangling bonds, or converts them into non-bonding electronic states (filled vs. empty bonds). Saturation can occur through rehybridization, as is observed for Si(111)(2×1) where sp² hybridization occurs, and the three-fold coordinated structure’s fourfold valence of the atoms is saturated through π-bonds.
2.2. **Semiconductor Surfaces**

2. Atomic relaxation tends to lower the surface energy, lowering occupied state energy and thus opening a band gap, leading to semiconducting characteristics, unless some surface states lie at the $E_F$. For example on buckled dimers of Si(100)(2×1) one atom is buckled upwards to occupy electrons on the bonding $\pi$-orbital and the other one is buckled downwards and $\pi^*$-orbital prominent on this atom is left empty. Without buckling, the energies on these orbitals would overlap, causing a metallic character.

3. The surface preparation determines which surface reconstruction will be observed, namely, the lowest energy structure kinetically accessible. For example, cleaved Si(111) reconstructs in a (2×1) structure, and turns into a lower energy (7×7) only after a high temperature annealing due to an activation barrier.
The three conditions above make up a solid basis for predicting behavior on elemental semiconductors. For compound semiconductors, one needs to take into account the different valence of the constituent atoms, \textit{i.e.}, atoms that have cationic or anionic nature on the surface.

4. Cations have higher energy dangling bond orbitals than the anions do, and hence, those of anions must be filled with the electrons available in the surface layers, leaving the cation dangling bonds empty, to retain the surface uncharged. This is known as the autocompensation rule, or electron counting rule, which is important in identifying III–V bonding at interfaces, and thus, also surface reconstructions. This simple rule, for example, successfully predicts three anion dimers and a single missing one in the III–V(100)(2\times4) surface unit cell. [39]

5. Dangling bonds rehybridize to lower the surface energy (or electrostatic repulsions), leading to particular atomic orientations and geometry observed for a particular atomic stoichiometry on the surface.

### 2.2.2 Surface and Interface States

As discussed previously, several properties of solid-state materials are consequences of their periodic crystalline nature. Therefore, a surface that breaks this symmetry has drastic implications on these properties. Specifically, to satisfy continuity conditions of the electron wavefunctions, the part of the wavefunction decaying into the solid and vacuum with its imaginary part, absent in the bulk, result in real energy eigenvalues that are not found in the bulk material (surface resonant or evanescent states). Therefore, additional energy states are an inherent property of the surface of a material. The real wavefunctions also decay into the vacuum. In practice, surface dangling bonds are taken into account in a tight-binding model utilized in solving the surface properties. This way, group III atoms are found to contain empty dangling bonds and group V filled ones in III–V compounds theoretically, dictated by their respective energy levels.

The markedly different geometry and occupation of bonds at the semiconductor surface generally also induces differences in the band structure, possibly resulting in significantly different electronic properties than predicted without any relaxation or reconstruction. The same is analogously true for semiconductor interfaces due to similar discontinuity as with vacuum, but the characteristics of the changes in properties are dictated by the constituents and specific atomic configurations of the heterostructure. [36]

Indeed, surface actually is a trivial type of interface of bulk material with vacuum, and interface states emerge in practice due to various effects such
as extrinsic defects, lattice mismatch, \textit{etc.} \cite{25} in addition to the intrinsic bonding induced states.

The surface or interface energy states that lie within the $E_g$ of the material can result in i) recombination centers or ii) charge trapping and consequently Fermi-level pinning (FLP). Non-radiative recombination through recombination centers will cause loss of effective current, transforming into heat. Effectively, this will increase power consumption or deteriorate \textit{e.g.} quantum efficiency of photonics devices. Charge trapping, on the other hand, has significant implications on the control of charge carrier density in the vicinity of surface layers.

A static charge building up on the surface or interface will leave a space charge region depleted of the carriers of the same sign, in order to retain charge neutrality. This is seen as band bending (internal $\mathcal{E}$) near the surface or interface, illustrated in Fig. 2.5 (a) as the offset of $E_F$ from VB and CB. When a sufficiently high amount of surface states are occupied by charge carriers, the surface potential $V(0)$ corresponds to the difference between the surface states’ energy and $E_F$ in the semiconductor bulk. Similarly, $V(z) = 0$ in bulk. The space charge $\varrho(z)$ consisting of donors, acceptors and charge carriers is then readily solved from Poisson’s equation:

$$\frac{d^2 V(z)}{dz^2} = \frac{-\varrho(z)}{\varepsilon_0 \kappa}, \quad \text{(2.12)}$$

where $\varepsilon_0$ is the vacuum permittivity and $\kappa$ the dielectric constant of the material. One can estimate the width of the depletion layer, $z_{\text{dep}}$, using the above constraints and considering the space charge to simply consist only of ionized impurities, \textit{i.e.}, for $n$-doped semiconductor $\varrho(z) = eN_d$ where $N_d$ is the donor impurity concentration. Then,

$$V(z) = \frac{e^2 N_d}{2\varepsilon_0 \kappa} (z - z_{\text{dep}})^2 \quad \text{(2.13)}$$

Consider, \textit{e.g.}, an $n$-doped GaAs sample with typical approximate parameters of $\kappa = 13$, $N_d = 1 \times 10^{18} \text{ cm}^{-3}$ and $V(0) = 0.7 \text{ eV}$ (near midgap) for surface acceptors, which is a typically observed separation between CBM and the Fermi-level at the surface due to surface or oxide interface states \cite{40–42}. This results in $z_{\text{dep}} \approx 30 \text{ nm}$ and a respective total space charge per area (that equals surface charge) of $\int_0^{z_{\text{dep}}} \varrho(z)dz = N_d z_{\text{dep}} \approx 3 \times 10^{12} \text{ e cm}^{-2}$.

Thus, the surface states can cause band bending near the surface, which is readily observed on the oxide interfaces as well. The states possess an energy and a donor or acceptor character. However, the effective bending is naturally much less if the amount of such surface states is significantly
diminished, since the space charge matching surface charge is significantly smaller, resulting in lowering of surface potential according to Eq. (2.12). However, when the surface charge reaches a value corresponding to potential at which the surface states lie, the increase or decrease in the amount of the charge has significantly smaller effect on the bending. This is illustrated in Fig. 2.5 (b) [36]. This effect is what was previously denoted FLP, which has significant implications concerning the switching from accumulation through depletion to inversion at the semiconductor surface. This can be understood in terms of how filling of the surface states affects the charge to be utilized for conduction near the surface. The extra number of states above which pinning occurs, are neutral in thermal equilibrium. However, when external voltage through e.g. an insulating film is applied to bring inversion charge carriers to the surface, they will merely occupy surface trap states and thus, do not contribute to conductivity until an undesirably high external voltage corresponds to a surface charge exceeding trap density. In other words, significant FLP is in practice equivalent to a high amount of surface/interface states that trap charge carriers and hinder e.g. switching in transistors.

From the above calculation and discussion, a critical concentration for
FLP can be interpreted to be in the range of $1 \times 10^{12} \text{cm}^{-2}$ for typical GaAs, which is quite small value considering that surface atom concentration is in the range of $1 \times 10^{15} \text{cm}^{-2}$ (0.1% of such surface/interface state causing atoms are enough to pin the $E_F$) [36]. Indeed, in practice, it is challenging to avoid FLP on III–V semiconductors, GaAs in particular [8, 43].

The surface states can be categorized as fast or slow states, depending on their respective time constants associated with the recombination. The surface terminology can be misleading, since slow states are often associated with border traps [44] within an oxide layer above, to which charge carriers need to tunnel through a barrier, and hence the long time constant. However, these interface related defect states or traps can play a significant role in, e.g., the band bending and pinning behavior described above. The surface/interface traps/states, or states at the intimate electrical constant of the semiconductor [44] are usually fast states which facilitate short time constants, or fast recombination at nonequilibrium conditions (under illumination causing excess carriers, for example). The recombination at semiconductor surface can be treated as SRH recombination, when additionally taking into account band bending and the associated changes in charge carrier concentrations. The surface recombination velocity $v_s$ is a fairly complex function of $c_p$ and $c_n$, trap state energy ($E_t$) and surface potential ($\phi_s = (E_F - E_{in})/e$), but depends linearly on $N_{ts}$ assuming no significant contribution to band bending (which is a good approximation for pinned bands) [25]:

$$v_s = \frac{N_{ts}\sqrt{c_p c_n (p_0 + n_0)}/2n_i}{\cosh [(E_t - E_i - e\phi_0)/k_BT] + \cosh [e(\phi_s - \phi_0)/k_BT]}$$

(2.14)

where

$$\phi_0 = \frac{k_BT}{2e} \ln (c_p/c_n).$$

(2.15)

The native oxide interface on GaAs(100) is naturally quite prone to surface recombination, and difference to epitaxial crystal interface is drastic; $v_s$ of $10 \text{cm s}^{-1}$ has been measured at lattice matched GaInP/GaAs interface and $1 \times 10^7 \text{cm s}^{-1}$ for the respective air/GaAs interface [45]. For chemically polished Si(100), surface state density can be less than $1 \times 10^{10} \text{cm}^{-2}$ and the corresponding $v_s 1 \times 10^3 \text{cm s}^{-1}$ [25].

2.3 Oxide–Semiconductor Stack

An oxide film deposited on top of semiconductor surface is an essential part of a variety of semiconductor devices. In particular, metal oxide semiconductor
(MOS) stack is a key component in FET technology, acting as a switch between conductive and non-conductive state of the semiconductor channel in MOSFETs [8, 46]. On the other hand, oxide layers on semiconductors can be utilized \textit{e.g.} as anti-reflection (AR) or high reflection (HR) coatings in photonics devices [v, xix], in memory devices [47], as functional parts of energy technology such as H$_2$O or CO$_2$ splitting [48, 49], and potentially as buffer layers for integration onto Si [50], along with various other ongoing research areas. Furthermore, a stable oxide film inherently protects the semiconductor crystal from atmospheric contaminants and reactions.

2.3.1 Thin Oxide Films

In this work, we have dealt mainly with an oxide layer that could be utilized as an interfacial layer produced on the semiconductor surface, remaining underneath a subsequently deposited additional oxide film. In electronics applications, FETs in particular, the purpose of the oxide film is to electrically isolate the semiconductor from the gate metal, while providing as high an $E$ through the oxide as possible when applying voltage to the gate, in order to efficiently modulate the charge concentration beneath the gate. That is, a high capacitance $C$ is required:

$$C = \varepsilon_0 \kappa \frac{A}{d}, \quad (2.16)$$

where $A$ is the area of the capacitor and $d$ the oxide thickness. MOSFET transistors originally consisting of SiO$_2$ on Si faced a major challenge with scaling as the oxide thickness $d$ needed to be scaled down to near monolayer range for attaining gate control, as this also results in an increase of tunneling current leakage through the oxide. Another parameter that enables scaling $C$ is $\kappa$, by changing the oxide material, which was a natural consequence in the progress of MOSFET technology (see also section 3.2.5).

Today’s standard MOSFETs contain an ALD produced high-$\kappa$ film, typically HfO$_2$, which offers large enough VBM and CBM offsets (3.3 eV and 1.4 eV) to Si, for leakage current suppression, and high $\kappa$-value of 16 to 25, as opposed to the corresponding value of SiO$_2$: 3.9 [4, 51]. Therefore, a HfO$_2$ film can be deposited 5 times as thick as a corresponding SiO$_2$ film with similar gate control (or effective oxide thickness, $EOT = \frac{3.9}{\kappa}d$). In addition, HfO$_2$ is thermodynamically highly stable, which is a necessity to prevent excessive formation of additional interfacial compounds such as SiO$_2$ along with metallic Hf and silicides. These are some of the various attributes required of a gate oxide, but \textit{e.g.} interface characteristics and other materials properties for high-$\kappa$ oxides are still being widely investigated [4].
One of the prospective materials concerning high-$\kappa$/III–V interfaces is $\text{Al}_2\text{O}_3$ due to its high thermal stability and properties related to interface formation. Indeed, the problematic properties of III–V native oxides can be alleviated with an ALD deposited $\text{Al}_2\text{O}_3$ due to the so-called ‘self-cleaning’ \cite{52}. The self-cleaning refers to the ability of an ALD precursor gas to scavenge oxygen from the native oxide as it adsorbs on the surface and this way leave a more abrupt interface, almost free of group III and V oxides. Depositing stacks or nanolaminates with alternating layers of e.g. $\text{Al}_2\text{O}_3$ and $\text{HfO}_2$ can further provide complementing benefits, i.e., self-cleaning, a very high crystallization temperature (for avoiding grain boundary formations), and diminished diffusion of O due to inclusion of $\text{Al}_2\text{O}_3$, while retaining high $\kappa$ of $\text{HfO}_2$ \cite{53, 54}, all of which can partly contribute to reduction of interface defect state density. Our interface investigations have included $\text{Al}_2\text{O}_3$ as the oxide layer due to its various beneficial properties that make $\text{Al}_2\text{O}_3$/III–V interface a prospective candidate to be utilized in various applications.

### 2.3.2 Oxide–Semiconductor Interface

As discussed previously, different compounds and different bonding states as well as disruption in symmetry give rise to energy levels that may lie within the band gap of the semiconductor analogously as at surfaces. Despite several models explaining general trends in formation of gap states at semiconductor interfaces (metal-induced gap state (MIGS) \cite{55} model, unified defect model (UDM) \cite{56}, disorder-induced gap state (DIGS) model \cite{57}, among other original and refined models \cite{58}), it seems that various effects need to be considered when considering a particular combination of materials, their interface, and the associated process flow. This emphasizes the importance of surface and interface chemistry studies of semiconductor crystals designed for operation in electronics. The importance of bulk termination at the interface is seen for example from comparison between GaAs(111)A (Ga-terminated) and GaAs(100) (Ga-As-terminated) $n$-MOSFETs \cite{58}: for GaAs(111)A the defect states are simply regarded due to missing anions (As) and for GaAs(100) mostly missing cations (Ga) at the interface. The defect state energy for missing anions is much closer to CBM which results in GaAs(111)A drastically outperforming GaAs(100) $n$-MOSFET. For $p$-MOSFETs the trend is exactly opposite, as one might expect.

The passivation, or suppression of charge carrier interaction with interface defect states, commonly relies on i) repelling of charge carriers from recombination centers with a field effect due to fixed charges in an oxide film, observed beneficial in for example Si solar cells with $\text{Al}_2\text{O}_3$ film \cite{59},
Table 2.2: Gibbs free energies $\Delta G$ for formation of various semiconductor and oxide compounds [62–64] in RT.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\Delta G$ (kJ mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al$_2$O$_3$</td>
<td>-1582.3</td>
</tr>
<tr>
<td>HfO$_2$</td>
<td>-1088.2</td>
</tr>
<tr>
<td>GaAs</td>
<td>-67.8</td>
</tr>
<tr>
<td>GaSb</td>
<td>-38.9</td>
</tr>
<tr>
<td>InAs</td>
<td>-53.6</td>
</tr>
<tr>
<td>InSb</td>
<td>-25.5</td>
</tr>
<tr>
<td>GaO (g)</td>
<td>253.5</td>
</tr>
<tr>
<td>Ga$_2$O</td>
<td>-315.5</td>
</tr>
<tr>
<td>Ga$_2$O$_3$</td>
<td>-999.7</td>
</tr>
<tr>
<td>GaAsO$_4$</td>
<td>-891.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\Delta G$ (kJ mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaSbO$_4$</td>
<td>-915.1</td>
</tr>
<tr>
<td>InO (g)</td>
<td>364.4</td>
</tr>
<tr>
<td>In$_2$O$_3$</td>
<td>-832.1</td>
</tr>
<tr>
<td>InAsO$_4$</td>
<td>-877.4</td>
</tr>
<tr>
<td>InSbO$_4$</td>
<td>-831.3</td>
</tr>
<tr>
<td>As$_2$O$_3$</td>
<td>-577.0</td>
</tr>
<tr>
<td>As$_2$O$_5$</td>
<td>-783.5</td>
</tr>
<tr>
<td>Sb$_2$O$_3$</td>
<td>-634.8</td>
</tr>
<tr>
<td>SiO (g)</td>
<td>-126.4</td>
</tr>
<tr>
<td>SiO$_2$</td>
<td>-856.3</td>
</tr>
</tbody>
</table>

ii) similar mechanism with a wide $E_g$ lattice matched overlayer (such as Al-GaAs or AlInP on GaAs) to act as a carrier confining barrier between the device active layers and the surface/interface in e.g. quantum well transistors or solar cells [7, 60], or iii) suppression of the defect states altogether. The means of pursuing point iii is optimizing the bonding and chemistry at the interface, which is often denoted chemical passivation. This is a favorable way of dealing with defect states, since it does not affect other device functionality by its nature, but it is, however, most complex issue to solve.

To encompass formation of favorable compounds at oxide interfaces, we consider their thermodynamic stability by comparing Gibbs free energies, $\Delta G$ for different compounds in a specific set of reagents. Any given reaction will tend to settle towards the compound which has the lowest $\Delta G$ in excess ambient of constituent materials (in RT). Although the $\Delta G$ as reference values are given for bulk compounds, they may serve as a guide for specific bond formation also for e.g. interfacial suboxides, since the reaction will have a tendency to end up in the compounds with lowest $\Delta G$. Table 2.2 lists some of the $\Delta G$ values for compounds related to this work. Here the importance of surface science approach is paramount for understanding atomic scale effects; although Ga–oxides have lower $\Delta G$, during RT oxidation, O with large electronegativity bonds initially to As atoms with the filled dangling bonds on GaAs(110) [61]. Generally, formation of oxides is thermodynamically a highly favorable process for III–V semiconductors. This is why oxidation of clean semiconductor surface is in practice impossible to completely avoid in atmospheric ambient conditions, making chemical passivation very difficult.

From binary III–V compounds, group III oxides are typically more stable
than group V and this effect is observed already at the initial stages of oxidation [65]. Many species have been suggested to give rise for defect states on III–V semiconductors, including both group III and V oxides, elemental group V species, and anti-site defects [66–69]. For example reaction $3O_2 + 4\text{GaAs} \rightarrow 2\text{Ga}_2\text{O}_3 + 4\text{As}$ will create additional metallic or elemental As at the interface in addition to trivalent Ga oxide. In practice this is seen as additional point defects arising from excess of As bonds. Indeed, the persistent FLP observed for GaAs has recently been suggested to be caused due to amphoteric nature of these As bonds, exhibited as saturated dangling bonds with an excess of electrons and as dimer bonds with excess of holes, both giving rise to midgap states for GaAs [40, 41, 70]. Despite being quite volatile, loosened group V atoms tend to get trapped underneath the III–V/oxide interface [67–69], which means rise of defect states considered responsible for the FLP at the interface without proper chemical passivation.

Chemical and structural effects of ALD produced films on III–V surface/interface have been widely under research in the recent years [52, 54, 67, 71–77]. The general consensus is that while the interface chemistry of ALD grown films offers prominent benefits, there are still various detrimental effects to consider in order to attain low interface defect densities. These include \textit{e.g.} out-diffusion of group III atoms compromising structural integrity [54, 74], formation of elemental/metallic group V species [67], and border traps within the film [78]. Furthermore, deposition and substrate preparation parameters need to be optimized in order to produce chemical species with inherently as low a defect state density as possible at the interface [67, 76, 79]. Models regarding the exact chemistry at the surface during film deposition are still being investigated and revised [77] and thus, there is still significant progress to be made for a more perfect understanding and control of the interface chemistry and structure.

It is considered that low oxidation state oxides of group III atoms at the interface produce a more ideal, even bulk-like [80] bonding and charge distribution, causing much less deterioration of carrier mobility or band structure than the corresponding high oxidation states [81, 82]. Many approaches have been used to create a layer between an oxide film and semiconductor that suppresses the natively occurring reactions at the interface, using the low oxidation state group III compounds, passivating interfacial layers, or \textit{ex situ} treatments with \textit{e.g.} sulfides [xvii, 54, 73, 80, 83–85]. Benchmark passivation method for GaAs(100) is UHV deposition of Ga$_2$O$_3$(Gd$_2$O$_3$) (GGO) oxide film that grows epitaxially in the first few monolayers with a very low $D_{it}$ of $< 1 \times 10^{11} \text{cm}^{-2} \text{eV}^{-1}$ [86]. The termination of bulk with appropriate chemical character and epitaxial fashion, \textit{i.e.}, preserving crystallinity at the interface is observed as of predominant importance [79].
2.3.3 Formation of Crystalline Oxides on III–V’s

A simple, yet highly efficient method to passivate III–V semiconductor surface with a crystalline oxide structure having the desirable characteristics [21] similar as for GGO was discussed previously (chapter 1). This method utilizes UHV conditions (see chapter 3) so that sample surface can be exposed to pure oxygen in a highly controlled manner. A significant advantage of creating such structure is in the requirement of simultaneous annealing of the semiconductor at above roughly 350°C. Evaporation of metallic As and Sb in UHV occur at this temperature due to their high vapor pressures. Hence, during crystalline oxide termination the surface is left free from much of the significant excess bond point defect species. Such species are thus easily avoided during subsequent thin film growth or other post treatments. In addition to inherent property of this method of avoiding defect state causing species, a stable oxide is formed on the surface. Therefore, excessive oxidation that could form defect states during subsequent film growth or brief atmospheric exposure between treatments can be avoided [21]. A potential improvement, or reduction in FLP using this structure is shown in Fig. 2.6 (a) and an example of crystalline In$_2$O$_3$ on InP atomic model in (b).

**Figure 2.6:** Schematic diagram of the effects of a passivating interfacial oxide layer on $n$-type semiconductor, reducing defect states and FLP (a) and an atomic model simulating such a structure (b).

**Thermal Oxidation of III–V in O$_2$ Ambient**

Other possible reactions will need to be also considered to examine the general applicability of such treatments; several mechanisms of interaction
2.3. Oxide–Semiconductor Stack

with O$_2$ are possible, depending on the experimental conditions. For example, it has been observed for Si(100) and (111) surfaces oxidized in elevated temperatures that oxygen can insert into the surface back-bonds, initializing subsurface oxide growth, or alternately oxidation can result in etching of Si due to creation of volatile suboxides such as SiO. These effects have been observed in UHV conditions and have a prominent dependence on the parameters used for treatments; i.e. different mechanisms observed at different regions of the whole parameter space: temperature (500 °C to 800 °C), pressure (1 × 10$^{-8}$ mbar to 1 × 10$^{-5}$ mbar) and oxidation time (i.e. exposure) [87–92]. For instance, while 1 × 10$^{-7}$ mbar O$_2$ pressure with annealing temperature of 650 °C etches Si(100) surface, similar treatment with 1 × 10$^{-6}$ mbar and 600 °C results in oxidized and in this sense passivated surface without prominent etching [93].

The effects for Si discussed above will also need to be taken into account for III–V materials. In investigations related to GaAs(100) oxide film dissociation it has been observed that As$_2$/As$_4$ and Ga$_2$O species start to desorb from the oxide film at temperatures of approximately 390 °C and 475 °C, respectively [94, 95]. Thus, the reaction mechanisms resulting in the compounds in Table 2.2 could also be prevented by loss of other species required in the reaction (in analogy to SiO $\rightarrow$ SiO$_2$). Namely, in the case for GaAs(100) oxide dissociation, As$_2$O$_5$ is not observed but Ga$_2$O is even though As$_2$O$_5$ has lower $\Delta G$. The combined effect of energetically favorable Ga$_2$O$_3$ formation and on the other hand tendency of low $\Delta G$ Ga$_2$O and As species to desorb suppresses further oxidation of As.

Ga$_2$O and In$_2$O are relatively volatile compounds compared to higher oxidation state bulk oxides or semiconductor crystals [80, 94, 96, 97]. Therefore, in conditions that could result in breaking of O-bonded Ga or In bonds to the semiconductor, desorption is likely to occur. This is particularly relevant in the case of oxidation carried out in elevated temperature, since the reaction chain might proceed through formation of such species. For example, in case of GaAs, if Ga is removed from the surface at elevated temperature as Ga$_2$O, the exposed As-rich surface layer will be also depleted of As as described previously, and thus, a net etching effect could occur.

Sticking probability of O$_2$ on different surface planes of GaAs is fairly similar and increases with temperature [98–100]. Furthermore, it has been observed for e.g. GaAs(100)c(8×2) that oxidation in elevated temperatures results much more prominently in dissociative adsorption [95, 101]. Thus, it can be concluded that observation of O$_2$ induced interaction at the surface in RT should indicate significant alteration in bonding at the surface in elevated temperatures. These issues are considered in this work.
III–V Crystalline Oxide Properties

Crystalline oxides through direct vacuum-based oxidation have been shown on AlGaN [102], GaN [103], InGaAs [76], InP, InAs, InSb, and on GaAs [21] and GaSb [1] through an intermediate In-deposition induced surface structure [104]. The present knowledge of III–V crystalline oxides has been recently briefly summarized in the introduction part of Ref. [xxx], which discusses the exact chemistry of $\text{Al}_2\text{O}_3$/crystalline oxide/InAs(100) interface and conclusions that should be drawn from its PES data.

Different reconstructions for the structures discussed can be observed depending on the process parameters, which is often interpreted as different O concentrations of the oxide [21, 76, 105]. In particular, atomic hydrogen treatment on InGaAs(100)(3×1)–O can shift the reconstruction to (3×2) [76], which could be related to changes at outermost bonds, while subsurface O remains (no (3×2) has been reported for clean InGaAs(100) surface). Indeed, several other studies also indicate [xii, 13, 105, 106] that such crystalline oxide structures extend to subsurface layers, where each O atom finds most group III atom bonds and thus, the most thermodynamically favorable bonding sites.

The benefits of the crystalline oxide structure in terms of low defect state density [22, 76, 102, 107] can be understood in terms of smaller disruption in the bulk periodicity compared to a totally amorphous interface structure, such as that of e.g. $\text{Al}_2\text{O}_3$. On the other hand, a more abrupt shift from bulk semiconductor to oxide [103] could work as inducing smaller probability for formation of defects. Thirdly, group III oxide bonds are favored in the structure, so that its chemistry induces less defect-state causing species. In particular, high oxidation state group V oxides observed after the oxidation treatment [75] could be a non-inherent property of the structure [xxx], and could likely thus be avoided with proper process parameters.

The stability of the structure [21] and benefits (along with crystalline interface nature) acquired that persist even after ALD growth [22] can be interpreted to be a complementary indication of the structure to really extend to subsurface; even if the topmost reconstruction is deteriorated during, e.g., ALD film growth [108], the subsurface passivating interface remains and protects the semiconductor crystal from defect state formation.

The investigations carried out on III–V crystalline oxides have mainly concentrated on (100) surface planes. However, taking into account the considerations discussed in the preceding sections, it is not at all apparent that similar outcome should be expected from similar oxidation of other surface planes of the same materials, which is crucial for utilization in non-planar device structures. This has been addressed for InSb in paper IV.
3 Experimental Methods

The methods used in this work for the investigation of materials mainly consist of surface science research tools. The quality and characteristics of clean semiconductor surfaces and after their well defined treatments with atomic precision are characterized with high reliability with these methods. After further processing alternative methods are required in order to visualize and understand the quality, performance and their origins from semiconductor device structures. The experimental methods utilized in this work are described in this chapter.

3.1 Ultrahigh Vacuum

Surface science tools in general require very clean working conditions, because i) the sample surface must not react with the ambient, or contamination rate needs to be very slow, ii) many of the methods rely on particle beams or uncovered delicate instruments that must not be interfered by the ambient. These requirements make it essential to remove atmospheric gases from the experimental setup.

3.1.1 Vacuum Ranges and Exposure Magnitudes

The concept of vacuum is generally applied to any enclosed volume with a pressure less than the atmospheric pressure, but the ranges are usually subdivided into rough, medium, high-vacuum (HV) and UHV ranges. UHV range is defined as volume with pressure less than $1 \times 10^{-7}$ mbar [109]. In typical modern UHV chambers, background pressures as low as $1 \times 10^{-11}$ mbar to $1 \times 10^{-10}$ mbar are reached.

Exposure experienced by the sample as a quantity is defined generally as experimental conditions that correspond to an impingement of a monolayer unit of adsorbate molecules on the solid sample surface from the gas phase, assuming a sticking probability $S$ of unity. $S$ is defined as the ratio of the amount of molecules adsorbed on a surface in a bound state to the total amount of molecules hitting the surface. It is straightforwardly seen
that in conventional experimental conditions [110] an exposure corresponding to one monolayer as described above (denoted Langmuir, 1 L) is achieved with a product of exposure time and pressure of \(1 \text{ L} = 1 \times 10^{-6} \text{Torr} \cdot \text{s} = 1.33 \times 10^{-6} \text{mbar} \cdot \text{s}\). The amount of Langmuirs in exposure therefore represents the upper limit of molecules adsorbed on the surface. For example, an exposure with a pressure of \(1 \times 10^{-7} \text{Torr}\) for 10 s results in a monolayer coverage (\(1 \times 10^{-7} \text{Torr} \cdot 10 \text{s} = 1 \times 10^{-6} \text{Torr} \cdot \text{s} = 1 \text{ L}\), if \(S = 1\), and sub-monolayer coverage if \(S < 1\). In practice, \(S < 1\), and usually it drops down as a function of exposure. Common initial \(S\) for gases found as impurities or reactants, such as H\(_2\) and O\(_2\) molecules, on atomically smooth III–V semiconductor surfaces are typically in the range of \(1 \times 10^{-5}\) to \(1 \times 10^{-3}\) [100], and thus \(1 \times 10^{-10} \text{mbar}\) range is more than sufficient for daily experiments on a freshly cleaned sample. In this pressure range, even if a sticking probability of 1 is assumed, it takes hours or days for a monolayer of surrounding gas molecules to adsorb on a sample [36, 111]. However, experimental conditions such as excitation sources like irradiation, Bayard–Alpert ion gauges [99] or sample temperature can have a significant effect on \(S\) and also the mechanisms involved [36, 101]. Minimum exposure needed for monolayer coverage is typically in the range of \(> 1000 \text{L}\) in RT. In HV range, this amount is reached within minutes or seconds.

### 3.1.2 Pumping

UHV is achieved by evacuating a closed volume with vacuum pumps. \(1 \times 10^{-10} \text{mbar}\) vacuum can be readily achieved using turbomolecular pumps, which have rotating wings that create collision paths for incoming molecules out of the vacuum volume. Turbomolecular pumps have a mechanical construction and they require a roughing pump to create a \(1 \times 10^{-3} \text{mbar}\) to \(1 \times 10^{-1} \text{mbar}\) vacuum in the outlet side for operation. This is typically done with a rotary-vane or dry scroll type roughing pump that operate by mechanically pressing a gas volume out in a cyclic fashion. To uphold the UHV and have a completely closed volume for a case of e.g. electrical breakdown, ion pumps are often utilized. Ion pump uses a high voltage and strong permanent magnets to create long trajectories for stray electrons in the volume. These electrons collide with stray molecules creating ions, which in turn hit and are immersed in the walls of the ion pump. The walls are typically additionally coated with Ti, that reacts with many gas molecules such as O\(_2\) and H\(_2\)O, completely removing their ability to desorb afterwards, which can happen with, e.g., noble gases such as Ar (memory effect). Ion pumps are typically equipped with Ti filaments so that
the coating can be refreshed by sublimation through running a high current, and thus, a high temperature in the filaments. [109, 111]

3.1.3 Pressure Measurement

In order to verify absence of impurity gases during and between measurements, as well as control the gas exposures, the chamber pressure needs to be constantly measured. In UHV, Bayard–Alpert (hot-filament) ion gauges are typically used, and this is the case in our UHV systems as well. In such a gauge, electrons are detached from a hot filament and accelerated towards a positively biased grid, which most of them bypass. The electrons ionize gas molecules which are collected at a wire at the center of the grid due to the negative potential in reference to the surrounding grid. The electric current resulting from electron flow, that compensates the positive charge of the collected ions, is measured and converted into pressure reading. The applicable range of Bayard–Alpert gauge is in the range of roughly $1 \times 10^{-10} \text{ mbar}$ to $1 \times 10^{-3} \text{ mbar}$ (depending on the filament current); too large pressure will break the filament, and on the lower end, accuracy is limited by the baseline caused by x-ray excitations due to photoelectric effect on the grid. [109, 111]

For higher pressure ranges, other types of gauges are necessary. Rough vacuum is often monitored with a Pirani gauge, which operates based on the variation in thermal conductivity of a wire element as a function of the ambient pressure. There are several types of operating principles with such gauges; in some gauges the temperature is kept constant while monitoring the heating power, or the heating power is kept constant while monitoring temperature or resistance/dissipated power in the wire. Typical pressures measurable with Pirani gauges range from $1 \times 10^{-3} \text{ mbar}$ to 100 mbar for sophisticated designs and from $1 \times 10^{-3} \text{ mbar}$ to 1 mbar for simpler ones [111].

3.2 Sample Treatment Facilities

UHV treatments and measurements for the investigated samples have been carried out in two separate systems, in which deposition tools, gas leak valves and research equipment are available.

3.2.1 Ion Bombardment, Sputtering

For cleaning semiconductor sample surfaces brought from ambient conditions, it is essential to have facilities for removing the stable native oxides without making significant roughening or corrugation on the surface
on atomic level. For this purpose, we have Ar$^+$ ion bombardment, \textit{i.e.}, sputtering gun which is used to efficiently remove atoms and clusters from the surface in a roughly layer-by-layer fashion. Methods such as secondary ion mass spectroscopy (SIMS) rely on depth profiling by analyzing the secondary ions produced by sputtering, and literature on such methods contains extensive descriptions of sputtering as a phenomenon [112]. In our sputter-cleaning, Ar gas is brought into the gun and ionized by thermal means using a filament, or, as in a cold-cathode by accelerating stray electrons inside a magnetic field which collide with Ar atoms producing an ion flux like an avalanche. Ions are then accelerated towards the sample with a 0.5 kV to 5.0 kV voltage, that gives high enough energy for the ions to remove surface atoms or small clusters. To be exact, the surface is left quite rough after only sputtering at RT. For mobilizing surface atoms to flatten the corrugations and relax the structure to atomic smoothness, sample is most often annealed after or also during the sputtering treatments (ion bombardment and annealing (IBA) or simultaneous ion bombardment and annealing (SIBA)). Annealing temperature depends on the sample material and possibly the surface reconstruction that is to be investigated.

A prolonged sputtering can have a marked effect on the surface corrugation on the atomic level (see \textit{e.g.} paper IV), and sputtering and annealing parameters need to be carefully optimized for optimal sample preparation conditions. For this reason, other methods have been used for producing atomically smooth real device structure surfaces in our studies (see section 4.4.1). Sputtering is used for cleaning surfaces whose properties were investigated in particular locally with \textit{e.g.} STM, or sputtered in a very similar fashion when doing comparison between a series of samples to suppress the effects arising from sputtering only.

### 3.2.2 Sample Annealing

Sample annealing is carried out by passing electrical current through an either tungsten (W) filament or pyrolytic boron nitride (PBN) heating element that is mounted to the sample stage. The heat is then transferred radiatively and by thermal conduction to the sample stage, sample holder and the sample itself. Temperature of the sample is controlled by checking the emitted IR spectrum with a pyrometer that converts it into temperature reading and the annealing power adjusted accordingly to achieve the desired temperature. This way, sample temperatures of up to approximately 600 $^\circ$C are achieved. Si and Ge samples are commonly annealed by passing a DC current through the sample [xv–xviii]. For Si, sample temperatures of above
1200°C are achieved this way, which is enough for very thoroughly decomposing and evaporating impurities and native oxides from the surface. Lower temperatures can be used to smoothen the atomic step-terrace structure on the surface by reducing the DC current.

### 3.2.3 Deposition

In order to investigate effects of different adsorbate layers on semiconductor surfaces, or to use them as templates for further treatments, deposition tools are required. Our tools for evaporating and depositing metals onto surfaces consist plainly of physical vapor deposition (PVD) modules with electrical UHV feedthroughs that have self-assembled W or Ta filament wrapped around a piece of metal such as Sn, Mg, In or Sb. Metals that have a high vapor pressure at relatively low temperatures are conveniently evaporated with such methods by just conducting current through the filament that brings heat to the metal by thermal conduction (where it touches the metal) and radiatively from outer parts of the filament. It is important to wrap a long enough wire so that heat is resistively produced in the wire, since too small filament will result in current flowing efficiently through mostly the metal piece, or also, significant proportion of the heat leaking into the feedthrough contacts, and thus non-efficient resistive production of heat. In the case of Sb, for example, efficient evaporation is achieved below its melting point, 631°C, and such materials are particularly suitable for such evaporators. However, for example In and Sn on the other hand have lower melting points (157°C and 232°C, respectively) than where efficient evaporation occurs, but such a high surface tension that molten metal is also well contained within a densely enough wrapped filament. Fig. 3.1 shows some vapor pressure versus temperature curves for low and high vapor pressure materials, that are due to this property particularly badly or well suitable as UHV materials, respectively, and some that are otherwise of specific interest concerning this work [113].

### 3.2.4 Gas Exposure

A significant advantage of using UHV is that the sample can be exposed to any gas with a line connected to the chamber in a highly controlled manner. This is achieved using a leak valve, which contains a hard needle pressed against a softer seal material with a pinhole, achieving an UHV seal when closed. When opened manually using lever with a very low transmission, a small gas leak can be introduced using high purity gas on the inlet side. The chamber is typically simultaneously pumped with a turbomolecular pump. The resulting equilibrium pressure is dictated by the net flow of gas molecules
Figure 3.1: Vapor pressures as a function of temperature for selected elements. [113]
from the leak valve, which is observed as a remainder pressure in the chamber due to non-instantaneous pumping. The extent of exposure can be chosen by using \textit{e.g.} higher pressure by opening the leak valve more, or selecting a longer time for exposure at a given pressure.

### 3.2.5 Atomic Layer Deposition

Semiconductor device structures very often require high quality insulating and/or protective films as a functional part of the device. ALD offers an ideal method for these purposes. ALD was originally invented by Suntola \textit{et al.} for synthesis of pinhole-free thin films that could withstand high electric fields in electroluminescent displays. The groundbreaking work was carried out from late 70’s to early 80’s [114–117], back when the method was known as atomic layer epitaxy (ALE).

ALD is a process where precursor molecules in gas phase are brought into the substrate containing volume as a pulse, so that they react with the surface (1st ‘half-reaction’), but not with each other. This treatment ensures a complete monolayer coverage of molecules, as long as the pulse time is enough to saturate the surface with the precursor. The gas remaining in the volume is then pumped or flushed (purged) with inert gas, such as $\text{N}_2$, and another precursor gas is brought into the volume. The molecules of the second precursor gas react strongly with the molecules left from the first precursor pulse on the surface (2nd ‘half-reaction’), and form a second complete monolayer of molecules. The remaining gas is again pumped or purged, and the first cycle is then complete with a single layer of constituent substance covering the surface. This sequence is repeated a desired number of cycles to form the desired layer thickness with atomic scale control of thickness due to the self-limited individual reactions [51, 118, 119]. A diagram of an ALD process is shown in Fig. 3.2.

Reaction irreversibility (desorption constant equals zero but adsorption constant does not), equal probability adsorption sites, self-limitedness, and lack of reactivity between similar molecules lead into desirable characteristics found in ALD [119]; equilibrium coverage of precursor molecules on the surface equals one irrespective of the pressure. Total coverage depends on typically relatively high adsorption constant, pressure, and time of exposure. To speed up the deposition process, magnitude of adsorption constant can be increased with temperature (according to the Arrhenius equation), but additional activation methods have also been examined [51, 118, 119]. The deposition parameters are controlled in such a way that ideally a single pulse always should result in a 100\% coverage, after which purge and the following pulse can be carried out.
Figure 3.2: Typical ALD process: 1) first pulse with precursor gas 1 and ‘self-cleaning’, p) purge and pumping of residual gas and volatile reaction products with inert carrier gas after layer saturation, 2) pulse with precursor gas 2, 3) pulse with precursor gas 1. Single cycle consists of steps 3), p), 2), p), and atomic resolution film thickness is achieved by repeating a predefined amount of cycles.

In order to deliver precursor gas to the sample, evaporation from liquid phase to a significantly higher pressure than the background in the chamber is necessary. For the Al$_2$O$_3$ deposition used in this work, precursor gases H$_2$O and trimethylaluminum [Al(CH$_3$)$_3$] (TMA) have high enough vapor pressure so that they are spontaneously evaporized in RT to over 10 mbar [120, 121] and transported to the surface conveniently in gas phase. Other precursors, such as tetrakis(dimethylamido)hafnium(IV) [Hf(N(CH$_3$)$_2$)$_4$] (TDMAH) for HfO$_2$ and tetrakis(dimethylamido)zirconium(IV) [Zr(N(CH$_3$)$_2$)$_4$] (TDMAZ) for ZrO$_2$ require heating to about 50°C to 80°C at the precursor bottle and inlet lines for 0.1 mbar to 1 mbar pressure, which has been observed as high enough for efficient growth [122].
3.2. Sample Treatment Facilities

The very nature of the ALD process gives the deposited films various beneficial properties: uniformity, conformity, highly thickness-controllable and pinhole-free film. In addition, the materials that can be deposited are dictated mainly only by selection of gases that can be used as precursors, and the deposition of a vast variety of materials is enabled, from insulators and metals to both crystalline and amorphous semiconductors. The properties such as permittivity, band alignments, conductivity, crystallinity and lattice constant, etc., necessary for a particular application is likely to be found among the reaction products of a vast selection of precursors, which are being intensely researched and developed. One of the main limitations of ALD technology concerning industrial utilization is somewhat strict temperature window for depositing any particular film. The growth rate in general is often another limitation, with typical deposition rates at about $2 \text{ nm min}^{-1}$ to $5 \text{ nm min}^{-1}$. [51, 118, 119]

The general model of ALD mechanism has been discussed above, but this will only apply in more or less limited conditions. For example, too low sample temperature can give rise to granular growth and too high result in desorption and decomposition, or, too low growth rate by altering the reactive surface sites. Different temperature can even change the inherent reaction mechanism, and, suitable precursors are needed to avoid the decomposition. Also, steric hindrance is a significant factor limiting the growth rate, since large molecules readily prevent adsorption of additional molecules into reactive surface sites by blocking their path, complicating the process. To be exact, the growth is very rarely monolayer-by-monolayer type (growth mode can even change during deposition), and after a low number of cycles, the surface chemistry is changed as the substrate surface sites are saturated, which can also affect the growth rate, and it commonly does. Taking the points noted (and more detailed discussion in e.g. [119]) into consideration, it is obvious that careful calibration is required for any ALD system. We have utilized H$_2$O and TMA as precursors since the ultimate quality of Al$_2$O$_3$ produced is well-known and an ideal-like, not too sensitive process is established in terms of processing parameter variation, and thus, the parameters are fairly easily considered and calibrated. Also instrumental factors such as chamber sidewall or heating filament temperatures can be calibrated when the optimal properties (uniformity, stoichiometry, growth rate) of the end-result are well-known. [119]

When transitioning from 65 nm to 45 nm technology node in integrated circuit (IC) and microelectronics industry, ALD, having established capability to deposit high-$\kappa$ oxides, was generally introduced in order to replace SiO$_2$ with HfO$_2$ as the gate oxide, when smaller equivalent oxide thickness was necessary but oxide scaling was started to impose major challenges. The
FinFET technology, first implemented at 22 nm technology node, required deposition of uniform, conformal, and pinhole-free oxide on high-aspect-ratio structures, which was a well-tailored task for ALD. Furthermore, the shrinking of dimensions has made relatively slow deposition, one of the drawbacks of ALD technology, quite redundant. The 3-dimensional structures established in FET technology have solidified the state of ALD as an industry standard method. The future prospects of ALD also look bright, since higher permittivity materials suitable for gate oxides are currently being explored, and ALD oxides have been shown to be particularly suitable for alternative channel materials, such as Ge and III–V’s. Even problematic III–V native-oxides causing Fermi-level pinning are inherently reduced by some ALD precursors, due to the self-cleaning [77], as described in section 2.3.1. Other prospective applications benefitting from ALD include e.g. thin film solar cells and solid oxide fuel cells [118].

From an applied point of view, the significant III–V/oxide interface chemistry needs to be considered integrated into a structure with an ALD film, since it has become a solid industry standard. Furthermore, concerning surface science UHV methods: in our studies it is reasonable to proceed from surface investigations to interfaces via in situ ALD whenever foreseeable benefits are attained, since ALD is a vacuum-compatible method and can thus be integrated into existing UHV systems. Strong correlation between fundamental and practical aspects of semiconductor–oxide interface physics are attained this way.

3.2.6 Surface Science Setup

Fig. 3.3 illustrates the laboratory’s older UHV equipment used for deposition and gas exposure treatments, sample annealing, STM, LEED, and x-ray photoelectron spectroscopy (XPS) characterization, and ALD. Sample can be transferred in situ via UHV to the ALD chamber. ALD chamber is normally in UHV pumped with a turbomolecular pump, and with a dedicated dry pump meant for harsh processing conditions (Edwards iXH100 [123]) during deposition. Pressure in ALD reactor is monitored with a Pirani gauge. Sample heating is done radiatively, or by heating the whole chamber with baking tapes up to about 150°C. Precursor lines include H₂O and TMA for Al₂O₃, TDMAH for HfO₂, and TDMAZ for ZrO₂ deposition. N₂ is used as purging gas through a mass flow controller, giving a purge pressure of 1 mbar. The whole system is self-assembled and programmed for maximum flexibility. The experimental setup is based on Omicron UHV STM 1 system [124].
3.2. Sample Treatment Facilities

A new STM system with option for cryogenic cooling (Fermi scanning probe microscopy (SPM) [125]) was installed during the work carried out in this dissertation and all such STM experiments were done using this system. This UHV system is illustrated in Fig. 3.4. LEED and sample preparation utilities (Sb, Sn, and In evaporators, O₂ leak valve and sample heating stage) were installed adjacent to the Fermi SPM chamber and are available for in situ treatments. Such a preparation chamber was necessary to implement in order to roughly analyze the sample surface with LEED before more time consuming STM measurements, and to make tailored studies according to any specific requirements. Successful epitaxial growths and oxidations have been carried out as discussed e.g. in paper IV. The advantage of this chamber construction is also in its versatility; filament assembly evaporators are easily exchanged, additional components are easily installed, the sample can be cooled with liquid nitrogen (LN2) or radiatively annealed and/or exposed to adsorbents while monitoring pressure, temperatures, and LEED.
Chapter 3. Experimental Methods

Figure 3.4: Omicron Fermi SPM based system at Materials research laboratory, University of Turku. Preparation chamber and related instrumentation was constructed as a part of this PhD work.

voltage. Furthermore, a custom computer software has been programmed and implemented to record the data, read through voltmeters and serial connections, and automatize e.g. annealing while monitoring LEED.

Both of the systems utilize similar pumping, pressure measurement and other UHV tools that were introduced earlier in sections 3.1 and 3.2, providing a stable and reproducible environment for experiments and measurements. The research methods of the systems will be discussed in sections 3.3, 3.4, and 3.5.

3.3 Photoelectron Spectroscopy

PES is based on the photoelectric effect, discovered by Hertz in 1887 and formulated by Einstein in 1905, which brought him Nobel prize later on [126]. The basis for PES was established in 1950’s and 1956 the first XPS spectra were measured by K. Siegbahn and his group in 1956 [126]. Today
PES is one of the most widely used tools for characterization of materials elemental surface composition and chemistry. A vast variety of literature is available on PES and XPS in particular [127–133], for a more detailed descriptions of these methods. A general description as well as aspects of PES related to this work will be described here. The significance of PES regarding semiconductor heterostructure chemistry and thin film research has been paramount in the progress so far, and it is one of the most widely used methods for this purpose [134].

When a photon with an energy of $h\nu$ is absorbed by an electron in a material, the excess energy after excitation is converted into kinetic energy $KE$ provided that $h\nu$ exceeds the binding energy $BE$ of the core-level and the work function in the spectrometer $\phi_{sp}$. This is the basic principle of photoelectric effect for core-level electrons in a material and is formulated on the basis of energy conservation as follows:

$$KE = h\nu - BE - \phi_{sp}. \quad (3.1)$$

Since $h\nu$ and $\phi_{sp}$ can be kept constant, the $BE$s at which electrons are present in a given material are straightforwardly seen by measuring the intensity distribution in the $KE$ interval between $\phi_{sp}$ and $h\nu$. This way, a photoelectron spectrum is obtained as shown in Fig. 3.5 and since core-level $BE$s are characteristic for each element, compositional surface analysis can be carried out using these peaks in the spectrum as a fingerprint of the material. $\phi_{sp}$ is usually calibrated by measuring a metal piece (such as Au, W, or Ta) in electrical contact with the instrument and the sample, and setting the emission edge at VB (i.e., the $E_F$ of the calibration piece and the sample) to 0 eV. Using this approach, the peak energies are always referred to the surface $E_F$ of the material.

The resolution of conventional PES is usually good enough for identifying also chemical states, which cause core-level shifts (CLSs) in the range of 1 eV for the core-level $BE$s of a given element due to charge redistribution and concurrent change in the screening from valence electrons [135]. Tabulated values for the peak $BE$s of elements and CLSs of different compounds are found in literature [129, 136]. Theoretical aspects and computational studies can further justify the analyses of complex materials systems, but several assumptions, such as contributions from initial and final state effects, and possibly simplifications of the systems under study might be necessary, due to complex and heavy computational tasks [xxxii, 135]. Furthermore, the intensity of a single peak is proportional to the amount of atoms from which it originates, which enables quantitative analysis. However, when considering different elements and different peak $BE$s, their photoionization
cross-sections, inelastic mean free paths (IMFPs), measurement geometry, detection efficiency, analyzer transmission, and possibly also photon flux will need to be taken into account, and they are often implemented as atomic sensitivity factors in conventional XPS data analysis. This is necessary when calculating the relative concentrations of different elements and compounds on the surface [129].

Basic principle of PES or XPS measurement is shown in Fig. 3.5. An anode material, typically Mg or Al, in an x-ray tube (1) is used to produce Mg or Al Kα x-rays (2) that are focused on a sample (3). Electrons (4) are emitted from the sample surface and collected into an electrostatic lens system (5). Lenses focus and accelerate or decelerate the electrons so that a given $KE$ signal intensity can be measured individually step by step. A single data point is acquired by filtering electrons through the HSA (6) with a given pass energy, which can be adjusted to tune intensity or resolution of the acquired signal [128]. The filtered set of electrons hit the multiplier tube (7), which gives the signal that is passed through the amplifier (8) and measurement electronics to the computer control unit (9).

Many other features besides photoelectron peaks contribute to the total

---

**Figure 3.5**: PES setup: x-ray tube (1), photon flux (2), sample (3), emitted electrons (4), electrostatic lenses (5), hemispherical analyzer (HSA) (6), multiplier tubes (7), signal amplifier (8), measurement electronics and computer (9)
3.3. Photoelectron Spectroscopy

spectrum that is measured. These features are related to interactions between the photoelectron or the excited atom and the surrounding valence charge of the material. Prominent peak features can arise due to Auger emission, which occurs if the core-level is filled with an electron from a higher energy level and this difference in energy is enough to excite another electron. Auger emission is a competing process with x-ray fluorescence, and more probable for lighter elements for each corresponding relaxation [137]. Due to the nature of Auger process, and in contrast to the photoelectron emission, the kinetic energies of Auger electrons are constant irrespective of the energy of the exciting radiation.

Electrons can lose their kinetic energy in a solid material in several ways. These effects are seen in the photoelectron spectrum as peaks or continuous distributions of intensity at lower kinetic energies from photoelectron peaks, corresponding to the energies that photoelectrons lose. Multiple energy loss events can occur for a single photoelectron, and thus multiple features with similar intervals in the spectrum can be observed. Some of the most important energy loss mechanisms are listed below [129].

- **Shake-up** process is related to excitation of a valence electron to a higher valence state from interaction with a photoelectron. These features are observed as prominent peaks or a continuous distribution depending on the band structure. In a **shake-off** process the valence electron leaves the host atom and thus, can acquire a continuous amount of energy. Shake-off satellites are therefore seen as much broader features. (energy losses from few eV’s to tens of eV’s)

- **Inelastic scattering** (0 eV–KE with KE corresponding to kinetic energy of photoelectron peak)

- **Surface and bulk plasmon excitation** features correspond to energies of the oscillation quanta of free electron plasma in a material. Due to the requirement of free electron plasma, plasmon features are prominently observed with metals or heavily doped semiconductors [129] (~10–50 eV) Surface plasmon features in photoelectron spectra can also be used for investigating surface nanostructure [138].

It is these energy loss features and mainly the inelastic scattering in general that makes PES such a surface sensitive technique. The kinetic energies of electrons in traditional PES correspond to IMFPs (distance the electron propagates on average without inelastic scattering) such that photoelectron peaks in the measured spectra correspond to electrons from the topmost atomic layers, approximately few nanometers from the surface. This is illustrated in Fig. 3.6. The attenuated differential proportion of the signal
intensity \( \frac{dI}{I} \), \( i.e. \), electron flux from a given depth is equal to the proportion of attenuation length \( \lambda_{al} \) it has propagated, \( \frac{dx}{\lambda_{al}} \):

\[
\frac{dI}{I} = \frac{dx}{\lambda_{al}},
\]

which is the definition of \( \lambda_{al} \). Integrating from a depth \( 0 \) with intensity \( I_0 \) to the surface with intensity \( I \) having propagated distance \( z \) leads into the well-known Beer–Lambert equation:

\[
I = I_0 \exp(-\frac{z}{\lambda_{al} \cos \theta})
\]

for electrons coming out of the sample at angle \( \theta \) with respect to the surface normal. From (3.3) one can see that from below \( 3\lambda_{al} \), we see only 5\% of the unattenuated signal with normal emission angle and homogeneous sample. \( 3\lambda_{al} \) is generally denoted the probing depth in PES. In literature, the \( \lambda_{al} \) discussed here is denoted attenuation length, effective attenuation length (EAL) or IMFP, that are slightly differently defined. It is most precise to use EAL, that takes into account also elastic scatterings and gives the exponential dependence of attenuation by its definition, unlike IMFP.
In practice, it is sometimes unpractical to use EAL for analysis as it depends on the overlayer thickness. In our investigations, it is sufficient to rely on IMFP quantity for qualitative approach, but it is to be noted that IMFP overestimates the actual attenuation length in practical measurement geometries by approximately 10% to 25% [139]. In general, attenuation coefficients should be used as guidelines rather than exact due to some lack of precision in any case [127]. First of all, although the IMFP–KE relation is denoted “universal”, it is easily observed from Fig. 3.6 that the attenuation is somewhat material dependent, and individual values should be computed (from e.g. TPP-2M equation [140]) instead of relying on the general trend. Inhomogeneous materials system could make the penetration depth estimation immediately much more complex, increasingly so if there is any intermixing of materials layers in the system. Furthermore, especially on the lowest IMFP values that are used for most surface sensitive analysis, the obtained measurements are most prone to errors due to non-homogeneous films (not grown exactly in layer-by-layer fashion assumed by the model). There are also material dependent effects that the attenuation model does not take into account, such as electron diffraction. However, the model can give valuable estimations when properly applied and taking these limitations carefully into account.

### 3.3.1 Laboratory Sources

The main drawback of conventional x-ray sources is the lack of high enough energy resolution for distinguishing the most delicate CLSs and surface core-level shifts (SCLSs) in the XPS spectra. For the most widely utilized XPS x-ray anode materials, Al and Mg, the Kα1,2 x-ray radiation peak energies are 1486.6 eV and 1253.6 eV and their natural widths are 0.85 eV and 0.70 eV, respectively. As a consequence, no narrower peaks than this can be obtained, no matter how good the instrumental resolution. This limits the analysis concerning separation of chemical states on the surface, since subtle SCLSs or chemical shifts can be less than e.g. 0.40 eV for semiconductor materials [II, 142, 143].

Concerning surface sensitive chemical identification, in particular for Ga, As, In and Sb, the 3d and 4d photoelectron peaks are the narrowest ones and fairly sensitive (i.e., intense) in the energy range of Al Kα and Mg Kα [129]. The peak width is usually smallest for the most shallow core-levels due to their relatively long core-hole lifetimes according to the Heisenberg uncertainty principle, which is why they are well suitable for CLS analysis. These peaks result in KEs of 700 eV to 1500 eV and thus IMFPs close to 1.0 nm (Fig. 3.6). Estimating from Eq. (3.3), 95% of the signal results
from the topmost 3 nm whereas topmost 0.5 nm, approximately 1 monolayer (ML) of atoms, gives about 40% of the signal.

Additional hindrance may arise due to i) x-ray ghost lines of \( \text{e.g.} \) O K\( \alpha \) radiation if the anode material is oxidized or otherwise contaminated, ii) x-ray satellites characteristic for each anode material caused by minor K\( \alpha_3 \), K\( \alpha_4 \), ..., and K\( \beta \) peaks characteristic at higher photon energies of 8 eV to 70 eV and intensity of 0.5% to 8.0% of the K\( \alpha_{1,2} \) main peak, iii) overlapping of Auger peaks, especially with single source x-ray guns \cite{129}. Monochromators can be used to diminish i) and ii), and to obtain narrower linewidth, but this comes with a price of significant losses in intensity.

### 3.3.2 Synchrotron Sources

To overcome many of the problematic issues described above, PES capabilities offered by synchrotron sources provide an elegant solution. Synchrotron radiation is produced by altering trajectory of high energy electron beam, which causes electrons to lose energy that is emitted as photons that are focused onto beamline endstations for experiments. Up to GeV range electron beam energies (electrons with near \( c \) velocity) are achieved by linear accelerator and subsequent booster ring construction, after which electron beam is injected into storage ring in which the energy is maintained by radio frequency (RF) electric field cavities. An example setup of a synchrotron storage ring and other facilities are illustrated in Fig. 3.7. 3rd generation synchrotron sources today are capable of producing radiation millions of times brighter than the sun (with undulators and bending magnets) and many more orders of magnitude brighter than conventional laboratory x-ray sources. Synchrotrons with 4th generation free-electron lasers are pushing this limit even further \cite{144}.

The immediate benefits of using synchrotron radiation in PES for solid samples are: 1) tremendous increase in the intensity of the photon flux, \( \text{i.e.,} \) orders of magnitude higher amount of excitations, producing very high signal to noise ratio with much shorter measurement times, 2) narrowed full width at half-maximum (FWHM) of the photon energy down to meV range through the use of undulators and monochromators (typical resolving power \( E/\Delta E = 5 \times 10^3–10^5 \) with \( h\nu = 100–1000 \text{eV} \) \cite{145}), 3) tunability of photon energy to i) achieve desired probing depth, or ii) alter to higher cross section and thus intensity of the core-level signal measured \cite{146,147}, or iii) obtain 3-dimensional band structure probing capability in ARPES. Other possibilities of synchrotron beamlines include time-resolved PES measurements due to short pulse lengths (10 ps range) and intervals (1 ns range) in electron bunches, photoelectron microscopy with coherent light, measurement of
orbital symmetries or molecular orientations of adsorbate molecules by utilizing the beam polarization [133, 148]. Synchrotron PES has been in a vital role in the recent developments regarding spatially resolved XPS, with the advantages of high brightness and resolution. The imaging modes of PES can be divided into photoemission electron microscopy (PEEM) or ‘spectroscopy’, in which the spatial distribution of photoelectrons is separated with electrostatic lenses, and scanning photoelectron emission microscopy (SPEM) or ‘microspectroscopy’, in which a very narrow x-ray photon beam spot is scanned over the sample and a spectrum is saved from each spot. Spot sizes down to 30 nm are standardly utilized in synchrotron sources, but even as small as 10 nm spot size has been achieved [133].

In addition to 2), also 1) and the intensity increasing factors in 3) contribute to the ultimate resolution achieved, since the high intensity reduces the need to use signal broadening high pass energies in the HSA of the endstation. Overall, synchrotron PES beamlines have an unsurpassed resolution giving prominent signal suitable for investigating, e.g., very thin structures
such as surface reconstructions or chemical effects induced by molecules with submonolayer coverages [149], or very early stages of atomic processes occurring during oxidation of surfaces, for example. Taking the factors discussed above into account, it is not surprising that synchrotron radiation is very widely utilized in the research of semiconductor heterostructures and thin films [134].

An important aspect regarding reactions at device interface is that they are not located on the topmost surface of the sample or device. Therefore as seen from Fig. 3.6 and Eq. (3.3) the probing of interface for traditional laboratory or synchrotron source energies (20 eV to 1500 eV) becomes close to impossible if topmost layer (e.g., oxide) thickness exceeds few nanometers. However, for probing bulk properties and buried interfaces with PES, hard x-ray photoelectron spectroscopy (HAXPES) with x-rays in the range of up to 10 keV offers a viable alternative [141, 150] and such facilities are nowadays available at some >3 GeV synchrotron facilities, such as HIKE endstation of KMC-1 beamline at BESSY II in Germany [151, 152].

3.3.3 PES Data Analysis

The natural shape of photoelectron peaks originates mainly from Lorentzian broadening of the excited core-hole lifetime and Gaussian broadening dictated by experimental conditions such as sample inhomogeneity, incident photon energy FWHM and instrumental precision [133]. As a consequence, the measured photoelectron peak shape is a convolution of Lorentzian and Gaussian peaks with total FWHM

$$\Delta E = \sqrt{\sum \Delta E_j^2},$$

(3.4)

where $\Delta E_j$ is the FWHM of an individual source of broadening. The total convolution is known as a Voigt-profile. Although such convolution gives the most exact basis for fitting photoelectron spectra, it requires heavy numerical computation due to lack of exact analytic expression of such peak shape [153, 154], and many XPS fitting softwares omit this option. Instead, peak shapes formed of sum or product forms of individual normalized Gaussian and Lorentzian peaks are widely utilized [155]. In contrast to wide utilization of Gaussian-Lorentzian product form [154, 155], it has also been suggested that sum form of a Voigt-type peak gives more reliable results than the corresponding product form [153]. However, such contradiction can arise due to loose constraints in the peak fitting, and giving reasonable constraints
(obtained from theoretical values or reference measurements) for e.g. spin-orbit splittings (SOSs) and relative peak intensity ratios of double peaks such as 2p or 3d can make such differences negligible.

In order to fit peaks in a spectrum with a non-zero asymmetrical baseline, one has to initially subtract a background baseline from the measured signal. Unfortunately, there are no universal, straightforwardly applicable PES spectrum background subtraction algorithms with physical meaning, and so background shapes that result in correct physical characteristics of the measured peaks are widely utilized to simplify the task [156]. Such are also the Shirley and offset Shirley background shapes [156, 157] utilized in this work. In these algorithms, the background intensity level is limited at the raw data intensity from beginning and end of the predefined investigated energy interval. The intensity level at a specific point between these is then dictated by the peak area proportions before and after this point. A predefined baseline is set and the resulting baseline computed iteratively [156].

After background subtraction, a number of peaks that reproduce the features in the spectrum are fitted into the envelope raw spectrum. Constraints with physical justifications are used for each peak. These include: 1) Gaussian/Lorentzian emphasis of the peak shape, 2) FWHM, 3) \( BE \) shifts between given peaks, commonly utilized for e.g. double peaks which have a solid theoretical number for a \( BE \) shift (SOS), 4) intensities and intensity ratios, utilized similarly for double peaks (peak branching ratio). The physical constraints and base knowledge about the sample significantly simplify the fitting task and can result in a quantification far simpler than would meet the eye. For example, when there are several overlapping double peaks, the degrees of freedom are halved by binding the double peaks with the theoretical restrictions, and if the chemical states present along with their \( BEs \) are known, the fit could simply be achieved with only peak intensities as free parameters, and surface stoichiometry quantified on this basis.

### 3.4 Low Energy Electron Diffraction

A high quality semiconductor surface is typically an atomically abrupt and stable, crystalline sheet, such as were discussed more in detail in chapter 2. The surface crystallinity and its characteristics are conveniently characterized with LEED. Electron diffraction as a phenomenon was first discovered soon after the wave nature of particles, perfectly consistent with their de Broglie wavelength, providing further evidence for the discoveries of modern physics of its time [137].
In LEED, electron beam hits a sample surface in normal incidence, and elastically scattering electrons create interference maxima in the directions fulfilling the Bragg condition for diffraction. Elastically scattering electrons have the same energy as the incident beam, and thus also the same magnitude of wavevector;

\[
E_{hk} = E_{in} \Leftrightarrow \frac{\hbar^2 \|k_{hk}\|^2}{2m_0} = \frac{\hbar^2 \|k_{in}\|^2}{2m_0} \Leftrightarrow \|k_{hk}\| = \|k_{in}\|, \tag{3.5}
\]

where “in” denotes the incident beam and “hk” a diffracted beam. The spheres in Fig. 3.8 illustrate Ewald’s spheres that represent the general

![Graphical representation](image)

**Figure 3.8:** Basic working principle of LEED with illustrative photographs of measurement of clean InSb(110) surface. In (a), the incident beam energy \(E_{in,1}\) is smaller, and \((1,0), (-1,0)\) spots are the farthest ones seen in x-direction. In (b), where \(E_{in,2} > E_{in,1}\) \(\Leftrightarrow \|k_{in,2}\| > \|k_{in,1}\|\), more reciprocal space spots are visible on the same screen. The \((0,0)\) spot is not perfectly centered due to slight tilt in the sample setup.
3.4. Low Energy Electron Diffraction

diffraction condition for wavevectors in the 3-dimensional reciprocal lattice. In the 2-dimensional case, \textit{i.e.}, sample surface, when there is no bulk periodicity (or periodicity only in infinitely spaced dimensions) in the surface normal direction, the reciprocal lattice consists of continuous rods instead of discretely separated points in this direction. This means that the 2-dimensional reciprocal lattice vectors, \( G_{hk} = h\mathbf{a}_1^* + k\mathbf{a}_2^* \), up to the reciprocal dimensions defined by the incident beam wavevector magnitude are always shown as a projection from the Ewald’s sphere. This is utilized in LEED by using the spherical screen (in order to facilitate the Ewald’s sphere symmetry) to gather the diffracted beams. When beam energy is increased, more spots are observed due to more dense distribution of diffracted beams in real space. In this way, the electron wavevector mediates a picture of the 2-dimensional reciprocal lattice in the real space. [137]

Fig. 3.9 illustrates the concept of a conventional LEED apparatus, such as has been used in this work, too. In the electron gun, electrons are emitted from a hot filament, and accelerated with a selected beam voltage towards the sample. Electron beam diffracts, and diffraction maxima, representing the reciprocal lattice, are seen as bright spots on spherically curved fluorescent screen with curve radius corresponding to distance from the sample. Separator grid at the sample potential is used to keep electron beams at their designated trajectories, suppressor grid to filter out inelastically scattered electrons, and fluorescent screen biased with high voltage to efficiently detect the beams and convert the signal into light. [137]

![Figure 3.9: Basic working concept of a conventional LEED equipment.](image)

Coherence length of LEED, \textit{i.e.}, the part of the beam size where electron de Broglie waves are in the same phase, is typically in the range of few tens of nanometers [110], meaning that separate domains of at least this size can be detected as LEED patterns on the screen. This also means that there
are specific limitations concerning the measurement; bright spots can be observed even if there would be a significant amount of cluster-like disorder on the surface, and on the other hand changes in features with higher density than between coherence length could result some reconstruction spots to be observed diffuse or non-existent [158]. For example, for GaSb(100)(4×3), a (1×3) LEED pattern is observed due to uniform rows in 4×direction, but adjacent rows being randomly shifted with respect to each other [159]. However, if a LEED reconstruction pattern is observed, domains of at least coherence length exist, and brighter spots always means a higher coverage, i.e., degree of ordering [158].

3.5 Scanning Tunneling Microscopy

STM is among the most precise instruments for determining the topography and local electronic structure of the surfaces of solid state materials. The first successful STM scans were reported by Binnig et al. [160, 161] in (1982), showing the famous Si(111)(7×7) reconstruction in real space in atomic scale, soon after this [162]. This was an important milestone for atomic scale physics and engineering of semiconductor surfaces and interfaces. G. Binnig and H. Rohrer were awarded a Nobel Prize for the invention of STM in 1986, reflecting its importance on the materials science and applied research. A huge number of research papers related on STM studies of semiconductor surfaces have been published since [163], from the early observation of clean planar Si(111)(7×7) to investigations of complex nanowire structures of today [164, 165]. This section provides working principles and basic theoretical aspects of STM, again as is appropriate regarding this work, and literature is available [163, 166–168] for a more profound description of the method.

STM takes advantage of the quantum mechanical property of particles having a finite probability to move through a potential barrier higher than their kinetic energy, i.e., tunneling. At a static setup with the free electron model, for a low bias voltage \( V_g \) between two electrodes separated by an insulating layer (e.g., vacuum) of width \( z \), a current \( I_t \) is generated [169, 170]:

\[
I_t \propto \frac{V_g}{z} \exp(-2k_dz),
\]

where \( k_d \) is the decay constant in the barrier. Everything apart from \( z \), i.e., separation of the tip apex from the probed position, staying constant during scanning, one can easily see the exponential dependence of total current \( I_t \) on the separation \( z \). With higher bias, Eq. (3.6) takes a more complex form with \( V_g \) dependence [170]. However, as \( V_g \) remains constant in a conventional
scan, we are interested only in behavior of $I_t$ with respect to $z$, which retains its similar, exponential dependence as in (3.6) for practical range of $V_g$. In STM, $z$ is typically in the 0.1 nm (or 1 Å) to 1 nm and $I_t$ in the pA to nA range. Change of 0.1 nm in $z$ easily produces an order of magnitude change in $I_t$, with a typical barrier height of 5 eV [168].

STM equipment consists of a sample stage and a scanner head holding a sharp tip, with apex that ideally has a tip radius in the atomic scale. The system needs to be electrically and mechanically vibration isolated, typically hanging from springs and resting on an eddy current dampening magnet construction. The construction typically has rough piezo elements for sample approach and movement with respect to the scanner head, and fine piezo tubes at the scanner head for moving the tip in $x$, $y$, and $z$ directions. A particular tip height $z$ with respect to the sample is sensed by measuring a non-zero $I_t$ with a given bias $V_g$, fulfilling the tunneling condition preset at some value in the pA to nA range. While having the tip at a tunneling distance, $x$ and $y$ coordinates can be changed pixel by pixel (and line by line), and parameters such as $z$ saved for each xy-point. With piezo-tubes utilized in STM systems, subatomic resolution is achieved, even in pm scale in $z$-direction due to significant changes in $I_t$ even with slight shifts in $z$ in this range. This creates the basis for the ultimate resolution that STM is capable of. The tip–sample separation can be kept nominally constant with a feedback loop implemented in the measurement electronics, and this is the most widely utilized, constant-current measurement mode. The benefits as compared to a constant-height mode are that the tip is reliably separated from the sample, avoiding collisions at the corrugations possibly present in the scanning area, and precision due to reliably keeping within the current range of the electronics setup for the same reason. A schematic diagram of STM measurement is shown in Fig. 3.10.

In constant-current mode, the resulting 3-dimensional scan profile is the topography map the tip has drawn, and this is identical to the sample surface topography if there are no directionally dependent features in DOS and the distribution is constant among all the atoms observed. This is true and the tunneling barrier can be symmetric for metals with s-orbital like DOS distribution at low $V_g$, but for semiconductors the situation is markedly different due to the typical sp$^3$ orbital hybridization for VB and CB, and the surface dangling bonds with filled or empty nature, as described in chapter 2 [168]. Similar restrictive approximations limit the validity of Tershoff-Hamann model introduced next.
3.5.1 Scanning Tunneling Spectroscopy

More exact theoretical model of the tunneling in STM was first described by Tersoff and Hamann [171, 172]. In the model, tip is brought sufficiently close to the surface so that their wave functions overlap, and the tunneling current is derived through first-order perturbation theory by using the tip potential as a perturbation in the Schrödinger equation. With different work functions for the tip and substrate, $\phi_{\text{tip}}$ and $\phi_{\text{sub}}$, respectively, one obtains an approximate expression when tip DOS is approximately featureless:

$$I_t \propto \int_0^{eV_g} \rho_{\text{sub}}(E) T(E, eV_g) dE.$$  \hspace{1cm} (3.7)

$\rho_{\text{sub}}$ is the local density of states (LDOS) of the substrate sample surface, and $T(E, eV)$ is the transmission probability of the electron:

$$T(E, eV_g) = \exp (-2k_d z),$$  \hspace{1cm} (3.8)

$$k_d = \sqrt{\frac{2m_0}{\hbar^2} \left( \frac{\phi_{\text{tip}} + \phi_{\text{sub}}}{2} - E + \frac{eV_g}{2} \right) + k_{\|}^2},$$  \hspace{1cm} (3.9)

and $k_d$ the decay constant over the vacuum, and $k_{\|}$ the wave vector perpendicular component, if other contributions than $\Gamma$ is to be taken into account. Here, we choose $E_F = 0$. It is noted that similar exponential $I_t$ to $z$ dependence follows from (3.7) as in (3.6). Consistent with the model by Tersoff and Hamann, the importance of considering STM images as a convolution...
of both topographic features and electronic properties has been observed using GaAs(110) as a case example; the difference in height of Ga and As atoms when imaging with STM is more than compensated by the difference in charge density when the surface states are probed separately from close to the VBM and CBM [172, 173].

Fig. 3.11 illustrates the contribution from the DOS to or from the tip to the substrate, as dictated by (3.7) with the additional factor of transmission probability \( T(E, eV_g) \). All of the states up to the energy \( eV_g \) contribute in \( I_t \), and thus, simplified analysis on LDOS at energy \( eV_g \) can be carried out by examining the differential conductance, \( \frac{dI_t}{dV_g} \):

\[
\frac{dI_t}{dV_g} \propto e\rho_{\text{sub}}(eV_g)T(eV_g, eV_g) + e\int_0^{eV_g} \rho_{\text{sub}}(E)\frac{d}{d(eV_g)}T(E, eV_g)dE. \tag{3.10}
\]

For example semiconductor \( E_g \) can be readily obtained from the inspection as in Eq. (3.10) from the width of zero-level signal as a function of \( eV_g \) around \( E_F \), that will correspond to the \( \rho_{\text{sub}} = 0 \), since \( T \neq 0 \). However, accurate inspection of LDOS becomes problematic because the above expression contains the exponential dependence of \( T \) on \( z \) and \( V_g \), creating a large background signal that tends to mask the LDOS features that are wished to be obtained from the derivative.

To overcome this problem, the above quantity is often numerically normalized to obtain \( \frac{dI_t}{dV_g} \), or equivalently \( d\ln I_t/d\ln V_g \), that is conventionally straightforwardly interpreted as LDOS obtained by scanning tunneling
spectroscopy (STS) [174]:

$$\frac{dI_t/dV_g}{I_t/V_g} = \frac{\rho_{\text{sub}}(eV_g) + \int_{0}^{eV_g} \rho_{\text{sub}}(E) \frac{d}{(eV_g)_{T(E,eV_g)}} T(E, eV_g) dE}{\frac{1}{eV_g} \int_{0}^{eV_g} \rho_{\text{sub}}(E) \frac{T(E,eV_g)}{T(eV_g,eV_g)} dE}.$$ (3.11)

Here, the transmission probabilities that bring the exponential dependence on $V_g$ and $z$, are observed as ratios $T(E, eV_g)/T(eV_g, eV_g)$, so that the strong dependences in fact cancel out. Thus, the expression in Eq. (3.11) results in a normalized (due to the denominator) quantity of LDOS with a varying background (due to the second term), which is slow especially at positive bias (probing empty states). Probing of filled states tends to get somewhat shadowed by the tip DOS since tunneling from the highest energy states (closest to $E_F$) has the highest probability and thus, proportionally highest contribution to the $I - V$ curve comes from the tip DOS matching VBM at a given $V_g$ (see Fig. 3.11) [174].

Utilizing the expressions above, we can investigate LDOS of a sample in the atomic resolution provided by STM by sweeping a voltage at a specific point of the scan area in a low bias range from $V_1$ to $V_2$. Numeric values for $I_t$ at each data point $V_g$ are all that is required for numerical computation of LDOS from (3.11). The above procedure is often carried out between a specified amount of data points and lines during a STM scan to obtain a spectroscopy map that can be used to laterally visualize LDOS in the scan area. This procedure is called current imaging tunneling spectroscopy (CITS). Voltage modulation with a lock-in amplifier can also be utilized to measure $dI_t/dV_g$ with high-frequency signal and low enough amplitude for low noise and high resolution [166, 168].

A setback of (3.11) for especially wide $E_g$ semiconductors is that near VBM and CBM, the denominator $I_t/V_g$ approaches 0 much faster than $dI_t/dV_g$, and the divergence creates a high level of noise in and near the band gap, where $I_t = 0$ [175]. Especially for highly doped semiconductors having $E_F$ in the vicinity of VBM or CBM, it might be impossible to distinguish highly relevant DOS features in this area. This problem is often avoided by the addition of a small positive offset $b$ into the denominator, or more elegantly:

$$\overline{I_t/V_g} = \sqrt{(I_t/V_g)^2 + b^2}.$$ (3.12)

$b$ is chosen such that further from the gap area $b < I_t/V_g$ and thus $b^2 \ll (I_t/V_g)^2$, so that the ratios of individual features in LDOS spectrum are not disturbed, but on the other hand, the noise suppression is efficient where $I_t/V_g \approx 0$ [176].
3.5. Scanning Tunneling Microscopy

Another potential weakness of STM and STS methods concerning semiconductor studies is tip-induced band bending (TIBB). This arises when there is significant resistivity in the probed area, that can result in a marked proportion of $V_g$ voltage dropping across the sample instead of $V_g$ drop locating completely in the vacuum between the tip and sample. This will create offsets in the energy values measured in the spectra, impeding the analysis. Sufficiently highly doped samples with low resistivity are thus preferred in STM measurements in order to minimize TIBB [163]. Tip modifications due to adsorbed molecules or so-called double-tip effect can also significantly contribute to the measured topography or STS spectra, underlining the significance of careful tip preparation and critical inspection of the measured data [177]. Indeed, even the $V_g$ can have a prominent effect on modifying the surface structure [177, 178].

3.5.2 Low Temperature STM

Low temperature scanning tunneling microscopy (LT-STM) instrumentation to be used in UHV at cryogenic temperatures have been developed since the early days of STM [179–181]. For semiconductor studies, the main benefits include stable conditions that enable e.g. atomic manipulation [182], greatly reduced thermal drift and thermal noise during scanning, and reduction in thermal excitation, i.e., diminished broadening of the energy states that are observed in STS, and thus higher resolution measurements in general [125, 180, 181, 183]. Any reaction kinetics is most straightforwardly altered by controlling and changing the temperature, making LT-STM a favorable method also for studying surface reaction kinetics and transient processes, freezing or stabilizing adsorbate molecules for studying their fundamental properties [184], or nature of surface constituents such as buckled dimer shifts on Si(100)(2×1) [32] and Ge(100)(2×1) [178, 185], but also other low temperature phenomena (magnetism, superconductivity, etc. [186, 187]) on the surface in atomic scale in general.

A wide selection from relatively simple to highly advanced commercial systems are nowadays provided by several manufacturers. A practical, simple example system is ScientaOmicron Fermi SPM [125], introduced in subsection 3.2.6. Low temperature at sample and SPM head is achieved with a continuous-flow cryostat by thermal conduction from the cryostat through cold finger during initial cooling and copper braids before and during measurement. Temperature at slightly above the cryogenic liquid temperature is achieved with a PID controlled counter-heater element. This procedure avoids unnecessarily long waiting times and thermal drifting with efficient stabilization, of less than 1 mK min$^{-1}$ with liquid helium (LHe) and
Chapter 3. Experimental Methods

2 mK min\(^{-1}\) with LN2. Corresponding stabilized sample temperatures in this system are about 15 K and 85 K. Contraction and changed response of piezo elements need to be taken into account in the measurement [167], and they are commonly implemented as predefined calibration setups.

3.6 Photoluminescence

PL is one of the most straightforwardly utilizable methods to investigate surface and interface quality and properties of semiconductors, and it has been used vastly in such studies [29, 84, 188] even though PL is commonly considered as a probe of bulk-like properties. Any electrical contacts or other special features are not required of the sample, and the measurement does not damage semiconductor sample material. PL is a method that is based on optical excitation of electron-hole pairs over \(E_g\) with higher energy photons and detection of the photons whose emission results from the eventual recombination. The emission from the sample is collected as a wavelength spectrum with a spectrometer. High purity semiconductor material gives a prominent peak at the wavelength corresponding to its intrinsic \(E_g\) due to CB–VB recombination. [29, 188] Fig. 3.12 shows a typical PL measurement setup with a grating monochromator spectrometer.

\[
\begin{align*}
\text{Laser} & \quad F & \quad C2 & \quad \text{L} & \quad \text{D} & \quad \text{Electronics} & \quad \text{Sample} \\
\text{Grating monochromator} \\
& \quad \text{Intensity} \\
\end{align*}
\]

\(\lambda\) (nm) 700 800 900 1000

C1 Sample

Figure 3.12: Schematic example PL measurement setup. C1 and C2 are choppers used for signal modulation and lock-in detection. F is a filter element, L focusing lens and D the detector, typically e.g. Si p-i-n, Ge diode or InGaAs unit, depending on the wavelengths measured. [29] Emission is decomposed into a wavelength spectrum in real space with a grating monochromator in this setup. The PL spectrum here with peak at about 860 nm is from a GaAs sample.
The optical absorption coefficients $\alpha$ for III–V materials are roughly around $1 \times 10^4 \, \text{cm}^{-1}$ for wavelengths at visible to near-ultraviolet (UV) [26] typically used in excitation lasers of PL setups (for example, Ar$^+$, frequency-doubled Nd:YAG, or HeNe laser at 488 nm to 633 nm, or, 2.0 eV to 2.5 eV), where the dominant mechanism is VB–CB and impurity state excitation [29]. Such absorption coefficients give a penetration depth $\delta$ of around 1 $\mu$m. Taking into consideration also the typical diffusion coefficients for e.g. GaAs with high crystal quality, in the range of 10 $\mu$m (subsection 2.1.2), it is easy to see that a significant proportion of the excited electron-hole pairs are able to diffuse to the surface, and thus, the surface of the sample contributes significantly to the overall recombination in case recombination centers have a markedly different density at the interface than in the bulk. Thus, PL is a useful method for probing these particular surface properties.

The relative peak intensities at different wavelengths give information about the energy levels through which recombination occurs, and also about the different recombination mechanisms. Most importantly concerning surface properties and quality, increasing non-radiative recombination related to surface states and defect states at different interfaces is straightforwardly observed as decreased PL. However, such states also tend to have an effect on bend bending as discussed in chapter 2. If differences in band bending are induced due to additional trap states, this can either promote surface recombination velocity $v_s$, or counterbalance it, according to Eq. (2.14). Therefore, interpretation is not always straightforward and complementary measurements are often required for profound analysis. Sophisticated methods and models based on PL to investigate the exact recombination mechanisms and $v_s$ exist, such as time-resolved photoluminescence (TRPL) [29, 188]. In our work, the straightforward approach of comparing PL intensities for samples with different surface treatments was utilized in combination with other methods to find out the qualitative differences that could be assigned to surface or interface properties.
4 Semiconductor Devices

The reason behind importance of semiconductor research is the utilization of these materials in electronics devices, from transistors and diodes to solar cells and lasers. This chapter gives an introduction to a small selected set of such components, with stress on the importance of semiconductor–oxide interface in particular. Firstly, short outline of $pn$-junction diode that works as a functional part of these devices is given.

4.1 $pn$-Junction Diodes

Tunability in photonic and electrical properties and coupling of these in different parts of the same solid structure are traditionally employed in semiconductor devices. One of the most fundamentally important components consisting of solid semiconductor structure is a $pn$-junction that is created between oppositely doped parts in a semiconductor. $E_F$ at both sides of this junction is equal since they are in electrical contact, meaning that the CB and VB are misaligned by energy that is equal to difference in separation of $E_F$ from CB or VB. The junction results in a depletion region surrounding the junction in a similar manner as in the case of surface depletion as the electrons are repelled from the $p$-side of the junction and holes vice versa, revealing a space charge region where charge carriers are depleted. This is observed as an internal built-in voltage $V_{bi} \cong k_B T \ln(N_a N_d/n_i^2)$ at the junction, depicted in Fig. 4.1. [23, 189]

At thermal equilibrium, no current flows as drift current exactly counterbalances the diffusion of charge carriers from one side to the other. When a forward external bias voltage $V_{ext}$ is applied (i.e., positive voltage to the $p$-side / negative to the $n$-side), the voltage drop in the circuit effectively occurs at the junction due to negligible resistivity elsewhere. As excess electrons gather on the $n$-side and holes on the $p$-side, space charge at depletion region is compensated, effectively lowering built-in voltage and decreasing depletion region width, and diffusion (net) current increases over the junction while drift current compensating the diffusion decreases. With reverse bias, thermally excited charge carriers are pulled away from each side of the junction, further depleting the region and thus, increasing also the built-in
voltage. Only small current can flow in the direction determined by the bias up to a breakdown threshold. Under non-breakdown for ideal diodes, the diode current $J_d$ dependence on the external bias $V_{\text{ext}}$ is characterized by:

\[
J_d = J_0 \left[ \exp \left( \frac{eV_{\text{ext}}}{k_B T} \right) - 1 \right],
\]

\[
J_0 = \frac{eD_n n_p 0}{L_{dn}} + \frac{eD_p p_n 0}{L_{dp}} = e \left( \frac{D_n}{L_{dn} N_a} + \frac{D_p}{L_{dp} N_d} \right) n_i^2
\]

and $n_i^2 \propto \text{DOS} \times \exp(-E_g/k_B T)$. Notably, $J_d$ has an exponential dependence on $V_{\text{ext}}$ and $-E_g$. $n_p 0$ and $p_n 0$ are the electron and hole concentrations at the $p$ and $n$-sides, respectively. [23, 189]

Diode-structure is utilized in e.g. LEDs, photodiodes, and transistors. While diode-characteristics of such devices are mostly independent of surface properties, the surface (defects, traps) can have a major impact on current flow, recombination or band bending, for example.

### 4.2 Photodiode Devices

Solar cells and photodetectors are effectively photodiodes with slightly different designs and operating conditions. In a photodiode, an internal $pn$-junction separates photoexcited minority charge carriers by the drift field within the depletion region, and they are then collected at Ohmic metal contacts, as illustrated in Fig. 4.2. $E_g$ of the material determines the minimum $h\nu$ of photons that excite charge carriers, the resulting short circuit drift current density $J_{\text{sc}}$ being linearly proportional to the photon flux.

The effect of photoexcited charge carriers within the cell is a spontaneous generation of current flow in the photodiode. Apart from this photocurrent
4.2. Photodiode Devices

Figure 4.2: Example pin photodiode structure with cross-sectional view of the device and band diagram. Photodetector pin is often placed between barrier layers presented here, and solar cell structure contains a front barrier denoted window layer and a back-surface field (BSF) with high doping but not necessarily different $E_g$.

term $J_{ph}$ the photodiode current-voltage behaves as a normal $pn$-diode:

$$J_{tot} = J_d - J_{ph} = J_0 \left[ \exp(\frac{eV_{ext}}{k_B T}) - 1 \right] - J_{ph}. \quad (4.3)$$

From here, it is straightforward to obtain voltage corresponding to the open circuit conditions, i.e., $J_{tot} = 0$:

$$V_{oc} = \frac{k_B T}{e} \ln(\frac{J_{ph}}{J_0} + 1). \quad (4.4)$$

Effects of different parameters on $J_{tot}$ and associated performance of photodiodes are shown in Fig. 4.3 (a), (b), and (c). $J_{ph}$ flows in the reverse direction of a photodiode, and therefore, it can be utilized to produce power up to the (forward) open circuit voltage $V_{oc}$, at which diffusion current compensates the photocurrent. The maximum power $P$ produced by a solar cell is determined by the voltage corresponding to a load where the product $P/A = V_{ext}J_{tot}$ is maximized (see Fig. 4.3 (c)). Higher $E_g$ materials result in an increased $V_{oc}$, but absorb smaller proportion of the solar spectrum, resulting in lower $J_{ph}$, whereas lower $E_g$ materials can absorb higher proportion of the solar spectrum, but excess $h\nu - E_g$ is lost as the charge carriers quickly relax. Thus, stacked cell structures (multi-junction tandem
Chapter 4. Semiconductor Devices

**Figure 4.3**: Photodiode current-voltage curves with the direction of $J_{ph}$ chosen as positive. Effect of band gap and photocurrent density $J_{ph}$ on the open circuit voltage $V_{oc}$ (however, note that higher $E_g$ material will generally give lower $J_{ph}$ for a broad spectrum of illumination) (a). Curves in the operating regime of photodetectors; low dark current and noise is desirable for maximum sensitivity (b). Effect of resistivity on solar cell operation; properly passivated surfaces and sidewalls are desirable to achieve a high maximum power operation (c).

Solar cells) with gradually lowering $E_g$ into the cell, connected by tunnel junctions are used for epitaxially grown materials, readily utilized in e.g. III–V solar cells to achieve significantly improved conversion efficiency. In such a structure, cell voltages build up accumulatively, but the produced current densities are matched to the same level. [23, 189] Proper passivation is necessary in order to avoid series resistance through non-radiative recombination at the Ohmic contacts and other interfaces, and to maximize shunt resistance, i.e., avoid short circuiting over the cell junctions (see Fig. 4.3 (c)). In particular, although wide $E_g$ barrier (window) layers are utilized to avoid diffusion of oppositely charged carriers at the outermost surface, it is difficult to avoid non-radiative recombination completely if recombination
centers are present. This is because part of the high-energy tail of the solar spectrum will also be absorbed, within the barrier layer. A BSF layer with a similar working principle but most often implemented by higher level of backside doping is conventionally embedded at the bottom of the cell.

Efficient conversion of energy is highly important to tap into the utilization of renewable energy sources, solar cells in this case. The record efficiencies of silicon solar cells are presently in the range of 27% while those of III–V tandem solar cells with concentrated light extend to 45% to 46% [190]. The theoretical maxima for single-junction (1.1 eV to 1.2 eV) and tandem solar cells with infinite junctions and graded $E_g$ profile are 32% to 33% and 68%, respectively, under atmospheric solar irradiation [191, 192]. These values can still be improved drastically to over 40% and 86% under concentrated illumination [192]. Non-radiative recombination can account of several percent losses in solar cell efficiency [191], indicating that optimal surface or interface at the semiconductor boundary is essential for achieving highest possible energy conversion. Field-effect passivation is spontaneously implemented and found highly advantageous in Si solar cells with ALD grown Al$_2$O$_3$ oxide film [59].

Opposite to a solar cell, a photodetector is commonly operated at reverse bias to accelerate photoexcited charge carriers thus reduce their transit time to have as fast response time as possible. On the other hand, increasing depletion layer width, its capacitance and thus also $RC$ time constant is reduced, enabling faster response. By having the detector structure confined within wide $E_g$ material, the slow response due to diffusion current of detector wavelength photoexcited carriers is avoided. Using such approaches, high bandwidth operation of a photodetector is obtained. Too high reverse bias will however result in charge carrier velocities approaching constant saturation velocity, while depletion layer width still increases. Thus, increasing bias beyond this can also increase the response time. At reverse bias conditions, the generated photocurrent is approximately linearly dependent on the photon flux, desirable for quantitative operation of a detector. The current response is offset by a dark current, at simplest case corresponding to $J_d$ in Eq. (4.3). [23]

Various figures of merit characterize the operation of photodetectors, such as noise equivalent power (NEP): the lowest power that can be detected equivalent to the noise at a given bandwidth; detectivity: proportional to the detector area and bandwidth and inversely proportional to the NEP with a given modulation frequency; responsivity (with straightforward proportionality to quantum efficiency): current that is generated from a given power of illumination at a given wavelength; and response speed. [23] It is obvious that as non-radiative recombination reduces current that can be
collected per photon and can act as a source of current leakage on the other hand, various figures of merit can be improved by optimizing surface properties. Indeed, surface passivation is highly necessary especially in novel IR detector structures consisting of nanowire structures with large surface area to bulk ratio [193].

4.3 Transistor Devices

4.3.1 Metal Oxide Semiconductor Field Effect Transistor

MOSFETs can be considered most important building block of digital technology due to low power consumption and high packing density complementary MOS (CMOS) logic. MOS devices rely on the capacitive effect of the stack to accumulate or deplete the interface of charge carriers with the same or opposite sign as the substrate doping (accumulation or inversion, respectively) with an external gate (G) voltage $V_G$ applied to the metal electrode. Fig. 4.4 (a) shows schematics of a planar $n$-channel MOSFET and switching from $V_G = 0$ V to inversion, enabling the flow of electrons from source (S) to drain (D), that are highly doped for Ohmic contacts and proper channel formation. In accumulation $V_G$ conditions, two $pn$-junctions with opposite internal biases are formed, preventing charge carrier flow in both directions. Thus, MOSFET can act as a switch or an amplifier through controlling the $V_G$ from accumulation to inversion. If the interface area spontaneously contains an inversion layer of charge carriers without any applied $V_G$, the FET is “normally-on”, or “depletion-mode” (d-mode) and a negative $V_G$ is needed for switching the device off or depleting the channel. Vice versa, if no spontaneous inversion occurs, the device is denoted “normally-off” or “enhancement-mode” (e-mode), and conduction or gate-modulation is achieved with $V_G > 0$ V. [25] The current flows in the immediate vicinity of the oxide-interface, making it the subject of a plethora of studies in the past decades (e.g. [4, 8, 194, 195] and references therein) and one of the motives of this work also.

The continuous scaling down engineering efforts of MOSFET dimensions have been related to gains in reduced channel resistance, smaller capacitance per device (while retaining the capacitive efficiency due to smaller dimensions in general) and thus faster switching along with lower switching energy (power-delay product), and higher packing density resulting in increased functionality per chip. [25, 196]

After a certain limit reached for a few years ago already, reducing the channel length has had also adverse effects on many of the properties. These
4.3. Transistor Devices

Figure 4.4: Schematics of some MOSFETs with band diagrams from S to D with applied V_D with and without V_G: (a) planar MOSFET and band diagram also from G to bulk (B); strong accumulation with V_G > 0 V, hindered slightly by FLP caused by trap states (b) passivated GAA MOSFET with good electrostatic integrity: high G barrier, sharp transition and thus low I_OFF, (c) heterostructure TFET and its switching mechanism showing degraded subthreshold swing (SS) (non-zero I_OFF) for unpassivated case. In reality, vertical nanowires are often fabricated to utilize abrupt interfaces of e.g. MBE grown structures.

include the power consumption as drain voltage V_D is reaching its fundamental limit while leakage current I_OFF is increasing (see below), carrier mobility, and intrinsic performance, due to short-channel effects (SCEs) that arise when the depletion widths of the source and drain areas become the same order of magnitude as the channel length. These effects include increased electric field in the channel rising over a critical field and velocity saturation,
hot carrier injection into the oxide film, and impact ionizations and possibly avalanche effect difficult to control with gate voltage, and drain-induced barrier lowering (DIBL). Some of these effects can be dealt with substrate doping, but at the expense of mobility degradation. Effectively SCE is seen as injection of charge carriers into the channel, which decreases the threshold voltage that onsets strong inversion, and increases $I_{\text{OFF}}$. This effect is made even worse by DIBL that is negligible in long channel devices. DIBL is observed also as increased $I_{\text{OFF}}$ and decreased threshold voltage when the $V_D$ is applied, since this extends the depletion region near the drain even further.

These effects are illustrated as a curved and lowered bands in the channel area in Fig. 4.4 (a). In the worst case, this results in a punch-through onset of the inversion current even without any applied $V_G$. Furthermore, in addition the degraded electrostatic integrity and dramatically increasing $I_{\text{OFF}}$, scaling down of planar Si MOSFETs from 65 nm would not have kept with the expected intrinsic device speed, which is why transition from planar MOSFET device architecture to silicon on insulator (SOI) devices, FinFETs and gate-all-around field effect transistors (GAA-FETs) has progressed quickly in the last decade. These architectures offer significantly improved electrostatic integrity (alleviated SCE and DIBL), due to better confinement of D/S junctions and depletion regions, and less device speed degrading parasitic capacitance. The ideal-like band diagram and higher gate barrier in Fig. 4.4 (b) illustrates schematically the improved electrostatic integrity. Oxide thickness also has a significant impact on the electrostatic control, and improvements made in this regard have been discussed in section 2.3.1. [197]

As mentioned in chapter 1, the present MOSFET technology is dominated by Si devices, yet there is an increasing need for III–V MOSFETs as the advancements gained from scaling down are at their tipping point due to the ultimate dimensions achievable. Even despite the oxide interface is the functional part in MOSFETs and therefore has played a significant role in the performance improvements so far, there is an increase in the significance of studying these interfaces due to i) scaling down and novel designs of the devices further increasing surface to bulk ratio, and ii) desirable characteristics of III–V materials but their naturally less ideal interface properties [8]. Indeed, indium-containing III–V MOSFETs in particular have shown great promise as high mobility channel devices with similar architectures as seen in Si technology [14, 198, 199].

Furthermore, it is to be noted that there is a theoretical limit to the increase of charge carriers ($\Delta n/n$) producing current in the D-S channel, $I_{\text{DS}}$, as a function of $V_G$ in a non-zero temperature. This can be demonstrated
by calculating carrier densities $n$ with different $E_F$’s below and above CBM with appropriate DOS and populations as in Eq. (2.2). Increasing the $E_F$ below or above a given energy (e.g., CBM) by about 60 meV at 300 K which corresponds to applying a gate voltage of 60 mV in RT will increase the population above this energy (in ideal conditions) 10-fold. This results in a figure of merit known as subthreshold swing (SS) representing how well the given voltage is utilized in the device operation, thermal excitations limiting it to approximately 60 mV per decade of increased $I_D$ $\propto n$. This limits the traditional $V_G$ on/off switching to $\gtrsim 0.5$ V at minimum. However, a different intrinsic switching mechanism has been presented in a TFET device, relying on band-to-band tunneling (BTBT) rather than transporting charge carriers at a single band [14–17, 200].

In a TFET, the highly oppositely doped S-D junction is reverse-biased and thus exhibits very low $I_{OFF}$, and when $V_G$ is applied, BTBT occurs between the source and the channel due to bending of the bands at the channel and empty states available for electrons to flow through. The transition between on/off states is very abrupt because the high energy tail of the Fermi-Dirac distribution of electrons have no states to occupy in the forbidden $E_g$ of the source, and the available electrons at energies lower than this can start tunneling only after CBM of the channel is lowered to source VBM level in order to occupy only allowed energy states. The tunneling mechanism is depicted in Fig. 4.4 (c). Furthermore, the tunneling probability is a complex function of the $E_g$, $m^*$, tunneling energy window (depending on $V_G$), and the screening tunneling length, which in turn depends on the transition region and its abruptness at the source-channel interface. Having many such fundamentally important characteristics desirable over those of Si, it is evident that III–V materials with atomically abrupt heterostructure interfaces are very potential candidates for TFET devices, especially for vertically grown nanowire TFETs [14, 200]. For optimal TFET operation, BTBT without trap-assisted tunneling (see Fig. 4.4 (c) upper pane), that could cause significant leakage current at subthreshold and thus degraded SS, is of paramount importance for steep SS, and again, here optimal semiconductor–oxide junction at various interfaces plays a significant role, although strict limits for defect state densities in order to avoid these effects have not been established.

Concerning the capability to modulate surface band bending with $V_G$, it is essential to estimate the energy levels of the defect/trap states, not just their total number, as discussed in chapter 2. Thus, a quantity known as interface state density $D_{it}$ has been defined as

$$D_{it} = \frac{dN_{ts}}{dE}$$ (4.5)
to estimate the differential magnitude of defect states at a given energy or at an energy interval by integrating over the corresponding interval. Note that gap state trap density $N_{ts}$ can be obtained from the above by integrating over the entire $E_g$ from VBM to CBM. However, it is of essence to retain low $D_{it}$ especially in the middle of $E_g$. This is a very fundamental figure of merit for any semiconductor–oxide interface passivation method and crucial for operation of especially MOSFETs.

$D_{it}$’s in the range of $1 \times 10^9 \text{cm}^{-2}\text{eV}^{-1}$ to $1 \times 10^{10} \text{cm}^{-2}\text{eV}^{-1}$ can be achieved for SiO$_2$/Si MOSFETs, establishing an ideal case benchmark for optimal switching with low SS and high $I_{ON}$ operation. In order to avoid FLP and other effects from interface traps, $D_{it}$ near midgap in the range of $1 \times 10^{11} \text{cm}^{-2}\text{eV}^{-1}$ or below should be achieved while total concentration of defect states at $1 \times 10^{12} \text{cm}^{-2}$ range pronounced at midgap already could pin the Fermi-level as discussed in section 2.2.2. $D_{it}$ in $1 \times 10^{12} \text{cm}^{-2}\text{eV}^{-1}$ range is typically observed for high-$\kappa$/III–V stacks with many different pre and/or post-treatment approaches, although low $1 \times 10^{11} \text{cm}^{-2}\text{eV}^{-1}$ $D_{it}$ stacks have also been demonstrated. [8] Thus, both fundamental understanding of interface properties as well as finding novel, simple methods for interface passivation are further still of great importance for future advancements in digital technology, highly governed by MOSFETs.

### 4.3.2 High Electron Mobility Transistor

Another device that benefits from an insulating film, due to carrier confinement and passivating characteristics is HEMT, even though its basic operation does not require an oxide film. HEMT is an FET device that includes an undoped channel to minimize ionized impurity scattering, and an adjacent wider $E_g$ layer, which is locally doped near the channel. The selective doping results in diffusion of charge carriers into the channel while remaining very close to the interface due to the Coulombic attractive force of the ionized dopant atoms in the wide $E_g$ material, and potential barrier at the high $E_g$ side on the other hand. The electrons are confined in such a narrow space that their energy levels are quantized perpendicular to the heterostructure, resulting in a two-dimensional electron gas (2-DEG). The carrier density determining the conductance is then that of the 2-DEG, which is controlled with the gate voltage. [25] Initially, HEMT operation was evidenced on AlGaAs/GaAs heterostructure [201], and basic principles of its device and band structure are shown in Fig. 4.5.

The intrinsic characteristics of a HEMT device structure aim at increasing $n_0$ without degrading mobility or saturation velocity. Performance can be further optimized by reducing the carrier transit time by lowering the gate
4.3. Transistor Devices

length, or increasing the achievable $I_{DS}$ by increasing gate width. Too short gate length might result in SCE, observed here as, e.g., current flow through the barrier layer. However, tuning the heterostructure can circumvent this problem and simultaneously bring other benefits. An InGaAs channel can be utilized in an AlGaAs/In$_x$Ga$_{1-x}$As/GaAs structure, provided that the channel layer is thin enough (5 nm to 20 nm for $x = 0.15$ to 0.35), so that lattice mismatch strain is elastically accommodated. A larger CB offset will result in significantly enhanced carrier confinement, in addition to other highly desirable electron transfer properties of InGaAs. For InP substrates, higher InAs content InGaAs can be utilized due to smaller lattice mismatch. Such structure is denoted pseudomorphic HEMT (pHEMT). [25] Alternatively, metamorphic HEMT (mHEMT) structured InGaAs can be grown, that has gradually increasing In content (i.e., lattice mismatch), eventually resulting in e.g. very high mobility In$_{0.53}$Ga$_{0.47}$As on GaAs, on which an InAlAs barrier can be used. [189]

Typical figures of merit for a HEMT include magnitudes of the (mutual) transconductance $g_m = dI_{DS}/dV_G$ (at constant $V_D$), cut-off frequency of current gain (where input and output current are equal), $f_T$, maximum oscillation frequency for a unity power gain, $f_{\text{max}}$, and a low noise figure in general. Unsurpassed performance in these properties have made III–V HEMTs standard low-noise microwave and millimeter wave components.
The application-specific structures include AlGaN/GaN HEMTs for high-power amplifiers due to the wide $E_g$ withstanding high temperatures and voltages and large sheet carrier concentration enabling high currents, and InGaAs channel devices for high-frequency applications. [25, 189]

AlGaN/GaN HEMTs surfaces have been successfully passivated with SiN [203] that reduces current collapse, dynamic on-resistance, DC-RF dispersion, among other detrimental effects from trapping. For InGaP surface, GaS passivation has been observed to result in reduced sheet resistivity and improved breakdown voltage due to less interface states (less depletion) [204]. In practical HEMT device structures also, SCE need to be reduced for optimal operation. This can be carried out by increasing S–G and G–D dimensions [205], but this will effectively induce parasitic resistance due to depletion effect caused by the surface states. A detrimental kink effect, namely abnormal current-voltage characteristics and transconductance dispersion, commonly observed in GaAs and InP based HEMT structures is onset by impact ionizations in the channel and consequent hole accumulation [206, 207], but the characteristics are generally considered to be dependent on interface state densities due to charge trapping and can be improved significantly with surface passivation [208], however, much depending on the exact device structure [209]. HEMTs have been shown viable for utilization in CMOS logic also [210, 211], mainly limited by unacceptably large leakage currents [6], and the necessity of barrier which results in large resistance [211]. These issues have been proposed to be alleviated by redesigning the structure to include a high-$\kappa$ oxide [6].

Approaches to reduce gate leakage and increase the voltage swing, in e-mode HEMTs in particular, limited by the Schottky gate barrier have been made by growing oxide films on pHEMTs and mHEMTs, with promising results [212–215]. Considering the discussion above, with the fact that MOS-HEMT is seen as increasingly desirable choice, it is evident that oxide interface passivation could bring various benefits for these technologies.

4.4 Device Fabrication

Manufacturing of semiconductor photonic and electronic devices consists of crystal growth, possibly including epitaxial growth of doped or undoped homo- or heterostructures by a growth method such as MBE, and subsequent processing and possible regrowth steps.
4.4.1 Molecular Beam Epitaxy

MBE is the most important growth method of epitaxial III–V semiconductor structures and offers precise control of several parameters for stacked, lattice-matched heterostructures. In MBE, material is deposited onto a heated sample or wafer substrate from effusion cells that typically contain evaporation sources for each separate element to be included in the composition of the grown layer. This is carried out in UHV to have as low a fraction of impurities as possible (∼1 × 10⁻⁹ concentration) in the bulk material, and a molecular (non-viscous) flux in the deposited beam, so that deposition can be stopped with quick operation of mechanical beam shutters. The substrate is heated at a temperature sufficient for surface migration of the adsorbed species (e.g., 500°C to 600°C for GaAs) so that they find and fill vacancy sites at the surface. Thus, a highly uniform and precise layer growth is achieved with good control over the composition, layer thicknesses and sub-monolayer abruptnesses. [216] Furthermore, diffuse or abrupt differences in doping levels relevant in e.g. tunnel or pn-junctions are also enabled by tuning the dopant flux during the growth [217, 218].

Indeed, industrially produced high mobility RF applications and high quality epitaxial heterojunctions in optoelectronics device structures of today are grown with MBE. In addition to well-defined growth of bulk semiconductor material, the trends in research and industry enabled by MBE include growth of low-dimensional structures such quantum dots (QDs) and quantum wells (QWs) as well as nanowires [217, 218]. A typical MBE setup is illustrated in Fig. 4.6.

Additional benefit of MBE is that various research tools (PES, STM, RHEED, SIMS, ...) can be implemented in the deposition system in the same or gate valve separated chambers in situ [216]. Furthermore, other treatments are enabled on ideally clean and smooth semiconductor surface since there is no need to break UHV between treatments, making MBE a versatile tool from semiconductor research to production [218]. In our studies for device structures grown with MBE, an As capping layer was deposited on the device surface. The topmost semiconductor layers are protected from atmospheric oxygen this way after taking the wafer out of the MBE system. As cap can be evaporated in a separate UHV system by heating the sample at 350°C to 400°C, enabling the use of a sample with close to an ‘as-grown’ surface for further experiments such as oxidation treatments.

4.4.2 Device Processing

After growing the stack that offers the device functionality, the sample or wafer is processed into complete devices, the exact procedures depending
on the particular desired outcome. Typical processes include photolithography, wet and/or dry etching, dielectric and/or metal deposition, and rapid thermal annealing (RTA). Methods used for III–V semiconductor processing have been widely investigated and a wide range of standard applications exist for these purposes [219]. The device arrays produced on a wafer or sample template are often diced and packaged (e.g., mounted and wired) for end-user utilization or characterization.

In photolithography, the wafer or sample piece is spin-coated with an appropriate thickness, polarity, profile and durability photoresist, illuminated with UV light through a photomask which sensitzes the photoresist, and the produced sensitized pattern is dissolved with a developer solution while other areas on the resist remain intact. The resulting open areas can then be dry or wet etched, or metal or dielectric film can be deposited, after which they are lift-off from the resist containing areas. Example processes are shown in Fig. 4.7. In Fig. 4.7 step 6, positive photoresist, suitable for etching treatments (such as MicroChemicals AZ6632 [220]) is used. In step 11, image reversal photoresist with an image reversal procedure (such as MicroChemicals AZ5214E [221]) is used, which produces good edge profile and is optimized for lift-off purposes. In practice, steps from 11 to 14 are produced with two different lithography steps: thick positive resist for oxide
Figure 4.7: A photovoltaic device process with each stage and following treatment sketched (not to scale). Sample 1: as-grown, 1 → 2: As cap removal, 2 → 3: crystalline oxidation, 3 → 4: ALD Al₂O₃ growth, 4 → 5: bottom metal deposition, 5 → 6: surface spin coating with photoresist, 6 → 7: resist patterning sensitization with UV light through a photomask and development, 7 → 8: oxides wet etching, 8 → 9: mesa isolation patterning through dry etching, 9 → 10: resist dissolving, 10 → 11: second resist spin coating, 11 → 12: front contact metal opening mask patterning, 12 → 13: oxide dissolving from contact metal areas, 13 → 14: metal deposition, 14 → 15: resist and residual metal lift-off, 15 → 16: RTA treatment.
etch and image reversal with slightly wider pattern for metal deposition. Thus, Fig. 4.7 is for illustrative purpose only.

The etching procedure and appropriate solutions/gases depend on the combination of materials to be etched, and a number of literature is available for a variety of purposes [222]. Wet chemical etching is appropriate when high selectivity is required, meaning that etching needs to effectively stop in some layer of the semiconductor structure. The downside of using chemical etching is in its isotropic nature or possibly preferential etching of some crystal planes over the others. These effects will also eventually result in etching from underneath the mask pattern (undercutting). Dry etching such as reactive ion etching (RIE) or inductively coupled plasma (ICP)-RIE can be utilized to obtain high aspect ratio and more anisotropic structures. The etching gas composition, etching RF power, pressure in the chamber and other plasma parameters can be tuned to give more of a vertically directed and physical nature. However, the more chemical effects dominate, the higher selectivity between different materials can be obtained. Therefore, the method and chemicals or gases utilized is a compromise and interplay between the desired etching result, and both dry and wet etching are often utilized in device processes [223].

Oxides and dielectrics can be deposited in several ways, each having their respective benefits and drawbacks. Some of the common methods are electron beam (e-beam) evaporation, sputter deposition, ALD, and plasma-enhanced chemical vapor deposition (PECVD). PECVD is the most commonly utilized method for this purpose in conventional device processes [224]. We have utilized ALD for the UHV \textit{in situ} film deposition; \textit{ex situ} dielectric depositions were not needed in these device processes.

For depositing metal electrodes or Ohmic contacts, sputter deposition and e-beam evaporation are the most common methods. Interface chemistry fine tuning after depositing metals or dielectrics can be done by post-deposition annealing (PDA) methods such as RTA in a dedicated apparatus. In a conventional RTA, sample is kept in an inert atmosphere, typically N$_2$, and heated radiatively with high temperature resistor filaments or lamps. This can \textit{e.g.} cause intermixing, doping and/or regrowth of different constituents at the interface. Such effects will have an impact on the band structure at the interface, and will thus affect the conductivity. A proper tailoring of PDA treatment is necessary for any metal stack for the highest possible conductivity (lowest resistivity, Ohmic contacts), stability and uniformity/morphology for device electrical contacts. A stack with Ni, Ge and Au results in low resistivity after RTA for both $n$ and $p$-type GaAs [225].
5 Summary and Conclusions

The work carried out as related to this thesis is summarized in this chapter. Backgrounds and interpretations made in each of the articles are briefly discussed on the basis introduced in chapters 2, 3, and 4.

5.1 Controlling Thermally Assisted Oxidation of GaSb(100)

Papers I and II were carried out to gain more detailed insight into thermal oxidation of GaSb(100), and how to control it. One of the important properties of GaSb from an application point of view is its high $\mu_p$, necessary for $p$-type MOSFETs in digital logic CMOS, and the narrow $E_g$ that can be utilized in IR photonics. As pointed out in several contexts for III–V’s, oxide interface properties for this material also are limiting its utilization in e.g. MOSFETs. By gaining insight to the events occurring at the initial stages of oxidation, one can draw conclusions and clarify the detailed effects that a given passivation method could provide.

We investigated effects of thermal oxidation on the atomic structure of the surface by utilizing STM/STS [I]. Estimating $E_g$ from $dU_t/dV_g$ STS measurements (see section 3.5.1), we could distinguish gap states induced by the oxidation, increasing for the point defects with extended dimensions. In particular, long exposure resulted in even a metallic character (non-zero LDOS at $E_F$). That is, even in initial stages of thermal oxidation, defective sites are observed although it was not entirely clear which atomic species cause such states. Earlier studies applied in particular to oxide/GaSb device interfaces and the observed STM topography characteristics suggest preliminarily that this is due to relieved Sb atoms present on the surface or formation of Ga–O. These are end-products of the thermodynamically driven processes in the conditions under investigation, and are found mainly responsible for thermally treated oxide/GaSb(100) $D_{it}$ [67, 68, 226–228] (see also table 2.2). Previously it has been observed that higher Sb-content reconstructions as a template for thin oxide film growth cause higher $D_{it}$ [69]. However, the lowest Sb-content reconstruction on GaSb(100) is $(4\times3)\alpha$ [69,
which contains only four Ga-atoms in the topmost atomic layer, and thus, avoiding surface Sb proves to be challenging in practice.

The specific initial bonding sites of O could not be completely resolved from STM studies, and home-laboratory XPS could not provide enough resolution or intensity from the topmost atomic layers to reliably characterize the surface chemistry with such a low O content. Thus, further investigations on the detailed chemistry of oxidation were done utilizing SR-PES data from measurements at MAX-lab, Lund University, Sweden (beamline I311) [II].

Detailed and systematic fitting of the high resolution spectra with delicate features revealed incorporation of oxygen into Ga-bonds, but referring to earlier PES studies on clean, reconstructed GaSb(100) [229, 230], we observed simultaneous increase in the Sb–Sb bonding, indicating metallic Sb formation. Ga–O bonding remains dominant for O in all stages of oxidation, indicating either insertion in the Ga–Sb bonds or Sb-substitution. However, no Sb-oxide components were observed at initial stages of thermal oxidation, likely indicating insertion into Ga back-bonds while simultaneously detaching some of the subsurface Sb. Only at later stages the outermost Sb–Sb bonds of the native reconstruction are modified while metallic Sb–Sb remains on the surface, suggesting the relative stability or continuous nucleation of metallic Sb after an initial Ga-bond saturation and subsequent Sb oxidation. It is to be noted that the thermal oxidation is initiated by formation of Ga–O bonds, and in the initial stages of oxidation investigated with STM/STS [I], gap states near CBM are observed while metallic-like features are observed after prolonged oxidation. Thus, the first effect could be related to Ga–O defect states and the latter one to Sb–Sb, evaporation of which is prevented in the case of extensive oxide coverage. Similar gap states are not observed with thermal oxidation of InSb(100), suggested to be related to remarkable lowering of empty In dangling bonds from VB to CB, while no parasitic Sb remains on the surface [105].

In paper I it was observed that defective sites are introduced by the nature of thermal oxidation and the chemical nature of these defect sites was inferred and confirmed in paper II. Based on these and earlier studies, it could be concluded that thermal oxidation will not produce a passivating film on GaSb(100). A way to circumvent this problem has been presented in paper I by depositing a submonolayer film of In on atomically smooth GaSb(100). Subsequent annealing of this structure produces GaSb(100)(4×2)-In which could preliminarily be assumed similar as InSb(100)(4×2), for which a crystalline oxide film can be introduced by thermal oxidation [21, 105]. Indeed, both In-terminated surface formation and subsequent oxidation was observed to occur in a very similar manner as for GaAs(100)(4×2)-In [21, 104], with the benefit of introducing much less Sb-rich reconstruction than
is possible to obtain for native GaSb(100). The treatment resulted in a stable, well-ordered crystalline (1×3) InO$_x$ phase in which no metallic features are observed in STS. Sb possibly detached in the oxidation process is therefore interpreted to be readily evaporated, or it is included in this stable phase, effectively suppressing any dangling, dimer, or metallic Sb bonding. As the initial native oxidation also introduces Ga–O bonds that seem to act as nucleation sites of the initial oxides, the InO$_x$ can also be interpreted to stabilize the crystal surface from forming a variety of possibly defect state causing bonding environments due to an amorphous film.

5.2 Introducing Crystalline Oxide on InGaAs(100) Photodetector Device Surface

With support from the previous results [I] to the already established knowledge on surface layer passivation with additionally deposited InO$_x$ layer [21], we tested this procedure in practice on a photodetector structure in paper III. This is the first test of crystalline oxidized III-V surface in an optoelectronic device. Reduction in series resistance was expected due to already proven reduction of $D_{it}$ with the treatment [22]. Accordingly wide area devices were fabricated in order to maximize the significance of front surface on the responsivity. To clarify the effects observed on atomic scale, a series of samples with similar treatments were investigated with PL and HAXPES. The device stack was grown with MBE. An In$_{0.55}$Ga$_{0.45}$As based IR detector structure [231] with a pseudomorphic In$_{0.25}$Ga$_{0.75}$As barrier layer was utilized, on which thermal oxidation was carried out in a similar way as for GaSb(100) [I] and GaAs(100) [21], and a subsequent protective in situ ALD Al$_2$O$_3$ was deposited. Device processing was carried out in a very similar manner as shown in Fig. 4.7. Indeed, responsivity showed a marked increase in responsivity specifically at lower wavelengths, stressing the fact that less recombination is observed for charge carriers excited close to the surface, in particular (see Eqs. (2.7) and (2.9), indicating pronounced absorption $\alpha$ and shorter penetration depth $\delta$ for higher $h\nu$, i.e., lower $\lambda$).

In order to elucidate the reasons behind benefits attained, we tested similar structures systematically on GaAs(100), on which surface properties are expected to be highly similar as on In$_{0.25}$Ga$_{0.75}$As. Depending on the amount of In deposited, different reconstructions were observed. ALD grown Al$_2$O$_3$ was utilized in a similar manner as for the device process. PL measurements revealed an interesting trend with a non-monotonic behavior of PL signal dependence on the amount of In. Highest increase was observed for an intermediate In content. An elementary interpretation would
suggestion the lowest amount of $D_{\text{r}}$ for this sample, but a detailed consideration (Eq. (2.14)) shows that recombination velocity $v_s$ and thus, PL signal intensity depends also on the nature of the trap states, their energy, and the corresponding band bending. That is, some reconstructions could alleviate band bending significantly more, equalizing the amount of oppositely charged carriers diffusing to the surface. In some special cases this could relate to increase in minority charge carrier non-radiative recombination.

For differentiating such effects, HAXPES was utilized, since the surface depletion width can be observed in the range of tens nanometers, whereas probing depths of conventional XPS or SR-PES are always in the range or less than few nanometers (see sections 2.2.2 and 3.3). Subtle differences were observed among the samples, showing highest amount of increased high and low $BE$ component emissions in As 3d spectra for the reference sample without InO$_x$. The low $BE$ components were interpreted to arise from dangling As bonds referring to the simultaneous oxidation of Ga [xii, 40, 41]. Furthermore, peak broadening at high $BE$ side was taken into account as another emission component, seen to vary systematically with surface sensitivity and the amount of As dangling bonds. This peak was seen significant even deep in the subsurface regions through varying the probing depth by changing $h\nu$. It was concluded that surface band bending, observable as 0.1 eV to 0.2 eV CLS [232] could be taken into account by fitting such an extra component, and that band bending along with As dangling bonds were significantly decreased with an interfacial InO$_x$ layer. The PL results could be tentatively explained this way. However, it could not be excluded that too high an amount of In could also create alternative recombination centers for charge carriers, and therefore, it is noted that an ideal crystal structure requires an optimal In-content.

5.3 General Applicability of III–V Crystalline Oxide — Case of InSb(111)B

After the successful utilization of an InO$_x$ passivation layer in a photodiode structure, we carried out investigations to see if a similarly crystalline structure could be demonstrated on III–V irrespective of the surface plane in paper IV. As discussed in chapter 4, several modern semiconductor devices utilize e.g. nanowire structures, imposing necessity to passivate several crystal faces for optimal device operation, making this investigation highly important from an applied point of view. In the study, we investigated effects of thermal oxidation on the crystallinity of InSb(111)B surface, relating to InSb(100) [21, 105]. It is to be noted, that observing similar chemistry
and bonding, not to speak about atomic configurations, for dissimilar surface planes of the same material is by no means given, as can be seen even by comparing e.g. elemental Si(100) and Si(111) surfaces (section 2.2.1). Furthermore, similar tests with similar results were carried out for ex situ cleaned samples, highlighting the applicability of this treatment in practice. In most of our investigations, InSb(111)B(3×3) was used as the starting surface, as it is readily produced by sputter-cleaning in UHV.

Clean InSb(111)B exhibits (2×2) and (3×3) reconstructions, but remarkably, they are also observed after thermal oxidations carried out in our studies, whereas RT oxidation results in dim (1×1) observed in LEED, suggesting O$_2$ interaction with the surface in elevated temperatures. For oxidized (3×3) we were able to show lowered LDOS (with methods as in section 3.5.1) in the $E_g$ area and clear modification of atomic features as compared to the clean surface with STM and STS. For resolving the modification of the structure on atomic scale, LT-STM was successfully utilized. It was not straightforward to observe structural or LDOS differences in oxidized (2×2) as compared to clean (2×2), but XPS was utilized here to show incorporation of O in the structure, bonded predominantly to In. Thus, we were able to present novel InSb(111)B(2×2)–O and InSb(111)B(3×3)–O phases that cannot be directly shown by periodicity, or diffractive methods only.

An interesting observation was made for surface that was worn out due to high temperature annealing and/or prolonged sputtering. Such surface still showed (3×3) although with slightly weakened LEED spots due to less smooth surface, showing facetted pits and step-edges on the atomic scale. For a pristine surface, (3×3) was observed for low dose and/or high temperature O$_2$ exposure (at least up to 5400 L at 400°C), and (2×2) for high dose (at least up to 130 000 L at 370°C). However, for worn out surface, no (2×2) could be introduced through oxidation. Instead, after prolonged oxidation, a rectangular semicommensurate surface lattice, denoted ($\frac{3}{2} \times \frac{3}{2}$)-sq was observed, simultaneously with (3×3) from low dose oxidation up to at least 50 000 L at 390°C. This phase was highly stable and the reconstruction was restored even after atmospheric exposure. Only small amount or no Sb was observed in this structure.

The novel semicommensurate structure is observed as quasicrystal-like, consisting of dimer-like protrusions in the unit cells that have tendency to have parallel orientation in adjacent unit cells, with deviation from this construction causing significantly disordered structure. We expect that (3×3)–O is the initial step of this structure, as they are observed simultaneously and can be seen in adjacent areas with STM. LDOS at the interface could not unfortunately be probed with STS due to the thickness of this film. In case the DOS at the interface remains comparable to the one observed for
Chapter 5. Summary and Conclusions

(3×3)–O, observation of this phase could prove highly important concerning oxide/III–V technology, as it is observed in facettted areas with atomic step-edges, identically to what is expected from nano-patterned device-structures. Furthermore, quasicrystalline thin films with such a heterogeneous character interface are experimentally a novel concept [233]; our findings could pave way for a wide range of fundamental investigations on oxide/semiconductor materials systems.

5.4 Conclusions

In this PhD work, effects of oxidation on III–V semiconductor surfaces have been investigated, and from an applied point of view, focused on the effect of InO$_x$ as a buffer layer grown prior to an insulating oxide film.

Individual achievements during this work can be highlighted concisely as follows:

- A UHV preparation chamber was constructed and integrated into an existing LT-STM system, and successfully utilized in the work.

- New oxidized III–V surface structures were found,
  - GaSb(100)(1×3)–InO,
  - InSb(111)B(2×2)–O,
  - InSb(111)B(3×3)–O,
  - InSb(111)B($\frac{3}{2} \times \frac{3}{2}$)sq–O

- Oxidized III-V surface treatment was introduced to an optoelectronic device for the first time.

In addition, a particular goal of the work presented here was in elucidating and identifying many of the effects that result from the thermal oxidation of III–V semiconductor surfaces. With the methods utilized, we have been able to support previous studies and propose the commonly observed defect state distribution to occur in case studies for GaSb(100) and GaAs(100) from group V atom dangling or metallic bonds in the vicinity of the interface. The oxidation of Ga cannot be neglected either, as the processes naturally occur simultaneously at the initial stages of oxidation, and therefore, as the Ga–O bonds act as the initial sites of oxidation causing an amorphous structure with presumably variety of bonding environments. This suggests the necessity of stabilizing the crystal structure at the outermost atomic layers. Indeed, we have been able to see a similar effect and simultaneously improve
the interface chemistry in real device structures by introducing an InO$_x$ buffer, which does stabilize the crystal structure in oxidizing environment.

A further step towards integration of the crystalline InO$_x$ in real device structures has been the observation of similar crystalline oxide phases on (111) crystal face with similar processing parameters. In other words, there is no absolute prerequisite for planar (100) surface to gain the benefits from this treatment, which is a key point for i.e. patterned structure passivation, and has thus far remained uncertain. Our results give reason to believe that the crystalline oxide passivation method, when properly tailored, is viable for utilization in nanowires, FinFETs, and IR detector or LED mesas among other modern III–V semiconductor device processes.

Several fundamentally interesting paths for new investigations are opened up from the work carried out in this thesis: i) creating and investigating epitaxial oxide QW structures for III–V’s by depositing and oxidizing adjacent layers in a stacked fashion, ii) characterizing atomic structures of III–V/oxide nanowires with STM, or, iii) investigation and growth of semiconductor materials on the semicommensurate oxide template, etc. All of these require more complicated fabrication processes than what has been done previously, but on the other hand, this work has provided much understanding on such issues. Key enablers for studies such as mentioned above are tapping into materials prospective for such technologies by introducing InO$_x$ layers, the processing capability and know-how of these structures, and possibility to use various crystal faces that could enable integration on Si substrates, for example. Such studies made possible by this work would be ambitious but fully within possibilities.

The work carried out opens opportunities for continuation of fundamental research but also a solid base for utilization of III–V crystalline oxide phases in real device structures. Thus, this work also markedly contributes to bridging the fields of surface science and semiconductor technology, both of which can significantly complement each other.
Bibliography


Bibliography


[214] Basu, S., Singh, P. K., Sze, P.-W., Wang, Y.-H., “Al0.2Ga0.8As/In0.15Ga0.85As MOSPHEMT with low temperature LPD-deposited Al2O3 as gate dielectric”, Solid-State Electron. 54, 763–768 (2010).


