DISSERTATION

Advanced Modeling of Strained CMOS Technology

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik
von

STEPHAN-ENZO UNGERSBÖCK

[redacted]

Wien, im April 2007

Kurzfassung

WICHTIGE Hersteller von Mikroprozessoren verbessern seit der Einführung der 90 nm-Technologien die Transporteigenschaften von Silizium durch den Einsatz von Techniken, die mechanische Spannungen im Transistorkanal induzieren. Die gängigsten dieser neuen Techniken, die entweder globale oder lokale Verspannungen am Silizium Wafer einprägen, werden in dieser Arbeit behandelt.

Während sich bisher theoretische Arbeiten fast ausschließlich mit biaxial verspanntem Siliziumsubstrat mit $\{001\}$ Orientierung beschäftigten, sind die physikalischen Ursachen der Verbesserung der Elektronenbeweglichkeit durch uniaxiale Verspannung noch nicht zur Gänze geklärt. In dieser Arbeit wird der Einfluss allgemeiner homogener Verspannung – beschrieben durch den Spannungstensor $\hat{\varepsilon}$ – auf den Bulk-Transport und die Oberflächenbeweglichkeit von Silizium untersucht. Als Grundlage der Analyse dient die Bandstruktur von verspanntem Silizium, welche einerseits numerisch mittels der nicht-lokalen empirischen Pseudopotentialmethode mit Spin-Bahn Kopplung und andererseits analytisch mit Hilfe der störungstheoretischen $\boldsymbol{k} \cdot \boldsymbol{p}$ Methode berechnet wird. Zusätzlich zur Verschiebung der Leitungsbandtäler relativ zueinander läßt sich aus den Berechnungen der Bandstruktur eine Änderung der effektiven Elektronenmassen ableiten. Diese kann auf die Aufhebung der Entartung der zwei untersten Leitungsbänder im X-Punkt zurückgeführt werden, die von einer Scherspannung hervorgerufen wird. Mittels der $\boldsymbol{k}\cdot\boldsymbol{p}$ Methode werden auch analytische Gleichungen zur Beschreibung der Massenänderung hergeleitet.

Die Beweglichkeit in verspanntem Silizium wird durch das Lösen der Boltzmanngleichung mit Hilfe einer Monte Carlo Methode berechnet, der sowohl eine analytische Bandbeschreibung als auch die numerisch berechnete Bandstruktur zugrunde gelegt werden kann. Es zeigt sich, dass nur im Fall von kleinen Scherverspannungen ($< 0.5\%$) die Beweglichkeit bei niedrigen Feldstärken aus Monte Carlo Simulationen mit einer analytischen Beschreibung der Bandstruktur mit den Simulationsergebnissen übereinstimmen, die die numerische Bandstruktur verwenden. Dies unterstreicht die Wichtigkeit die gesamte Bandstruktur einer genauen Modellierung des Ladungsträgertransports in verspannten Silizium zugrunde zu legen.

Durch das Aufkommen der *Hybrid-Orientation* Technologie haben die Substratorierungen $\{110\}$ und $\{111\}$ an Bedeutung gewonnen. Da eine mechanische Verspannung zu einer Anisotropie der Ladungsträgerbeweglichkeit führt, muss die Verspannung auf die Richtung des Kanals und die Substratoberfläche abgestimmt werden, um den größten Beweglichkeitszuwachs zu erreichen. Um die Auswirkungen der Verspannung auf die Oberflächenbeweglichkeit zu bestimmen, wird die Subbandstruktur durch selbstkonsistentes Lösen der Schrödinger-Gleichung und der Poisson-Gleichung berechnet. Im zweidimensionalen Elektronengas bewirkt die Scherspannung eine energetische Verschiebung der Subbandleitern und eine Änderung der effektiven Massen. Beweglichkeitssimulationen zeigen, dass durch die inhärente Aufspaltung der Subbänder bei starker geometrischer Quantisierung, wie etwa bei Dünnschichttransistoren, die Änderung der effektiven Masse gegenüber der Verschiebung der Subbandleitern an Bedeutung gewinnt.

Des Weiteren wird in dieser Arbeit der Einfluss des Pauli-Verbots auf den Elektronenladungsträgertransport untersucht. Dies ist vor allem im zweidimensionalen Elektronengas von Bedeutung, wo das Fermi-Niveau bei starker Inversion deutlich über dem Minimum des untersten Leitungsbandes liegt.

Abstract

STARTING with the introduction of the 90 nm CMOS technology node many important microprocessor manufacturers improve the transport characteristics of silicon using techniques that induce strain in the transistor channel. The most common techniques that produce strain globally or locally on the wafer are being reviewed in this work.

While most theoretical work has been performed for biaxially strained Si with $\{001\}$ substrate orientation, a thorough theoretical analysis of electron mobility enhancement in arbitrarily strained Si is missing. In this work the effect of a general homogeneous strain – described by the strain tensor $\hat{\varepsilon}$ – on the transport in bulk Si and Si inversion layers is analyzed. The band structure of strained Si is calculated numerically using the empirical pseudopotential method with nonlocal and spin-orbit corrections. The results of band structure calculations reveal that shear strain changes the effective electron masses in addition to the splitting of the six Δ-valleys. The effective mass change can be attributed to the lifting of the degeneracy of the two lowest conduction bands at the X-point due to shear strain. Using the $\boldsymbol{k}\cdot\boldsymbol{p}$ method analytical expressions for the effective mass change are derived.

The transport properties of strained Si are investigated by solving the semiclassical Boltzmann equation using the Monte Carlo (MC) method employing fullband and analytical band models. The low-field electron mobility resulting from MC simulations using an analytical description of the electron bands and the fullband description coincide only for not too high shear strain ($< \pm 0.5\%$). At larger shear strain the band deformation is so pronounced, that fullband modeling is required.

The *hybrid-orientation* technology combines different silicon substrate orientations and channel directions on the same wafer and can be used in conjunction with strain techniques. Since strain yields an anisotropic mobility, the proper channel direction and substrate orientation have to be chosen to obtain the maximum mobility enhancement. The effect of strain on the inversion layer mobility of electrons is investigated by calculating the subband structure using a self-consistent Schrödinger-Poisson solver. In the two-dimensional electron gas shear strain results in a shift of the subband ladders and a change of the effective electron masses. From simulations of the effective electron mobility in ultra-thin-body MOSFETs it can be concluded that the change of the effective masses is the dominant effect leading to mobility enhancement, since the strong geometrical confinement yields a large intrinsic splitting of the subband ladders, such that an additional splitting induced by strain has only a little effect.

The effect of degeneracy both on the phonon-limited mobility and the effective mobility including surface-roughness scattering is studied using a new MC algorithm developed in this work. By comparison with results from MC simulations where the Pauli exclusion principle is neglected, it is shown that a correct treatment of degenerate carrier statistics of the 2DEG of Si inversion layers is important.

Danksagung

VIELE Menschen haben am Gelingen dieser Arbeit einen wesentlichen Anteil geleistet. Meinen Dank möchte ich vor allem Prof. Dr. Siegfried Selberherr und Prof. Dr. Erasmus Langer ausdrücken, die es mir ermöglichten an ihrem Institut zu promovieren, an dem ich eine ausgezeichnete Infrastruktur vorfand.

Großer Dank gilt meinem Betreuer Prof. Hans Kosina, von dessen profundem Fachwissen ich sehr profitierte, und der sich trotz der Vielzahl seiner Aufgaben immer Zeit für mich genommen hat. Er hat diese Arbeit durch zahlreiche Anregungen wesentlich verbessert.

Fachliche Unterstützung fand ich auch bei Prof. Dr. Tibor Grasser, Dr. Mihail Nedjalkov, Dr. Andreas Gehring und bei Dr. Viktor Sverdlov, die mich durch ihre Kompetenz und ihren Arbeitseifer beeindruckten. Des weiteren profitierte ich von den Vortragsreihen der Gastprofessoren Prof. Dr. Christoph Jungemann und Prof. David Pulfrey.

Die ganzen Jahre hinweg wurde ich außerdem von einer Kollegenschaft begleitet, die ihresgleichen sucht. Rund um die Uhr fand ich hilfsbereite Kollegen vor, die mir immer zur Seite standen. Zuallererst sei an dieser Stelle meinem langjährigen Zimmerkollegen Robert Entner gedankt. Mit ihm verbindet mich eine Freundschaft, die hoffentlich noch viele Jahre währen wird, auch, wenn sich unsere beruflichen Wege jetzt trennen. Das gleiche gilt für meine Kollegen Oskar Baumgartner, Johann Cervenka, Siddhartha Dhar, Klaus Dragosits, Philipp Hehenberger, Stefan Holzer, Gerhard (Monte-)Karlowatz, Markus Karner, Gregor Meller, Vassil Palankovski, Alireza Sheikholeslami, Mahdi Pourfath, Martin Vasicek, Martin Wagner, Paul und Stephan Wagner, Wilfried Wessner und Robert Wittmann.

Schließlich sei meiner Famile gedankt. Sie war immer für mich da und gab mir dadurch mehr Kraft als sie wahrscheinlich vermutet.

Zuletzt möchte ich mich bei meiner Freundin Alexandra bedanken. Sie schenkte mir letztes Jahr einen wundervollen Sohn und unterstützte mich immer mit viel Liebe und Geduld.

Viele Personen, die mir bei dieser Arbeit und während meiner Zeit am Institut für Mikroelektronik ebenfalls halfen, sind ungenannt geblieben. Auch ihnen sei an dieser Stelle gedankt. Ich freue mich, dass mich eine so erstaunlich große Anzahl von Menschen unterstützt.

Contents

List of Abbreviations and Acronyms

ABMC	...	Monte Carlo simulation based on analytical band description
nMOS	...	n-channel MOS
pMOS	...	p-channel MOS
BTE	...	Boltzmann transport equation
CESL	...	Contact etch stop liner
CMOS	...	Complementary MOS
DSL	...	Dual stress liner
EPM	...	Empirical pseudopotential method
FBMC	...	Monte Carlo simulation based on full band description
FET	...	Field-effect transistor
HOT	...	Hybrid orientation technology
ITRS	...	International Technology Roadmap for Semiconductors
MC	...	Monte Carlo
MOS	...	Metal-oxide-semiconductor
MOSFET	...	MOS field-effect transistor
RTA	...	Relaxation-time approximation
SEG	...	Selective epitaxial growth
SGOI	...	SiGe on insulator
SIA	...	Semiconductor Industry Association
SMT	...	Stress memorization technique
SOI	...	Silicon on insulator
SSGOI	...	Strained Si on SiGe on insulator
SSDOI	...	Strained Si directly on insulator

STI	...	Shallow trench isolation
TCAD	...	Technology computer-aided design
UTB	...	Ultra-thin-body
VLSI	...	Very large scale integration

List of Symbols

Notation

x	...	Scalar
$\boldsymbol{x}$	...	Vector
$\hat{x}$	...	Tensor
$\mathcal{X}$	...	Operator
$\underline{A}$	...	Matrix
A_{ij}	...	Elements of the matrix $\underline{A}$
$\boldsymbol{x} \cdot \boldsymbol{y}$	...	Scalar product
$[hk\ell]$	...	Miller indices to specify a crystal direction
$\langle hk\ell \rangle$	...	Miller indices to specify equivalent crystal directions
$(hk\ell)$	...	Miller indices to specify a crystal plane
$\{hk\ell\}$	...	Miller indices to specify equivalent crystal planes

Physical Quantities

Symbol	Unit	Description
$\mathcal{Q}(f)$	s^{-1}	Collision operator
e_{ij}, γ_{ij}	1	Engineering strain component (i, j)
ε_{ij}	1	Component (ij) of the strain tensor
σ_{ij}	GPa	Component (ij) of the stress tensor
C_{ijkl}	GPa	Component $(ijkl)$ of the elastic stiffness tensor
c_{ij}	GPa	Component (ij) of the contracted stiffness tensor
S_{ijkl}	GPa^{-1}	Component $(ijkl)$ of the elastic compliance tensor
s_{ij}	GPa^{-1}	Component (ij) of the contracted compliance tensor
D_n	$\mathrm{m}^2\mathrm{s}^{-1}$	Electron diffusion coefficient
E	eV	Energy
E_f	eV	Fermi energy
E_g	eV	Band gap energy
$\boldsymbol{E}$	Vm^{-1}	Electric field
$\boldsymbol{F}$	N	Force
$f(\boldsymbol{r}, \boldsymbol{k}, t)$	1	Distribution function
ϕ	V	Electrostatic potential
g	$\mathrm{m}^{-3}\mathrm{eV}^{-1}$	Density of states
$\hbar\omega$	eV	Phonon energy
k	m^{-1}	Wave number
$\boldsymbol{k}$	m^{-1}	Wave number vector
k_f	m^{-1}	Radius of Fermi sphere
κ_diel	$\mathrm{AsV}^{-1}\mathrm{m}^{-1}$	Dielectric permittivity of dielectric layer
κ_sc	$\mathrm{AsV}^{-1}\mathrm{m}^{-1}$	Dielectric permittivity of semiconductor
μ_n	$\mathrm{m}^2\mathrm{V}^{-1}\mathrm{s}^{-1}$	Electron mobility
μ_eff	$\mathrm{m}^2\mathrm{V}^{-1}\mathrm{s}^{-1}$	Effective electron mobility
μ_p	$\mathrm{m}^2\mathrm{V}^{-1}\mathrm{s}^{-1}$	Hole mobility
m	kg	Mass
n	m^{-3}	Electron concentration
n_i	m^{-3}	Intrinsic concentration
N_D	m^{-3}	Concentration of donors
N_A	m^{-3}	Concentration of acceptors
Ψ	$\mathrm{m}^{-1/2}$	Wave function
$\mathbf{r}$	m	Space vector

Symbol	Unit	Description
t	s	Time
T	K	Temperature
$\mathcal{T}$	eV	Operator of kinetic energy
$\boldsymbol{v}$	ms^{-1}	Velocity vector

Constants

h	...	Planck's constant	$6.6260755 \times 10^{-34}\,\mathrm{Js}$
$\hbar$	...	Reduced Planck's constant	$h/(2\,\pi)$
$\mathrm{k_B}$	...	Boltzmann's constant	$1.380662 \times 10^{-23}\,\mathrm{JK}^{-1}$
q	...	Elementary charge	$1.6021892 \times 10^{-19}\,\mathrm{C}$
$\mathrm{m_0}$	...	Electron rest mass	$9.1093897 \times 10^{-31}\,\mathrm{kg}$
κ_0	...	Dielectric constant of vacuum	$8.8541878 \times 10^{-12}\,\mathrm{AsV}^{-1}\mathrm{m}^{-1}$

Chapter 1

Introduction

For more than two decades the rapid progress in complementary metal-oxide-semiconductor (CMOS) technology was accompanied by a tremendous pace of scaling, leading to an enormous increase of the speed and functionality of electronic devices. This trend is expected to continue in the coming decade as predicted and institutionalized by the International Technology Roadmap for Semiconductors [SIA06].

The end of scaling has often been predicted in the past, but engineers' ingenuity has regularly proven wrong these predictions [Haensch06]. Two remarkable failures involve a predicted limit for spatial resolutions of about 400 nm imposed by lithography [Hoeneisen72, Wallmark75] and a lower bound of 3 nm for the gate oxide thickness, below which an unacceptable gate leakage current was expected [Hu96, Stathis98].

However, when examining high-performance devices in recent technology generations, early signs of scaling limits can be seen. Even though potential brick walls for scaling have been overcome, it is becoming increasingly difficult to meet metal-oxide-semiconductor field-effect transistor (MOSFET) performance gains with reasonable device leakage. To obtain the projected performance enhancement of 30% per generation, device designers have been forced to relax the device subthreshold leakage from one nA/µm for the 250 nm node to several hundreds of nA/µm for the 65 nm node [Chang03]. Consequently, the gate leakage current constitutes a significant portion of the power budget of microprocessors. Another critical scaling issue involves the increase of the source-drain series resistance resulting from the need for ultra-shallow p-n junctions in the source-drain region [Skotnicki05]. To keep the source-drain series resistance at a reasonable fraction of the total channel resistance (~10%), several alternative MOSFET structures have been proposed, such as non-overlapped gate structures, which do not require ultra-shallow source-drain junctions [Boeuf01, Lee02], or structures with metallic source and drain electrodes to minimize the series resistance [Connelly03, Fritze04, Zhu04].

In the present time the scaling of the gate oxide thickness has halted [Haensch06]. Advanced single gate structures, such as FinFETs, ultra-thin-body (UTB) MOSFETs, or multiple gate MOSFETs might provide a path to scaling CMOS to the end of the ITRS roadmap. Additionally, the material properties of a device are now being included in the scaling analysis. For example, the permittivity constant of the gate insulator has only slightly participated in scaling in the past. This is expected to change with the introduction of metal gates and hafnium-based high-κ

dielectrics in the 45 nm technology node, which will be in mass production this year [Intel07].

Another parameter not included in scaling in the past was the mobility of the channel material. This fact was especially critical, since mobility tends towards lower values due to higher vertical fields in down-scaled MOSFET devices. With the 90 nm technology node *strain techniques* have been introduced that efficiently increased the transistor drive current by enhancing the mobility of carriers in the channel [Rim02, Thompson04].

Strain engineering is a key element in current CMOS technologies and is widely believed to take a key position also in the future, since the benefits caused by mobility enhancement in the channel are comparatively big [Thompson06]. Furthermore, it has been demonstrated that this technology can also be used in non-classical CMOS structures [Rim03, Andrieu06]. For example, the advantages of a UTB MOSFET, such as the better electrostatic control of the channel by the gate in the off-state, can be combined with enhanced carrier transport provided by strain. Therefore, increasing emphasis is put on this technology to enhance chip performance. Since the maximum performance enhancement this new technology can deliver is still not known, it is convenient to do numerical simulations with accurate physical models in order to gain a better understanding of the underlying mechanisms responsible for the observed mobility enhancement. In this thesis, the effect of strain on the band structure of silicon (Si) and the electron mobility enhancement in bulk MOSFETs and UTB MOSFETs is analyzed numerically.

The thesis is organized as follows: In **Chapter 2** the reader is introduced to the short but versatile history of CMOS strain engineering. After reviewing the most important global and local strain techniques, several state-of-the-art technologies are outlined. In **Chapter 3** the bulk band structure of strained cubic semiconductors is discussed. A $\boldsymbol{k}\cdot\boldsymbol{p}$ method capturing the effect of strain on the lowest conduction band of Si is developed. Additionally, the adaptation of the empirical pseudopotential method (EPM) including spin orbit coupling to incorporate strain effects is presented. In **Chapter 4** the subband structure of strained Si inversion layers formed at Si-SiO$_2$ interfaces with various substrate orientations is discussed. The subband energies and the wave functions of Si inversion layers are obtained by a self-consistent solution of the Schrödinger equation and the Poisson equation. **Chapter 5** is dedicated to the numerical modeling of mobility. The transport properties of strained Si are investigated by solving the semiclassical Boltzmann equation using the Monte Carlo (MC) method. The models for the scattering mechanisms used in the MC simulations of bulk Si and Si inversion layers are given. Additionally, a MC method that includes degeneracy effects for small driving fields is presented. Simulation results are summarized in **Chapter 6**. The analytical band structure obtained from the $\boldsymbol{k}\cdot\boldsymbol{p}$ theory is compared to numerical calculations using the EPM. The subband structure of strained Si inversion layers is presented for various strain configurations and substrate orientations. The simulation results for the bulk electron mobility of strained Si are compared to measurements and to phenomenological mobility models based on the piezoresistance coefficients. The electron mobility is calculated using the MC method employing fullband (FBMC) and analytical band models (ABMC). A comparison of the results allows the extraction of the limits of validity of the analytical band model. Finally, the effect of strain and degeneracy on the effective mobility in Si inversion layers for (001) and (110) substrate orientation is presented and the influence of the Si body thickness on the effective mobility in ultra thin body MOSFETs is discussed.

Chapter 2

Strain Engineering

The influence of strain on the intrinsic mobility of Si was first investigated in the early 1950's [Hall51, Smith54]. While this effect was not exploited initially, the idea was revived in the early 1990's [Fitzgerald91]. In 1992 it was first demonstrated that n-channel MOSFETs on a strained Si substrate exhibit a 70% higher effective mobility (μ_{eff}) than those on unstrained substrates [Welser92, Welser94]. Ever since semiconductor industry has adopted several different technologies to introduce strain in the channel of a MOSFET.

Strain technologies are based on mechanically stretching and/or compressing the Si crystal lattice by various means. Innovative techniques introducing stress in the Si channel have been developed, which require only small modifications of some process steps, thus keeping the additional costs small. At the same time the integrability of strained Si in the CMOS manufacturing process flow is retained.

A classification of strain techniques into two categories is possible. Strain is introduced across the entire substrate in *global* strain techniques, whereas *local* techniques induce strain in selected regions of the wafer. Some of the most prominent strain technologies that are currently used in industry are given in Figure 2.1. A key challenge of all technologies is their ability to be integrated into the CMOS manufacturing processes and to avoid significant increase in processing costs. In this chapter several strain technologies and the current progress in high-mobility strained MOSFETs are reviewed.

2.1 Global Strain Techniques

Most of the pioneering work on strained Si was focused on biaxial global strain generated by epitaxial growth of a thin Si layer on a relaxed SiGe virtual substrate [Welser92, Welser94]. Because of the lattice mismatch between Si and SiGe, the lattice of the Si layer is biaxially tensile strained in the plane of the interface. On {001} oriented substrates this deformation results in enhanced carrier transport in the strained Si layer, and mobility enhancements of 110% for electrons and 45% for holes have been demonstrated on sub-100 nm strained Si MOSFETs [Rim02].

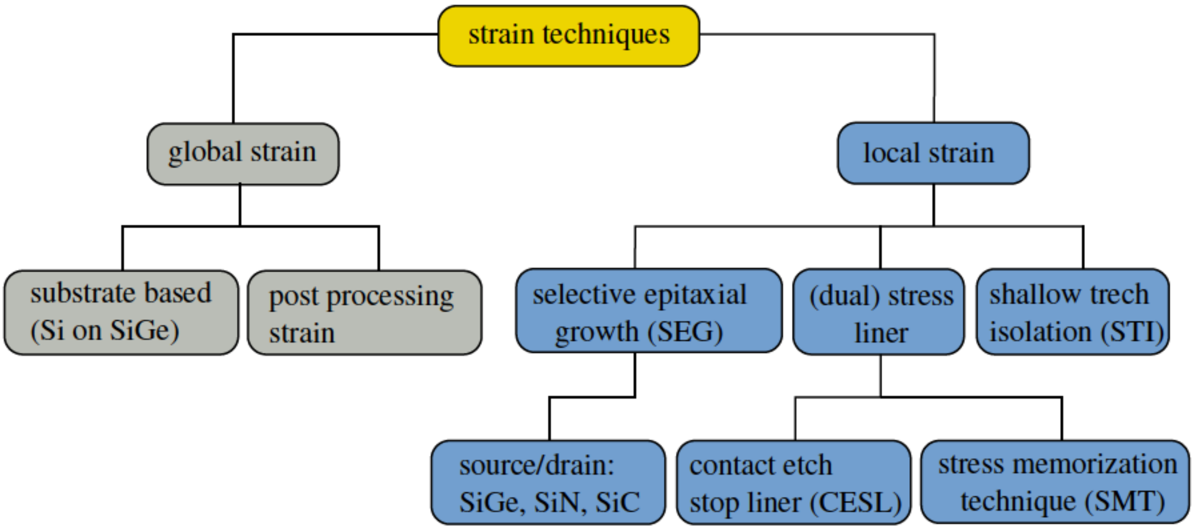


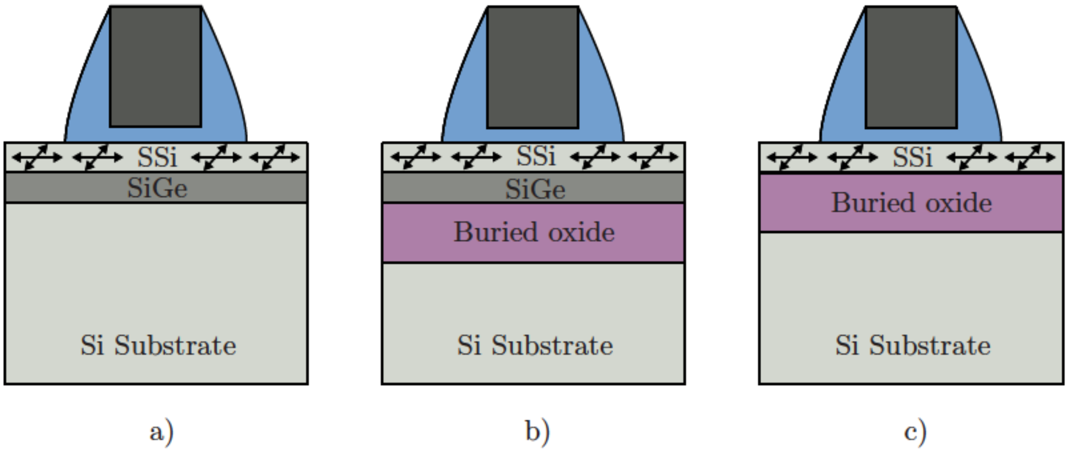
Figure 2.1: Classification of strain technologies.

Using layer transfer and wafer bonding techniques, global strain can also be integrated on SOI substrate (see Figure 2.2). Electron and hole mobility enhancement comparable with the enhancement of wafers without the insulating layer were observed in ultra-thin strained Si layers on SiGe on insulator (SSGOI, Figure 2.2b) [Cheng01, Ghyselen04, Sadaka04, Andrieu06] and strained Si layers directly on insulator (SSDOI, Figure 2.2c) [Huang01, Rim03].

Technologies using ultra thin strained Si directly on insulator (SSDOI) are especially promising, since in those structures the SiGe layer is eliminated before transistor fabrication, hence critical process-integration problems related to the SiGe layer can be avoided. Among the difficulties of SSGOI are:

- An enhanced diffusion rate of dopants (B, As) in SiGe [Uppal04].
- Significant device self-heating is observed in SSGOI devices because of the lower thermal conductivity of SiGe [Rim00].
- SiGe strain relaxation induces a high density of defects in the neighboring strained Si layer [Fitzgerald91].
- The thermal budget window for SGOI is smaller due to Ge up-diffusion from the underlying SiGe layer [Cheng01].

Lately, an alternative back-end approach was proposed, where strain is introduced after the wafer has been completely processed [Belford01, Haugerud03]. The process starts with thinning the processed wafer to less than 10 μm, transferring it to a polymer film and then mechanically straining the Si membrane. This technique allows the introduction of biaxial and uniaxial strain parallel to the substrate surface without inducing defects in the Si layer. Within the elastic limit, the wafer can be safely mechanically strained and then bonded to a final substrate. At ultra-low strain levels (0.031%), holes in p-channel MOSFETs showed an increase in effective mobility μ_{eff} of 14.35% and an enhancement in saturation current of 14.56%. An improvement



a) b) c)

Figure 2.2: Schematic plot of strained Si MOSFETs using global strain: a) strained Si on SiGe on bulk wafer; b) strained Si on SiGe on insulator (SSGOI); c) strained Si directly on insulator (SDGOI).

in mobility of 18.49% and in saturation current of 18.05% was observed for n-channel MOSFETs strained by 0.052% [Haugerud03]. This relatively new method is promising as it is cost effective, however some yield and reliability issues have to be resolved before it can be used for full-scale IC manufacturing [Bera06].

A major drawback common to all global strain techniques for CMOS technology is that they can provide only one type of strain. Since the mobilities of electrons and holes are differently affected by strain, a global strain configuration, for example, compressive biaxial strain, can be beneficial for p-channel MOSFETs, but deteriorate the n-channel MOSFET performance. This problem is circumvented by local strain techniques, which are able to provide different strain patterns in n-channel- and p-channel MOSFETs.

2.2 Local Strain Techniques

Various CMOS fabrication processes can be exploited to induce strain in the transistor channel (see Figure 2.1). Since these processes generate local strain that depends on the position in the MOSFET channel, techniques based on process-induced strain are frequently called local strain techniques.

Starting from the late 1990's effects related to local stress arising from various process steps on the performance of MOSFETs were investigated. It was found that mainly the following process steps are responsible for stress in the transistor channel:

- shallow trench isolation (STI) [Scott99, Matsumoto02]
- silicidation at the source/drain region [Steegen99]
- nitride contact etch stop liners (CESL) [Ito00, Shimizu01]

Even though process-induced strain initially was not able to provide as large strain levels as global strain, the local techniques enjoy three main advantages: (i) strain can be independently tailored to optimize performance enhancement for both n-channel- and p-channel MOSFETs, (ii) the threshold voltage shift is smaller in uniaxially stressed MOSFETs [Lim04], and (iii) local stress techniques are cheaper and more compatible with standard CMOS technology [Khamankar04].

The main challenge is to optimize process modules so as to maximize beneficial effects from stressors while minimizing negative side effects. A drawback of process-induced strain techniques is their strong device geometry dependence, making their scaling behavior less predictable [Eneman05].

Four important stress-transfer techniques will be discussed in the following: (i) the contact etch stop liner technique (CESL); (ii) the stress memorization technique (SMT); (iii) a technique based on selective epitaxial growth (SEG) of the source/drain regions, and (iv) stress from shallow trench isolation (STI).

2.2.1 Contact Etch Stop Liner Technique

In this technique a highly stressed liner is uniformly deposited on top of the wafer after the silicide formation, transferring the stress from the liners to the channel. The stress-transfer depends on the thickness and the material properties of the liner [Ito00]. Tensile liners have been reported to improve the saturated drive current of n-channel and p-channel MOSFETs by 11% and 20%, respectively [Yang04].

If one single liner is used, only one type of stress can be introduced in both the n-channel and the p-channel MOSFET. As a consequence, similar to global strain techniques, just one transistor type is enhanced while the other might even be degraded. In order to achieve performance enhancement in both n-channel- and p-channel MOSFETs, two types of stress liners have to be processed. In a dual stress liner (DSL) process a highly compressive nitride is deposited on top of the p-channel MOSFET, whereas a highly tensile nitride is deposited on top of the n-channel MOSFET [Sheraw05, Yang04].

Si_3Ni_4 layers with more than 2.0 GPa tensile and 2.5 GPa compressive stress have recently been reported. These films introduce more than 1.0 GPa stress in the MOSFET channel [Arghavani06] being comparable in magnitude to stress-induced by selective epitaxial growth techniques (see Section 2.2.3).

2.2.2 Stress Memorization

In a process using the stress memorization technique the conventional spike anneal for dopant activation is performed after the deposition of a tensile stressor capping layer [Khamankar04, Ota02, Chen04, Horstmann05]. This layer is subsequently removed before the salicide process. Even though the stressor nitride layer is removed from the final structure, the stress has been transferred from the nitride film to the channel during annealing and is memorized during re-crystallization of source/drain and the poly gate amorphized layers. Improvements of the on-current up to 15% were reported for n-channel MOSFETs using this technique [Chan05].

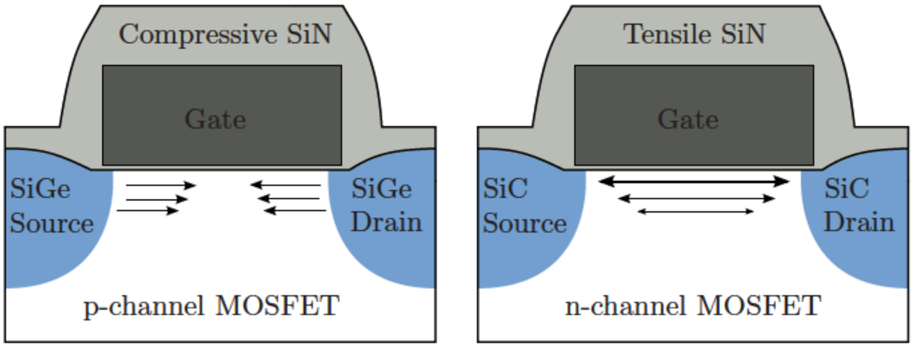


Figure 2.3: Combination of epitaxial source/drain regions and stress liner to induce compressive uniaxial stress for the p-channel MOSFET and tensile uniaxial stress for the n-channel MOSFET.

2.2.3 Selective Epitaxial Growth Technique

Uniaxial stress can be introduced into the channel of a MOSFET by selectively growing a local epitaxial film in the source and drain regions of the transistor. Large uniaxial stress can be created using this technique, originating from the difference of the lattice constants in the channel region and source/drain region. For this purpose, first the source and drain regions are etched, creating a recess area. Afterwards, this recess is filled by epitaxy. Alternatively, the epitaxial film can be directly grown on top of source and drain without previously etching these regions [Chidambaram06].

The level of stress depends on the thickness of the epitaxial layer and the lattice constant mismatch. Epitaxially grown SiGe is used to induce uniaxial compressive stress in p-channel MOSFETs [Thompson04, Eneman05, Horstmann05, Ouyang05, Bai04, Ouyang05, Zhang05]. In the same manner, tensile stress can be induced in the n-channel MOSFET by using $Si_{1-x}C_x$ stressors with molefractions around 1 % [Ang04]. $Si_{0.99}C_{0.01}$ was found to induce the same amount of stress as $Si_{0.75}Ge_{0.25}$ [Ang05]. It was recently reported that for n-channel MOSFETs the implementation of the $Si_{0.987}C_{0.013}$ source/drain regions provide significant drive current enhancement of up to 50% at a gate length of 50 nm [Chui07].

2.2.4 Stress from Shallow Trench Isolation

In sub-micron technologies with small active area, the mechanical stress induced from shallow trench isolation (STI) cannot be neglected [Wang03]. STI can induce lateral (parallel to the channel) and transversal mechanical stress (in the width direction of transistor). A reduction of the transistor width amplifies the compressive stress in the channel induced by STI, which was shown to yield an increase of the hole mobility [Chan03].

2.3 Strain Technologies used in High Volume Production

Since the local strain approach turned out to be more promising for industrial applications, the first strain technologies used in high volume production were developed on the basis of uniaxial process-induced stress.

Starting from the 90 nm node, companies such as IBM [Ouyang05], Intel [Thompson04], Texas Instruments [Chidambaram06], and Freescale [Zhang05], incorporated the selective epitaxial growth technique to transfer uniaxial compressive stress into the Si channel by growing a local epitaxial film of SiGe in the source and drain region of p-channel MOSFETs. Depending on the proximity of the SiGe to the channel and the Ge content, 500–900 MPa stress is created in the channel [Mohta05]. Using this technique impressive saturation drain current enhancement up to 20%–25% have been demonstrated for p-channel MOSFETs [Thompson04, Zhang05]. A tensile Si nitride-capping layer is used to introduce tensile uniaxial strain into the n-channel MOSFET, which enhanced the drive current by 10% [Thompson04].

Alternatively, a less complex technique was jointly developed by AMD and IBM [Yang04]. Compressive and tensile capping layers grown on top of the transistors were used as local stressors instead of epitaxial films in the source and drain region [Pidin04, Sheraw05]. Si nitride (Si_3N_4) was found to produce either tensile or compressive strain depending on its deposition conditions. During the process, a highly tensile Si nitride layer is first deposited using thermal chemical vapor deposition over the entire wafer. This layer is afterwards etched away selectively from the p-MOS active areas. A compressive Si_3Ni_4 layer is then deposited using plasma-assisted chemical vapor deposition. Then the nitride layer on the n-channel transistors is etched away, resulting in wafers with n-channel transistors under tensile and p-channel transistors under compressive uniaxial strain. This dual stress liner (DSL) approach resulted in drive current enhancement of 11% (20%) for n-channel (p-channel) MOSFETs [Yang04].

In future CMOS technology nodes various strain techniques may be combined to yield even larger strain levels, as shown in Figure 2.3. P-channel MOSFETs with selective SiGe epitaxial layers providing compressive strain, and n-channel MOSFETs that are uniaxially strained by tensile cap films have been successfully combined on the same wafer recently [Jan05]. An optimized stress integration on SOI CMOS was presented by [Horstmann05], where an embedded SiGe process and a compressively stressed liner film were used to induce compressive strain in the p-channel MOSFET, whereas a stress memorization process and a tensile stressed liner film were used to induce tensile strain in the n-channel MOSFET. With optimization, the different stress techniques were shown to be highly compatible and additive to each other, improving p-channel and n-channel MOSFET saturation drive current by 53% and 32%, respectively.

2.3.1 Hybrid Orientation Technology

An alternative approach yielding mobility improvement in Si exploits the dependence of the carrier mobility in Si inversion layers on the crystal orientation and on the current flow direction. For example, for holes the mobility is 2.5 times higher for (110) surface orientation than for standard (001) orientation depending on the applied effective vertical field. In the *hybrid orientation technology* (HOT), which is based on wafer bonding techniques and selective epitaxy, the larger carrier mobility of holes for (110) oriented substrate is exploited to enhance the performance of p-channel MOSFETs [Ouyang05, Yang06, Sheraw05]. HOT seems promising because processes are directly compatible with existing CMOS technology and strain engineering.

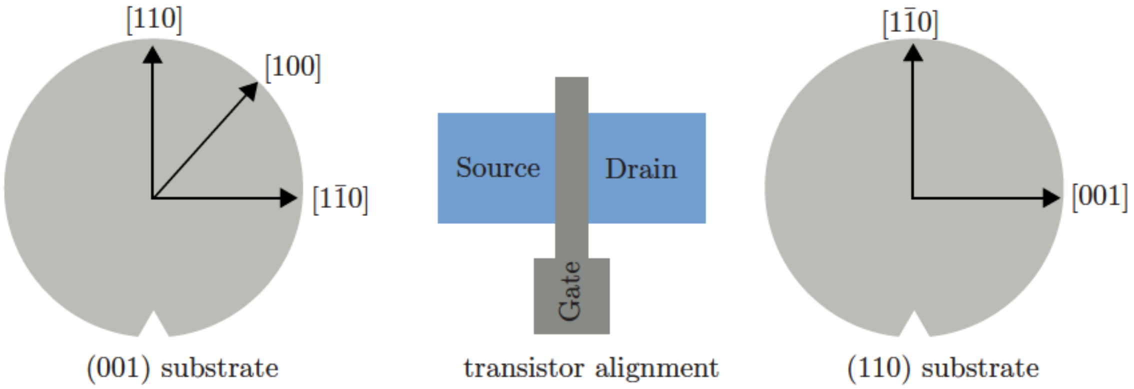


Figure 2.4: Crystallographic directions on the (001) and the (110) substrate. Conventionally, MOSFET channels are aligned along the ⟨110⟩ direction on (001) substrate. Highest electron mobility is obtained on (001) substrate, highest hole mobility on (110) substrate with channel direction ⟨110⟩.

Chapter 3

Strain Effects on the Bulk Band Structure

THE band structure describes the states of energy in the crystal momentum space that electrons and holes are allowed to have. It presents the electronic dispersion relation under the influence of the potential of the solid. The band structure determines several important characteristics, in particular its electronic and optical properties.

In this chapter, first the basic physical definitions are introduced, such as the strain and stress tensors and how they are related in cubic semiconductors (Section 3.1 to Section 3.3).

The vast majority of semiconductors of interest for electronic and optoelectronic applications have a diamond structure with an underlying face centered cubic (fcc) lattice consisting of two atoms per basis. The basic properties of the diamond structure and the band structure of relaxed Si are presented in Section 3.4.

The symmetry of the crystal in real space is directly reflected in the symmetry of the band structure in momentum space. Using group-theoretic methods many properties of the band structure can be obtained from these symmetries. Section 3.5 is devoted to the effect of strain on the symmetry of the diamond structure. Special focus is put on the consequence of the strain-induced reduction of symmetry for band structure calculations.

Hereafter, three different methods of calculating the effect of strain on the band structure are presented:

- Historically, the first approach was developed by Bardeen and Shockley and is known as the *deformation potential theory*. The perturbation caused by strain is attributed to an additional Hamiltonian which is linearly proportional to the deformation potential operator and to strain. First order perturbation theory is used to calculate the effect of strain on the band structure and analytical expressions for the strain-induced energy shifts of the conduction- and valence-bands can be obtained. The salient features of this theory and relevant results are summarized in Section 3.6.
- Strain does not only shift the conduction and valence bands, but also changes the curvature

of the energy bands. In Section 3.7 the main features of the $\boldsymbol{k}\cdot\boldsymbol{p}$ method are briefly outlined, as this method is able to capture the deformation of the shape of the energy bands under strain. It is frequently used to model the influence of strain on the valence bands. On the contrary, the impact of strain on the curvature of the conduction band has often been neglected. In Section 3.7.2 a $\boldsymbol{k}\cdot\boldsymbol{p}$ analysis capturing the effect of strain on the lowest conduction band of Si is developed. It shows that shear strain can alter the curvature of the conduction bands and thus can reduce the effective masses of electrons.

- The *empirical pseudopotential method* (EPM) [Chelikowsky76] including nonlocal effects and spin orbit coupling is frequently used to calculate the band structure of semiconductors, as this method is efficient and requires only a small number of parameters. These are usually calibrated to match energy gaps and effective masses determined from experiments and are available for a large set of materials [Yu03]. The method can very naturally be adapted to incorporate strain effects and has been used to investigate the band structure of biaxially strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ grown on $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ for various surface orientations [Fischetti96a, Van de Walle86, Rieger93]. In Section 3.8 it is shown how a general strain tensor can be taken into account by this method for band structure calculation. Special focus is put on the orthorhombic distortion of the crystal structure resulting from uniaxial stress along the [110] direction. This type of stress is used in practice to enhance the electron mobility in n-channel MOSFETs.

3.1 Strain

In this section the basic expressions and notations for strain in cubic crystalline solids are established. A more detailed analysis can be found in textbooks [Bir74, Kittel96, Singh93]. Starting point for the definition of strain in a system is a set of orthonormal vectors $\boldsymbol{x}, \boldsymbol{y}$ and $\boldsymbol{z}$ embedded in an unstrained solid. These vectors are distorted to $\boldsymbol{x}', \boldsymbol{y}'$, and $\boldsymbol{z}'$ under the influence of a uniform deformation

$$\begin{aligned}\boldsymbol{x}' &= (1+\epsilon_{xx})\boldsymbol{x} + \epsilon_{xy}\boldsymbol{y} + \epsilon_{xz}\boldsymbol{z}\ ,\\ \boldsymbol{y}' &= \epsilon_{yx}\boldsymbol{x} + (1+\epsilon_{yy})\boldsymbol{y} + \epsilon_{yz}\boldsymbol{z}\ ,\\ \boldsymbol{z}' &= \epsilon_{zx}\boldsymbol{x} + \epsilon_{zy}\boldsymbol{y} + (1+\epsilon_{zz})\boldsymbol{z}\ ,\end{aligned} \tag{3.1}$$

where the coefficients ϵ_{ij} define the deformation of the system.

For a uniform deformation of a body point originally located at $\boldsymbol{r} = x\boldsymbol{x} + y\boldsymbol{y} + z\boldsymbol{z}$ the displacement to $\boldsymbol{r}' = x\boldsymbol{x}' + y\boldsymbol{y}' + z\boldsymbol{z}'$ is defined as

$$\begin{aligned}\boldsymbol{R} \equiv \boldsymbol{r}' - \boldsymbol{r} &= x(\boldsymbol{x}' - \boldsymbol{x}) + y(\boldsymbol{y}' - \boldsymbol{y}) + z(\boldsymbol{z}' - \boldsymbol{z})\\ &= (x\epsilon_{xx} + y\epsilon_{yx} + z\epsilon_{zx})\boldsymbol{x}\\ &+ (x\epsilon_{xy} + y\epsilon_{yy} + z\epsilon_{zy})\boldsymbol{y}\\ &+ (x\epsilon_{xz} + y\epsilon_{yz} + z\epsilon_{zz})\boldsymbol{z}\ .\end{aligned} \tag{3.2}$$

More generally, the displacement for a non-uniform deformation can be defined by introducing a position dependent vector function $\boldsymbol{u}(\boldsymbol{r})$,

$$\boldsymbol{R}(\boldsymbol{r}) = u_x(\boldsymbol{r})\boldsymbol{x} + u_y(\boldsymbol{r})\boldsymbol{y} + u_z(\boldsymbol{r})\boldsymbol{z}\ . \tag{3.3}$$

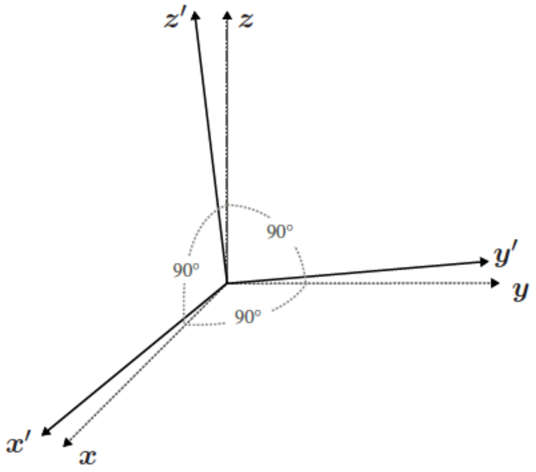


Figure 3.1: Deformation acting on a orthonormal coordinate system $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ yielding the strained axes $\boldsymbol{x}', \boldsymbol{y}'$, and $\boldsymbol{z}'$.

The vector function $\boldsymbol{u}$ can be related to the local strain, when taking the origin of $\boldsymbol{r}$ close to the region of interest. A Taylor series of $\boldsymbol{R}$ up to the first order, using $\boldsymbol{R}(\mathbf{0}) = \mathbf{0}$, and a comparison of the coefficients yields:

$$\begin{aligned} \epsilon_{xx} &= \frac{\partial u_x}{\partial x} & \epsilon_{yx} &= \frac{\partial u_x}{\partial y} & \epsilon_{zx} &= \frac{\partial u_x}{\partial z} \\ \epsilon_{xy} &= \frac{\partial u_y}{\partial x} & \epsilon_{yy} &= \frac{\partial u_y}{\partial y} & \epsilon_{zy} &= \frac{\partial u_y}{\partial z} \\ \epsilon_{xz} &= \frac{\partial u_z}{\partial x} & \epsilon_{yz} &= \frac{\partial u_z}{\partial y} & \epsilon_{zz} &= \frac{\partial u_z}{\partial z} \end{aligned} \tag{3.4}$$

Thus, the displacements can be written as $\epsilon_{ij} = \frac{\partial u_j}{\partial x_i}$.

Usually, instead of the displacements ϵ_{ij}, the strain tensor ε_{ij} is used to quantify the the deformation of an body in three dimensions. In the case of small deformations, the strain tensor is known as the *Green tensor* or *Cauchy's infinitesimal strain tensor*, which is defined as

$$\varepsilon_{ij} = \frac{1}{2}(\epsilon_{ij} + \epsilon_{ji}) \ . \tag{3.5}$$

The diagonal coefficients ε_{ii} of this symmetric tensor define the relative change in length in direction $\boldsymbol{x}_i$, and the off-diagonal terms ε_{ij} $(i \neq j)$ are the shear strains, which are related to the angular distortions.

Frequently, in literature the engineering strains e_{ij} are used which are related to the components of the strain tensor via

$$\begin{pmatrix} e_{xx} & e_{xy} & e_{xz} \\ e_{xy} & e_{yy} & e_{yz} \\ e_{xz} & e_{yz} & e_{zz} \end{pmatrix} = \begin{pmatrix} \varepsilon_{xx} & 2\varepsilon_{xy} & 2\varepsilon_{xz} \\ 2\varepsilon_{xy} & \varepsilon_{yy} & 2\varepsilon_{yz} \\ 2\varepsilon_{xz} & 2\varepsilon_{yz} & \varepsilon_{zz} \end{pmatrix} \ . \tag{3.6}$$

Here, sometimes the off-diagonal terms of the engineering strains are denoted as $\gamma_{ij} = e_{ij} = 2\varepsilon_{ij}$. Using the one suffix Voigt notation the six components of the strain tensor can be arranged in

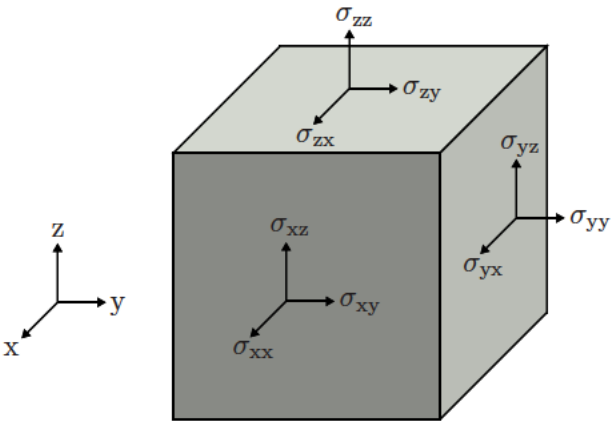


Figure 3.2: Stress components acting on infinitesimal cube.

a vector

$$(\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}, 2\varepsilon_{yz}, 2\varepsilon_{xz}, 2\varepsilon_{xy}) = (e_1, e_2, e_3, e_4, e_5, e_6) \ . \tag{3.7}$$

This notation is convenient when writing Hooke's law in cubic semiconductors as will be seen later.

3.2 Stress

Within a body the forces that an enclosed volume imposes on the remainder of the body must be in equilibrium with the forces upon it from the remainder of the body. This principle was first enunciated by Cauchy and provides a path to calculate stress. The stress at a point may be determined by considering a small element of the body enclosed by areas ΔA, on which forces $\Delta \boldsymbol{F}$ act. By making the element infinitesimally small, the stress vector σ is defined as the limit

$$\boldsymbol{\sigma} = \lim_{\Delta A \to 0} \frac{\Delta \boldsymbol{F}}{\Delta A} = \frac{\mathrm{d}\boldsymbol{F}}{\mathrm{d}A} \ . \tag{3.8}$$

From Figure 3.2 one can observe that the force acting on a plane can be decomposed into a force within the plane, the shear components, and one force perpendicular to the plane, the normal component. The shear stress can be further decomposed into two orthogonal force components giving rise to three total stress components acting on each plane.

Thus, the stress on an infinitesimally small element can be characterized by nine stress components in total, namely three normal and six shear components. These components form the stress tensor

$$\underline{\sigma} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_{zz} \end{pmatrix} , \tag{3.9}$$

where the shear stress components across the diagonal are identical as a result of static equilibrium

$$\sigma_{xy} = \sigma_{yx} \ , \quad \sigma_{yz} = \sigma_{zy} \ , \quad \sigma_{zx} = \sigma_{xz} .$$

	Silicon	Germanium	Units
c_{11}	166.0	126.0	GPa
c_{12}	64.0	44.0	GPa
c_{44}	79.6	67.7	GPa

Table 3.1: Elastic stiffness constants of Si and Ge [Levinshtein99].

3.3 Stress-Strain Relations

A mathematical expression of the stress-strain relation for the elastic deformation of materials was first suggested by Robert Hooke

$$\boldsymbol{F} = k\boldsymbol{u} \ . \tag{3.10}$$

Here, $\boldsymbol{F}$ is the applied force, $\boldsymbol{u}$ is the deformation of the elastic body subjected to the force $\boldsymbol{F}$, and k is the material dependent spring constant. Cauchy generalized Hooke's law for three dimensional elastic bodies

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl} \ , \tag{3.11}$$

where C_{ijkl} is the elastic stiffness tensor of order four, which contains 81 entries. The number of components can be reduced invoking symmetry arguments [Kittel96]. For a cubic semiconductor such as Si, Ge or GaAs, there are only three independent components, namely c_{11}, c_{12}, and c_{44}. The elastic stiffness constants for Si and Ge are given in Table 3.1.

Exploiting the symmetry of a cubic semiconductor the elastic stiffness tensor can be written as a 6 × 6 matrix, and generalized Hooke's law reduces to a set of six equations

$$\begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0 \\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{44} \end{pmatrix} \cdot \begin{pmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ 2\varepsilon_{yz} \\ 2\varepsilon_{xz} \\ 2\varepsilon_{xy} \end{pmatrix} . \tag{3.12}$$

If the stresses are known, the values for the strains are to be determined by inversion of (3.11). Introducing the elastic compliance tensor S_{ijkl}, the inverted equation becomes in the index notation

$$\varepsilon_{ij} = S_{ijkl}\sigma_{kl} \ , \tag{3.13}$$

or in matrix form

$$\begin{pmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ 2\varepsilon_{yz} \\ 2\varepsilon_{xz} \\ 2\varepsilon_{xy} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{11} & s_{12} & 0 & 0 & 0 \\ s_{12} & s_{12} & s_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & s_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & s_{44} \end{pmatrix} \cdot \begin{pmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{xz} \\ \sigma_{xy} \end{pmatrix} . \tag{3.14}$$

The elastic compliance constants s_{ij} are related to the elastic stiffness constants c_{ij} via

$$\begin{aligned} s_{11} &= \frac{c_{11}+c_{12}}{c_{11}^2+c_{11}c_{12}-2c_{12}^2} \,, \\ s_{12} &= \frac{-c_{12}}{c_{11}^2+c_{11}c_{12}-2c_{12}^2} \,, \quad \text{and} \\ s_{44} &= \frac{1}{c_{44}} \,. \end{aligned}$$

Note that the stiffness constants are traditionally represented by the symbol c_{ij}, while s_{ij} is reserved for the compliance constants.

3.3.1 The Miller Index Notation

The Miller indices can be used to specify directions and planes in a crystal [Ashcroft76, Kittel96]. The Miller indices of a plane are defined in the following way: First, three lattice vectors have to be defined. For cubic crystal systems, the lattice vectors are chosen along the edges of the crystallographic unit cell (unit cube). Any crystal plane intercepts the axes in certain points. The Miller indices are the ratios of these points and are given as a triplet of integer values $(hk\ell)$. A Miller index 0 means that the plane is parallel to the respective axis. Negative indices are indicated with a bar written over the number.

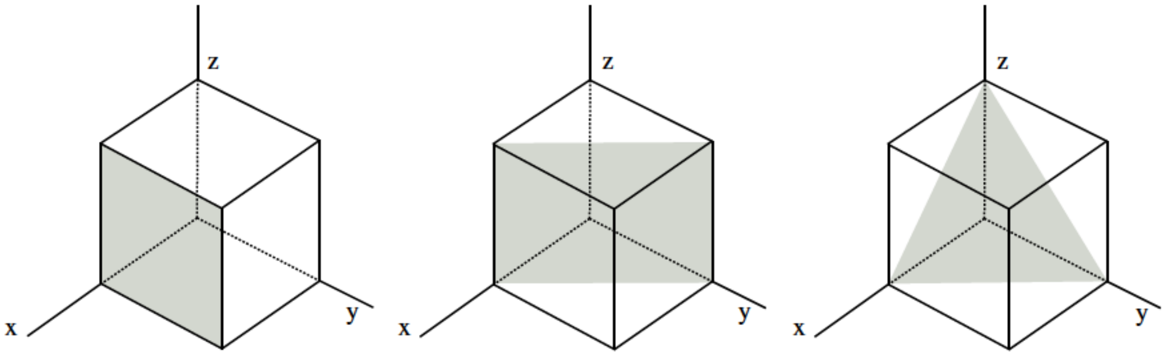
In the notation of [Ashcroft76], $[hk\ell]$ with square brackets instead of round brackets, denotes a direction in the basis of the lattice vectors. The notation $\{hk\ell\}$ denotes all planes that are equivalent to $(hk\ell)$ by the symmetry of the crystal. Similarly, the notation $\langle hk\ell\rangle$ denotes all directions that are equivalent to $[hk\ell]$ by symmetry.

In cubic crystal systems the Miller indices of a plane are the same as those of the direction perpendicular to the plane.

3.3.2 Strain Resulting from Uniaxial Stress

This section describes the calculation of the strain tensor resulting from a uniaxial stress of magnitude P along an arbitrary direction.

Analysis begins by adopting a coordinate system (x', y', z') in which the x' axis is parallel to the stress direction. This system is related to the coordinate system (x, y, z) of the primary



(100) (110) (111)

Figure 3.3: Three planes in the cubic system along with their Miller indices.

crystallographic axes of the semiconductor by a rotation $\underline{U}$

$$\underline{U}(\phi) = \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix} , \tag{3.15}$$

where θ and ϕ are the polar and azimuthal angles of the stress direction relative to the crystallographic coordinate system. In the primed coordinate system, the stress tensor has only one non-zero component, $\sigma_{xx}{}' = P$. The stress tensor in the crystallographic system can be calculated from

$$\sigma_{ij} = U_{\alpha i} U_{\beta j} \sigma'_{\alpha\beta} \ . \tag{3.16}$$

If uniaxial stress is applied along one of the directions [100], [110], [111], and [120], the related stress tensors in the principal system become:

$$\underline{\sigma}_{[100]} = \begin{pmatrix} P & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \underline{\sigma}_{[110]} = \begin{pmatrix} P/2 & P/2 & 0 \\ P/2 & P/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\underline{\sigma}_{[111]} = \begin{pmatrix} P/3 & P/3 & P/3 \\ P/3 & P/3 & P/3 \\ P/3 & P/3 & P/3 \end{pmatrix} \qquad \underline{\sigma}_{[120]} = \begin{pmatrix} P/5 & 2P/5 & 0 \\ 2P/5 & 4P/5 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.17}$$

By inserting 3.17 in 3.14 the strain tensors can be calculated:

$$\varepsilon_{[100]} = P\begin{pmatrix} s_{11} & 0 & 0 \\ 0 & s_{12} & 0 \\ 0 & 0 & s_{12} \end{pmatrix} \qquad \varepsilon_{[110]} = \frac{P}{2}\begin{pmatrix} s_{11}+s_{12} & s_{44}/2 & 0 \\ s_{44}/2 & s_{11}+s_{12} & 0 \\ 0 & 0 & 2s_{12} \end{pmatrix}$$

$$\varepsilon_{[111]} = \frac{P}{3}\begin{pmatrix} s_{11}+2s_{12} & s_{44}/2 & s_{44}/2 \\ s_{44}/2 & s_{11}+2s_{12} & s_{44}/2 \\ s_{44}/2 & s_{44}/2 & s_{11}+2s_{12} \end{pmatrix} \qquad \varepsilon_{[120]} = \frac{P}{5}\begin{pmatrix} s_{11}+4s_{12} & s_{44} & 0 \\ s_{44} & s_{12}+4s_{11} & 0 \\ 0 & 0 & 5s_{12} \end{pmatrix} \tag{3.18}$$

In the same manner the strain tensor resulting from a uniaxial stress in general directions $[hk\ell]$ can be obtained by applying the proper coordinate transformation (3.15).

3.3.3 Strain Resulting from Epitaxy

In techniques using global strain, biaxial strain is induced in the Si layer by epitaxial growth on a relaxed virtual substrate with different lattice constant. The strain tensor depends on the mismatch of the lattice constants and on the substrate orientation. The strain in the plane of the hetero-interface can be determined from the lattice mismatch

$$\varepsilon_{||} = \frac{a_s - a_0}{a_0}\ , \tag{3.19}$$

where $a_0 = 5.431$ Å is the lattice constant of Si [Levinshtein99] and a_l that of the substrate layer.

The strain tensor for arbitrary substrate orientations can be calculated according to [Hinckley90] and is given here for some frequently used substrate orientations:

$$\varepsilon_{(001)} = \varepsilon_{||}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\dfrac{2c_{12}}{c_{11}} \end{pmatrix}$$

$$\varepsilon_{(110)} = \varepsilon_{||}\begin{pmatrix} \dfrac{2c_{44}-c_{12}}{c_{11}+c_{12}+2c_{44}} & -\dfrac{c_{11}+2c_{12}}{c_{11}+c_{12}+2c_{44}} & 0 \\ -\dfrac{c_{11}+2c_{12}}{c_{11}+c_{12}+2c_{44}} & \dfrac{2c_{44}-c_{12}}{c_{11}+c_{12}+2c_{44}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\varepsilon_{(111)} = \varepsilon_{||}\begin{pmatrix} \dfrac{4c_{44}}{c_{11}+2c_{12}+4c_{44}} & -\dfrac{c_{11}+2c_{12}}{c_{11}+2c_{12}+4c_{44}} & -\dfrac{c_{11}+2c_{12}}{c_{11}+2c_{12}+4c_{44}} \\ -\dfrac{c_{11}+2c_{12}}{c_{11}+2c_{12}+4c_{44}} & \dfrac{4c_{44}}{c_{11}+2c_{12}+4c_{44}} & -\dfrac{c_{11}+2c_{12}}{c_{11}+2c_{12}+4c_{44}} \\ -\dfrac{c_{11}+2c_{12}}{c_{11}+2c_{12}+4c_{44}} & -\dfrac{c_{11}+2c_{12}}{c_{11}+2c_{12}+4c_{44}} & \dfrac{4c_{44}}{c_{11}+2c_{12}+4c_{44}} \end{pmatrix} \tag{3.20}$$

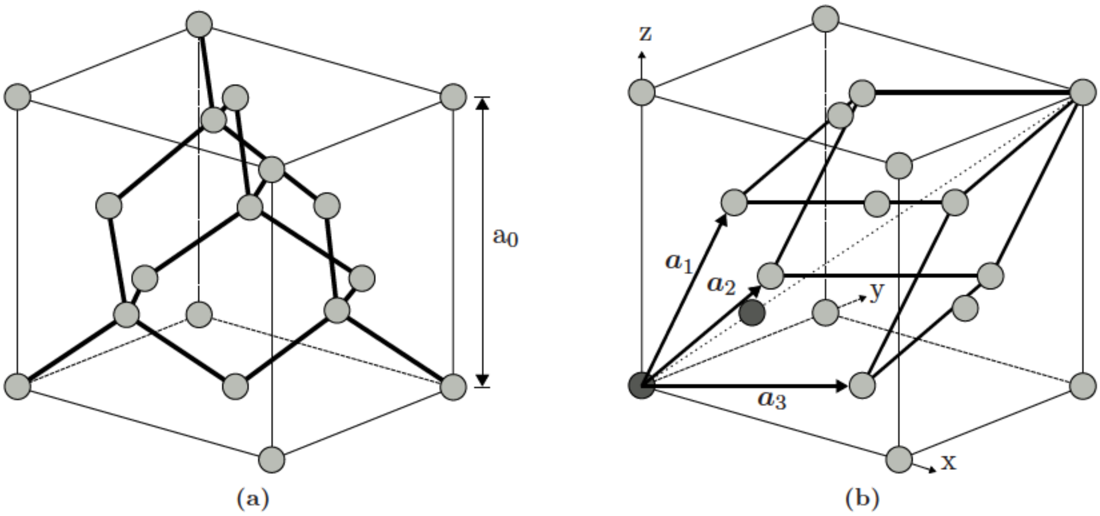


Figure 3.4: (a) Crystallographic unit cell (unit cube) of the diamond structure. (b) The primitive basis vectors of the face centered cubic (fcc) lattice and the two atoms forming the basis are highlighted.

Here, the elastic stiffness constants, given in Table 3.1, were used. The strain tensor contains shear components when the Si layer is grown on a (110) and (111) oriented substrate. These components can take large values. For example, the magnitude of the shear component ε_{xy} is larger that 3%, when growing Si on (110) oriented Ge.

3.4 Basic Properties of the Diamond Structure

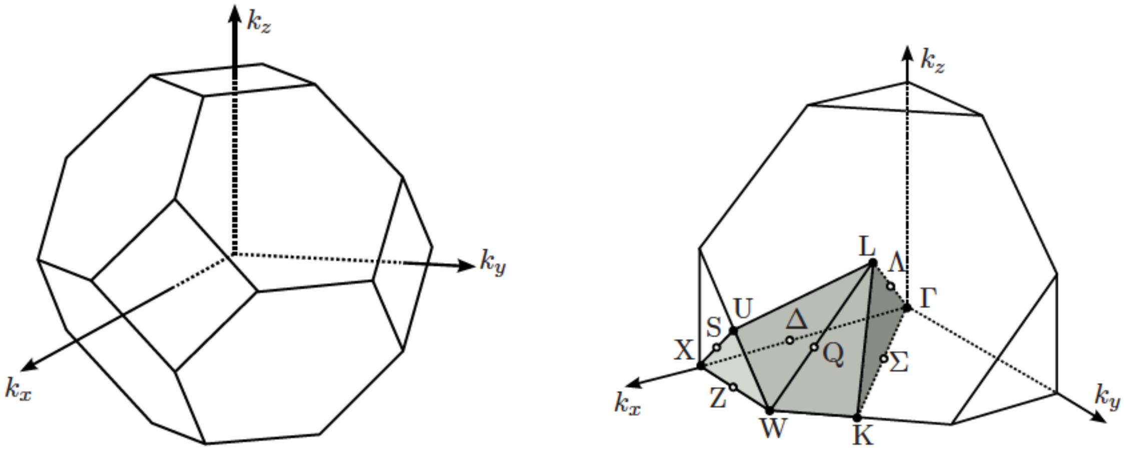
The structure depicted in Figure 3.4 consists of two basis atoms and may be thought of as two inter-penetrating face centered cubic (fcc) lattices, one displaced from the other by a translation of $\frac{a_0}{4}(1,1,1)$ along a body diagonal.

In cubic semiconductors such as Si or Ge the two atoms of the basis are identical and the structure is called the diamond structure. If the two basis atoms are different, the structure is called the zinc-blende structure. Many III-V semiconductors such as GaAs, AlsAs, InAs, or InP are of zinc-blende type.

In Figure 3.4 the diamond structure is depicted. The primitive basis vectors and the two atoms at $(0,0,0)$ and $\frac{a_0}{4}(1,1,1)$ are highlighted in Figure 3.4b. The basis vectors of the direct Bravais lattice are

$$\boldsymbol{a}_1 = \frac{a_0}{2}\begin{pmatrix}0\\1\\1\end{pmatrix}, \quad \boldsymbol{a}_2 = \frac{a_0}{2}\begin{pmatrix}1\\0\\1\end{pmatrix}, \quad \text{and} \quad \boldsymbol{a}_3 = \frac{a_0}{2}\begin{pmatrix}1\\1\\0\end{pmatrix}, \tag{3.21}$$

where a_0 denotes the lattice constant of the relaxed lattice. The lattice is invariant under



(a) (b)

Figure 3.5: (a) First Brillouin zone of the fcc lattice. (b) The inset gives the locations of certain symmetry points and symmetry lines in the BZ.

translations involving lattice vectors of the form

$$\boldsymbol{R}_{ijk} = i\boldsymbol{a}_1 + j\boldsymbol{a}_2 + k\boldsymbol{a}_3 \ , \tag{3.22}$$

where i, j and k are integers. In the unstrained lattice the distance from the basis atom at $\frac{a_0}{4}(1, 1, 1)$ to its four nearest neighbors is given by $a_0 \cdot \sqrt{3}/4$.

The basis vectors of the reciprocal lattice $\boldsymbol{b}_i$ are obtained from the relation $\boldsymbol{b}_i \cdot \boldsymbol{a}_j = 2\pi\delta_{ij}$ as

$$\boldsymbol{b}_1 = \frac{2\pi}{a_0}\begin{pmatrix}-1\\1\\1\end{pmatrix} \ , \quad \boldsymbol{b}_2 = \frac{2\pi}{a_0}\begin{pmatrix}1\\-1\\1\end{pmatrix} \ , \quad \text{and} \quad \boldsymbol{b}_3 = \frac{2\pi}{a_0}\begin{pmatrix}1\\1\\-1\end{pmatrix} \ . \tag{3.23}$$

General reciprocal lattice vectors are of the form

$$\boldsymbol{G}_{lmn} = l\boldsymbol{b}_1 + m\boldsymbol{b}_2 + n\boldsymbol{b}_3 \ , \tag{3.24}$$

where l, m, and n are integers.

The first Brillouin zone (BZ) represents the central (Wigner-Seitz) cell of the reciprocal lattice. It contains all points nearest to the enclosed reciprocal lattice point. The boundaries of the first BZ are determined by planes which are perpendicular to the reciprocal lattice vectors pointing from the center of the cell to the 14 lattice points nearest to the origin of the cell at their midpoints. The 14 faces are

$$|k_x| + |k_y| + |k_z| = \frac{3}{2}\frac{2\pi}{a_0} \ , \quad |k_x| = \frac{2\pi}{a_0} \ , \quad |k_y| = \frac{2\pi}{a_0} \ , \quad \text{and} \quad |k_z| = \frac{2\pi}{a_0}. \tag{3.25}$$

In Figure 3.5a the first BZ is depicted. It has a volume Ω_0 of $4(2\pi/a_0)$ [Singh93].

Due to the translational invariance of the lattice the wave functions and the energy bands are periodic in the reciprocal space and it is sufficient to consider only the first BZ for band structure calculations [Yu03].

The diamond structure is invariant not only under translations, but also under several other symmetry operations such as reflections, rotations, or inversion. These symmetry operations are usually denoted as point operations, since they leave at least one point of the lattice invariant, which is not the case for translations. The set of all point operations for a particular crystal structure forms a group which is denoted as point group. The point group of the diamond structure has 48 symmetry elements which are reflected in the symmetry of the first BZ. A quick examination (see Figure 3.5) shows that the BZ is invariant under various rotations, for example 90° rotations about the $\boldsymbol{k}_x$, $\boldsymbol{k}_y$, and $\boldsymbol{k}_z$ axes and under reflections through certain planes. A detailed examination of all 48 point symmetries of the unstrained diamond structure will be given in Section 3.5.2.

The point symmetries of the crystal structure are mirrored in the crystal potential, and hence in the one-particle Hamiltonian used for band structure calculations. Two important consequences for the electron band structure arise:

- Wave functions can be expressed in such a form that they have definite transformation properties under symmetry operations of the crystal. Certain matrix elements of operators can be shown to vanish and selection rules can be deduced, when classifying the wave functions according to their symmetry.
- The energy bands possess the full point symmetry of the point group of the crystal [Yu03, Nowotny98]. The symmetries can be exploited to restrict the band structure calculation to a fraction of only 1/48th of the first BZ. This fraction is called irreducible wedge of the BZ.

The irreducible wedge of the fcc lattice is depicted in Figure 3.5b. It has six corners

$$\Gamma = \frac{2\pi}{a_0}\begin{pmatrix}0\\0\\0\end{pmatrix}, \qquad X = \frac{2\pi}{a_0}\begin{pmatrix}1\\0\\0\end{pmatrix}, \qquad L = \frac{2\pi}{a_0}\begin{pmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{pmatrix},$$
$$W = \frac{2\pi}{a_0}\begin{pmatrix}1\\\frac{1}{2}\\0\end{pmatrix}, \qquad K = \frac{2\pi}{a_0}\begin{pmatrix}\frac{3}{4}\\\frac{3}{4}\\0\end{pmatrix}, \qquad U = \frac{2\pi}{a_0}\begin{pmatrix}1\\\frac{1}{4}\\\frac{1}{4}\end{pmatrix}, \tag{3.26}$$

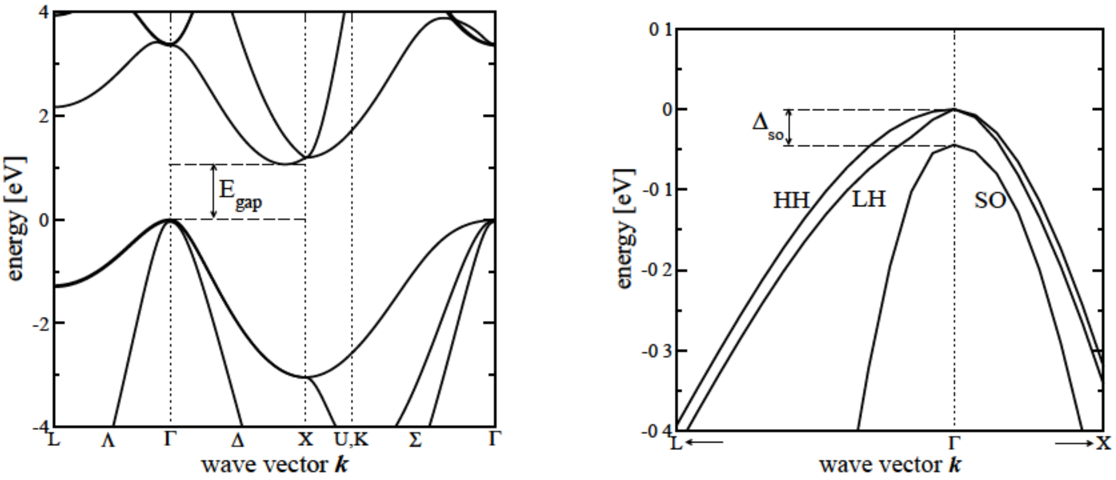
connected via symmetry lines

$$\Lambda : \Gamma \to L, \qquad \Delta : \Gamma \to X, \qquad S : X \to U, K,$$
$$\Sigma : \Gamma \to U, K, \qquad Q : L \to W, \qquad Z : X \to W. \tag{3.27}$$

By convention upper case Greek letters are used for points and lines inside the first BZ, whereas roman letters denote points and lines on the surface of the first BZ. Here, the points U and K can be interchanged, since they are equivalent in the reciprocal space (see Section 3.5).

3.4.1 Band Structure of Relaxed Si

The band structure describes the variation of the energy E with the wave-vector $\boldsymbol{k}$. The valence bands contain the last filled energy levels at $T = 0$ K, whereas the conduction bands are empty



(a) (b)

Figure 3.6: (a) Band structure of Si calculated using the pseudopotential method. The edges of the valence bands are located at the Γ point, the minimum of the lowest conduction band lies on the Δ symmetry line close to the X point. (b) Valence band structure with heavy hole (HH), light hole (LH) and split-off (SO) band along the Λ and the Δ direction.

at $T = 0$ K. The band gap E_{gap} separates the conduction band from the valence band. The band structure is usually visualized by plotting $E_n(\boldsymbol{k})$ on symmetry lines, where n denotes the band index. In Figure 3.6 the band structure of Si is plotted on the symmetry lines given in (3.27).

The band structure close to the conduction band edge can be approximated by ellipsoidal energy surfaces and a parabolic energy dispersion $E_n(\boldsymbol{k})$. In Si the conduction band edge is located near the zone boundary X points along the Δ symmetry lines. For the conduction band valley at $\boldsymbol{k}_{\mathrm{min}} = \frac{2\pi}{a_0}(0, 0, 0.85) = (0, 0, k_{\mathrm{min}})$ the energy dispersion reads

$$E(\boldsymbol{k}) = \frac{\hbar^2 (k_z - k_{\mathrm{min}})^2}{2 m_{\mathrm{l}}} + \frac{\hbar^2 (k_x^2 + k_y^2)}{2 m_{\mathrm{t}}} , \tag{3.28}$$

where $m_{\mathrm{l}} = 0.916\mathrm{m}_0$ is the longitudinal and $m_{\mathrm{t}} = 0.19\mathrm{m}_0$ the transverse mass of Si [Singh93].

Due to the point symmetry of the fcc lattice the six directions [100], [$\bar{1}$00],[010],[0$\bar{1}$0],[001], and [00$\bar{1}$] are equivalent. Consequently, there are six conduction band valleys. The constant energy surfaces of all six equivalent valleys along the principal axes ⟨100⟩ are shown in Figure 3.7. Since electron transport in unstrained Si involves the electrons moving in all of the six valleys, it shows little anisotropy, even though there is strong anisotropy in each valley.

The valence band edge is located at the Γ point, where the heavy hole (HH) and light hole (LH) band are degenerate. The split-off band (SO) is very close, since the split-off energy is only 44 meV in Si. For very small energies constant energy surfaces can be approximated by

$$E(\boldsymbol{k}) = -Ak^2 \mp \sqrt{B^2 k^4 + C^2 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)} . \tag{3.29}$$

This is the warped band approximation [Yu03]. The warping occurs along the [100] and [111] directions because of the cubic symmetry of the crystal.

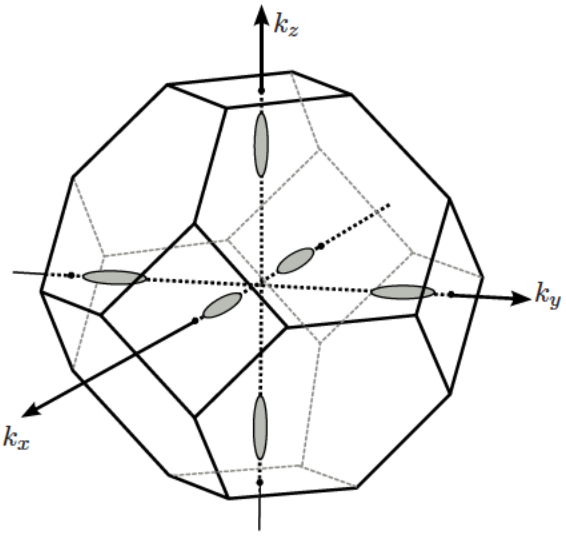


Figure 3.7: Constant energy surfaces for the lowest conduction band of Si. In unstrained Si the six valleys along the three principal axes are equivalent.

3.5 Effect of Strain on Crystals with Diamond Structure

Application of strain to a crystal reduces its symmetry. The Bravais lattice basis vectors $\boldsymbol{a}_i{}'$ of a crystal under homogeneous strain are obtained by deforming the vectors $\boldsymbol{a}_i$ of the unstrained crystal

$$\boldsymbol{a}_i{}' = (\mathbb{1} + \hat{\varepsilon}) \cdot \boldsymbol{a}_i \ , \tag{3.30}$$

where $\hat{\varepsilon}$ denotes the strain tensor [Bir74]. Not only is the volume Ω_0' of the strained primitive unit cell modified

$$\Omega_0' = \Omega_0(1 + \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) \ , \tag{3.31}$$

but also the type of the Bravais lattice may change. The new Bravais lattice is then characterized by basis vectors which may be different from (3.30). However, the volume of the new primitive cell always differs from Ω_0 by a factor proportional to the strain, given in (3.31) if the hydrostatic strain component, $\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}$, is nonzero. The Brillouin zone (BZ) of the strained crystal, being independent of the specific choice of basis vectors, can always be obtained by a suitable deformation of the BZ of the unstrained crystal [Bir74].

Generally, the strain tensor (3.5) determines only the deformation of the primitive cell as a whole, but the relative displacements of the atoms within the cell differ for the atoms forming the basis. This has no impact on the lattice symmetry but has to be accounted for in band structure calculations. In Section 3.8.2 it is shown how to take this internal displacement into account properly.

Points and lines of the BZ can be classified according to their symmetry. For each lattice there are symmetry operations that leave the lattice and hence the BZ invariant. These operations bring the BZ into coincidence with itself. In such an operation a specific vector $\boldsymbol{k}$ is not necessarily projected onto itself or $\boldsymbol{k} + \boldsymbol{G}_{lmn}$. For each vector $\boldsymbol{k}$ there exists a set of symmetry operations α fulfilling

$$\alpha\boldsymbol{k} = \boldsymbol{k}\ , \quad \text{or} \quad \alpha\boldsymbol{k} = \boldsymbol{k} + \boldsymbol{G}_{lmn}\ . \tag{3.32}$$

The number of symmetry operations α depends on the wave vector $\boldsymbol{k}$ and is denoted as the symmetry group $P(\boldsymbol{k})$ of the respective $\boldsymbol{k}$ vector

$$P(\boldsymbol{k}) = \{\alpha \,|\, \alpha\boldsymbol{k} = \boldsymbol{k} + \boldsymbol{G}_{lmn}\}\ . \tag{3.33}$$

Since the unity operation E always fulfills (3.33), the symmetry group $P(\boldsymbol{k})$ contains one element at least ($\alpha = \mathrm{E}$). If the symmetry group of a given vector $\boldsymbol{k}$ contains more elements than the symmetry group of neighboring points, this specific vector $\boldsymbol{k}$ is referred to as symmetry point.

In Figure 3.5b the symmetry points (filled circles) and symmetry lines (open circles) of the fcc lattice are shown. Strictly, the given points K and U are not symmetry points according to the definition of (3.33), since they have the same symmetry as the points along the symmetry lines Σ and S. Hence, they can be included in these symmetry lines.

Since the center point Γ of the BZ is mapped onto itself at any point operation α of the crystal lattice, all symmetry operations of the lattice are included in the point group $P(\Gamma)$. Thus, this group determines the shape and volume of the irreducible wedge. The number of symmetry elements of $P(\Gamma)$ determines the volume of the irreducible wedge as [Nowotny98]

$$\Omega_{\text{irred}} = \Omega_{\text{BZ}}/|P(\Gamma)|\ , \tag{3.34}$$

where $|P(\Gamma)|$ is the number of elements of the symmetry group $P(\Gamma)$.

3.5.1 Hierarchy of systems

The point group of the unstrained diamond structure is denoted by O_h according to the Schönfließ notation, and contains 48 symmetry elements listed in Table 3.2 [Yu03]. The symmetry elements are given in terms of rotations and rotations followed by the inversion operation ($\mathrm{I}\boldsymbol{k} = -\boldsymbol{k}$). The following notation for the point operations is chosen:

E	unity operation
n_k^+	clockwise rotation of angle $2\pi/n$ around axis $\boldsymbol{e}_\mathrm{k}$
n_k^-	counter-clockwise rotation of angle $2\pi/n$ around axis $\boldsymbol{e}_\mathrm{k}$
I	inversion
$\bar{n}_k^+$	clockwise rotation of angle $2\pi/n$ around axis $\boldsymbol{e}_\mathrm{k}$ followed by inversion
$\bar{n}_k^-$	counter-clockwise rotation of angle $2\pi/n$ around axis $\boldsymbol{e}_\mathrm{k}$ followed by inversion

The rotation axes $\boldsymbol{e}_\mathrm{k}$ are grouped into five classes:

$\boldsymbol{e}_i = (1,0,0),\ (0,1,0),\ (0,0,1)$
$\boldsymbol{e}_{i''} = (0,1,0),\ (\sqrt{3},-1,0),\ (-\sqrt{3},-1,0)$
$\boldsymbol{e}_j = (1,1,1),\ (-1,-1,1),\ (1,-1,-1),\ (-1,-1,-1)$
$\boldsymbol{e}_p = (1,1,0),\ (-1,1,0),\ (1,0,1),\ (0,1,1),\ (-1,0,1),\ (0,-1,1)$
$\boldsymbol{e}_s = (1,1,0),\ (-1,1,0)$

point group	symmetry elements	$\lvert P(\Gamma)\rvert$	stress direction
O_h	E 2_i 3_j^+ 3_j^- 2_p 4_i^+ 4_i^- I $\bar{2}_i$ $\bar{3}_j^+$ $\bar{3}_j^-$ $\bar{2}_p$ $\bar{4}_i^+$ $\bar{4}_i^-$	48	unstrained
D_{4h}	E 4_z^+ 4_z^- 2_i 2_s I $\bar{4}_z^+$ $\bar{4}_z^-$ $\bar{2}_i$ $\bar{2}_s$	16	stress along $\langle 100\rangle$
D_{3d}	E 3_z^+ 3_z^- $2_{i''}$ I $\bar{3}_z^+$ $\bar{3}_z^-$ $\bar{2}_{i''}$	12	stress along $\langle 111\rangle$
D_{2h}	E 2_i I $\bar{2}_i$	8	stress along $\langle 110\rangle$
C_{2h}	E 2_z I $\bar{2}_z$	4	stress along $\langle 120\rangle$
S_2	E I	2	other directions

Table 3.2: Point group and symmetry elements of strained lattices that originate when stress is applied along various high symmetry directions to an initially cubic lattice O_h. The Schönfließ symbol is used to specify the point group. $|P(\Gamma)|$ denotes the number of elements of the point group.

Of all the point symmetry groups $P(\Gamma)$ of the crystal lattice, the group O_h possesses the highest symmetry. The symmetry group $P(\Gamma')$ of the Bravais lattice of the strained crystal is a subgroup of the symmetry group $P(\Gamma')$ of the unstrained crystal and does not generally belong to the same crystal class as O_h. It contains only those symmetry elements which are preserved under strain. The effect of a homogeneous strain on the symmetry of the Bravais lattice depends on the specific form of applied strain.

By successively lowering the symmetry, one can go from O_h to point group S_2 on two distinct paths [Bir74]:

$$O_h \rightarrow D_{4h} \rightarrow D_{2h} \rightarrow C_{2h} \rightarrow S_2 \tag{3.35}$$

$$O_h \rightarrow D_{3d} \rightarrow C_{2h} \rightarrow S_2 \ . \tag{3.36}$$

The symmetry elements of the five specified point groups $P(\Gamma')$, namely $D_{4h}, D_{3d}, D_{2h}, C_{2h}, S_2$, which are subgroups of O_h are given in Table 3.2. The symmetry can be lowered by distorting the crystal by applying uniaxial stress. If, for example, stress is applied along a fourfold axis, the point group O_h reduces to D_{4h}, which contains only 16 symmetry elements. In Table 3.2 five directions of uniaxial stress are given that yield a direct $O_h \rightarrow P(\Gamma')$ symmetry reduction for any of the five subgroups of O_h.

From (3.34) it can be concluded that the higher the point symmetry of the crystal lattice, the smaller is the volume of the irreducible wedge. In the following, the symmetries of the band structure and the shape of the irreducible wedge in terms of the irreducible wedge of the unstrained crystal are given for the crystal systems $O_h, D_{4h}, D_{3d}, D_{2h}, C_{2h}$, and S_2.

3.5.2 O_h symmetry

The point group O_h contains 48 symmetry elements: listed in

- the unity operation,
- clockwise rotation of 180° about the principal axes $\boldsymbol{e}_i$ (three operations),

- clockwise and counter-clockwise rotation of 120° about the four space diagonals (eight operations),
- clockwise rotation of 180° about the axes $\boldsymbol{e}_p$ representing the diagonals on the faces of a unit cube (six operations),
- clockwise and counter-clockwise rotations of 90° about the principal axes $\boldsymbol{e}_i$ (six operations).

In total these are 24 symmetry operations. Since the symmetry is preserved if any of these operations are followed by the inversion operation, there are 48 symmetry operations in total.

The 48 point symmetries are reflected in the energy bands and can be shown to leave the energy bands invariant under the following eight reflections

$$E_n(k_x, k_y, k_z) = E_n(|k_x|, |k_y|, |k_z|) \tag{3.37}$$

and six permutations

$$\begin{aligned} &E_n(k_x, k_y, k_z) = E_n(k_x, k_z, k_y) = E_n(k_y, k_x, k_z) = \\ &E_n(k_y, k_z, k_x) = E_n(k_z, k_x, k_y) = E_n(k_z, k_y, k_x) \ . \end{aligned} \tag{3.38}$$

A possible choice for the irreducible wedge of unstrained Si is depicted in Figure 3.5b. It should be noted that this specific choice is not unique, and other shapes for the irreducible wedge, reflecting the symmetries of O_h, can be found [Stanley98]. The volume of the specified irreducible wedge is $\Omega_{\mathrm{BZ}}/48$ in accordance with the relation (3.34).

3.5.3 D_{4h} symmetry

If stress is applied along a fourfold axis $\boldsymbol{e}_i$ of a cubic lattice of symmetry class O_h, the unit cube becomes a square cuboid (rectangular parallelepiped on a square base), representing the Bravais parallelepiped belonging to symmetry class D_{4h} [Bir74]. Note, that D_{4h} is a member of the tetragonal crystal system. A similar symmetry reduction is observed, if biaxial strain is present in a {001} plane. According to Table 3.2 the point group D_{4h} has 16 symmetry elements, since only one fourfold axis and no threefold axis remains.

The strain tensor in the principal system yielding a $O_h \rightarrow D_{4h}$ symmetry reduction has non-zero elements in the diagonal (e.g. $\varepsilon_{11} = \varepsilon_{22} \neq \varepsilon_{33}$), whereas all off-diagonal elements vanish.

The symmetry operations yield invariance of the energy bands under reflections

$$E_n(k_x, k_y, k_z) = E_n(|k_x|, |k_y|, |k_z|) \ . \tag{3.39}$$

and the invariance of the energy bands under the permutation of the indices perpendicular to the direction of stress. If stress is applied along [100], the energy bands are invariant under the permutation

$$E_n(k_x, k_y, k_z) = E_n(k_x, k_z, k_y) \ . \tag{3.40}$$

From (3.39) it follows that the bands have to be calculated only in the first octant of the BZ. The additional symmetry of (3.40) can be exploited to further reduce the volume of the irreducible

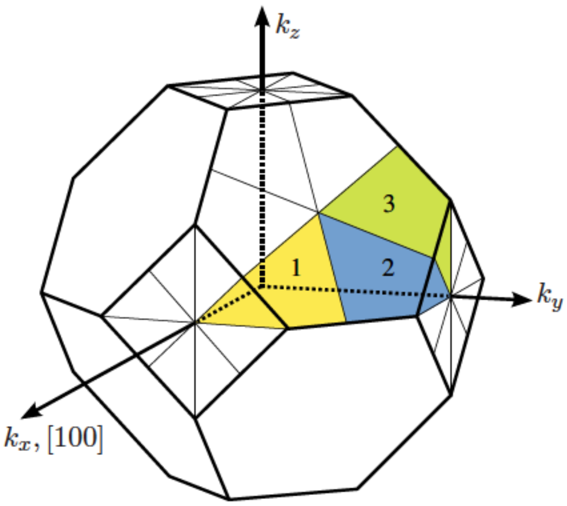


Figure 3.8: Irreducible wedge of the first BZ of a diamond structure stressed along direction [100].

wedge by a factor of two. For band structure calculation a volume for the irreducible wedge can be chosen which combines three irreducible wedges of the relaxed crystal in the limit of vanishing strain. These wedges are depicted in Figure 3.8. They are labeled with a number and can be transformed into the first irreducible wedge of the relaxed crystal by a symmetry operation

$$T_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad T_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{3.41}$$

The volume of the irreducible wedge depicted in Figure 3.8 is $\Omega_{\mathrm{BZ}}/16$ and thus complies with relation (3.34).

3.5.4 D_{3d} symmetry

Under uniaxial stress along the threefold symmetry axes $\boldsymbol{e}_j$ representing the cube diagonals, the cube of the O_h lattice becomes a rhombohedron with an arbitrary angle between the edges. The cubic fcc lattice is transformed to a primitive rhombohedral lattice ($\Gamma' = \Gamma_{rh}$) belonging to the crystal class D_{3d} of the trigonal (= rhombohedral) system [Bir74].

Apart from the trivial unity transformation and inversion, the symmetry operations of D_{3d} include threefold rotations about the direction of stress and twofold rotations around three twofold axes perpendicular to the threefold axis. All twelve symmetry operations and the resulting invariances of the energy band structure are listed in Table 3.3. Note that in this table the Miller index notation of the cubic lattice is chosen to specify the directions of the axes of rotation, whereas the directions used in Table 3.2 are given using the Bravais basis vectors of the trigonal

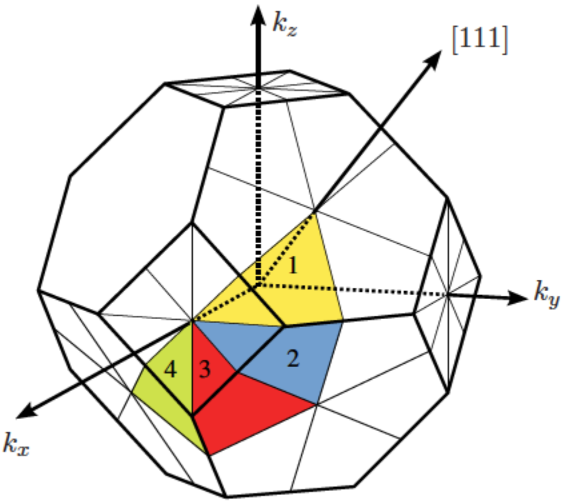


Figure 3.9: Irreducible wedge of the first BZ of a diamond structure stressed along direction [111].

lattice:

$$\boldsymbol{a}_1 = \begin{pmatrix} \frac{a}{2\sqrt{3}} \\ -\frac{a}{2} \\ \frac{c}{3} \end{pmatrix}, \quad \boldsymbol{a}_2 = \begin{pmatrix} \frac{a}{2\sqrt{3}} \\ \frac{a}{2} \\ \frac{c}{3} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{a}_3 = \begin{pmatrix} \frac{-a}{\sqrt{3}} \\ 0 \\ \frac{c}{3} \end{pmatrix}. \tag{3.42}$$

These basis vectors are equivalent to those of the fcc lattice when $a = (\sqrt{3}a_0)/2$ and $c = (3a_0)/2$.

From (3.18) it can be seen that the strain tensor resulting from stress along $\langle 111 \rangle$, which yields a $O_h \rightarrow D_{3d}$ symmetry reduction, contains equal off-diagonal components $\varepsilon_{12} = \varepsilon_{13} = \varepsilon_{23}$ and equal diagonal components $\varepsilon_{11} = \varepsilon_{22} = \varepsilon_{33}$. Note that biaxial strain in the $\{111\}$ plane also yields $\varepsilon_{12} = \varepsilon_{13} = \varepsilon_{23}$, and $\varepsilon_{11} = \varepsilon_{22} = \varepsilon_{33}$, hence the same symmetry reduction occurs.

symmetry operation	energy	symmetry operation	energy
E	$E_n(k_x, k_y, k_z)$	I	$E_n(-k_x, -k_y, -k_z)$
$3^-_{[111]}$	$E_n(k_z, k_x, k_y)$	$\bar{3}^-_{[111]}$	$E_n(-k_y, -k_z, -k_x)$
$3^+_{[111]}$	$E_n(k_y, k_z, k_x)$	$\bar{3}^+_{[111]}$	$E_n(-k_y, -k_z, -k_x)$
$2_{[\bar{1}10]}$	$E_n(-k_y, -k_x, -k_z)$	$\bar{2}_{[\bar{1}10]}$	$E_n(k_y, k_x, k_z)$
$2_{[\bar{1}01]}$	$E_n(-k_z, -k_y, -k_x)$	$\bar{2}_{[\bar{1}01]}$	$E_n(k_z, k_y, k_x)$
$2_{[0\bar{1}1]}$	$E_n(-k_x, -k_z, -k_y)$	$\bar{2}_{[0\bar{1}1]}$	$E_n(k_x, k_z, k_y)$

Table 3.3: Symmetry operations leaving the band structure invariant in the D_{3d} class. The direction of the rotation axes are given using the Miller index notation of the cubic lattice.

The twelve involved symmetry operations give rise to a volume of the irreducible wedge of

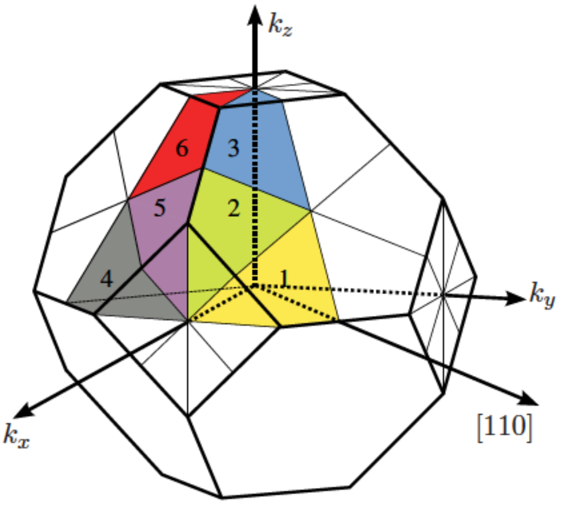


Figure 3.10: Irreducible wedge of the first BZ of a diamond structure stressed along direction [110].

$\Omega_{\mathrm{BZ}}/12$, which can be mapped onto four irreducible wedges of the unstrained crystal as shown in Figure 3.9. The four wedges can be transformed into the first wedge by the symmetry operations

$$\underline{T}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ \underline{T}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \ \underline{T}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \text{ and } \underline{T}_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}. \tag{3.43}$$

3.5.5 D_{2h} symmetry

The Bravais cube of the crystal class O_h can be converted to a parallelepiped of the orthorhombic system belonging to D_{2h} in two ways [Bir74]:

1. Dilatation or compression along two of the three fourfold axes $\boldsymbol{e}_i$. This results in a right parallelepiped with rectangular faces (cuboid). Of the five twofold axes $\boldsymbol{e}_i$ and $\boldsymbol{e}_s$ of D_{4h}, only the three $\boldsymbol{e}_i$ along the edges of the parallelepiped remain. This symmetry reduction can be achieved by applying stress of different magnitude along two of the three equivalent $\langle 100 \rangle$ directions, simultaneously. In this case, the strain tensor is given by $\varepsilon_{11} \neq \varepsilon_{22} \neq \varepsilon_{33}$ and contains vanishing off-diagonal components.

2. The deformation originates from shearing the unit cube, thus altering the angles between the basis vectors. The result is a rectangular parallelepiped with rhombic base, which is also invariant under D_{2h}. Of the original five twofold axes $\boldsymbol{e}_i$ and $\boldsymbol{e}_s$ only two (diagonals of the base) remain. This type of lattice results, when uniaxial stress is applied along $\langle 110 \rangle$

or from biaxial strain in a $\{110\}$ plane. The strain tensor has the form

$$\underline{\varepsilon} = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & 0 \\ \varepsilon_{12} & \varepsilon_{11} & 0 \\ 0 & 0 & \varepsilon_{33} \end{pmatrix} , \tag{3.44}$$

where the components of the strain tensor can be related to stress according to (3.18).

This group has only eight symmetry elements (given in Table 3.2). The irreducible wedge with a volume of $\Omega_{\mathrm{BZ}}/8$ can be mapped onto six wedges of the unstrained lattice in the limit of vanishing strain.

When dilating or compressing along two of the three fourfold axes $\boldsymbol{e}_i$ (case 1), any octant of the BZ can be chosen as the irreducible wedge. In the presence of uniaxial stress along [110] (case 2) a possible choice for the irreducible wedge is depicted in Figure 3.10. The six wedges labeled in Figure 3.10 can be transformed into the first wedge by the transformations

$$\underline{T}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \underline{T}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} , \underline{T}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} ,$$

$$\underline{T}_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \underline{T}_5 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} , \underline{T}_6 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} .$$

3.5.6 C_{2h} symmetry

Higher symmetry reduction results from deforming the base of the Bravais parallelepiped of the orthorhombic system so that the angle between its edges is changed. In this way the invariant parallelelided of the system C_{2h} is obtained from the cubic lattice O_h [Bir74]. It contains four symmetry operations given in Table 3.2 with only one twofold symmetry axis.

Uniaxial stress in [120] direction can achieve this kind of symmetry reduction. The strain tensor has three different nonzero diagonal components and one off-diagonal component

$$\underline{\varepsilon} = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & 0 \\ \varepsilon_{12} & \varepsilon_{22} & 0 \\ 0 & 0 & \varepsilon_{33} \end{pmatrix} . \tag{3.45}$$

A irreducible volume can be chosen according to Figure 3.11. The volume can be mapped onto twelve wedges of the unstrained lattice in the limit of vanishing strain. The twelve wedges are labeled in Figure 3.11 and can be transformed into the first wedge by proper transformations.

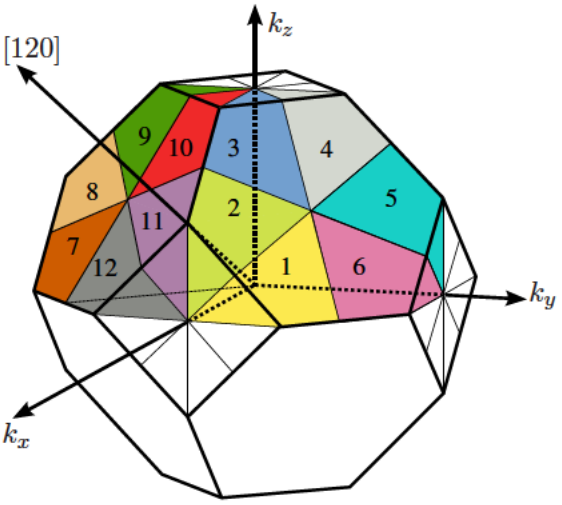


Figure 3.11: Irreducible wedge of the first BZ of a diamond structure stressed along [120].

3.5.7 S_2 symmetry

Under stress along directions other than those given in Table 3.5 the crystal loses all rotational symmetries and is invariant only under inversion. The resulting crystal class is S_2 and half of the BZ must be chosen as irreducible volume for band structure calculation.

3.6 Linear Deformation Potential Theory

Deformation potential theory was originally developed by Bardeen and Shockley [Bardeen50] and generalized by Herring and Vogt [Herring56]. In an extensive study Bir and Pikus showed later how to calculate strain effects on the band structure of various semiconductors on the basis of deformation potentials by applying laws of group theory [Bir74]. In the following a brief outline of this theory is given before it will be applied to calculate strain-induced energetic shifts of the conduction band valleys of cubic semiconductors.

An additional Hamiltonian attributed to strain, $\mathcal{H}(\hat{\varepsilon})$, is defined and its effect on the band structure is calculated using first order perturbation theory. The matrix elements of the perturbation Hamiltonian are given by

$$\{\mathcal{H}(\hat{\varepsilon})\}_{ij} = \sum_{\alpha,\beta=1}^{3} \mathcal{D}_{ij}^{\alpha\beta} \varepsilon_{\alpha\beta} \ , \tag{3.46}$$

where $\varepsilon_{\alpha\beta}$ denotes the $(\alpha\beta)$ component of the strain tensor, and $\mathcal{D}^{\alpha\beta}$ is the deformation potential operator which transforms under symmetry operations as a second rank tensor [Hinckley90]. The additional subscripts (i,j) in $\mathcal{D}_{ij}^{\alpha\beta}$ denote the matrix element of the operator $\mathcal{D}^{\alpha\beta}$. From the symmetry of the strain tensor with respect to α and β it follows $\mathcal{D}^{\alpha\beta} = \mathcal{D}^{\beta\alpha}$, such that at maximum six independent deformation potential operators can exist.

In cubic semiconductors the edges of the conduction and the valence band are located on symmetry lines which are reflected in symmetries of the energy band structure and of the basis states. Due to the symmetry of the basis states the deformation potential operator of a particular state can be described in terms of two or three deformation potential constants [Herring56, Singh93]. The shift of the band edges can be calculated from these deformation potential constants.

In principle the deformation potential constants can be determined numerically using the empirical pseudopotential method (see Section 3.8) or from ab initio calculations. However, it is common practice to fit deformation potentials to measurements using electrical, optical, microwave techniques or by analyzing stress-induced indirect absorption edges. Although measurements seem to verify theoretical predictions, the deformation potentials obtained by different methods deviate from each other and different values can be found in literature [Fischetti96a].

3.6.1 Strain-Induced Conduction Band Splitting

In this section we give expressions for the strain-induced energy shifts of the nondegenerate energy levels of the conduction band edges of the cubic crystal class O_h.

At the Γ and L point, and along the Δ symmetry line, the deformation potential operators $\mathcal{D}^{\alpha\beta}$ are scalars and given by one ore two independent constants. Neglecting the strain-induced splitting of the degenerate conduction bands Δ_1 and $\Delta_{2'}$ at the X point for the moment, the energy shifts of the conduction band edge of valleys along the $\langle 100\rangle$ and $\langle 111\rangle$ direction can be calculated from two independent deformation potential constants [Balslev66]

$$\delta E_0^{v_i} = \Xi_\mathrm{d}^v \operatorname{Tr}(\underline{\varepsilon}) + \Xi_\mathrm{u}^v \boldsymbol{a}_i^\mathrm{T} \underline{\varepsilon} \boldsymbol{a}_i \ . \tag{3.47}$$

Here, Ξ_d^v denotes the dilatation- and Ξ_u^v the uniaxial deformation potential constant for a valley of type $v = \Delta, L$, and $\boldsymbol{a}_i$ is a unit vector parallel to the $\boldsymbol{k}$ vector of valley i. The valley shift of the Γ_2' conduction band minimum can be obtained from a single deformation potential constant

$$\delta E_0^{\Gamma} = \Xi_\mathrm{d}^{\Gamma} \operatorname{Tr}(\underline{\varepsilon}) \ . \tag{3.48}$$

The valley splitting from uniaxial stress along any direction can be obtained from the strain tensor using the relations above. The strain tensors resulting from uniaxial stress are discussed in Section 3.3.2. The analytical expressions for the energy shifts of the conduction band valleys for three stress directions [100], [110], and [111] are given in in Table 3.4.

3.6.2 Strain-Induced Lifting of Degeneracy at X point

In the case of degenerate bands strain does not only shift the band as a whole, but may also split bands as a result of partial or complete removal of degeneracy upon the reduction of symmetry. In the diamond crystal structure, the lowest two conduction bands Δ_1 and $\Delta_{2'}$ touch at the zone boundary X due to a special symmetry of the diamond structure, namely the presence of three glide reflection planes, given by $x = a_0/8, y = a_0/8$, and $z = a_0/8$ [Yu03]. For example, the plane $z = a_0/8$ is a glide plane since diamond is invariant under a translation by $\frac{a_0}{4}(1, 1, 0)$ followed by a reflection on this plane. Whenever the strain tensor in the crystal system contains a shear component ε_{xy} (for example, as a result from stress along the [110] direction), the strained lattice belongs to an orthorombic crystal system (see Section 3.5.5). The shear component removes the glide reflection plane $z = a_0/8$ and consequently the degeneracy of the two lowest conduction

stress direction	valley	valley direction	$\delta E/P$
[100]	Δ	[100]	$\Xi_\mathrm{d}^\Delta(s_{11}+2s_{12})+\Xi_\mathrm{u}^\Delta s_{11}$
	Δ	[010][001]	$\Xi_\mathrm{d}^\Delta(s_{11}+2s_{12})+\Xi_\mathrm{u}^\Delta s_{12}$
	L	$[111][11\bar{1}][1\bar{1}1][\bar{1}11]$	$\Xi_\mathrm{d}^\mathrm{L}(s_{11}+2s_{12})+\Xi_\mathrm{u}^\mathrm{L}/3(s_{11}+2s_{12})$
	Γ	[000]	$\Xi_\mathrm{d}^\Gamma(s_{11}+2s_{12})$
[110]	Δ	[100][010]	$\Xi_\mathrm{d}^\Delta(s_{11}+2s_{12})+\Xi_\mathrm{u}^\Delta/2(s_{11}+s_{12})$
	Δ	[001]	$\Xi_\mathrm{d}^\Delta(s_{11}+2s_{12})+\Xi_\mathrm{u}^\Delta s_{12}$
	L	$[111][11\bar{1}]$	$\Xi_\mathrm{d}^\mathrm{L}(s_{11}+2s_{12})+\Xi_\mathrm{u}^\mathrm{L}/3(s_{11}+2s_{12}+s_{44})$
	L	$[\bar{1}11][1\bar{1}1]$	$\Xi_\mathrm{d}^\mathrm{L}(s_{11}+2s_{12})+\Xi_\mathrm{u}^\mathrm{L}/3(s_{11}+2s_{12}-s_{44})$
	Γ	[000]	$\Xi_\mathrm{d}^\Gamma(s_{11}+2s_{12})$
[111]	Δ	[100][010][001]	$\Xi_\mathrm{d}^\Delta(s_{11}+2s_{12})+\Xi_\mathrm{u}^\Delta/3(s_{11}+2s_{12})$
	L	[111]	$\Xi_\mathrm{d}^\mathrm{L}(s_{11}+2s_{12})+\Xi_\mathrm{u}^\mathrm{L}/3(s_{11}+2s_{12}+2s_{44})$
	L	$[\bar{1}11][1\bar{1}1][11\bar{1}]$	$\Xi_\mathrm{d}^\mathrm{L}(s_{11}+2s_{12})+\Xi_\mathrm{u}^\mathrm{L}/3(s_{11}+2s_{12}-2/3s_{44})$
	Γ	[000]	$\Xi_\mathrm{d}^\Gamma(s_{11}+2s_{12})$

Table 3.4: Strain-induced energy shifts of the conduction band valleys of cubic semiconductors when uniaxial stress P is applied along three high symmetry directions. The energy shifts are divided by P.

bands Δ_1 and $\Delta_{2'}$ at the symmetry points $X = \frac{2\pi}{a_0}(0,0,\pm 1)$ is lifted [Hensel65, Bir74] (compare Section 3.5.5). It should be noted that the glide reflection symmetry is preserved in biaxially strained Si layers grown on $\{001\}$ $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ substrate, as well as in Si uniaxially strained/stressed along a fourfold rotation axis $\langle 100\rangle$.

From $\boldsymbol{k}\cdot\boldsymbol{p}$ theory (see Section 3.7) including terms of third order Bir and Pikus found that by lifting the degeneracy at a zone boundary X point a comparatively large change in the energy dispersion of the conduction band minimum located close to this X point is induced [Bir74]. In Si this effect was verified experimentally by Hensel and Hasegawa [Hensel65], who measured the change in effective mass for stress along $\langle 110\rangle$, and by Laude [Laude71] who measured the indirect exciton spectrum.

Equation (3.47) has to be modified as follows when taking into account the lifting of the degeneracy of the two lowest conduction bands Δ_1 and $\Delta_{2'}$ at the X points $\frac{2\pi}{a_0}(0,0,\pm 1)$ [Hensel65]

$$\begin{pmatrix} \delta E_0 & \delta E_1 \\ \delta E_1 & \delta E_0 \end{pmatrix} \begin{pmatrix} \xi \\ \hat{\xi} \end{pmatrix} = \delta E \begin{pmatrix} \xi \\ \hat{\xi} \end{pmatrix}, \tag{3.49}$$

where

$$\delta E_0 = \Xi_\mathrm{d}^\Delta \,\mathrm{Tr}(\varepsilon) + \Xi_\mathrm{u}^\Delta \varepsilon_{zz} \;, \tag{3.50}$$

$$\delta E_1 = \Xi_{u'} e_{xy} = 2\Xi_{u'} \varepsilon_{xy} \;. \tag{3.51}$$

Here, a new deformation potential $\Xi_{u'}$ is introduced. The solutions of the eigenvalue problem (3.49) are

$$\delta E = \delta E_0 \pm \delta E_1 \quad \text{for} \quad \hat{\xi} = \pm \xi \ . \tag{3.52}$$

Thus, at the X points $\frac{2\pi}{a_0}(0, 0, \pm 1)$ the bands shift by an amount of δE_0, which is the shift given by (3.47). Additionally, the degeneracy is lifted with a splitting of $2\delta E_1$. From (3.51) it follows that the splitting is proportional to the shear strain ε_{xy} is given by

$$(E_{\Delta_1} - E_{\Delta_{2'}})\Big|_{X_{[001]}} = 2\delta E_1 = 4\Xi_{u'}\varepsilon_{xy} \ . \tag{3.53}$$

For the shear deformation potential $\Xi_{u'}^X$ a value of 5.7 ± 1 eV [Hensel65] has been predicted from cyclotron resonance experiments. From measurements of the indirect exciton spectrum of Si a similar value (7.5 ± 2 eV) has been obtained [Laude71].

Figure 3.12 shows the splitting for three levels of strain ε_{xy}. The splitting is very pronounced even for relatively small strain ($< 1\%$). One can also observe that the Δ_1 conduction band is deformed in the vicinity of the symmetry points $X = \frac{2\pi}{a_0}(0, 0, \pm 1)$ due to the lifting of the degeneracy.

In unstrained Si the constant energy surfaces of the six conduction band valleys have a prolate ellipsoidal shape, where the semi-axes are characterized by m_l and m_t, denoting the longitudinal and transverse electron masses, respectively. The minima of the three valley pairs are located along the three equivalent $\langle 100 \rangle$ directions and have the same energies (see Section 3.4.1).

From Figure 3.12 it can be observed that a non-vanishing shear component ε_{xy} in the strain tensor affects the energy dispersion of the lowest conduction band in three ways:

1. The band edge energy of the valley pair oriented along the [001] direction moves down with respect to the four valleys oriented along [100] and [010].
2. Since the shape of the conduction band around its minima is considerably deformed, the effective mass of the valley pair along [100] is expected to change as ε_{xy} grows.
3. The positions of the conduction band minima along [001] move towards the zone boundary X points at $\frac{2\pi}{a_0}(0, 0, \pm 1)$.

Within the presented model ε_{xy} has no effect on the conduction bands near the zone boundaries $X = \frac{2\pi}{a_0}(1, 0, 0)$ and $X = \frac{2\pi}{a_0}(0, 1, 0)$ (compare Figure 3.12). A nonzero component ε_{xz} and ε_{yz}, however, will lift the degeneracy at $X = \frac{2\pi}{a_0}(0, \pm 1, 0)$ and $X = \frac{2\pi}{a_0}(\pm 1, 0, 0)$, respectively.

If the splitting of the conduction bands is different at the different zone boundaries (for example, $\varepsilon_{xy} \neq \varepsilon_{xz} \neq \varepsilon_{yz}$), the conduction band minima along the $\langle 001 \rangle$ axes have different energies. This may result in a repopulation between the six conduction band valleys. Note that such an effect cannot be explained via equation (3.47) alone, where a possible lifting of the degeneracy at the X point induced by shear strain is neglected and application of shear strain yields no valley repopulation.

An analytical expression for the valley shift along the Δ direction can be derived using a degenerate $\boldsymbol{k}\cdot\boldsymbol{p}$ theory at the zone boundary X point [Bir74, Hensel65]. A shear strain ε_{xy} causes an

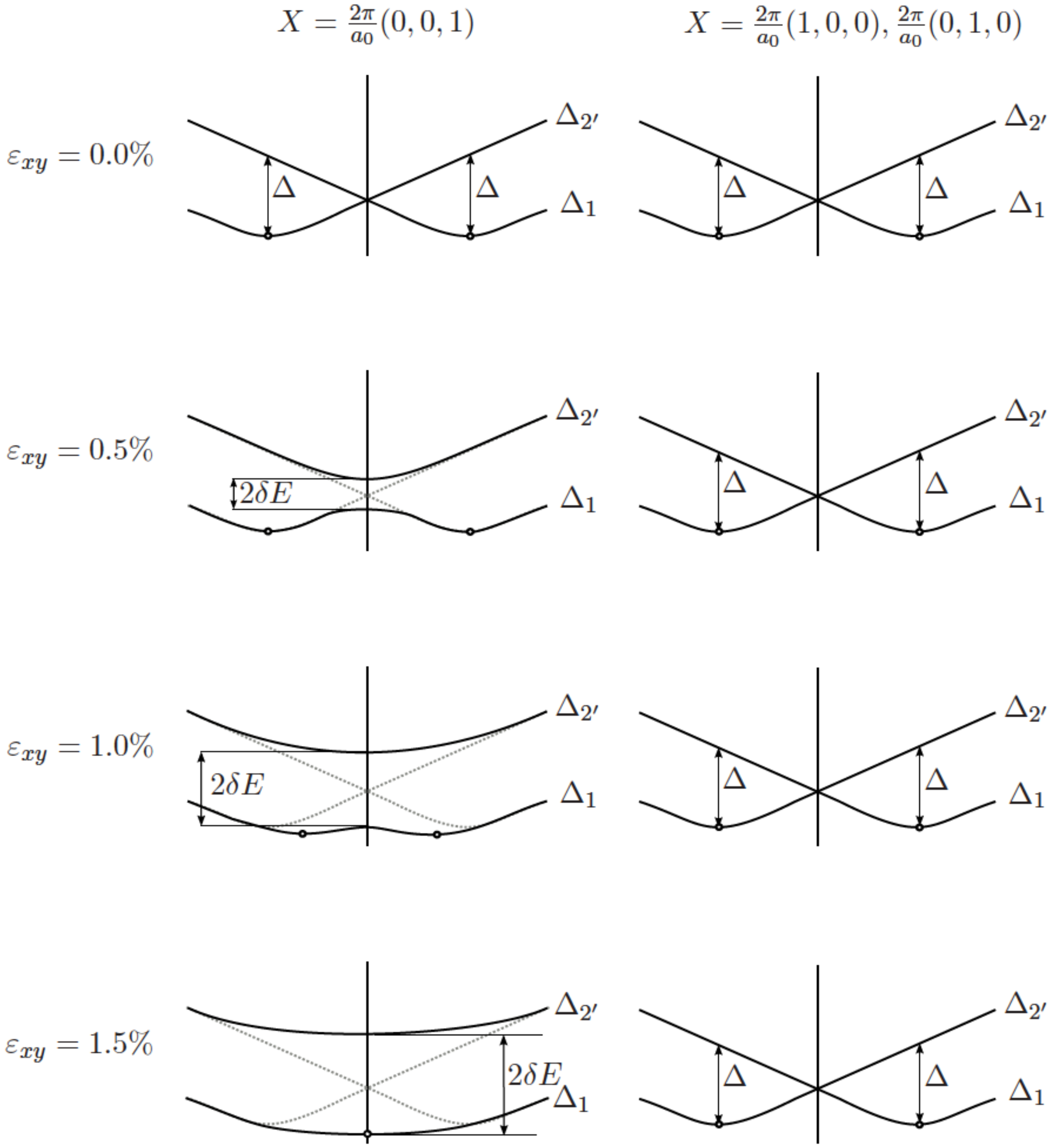


Figure 3.12: Structure of the Δ_1 and $\Delta_{2'}$ conduction bands near the zone boundary X points. The solid lines indicate a splitting of the bands due to shear strain ε_{xy} for three levels of strain. Δ denotes the band separation of unstrained Si at the conduction band edge $\boldsymbol{k}_{\mathrm{min}} = \frac{2\pi}{a_0}(0,0,0.85)$. It can be observed that the minimum of the conduction band Δ_1 along [001] (highlighted by an open circle) moves towards the zone boundary as ε_{xy} increases. The conduction bands along [100] and [010] are not affected by ε_{xy}.

energy shift between the conduction band valleys along [001] with respect to the valleys along [100] and [010]. The shift is given by

$$\delta E_{\text{shear}} = \begin{cases} -\frac{\Delta}{4}\kappa^2 {\varepsilon_{xy}}^2 & , \quad |\varepsilon_{xy}| < 1/\kappa \\ -\frac{\Delta}{4}\left(2\kappa|\varepsilon_{xy}| - 1\right) & , \quad |\varepsilon_{xy}| > 1/\kappa \end{cases} \tag{3.54}$$

Here, a dimensionless parameter $\kappa = (4\Xi_{u'})/\Delta$ has been introduced, and Δ is the band separation between the two lowest conduction bands at the conduction band edge

$$\Delta = (E_{\Delta_{2'}} - E_{\Delta_1})\Big|_{\boldsymbol{k}=\boldsymbol{k}_{\text{min}}} \ . \tag{3.55}$$

The position of the band edge in the unstrained lattice is $\boldsymbol{k}_{\text{min}} = \frac{2\pi}{a_0}(0,0,0.85)$ (compare Figure 3.12). The derivation of (3.54) is given in Section 3.7.2 after outlining the basic framework of the $\boldsymbol{k}\cdot\boldsymbol{p}$ method.

3.6.3 Strain-Induced Valence Band Splitting

The deformation potential theory for the valence bands is different from that for the conduction bands because of the degeneracy of the valence bands at the valence band maximum. The operators $\mathcal{D}^{\alpha\beta}$ are no longer scalars. Instead, they can be expressed as 3×3 matrices. Due to symmetry the six independent operators have only three independent entries, usually labeled l, m, n or a, b, d, depending on the used basis for the eigenfunctions [Cardona66].

In the basis $|x, s\rangle$, $|y, s\rangle$, $|z, s\rangle$ with s denoting the spin state ($\uparrow$ for $+z$) and ($\downarrow$ for $-z$), the matrix of the perturbation Hamiltonian is

$$\underline{H}_{strain} = \begin{array}{cc} \begin{pmatrix} \underline{H} & \underline{0}_{3\times3} \\ \underline{0}_{3\times3} & \underline{H} \end{pmatrix} & \begin{array}{l} |\uparrow\rangle \\ |\downarrow\rangle \end{array} \end{array} , \tag{3.56}$$

with $\underline{H}$ denoting the 3×3 matrix

$$\underline{H} = \begin{array}{cc} \begin{pmatrix} l\varepsilon_{xx} + m(\varepsilon_{yy} + \varepsilon_{zz}) & n\varepsilon_{xy} & n\varepsilon_{zx} \\ n\varepsilon_{xy} & l\varepsilon_{yy} + m(\varepsilon_{zz} + \varepsilon_{xx}) & n\varepsilon_{yx} \\ n\varepsilon_{zx} & n\varepsilon_{yz} & l\varepsilon_{zz} + m(\varepsilon_{xx} + \varepsilon_{yy}) \end{pmatrix} & \begin{array}{l} |x\rangle \\ |y\rangle \\ |z\rangle \end{array} \end{array} . \tag{3.57}$$

In Section 3.6.1 and Section 3.6.2 it was shown that from deformation potential theory simple analytical expressions can be derived for the conduction band shifts induced by an arbitrary strain tensor ε_{ij}. For the valence band the expressions for the strain-induced shifts of the heavy-hole, light-hole, and split-off band are more complex, which limits their practical use [Balslev66].

3.7 The $\boldsymbol{k}\cdot\boldsymbol{p}$ method

Using the $\boldsymbol{k}\cdot\boldsymbol{p}$ method [Bir74, Yu03] one can obtain analytical expressions for the band dispersion and the effective masses. It allows the extrapolation of the band structure over the entire

Brillouin zone from the energy gaps and optical matrix elements at the zone center. While the $\boldsymbol{k}{\cdot}\boldsymbol{p}$ theory has been frequently used to model the valence band of semiconductors, we will additionally apply it to model the impact of strain on the conduction band minimum.

The $\boldsymbol{k}{\cdot}\boldsymbol{p}$ method can be derived from the one-electron Schrödinger equation

$$\mathcal{H}\phi_n(\boldsymbol{r}) = \left(\frac{p^2}{2m} + V(\boldsymbol{r})\right)\phi_n(\boldsymbol{r}) = E_n\phi_n(\boldsymbol{r})\ . \tag{3.58}$$

Here $\mathcal{H}$ denotes the one-electron Hamilton operator and $V(\boldsymbol{r})$ the periodic lattice potential. The wavefunction of an electron in an eigenstate labeled n and its energy are denoted by $\phi_n(\boldsymbol{r})$ and E_n, respectively. In a periodic potential Bloch's theorem applies, and the solutions of (3.58) can be expressed as

$$\phi_{n\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}u_{n\boldsymbol{k}}(\boldsymbol{r})\ , \tag{3.59}$$

where n is the band index, $\boldsymbol{k}$ a wave vector, and $u_{n\boldsymbol{k}}(r)$ has the periodicity of the lattice. Assuming that the potential $V(\boldsymbol{r})$ is local[1], one can substitute $\phi_{n\boldsymbol{k}}(\boldsymbol{r})$ into (3.58) to obtain an equation for $u_{n\boldsymbol{k}}(\boldsymbol{r})$

$$\left(\frac{p^2}{2m} + V(\boldsymbol{r}) + \frac{\hbar\boldsymbol{k}\cdot\boldsymbol{p}}{m}\right)u_{n\boldsymbol{k}}(\boldsymbol{r}) = \left(E_{n\boldsymbol{k}} - \frac{\hbar^2k^2}{2m}\right)u_{n\boldsymbol{k}}(\boldsymbol{r})\ . \tag{3.60}$$

Considering any fixed wavevector $\boldsymbol{k} = \boldsymbol{k}_0$, the above equation yields a complete set of eigenfunctions $u_{n\boldsymbol{k}_0}$, which completely span the space of lattice periodic functions in the real space. Hence, the wavefunction $\phi_{n\boldsymbol{k}}(\boldsymbol{r})$ at $\boldsymbol{k}$ can be expanded in terms of $u_{n\boldsymbol{k}_0}$

$$\phi_{n\boldsymbol{k}}(\boldsymbol{r}) = \sum_{n'} C_{n,n'}(\boldsymbol{k},\boldsymbol{k}_0)e^{i\boldsymbol{k}\boldsymbol{r}}u_{n'\boldsymbol{k}_0}\ . \tag{3.61}$$

Once, $E_{n\boldsymbol{k}_0}$ and $u_{n\boldsymbol{k}_0}$ are known, the functions $\phi_{n\boldsymbol{k}}(\boldsymbol{r})$ and the eigenenergies $E_{n\boldsymbol{k}}$ at any $\boldsymbol{k}$ vector $\boldsymbol{k}_0 + \Delta\boldsymbol{k}$ in the vicinity of $\boldsymbol{k}_0$ can be obtained by treating the term $\hbar\Delta\boldsymbol{k}\cdot\boldsymbol{p}/m$ in (3.60) as a perturbation. Either degenerate or nondegenerate perturbation theory has to be used. The method has been first applied by Seitz [Seitz35] and was later extended to study the band structure of semiconductors [Luttinger55, Kane56, Cardona66].

This method for calculating the band structure is known as the $\boldsymbol{k}{\cdot}\boldsymbol{p}$ method. It works best for small $\Delta\boldsymbol{k}$ and can be applied to calculate the band structure near any given point $\boldsymbol{k}_0$ provided that the matrix elements of $\boldsymbol{p}$ between the wavefunctions (or the wavefunctions themselves) and the energies at $\boldsymbol{k}_0$ are known. When using a sufficiently large number of $u_{n\boldsymbol{k}_0}$ to approximate a complete set of basis functions, the band structure over the entire first Brillouin zone can be calculated by diagonalizing (3.60) numerically [Cardona66].

In the following, a nondegenerate $\boldsymbol{k}{\cdot}\boldsymbol{p}$ theory will be used to derive the band dispersion and the effective masses for the nondegenerate conduction band of Si. To analyze the effect of shear strain on the lowest conduction band Δ_1, a degenerate $\boldsymbol{k}{\cdot}\boldsymbol{p}$ method is adopted in Section 3.7.2, since the Δ_1 conduction band is expanded around the symmetry point X (zone boundary) where it touches the $\Delta_{2'}$ conduction band.

---

[1] It depends only on one spatial coordinate r.

3.7.1 Effective Electron Mass in Unstrained Si

The conduction band minima of Si lie on the $\langle 001\rangle$ axes at points $\boldsymbol{k}_{\mathrm{min}}$ distant $0.15\frac{2\pi}{a_0}$ from the X symmetry points. From the knowledge of the eigenenergies $E_{n\boldsymbol{k}_0}$ and the wavefunctions $u_{n\boldsymbol{k}_0}$ at the conduction band minima $\boldsymbol{k}_0$, the eigenvalues $E_{n\boldsymbol{k}}$ at neighboring points $\boldsymbol{k}$ can be expanded to second order in k_i in terms of the unperturbed wavefunctions and eigenenergies using nondegenerate perturbation theory

$$E_{n\boldsymbol{k}} = E_{n\boldsymbol{k}_0} + \frac{\hbar^2}{2\mathrm{m}_0}\sum_{ij} k_i\delta_{ij}k_j + \frac{\hbar^2}{\mathrm{m}_0^2}\sum_{ij}\sum_{n'\neq n} k_i\frac{\langle u_{n\boldsymbol{k}_0}|p_i|u_{n'\boldsymbol{k}_0}\rangle\langle u_{n'\boldsymbol{k}_0}|p_j|u_{n\boldsymbol{k}_0}\rangle}{E_{n\boldsymbol{k}_0}-E_{n'\boldsymbol{k}_0}}k_j\ . \tag{3.62}$$

Here, we used the index notation $\sum_{i=1}^{3} k_i p_i$ for $\boldsymbol{k}\cdot\boldsymbol{p}$ and Dirac's notation for the matrix elements

$$\langle u_{n\boldsymbol{k}_0}|p_j|u_{n'\boldsymbol{k}_0}\rangle = \frac{1}{\Omega}\int_{\Omega}\mathrm{d}\boldsymbol{r}\,u_{n\boldsymbol{k}_0}\frac{\hbar}{i}\frac{\partial}{\partial x_j}u_{n'\boldsymbol{k}_0}\ . \tag{3.63}$$

Linear terms in k_i vanish because $E_{n\boldsymbol{k}_0}$ has been assumed to be a minimum. The dispersion relation (3.62) can be rewritten in terms of the effective mass tensor $m^*_{n,ij}$ of band n

$$\frac{1}{m^*_{n,ij}} = \frac{1}{\mathrm{m}_0} + \frac{2}{\mathrm{m}_0^2}\sum_{n'\neq n}\frac{\langle u_{n\boldsymbol{k}_0}|p_i|u_{n'\boldsymbol{k}_0}\rangle\langle u_{n'\boldsymbol{k}_0}|p_j|u_{n\boldsymbol{k}_0}\rangle}{E_{n\boldsymbol{k}_0}-E_{n'\boldsymbol{k}_0}}\ . \tag{3.64}$$

In crystals with diamond structure, the effective mass tensor for the lowest conduction band Δ_1 is diagonal and can be characterized by two masses. For the [001] valley one obtains in the principal coordinate system

$$\frac{1}{m_{\mathrm{l}}} = \frac{1}{\mathrm{m}_0} + \frac{2}{\mathrm{m}_0^2}\sum_{n'\neq\Delta_1}\frac{|\langle u_{\Delta_1\boldsymbol{k}_0}|p_z|u_{n'\boldsymbol{k}_0}\rangle|^2}{E_{\Delta_1\boldsymbol{k}_0}-E_{n'\boldsymbol{k}_0}}\ , \tag{3.65}$$

$$\frac{1}{m_{\mathrm{t}}} = \frac{1}{\mathrm{m}_0} + \frac{2}{\mathrm{m}_0^2}\sum_{n'\neq\Delta_1}\frac{|\langle u_{\Delta_1\boldsymbol{k}_0}|p_x|u_{n'\boldsymbol{k}_0}\rangle|^2}{E_{\Delta_1\boldsymbol{k}_0}-E_{n'\boldsymbol{k}_0}}\ , \tag{3.66}$$

where Δ_1 denotes the band index n of the lowest conduction band. Thus, the energy dispersion (3.62) can be written in the form of (3.28).

The derived equations show that because of the coupling between electronic states in different bands via the $\boldsymbol{k}\cdot\boldsymbol{p}$ term, an electron in a solid has a mass different from that of a free electron. The coupling terms depend on two factors

1. The separation in energy between two bands n and n' determines the relative importance of the contribution of band n' to the effective mass of band n. The bigger the energetic gap between two bands the smaller is the effect on the effective mass.
2. The matrix element theorem [Tinkham64] can be used to find all bands n' that have nonzero matrix elements $\langle u_{n\boldsymbol{k}_0}|\boldsymbol{k}\cdot\boldsymbol{p}|u_{n'\boldsymbol{k}_0}\rangle$ by applying group theoretical considerations to determine all possible symmetries $u_{n'\boldsymbol{k}_0}$ can have.

Using the empirical pseudopotential method for band structure calculations (see Section 3.8) it is possible to numerically evaluate the matrix elements and hence to obtain the effective masses from (3.64).

3.7.2 Strain Effect on the Si Conduction Band Minimum

Having discussed strain-induced shifts of the conduction bands in Section 3.6.1, here the effect of strain on the effective electron mass in the lowest conduction band is analyzed.

Equation (3.47) neglects the lifting of the degeneracy of the lowest two conduction bands induced by shear strain and describes the energy shift of the conduction bands *as a whole*. Apart from the direction of the wavevector $\boldsymbol{k}$ indicating the location of the valley no information is required in (3.47) to determine the shift of the valley minima. As a consequence, the shift of a valley is independent of the exact value of the wavevector $\boldsymbol{k}$, and all $\boldsymbol{k}$ points belonging to a particular valley experience the same shift. As the effective mass relates to the curvature of the energy band, which is not changed by an overall shift in energy, equation (3.47) implies that the effective electron mass is not affected by strain.

There is clear experimental evidence that shear strain changes the effective masses of electrons in the lowest conduction band [Hensel65] and also the exciton spectrum of Si [Laude71]. In order to explain these experiments, the splitting of the two lowest conduction bands at the X symmetry point induced by shear strain (see Section 3.6.2) has to be taken into account. From (3.52) the lifting of the degeneracy at the X point can be calculated using the deformation potential constant $\Xi_{u'}$. However, since (3.52) is only valid for the symmetry point X, it cannot be used to predict the effect of strain on the valley minima $\boldsymbol{k}_{\mathrm{min}}$. To determine the change of the effective electron mass under shear strain, a degenerate $\boldsymbol{k}\cdot\boldsymbol{p}$ theory around the symmetry point X must be applied, since the two conduction bands Δ_1 and $\Delta_{2'}$ are degenerate in the unstrained lattice.

Using the theory of invariants [Luttinger56] Bir and Pikus [Bir74] determined a suitable choice of matrices describing the Hamiltonian at the points $X = \frac{2\pi}{a_0}(0,0,\pm1)$

$$\mathcal{H}(\hat{\varepsilon},\boldsymbol{k}) = \lambda\mathbb{1} + \underline{\sigma}_x(A_3k_xk_y + D_3\varepsilon_{xy}) + A_4\underline{\sigma}_zk_z \ . \tag{3.67}$$

Here,

$$\lambda = A_1k_z^2 + A_2(k_x^2 + k_y^2) + D_1\varepsilon_{zz} + D_2(\varepsilon_{xx} + \varepsilon_{yy}) \ , \tag{3.68}$$

A_1 to A_2 are scalar constants and $\underline{\sigma}_x$ and $\underline{\sigma}_z$ denote the spinor matrices given by

$$\underline{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \underline{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{3.69}$$

The scalars D_1, D_2, and D_2 are related to the deformation potential constants Ξ_u, Ξ_d, and $\Xi_{u'}$ via

$$D_1 = \Xi_u + \Xi_d \ , \tag{3.70}$$

$$D_2 = \Xi_d \ , \tag{3.71}$$

$$D_3 = 2\Xi_{u'} \ . \tag{3.72}$$

The energy dispersion of the first and the second conduction band can be determined as the eigenvalues of (3.67)

$$E_{\pm}(\hat{\varepsilon},\boldsymbol{k}) = \lambda \pm \sqrt{A_4^2k_z^2 + (2\Xi_{u'}\varepsilon_{xy} + A_3k_xk_y)^2} \ , \tag{3.73}$$

where E_- describes the dispersion of Δ_1 and E_+ that of $\Delta_{2'}$.

Assuming that this expansion around the X point is valid up to the minimum of the lowest conduction band at $\boldsymbol{k}_{\min} = \frac{2\pi}{a_0}(0,0,\pm 0.85)$, the constant A_4 can be related to A_1

$$0 = \frac{\partial E_-(\hat{\varepsilon}=0,\boldsymbol{k})}{\partial k_z}\Big|_{\boldsymbol{k}_{\min}} = 2A_1k_0 + \frac{A_4^2k_0}{\sqrt{A_4^2k_0^2}} \ . \tag{3.74}$$

Here, $k_0 = 0.15\frac{2\pi}{a_0}$ denotes the distance of the conduction band minimum of unstrained Si from the X point. From (3.74) the magnitude of A_4 can be determined

$$|A_4| = 2A_1k_0 \ . \tag{3.75}$$

In the following the effect of shear strain on the shape of the lowest conduction band is derived by two different methods.

Energy Dispersion of the Conduction Band Minimum of Strained Si: Method 1

The effect of strain on the spectrum near the lowest conduction band edge is estimated by expanding the root in (3.73) and neglecting terms proportional to ε^2 and k^4. The spectrum is expanded near the minimum of the conduction band Δ_1 at $\boldsymbol{k}_{\min} = 2\pi/a_0(0,0,0.85)$ of the *unstrained lattice*.

$$E(\hat{\varepsilon},\boldsymbol{k}) = A_1(k_z-k_0)^2 + A_2(k_x^2+k_y^2) + \Xi_u\varepsilon_{xx} + \Xi_d(\varepsilon_{xx}+\varepsilon_{yy}+\varepsilon_{zz}) - \frac{2\Xi_{u'}A_3\varepsilon_{xy}}{|A_4|k_0}k_xk_y \ . \tag{3.76}$$

By comparing this equation with (3.28) the constants A_1 and A_2 are given by

$$A_1 = \frac{\hbar^2}{2m_\mathrm{l}} \ , \tag{3.77}$$

$$A_2 = \frac{\hbar^2}{2m_\mathrm{t}} \ , \tag{3.78}$$

which leaves only A_3 undetermined.

At zero shear strain, the splitting between the two lowest conduction bands, which is denoted as Δ (see (3.74)), can be related to A_4 by evaluating (3.73) at $\boldsymbol{k} = \boldsymbol{k}_{\min}$

$$\Delta = 2|A_4|k_0 \ . \tag{3.79}$$

Thus, the last term of (3.76) can be written as

$$\frac{4A_3\Xi_{u'}}{\Delta}\varepsilon_{xy}k_xk_y \ . \tag{3.80}$$

Since this term is proportional to ε_{xy} and k_xk_y, it describes a change in effective mass proportional to strain. A $\boldsymbol{k}\cdot\boldsymbol{p}$ theory capable of describing the change in the effective mass due to strain must contain third order terms proportional to $\hat{\varepsilon}k^2$. It was shown by Bir and Pikus [Bir74] that the dominating $\hat{\varepsilon}k^2$ correction to the spectrum $E(\hat{\varepsilon},\boldsymbol{k})$ of the lowest conduction band at $\boldsymbol{k}_{\min}$ is

$$\delta E_{\hat{\varepsilon}k^2} = \frac{4\hbar^2\Xi'_{u'}}{m'\Delta}\varepsilon_{xy}k_xk_y \ , \tag{3.81}$$

where

$$\frac{1}{m'} = \frac{2}{\mathrm{m}_0^2} \sum_{n \neq \Delta_1} \frac{\langle \Delta_1 | p_x | n \rangle \langle n | p_y | \Delta_{2'} \rangle}{E_n - E_{\Delta_1}} \ , \tag{3.82}$$

and $\Xi'_{u'}$ denotes the deformation potential constant $\Xi_{u'}$ evaluated at $\boldsymbol{k}_{\mathrm{min}}$.

The coefficient A_3 can now be obtained by comparing (3.80) with (3.81). Assuming that the deformation potential constant $\Xi_{u'}$ is the same at both points X and $\boldsymbol{k}_{\mathrm{min}}$, that is $\Xi_{u'} = \Xi'_{u'}$, A_3 is given by

$$A_3 = \frac{\hbar^2}{m'} \ , \tag{3.83}$$

and the energy dispersion (3.76) around the conduction band edge becomes

$$E(\hat{\varepsilon}, \boldsymbol{k}) = \frac{\hbar^2 (k_z - k_{\mathrm{min}})^2}{2 m_{\mathrm{l}}} + \frac{\hbar^2 (k_x^2 + k_y^2)}{2 m_{\mathrm{t}}} - \frac{4 \hbar^2 \Xi_{u'} \varepsilon_{xy}}{m' \Delta} k_x k_y \ . \tag{3.84}$$

Here, the diagonal entries of the strain tensor were assumed to be zero, since they do not cause a change in the effective mass in this approximation. The impact of shear strain ε_{xy} on the effective masses becomes clearer when changing the coordinate system

$$x' \to [110] \qquad y' \to [1\bar{1}0] \qquad z' \to [001] \ .$$

This coordinate system is rotated 45° about the z axis with respect to the principal coordinate system, thus,

$$k_{x'} = \frac{k_x + k_y}{\sqrt{2}} \qquad k_{y'} = \frac{k_x - k_y}{\sqrt{2}} \qquad k_{z'} = k_z \ . \tag{3.85}$$

In the rotated coordinate system the effective mass tensor is diagonal

$$\begin{aligned} E(\hat{\varepsilon}, \boldsymbol{k}') &= \frac{\hbar^2 (k_{z'} - k_{\mathrm{min}})^2}{2 m_{\mathrm{l}}} + \frac{\hbar^2 k_{x'}^2}{2 m_{\mathrm{t}}} \left(1 + \frac{2 \Xi_{u'} m_{\mathrm{t}}}{m' \Delta} \varepsilon_{xy} \right) + \frac{\hbar^2 k_{y'}^2}{2 m_{\mathrm{t}}} \left(1 - \frac{2 \Xi_{u'} m_{\mathrm{t}}}{m' \Delta} \varepsilon_{xy} \right) \\ &= \frac{\hbar^2 (k_{z'} - k_{\mathrm{min}})^2}{2 m_{\mathrm{l}}} + \frac{\hbar^2 k_{x'}^2}{2 m_{\mathrm{t},x'}(\varepsilon_{xy})} + \frac{\hbar^2 k_{y'}^2}{2 m_{\mathrm{t},y'}(\varepsilon_{xy})} \ . \end{aligned} \tag{3.86}$$

Two transverse masses occur

$$m_{\mathrm{t},x'} = m_{\mathrm{t}} (1 + \eta \kappa \varepsilon_{xy})^{-1} \ , \tag{3.87}$$

$$m_{\mathrm{t},y'} = m_{\mathrm{t}} (1 - \eta \kappa \varepsilon_{xy})^{-1} \ , \tag{3.88}$$

that depend strain. Here, two parameters

$$\kappa = 4 \Xi_{u'} / \Delta \quad \text{and} \quad \eta = m_{\mathrm{t}} / (2 m') \tag{3.89}$$

have been introduced.

Within this approximation the effect of shear strain on the transverse masses of Si can be modeled. It was developed by Bir and Pikus [Bir74] and Hensel [Hensel65]. In the following, a more rigorous model is presented, which is also able to predict the effect of shear strain on the longitudinal mass, the splitting between conduction band valleys, and the change of position of the conduction band minimum.

Energy Dispersion of the Conduction Band Minimum of Strained Si: Method 2

The above formulae have been derived, assuming that the conduction band minima are located at $\boldsymbol{k}_{\min} = 2\pi/a_0(0, 0, \pm 0.85)$, which is a good approximation only for small shear strain. As can be seen from Figure 3.12 the minimum of the conduction band is expected to move towards the X point as the strain-induced splitting between the conduction band becomes larger. As a direct consequence, the conduction bands shape is deformed and the previous assumption $\boldsymbol{k}_{\min} = 2\pi/a_0(0, 0, 0.85)$ is not satisfied.

Thus, a more general model for the effect of shear strain on the effective masses needs to be developed, which takes the movement of the conduction band minimum as a function of strain into account. The effective masses are subsequently evaluated at the position of the conduction band minimum $\boldsymbol{k}_{\min}(\varepsilon_{xy})$.

The position of the minimum can be found from (3.73) by setting $k_x = k_y = 0$:

$$E_- = \frac{\hbar^2}{2m_\mathrm{l}} k_z^2 - \sqrt{\frac{\hbar^4 k_0^2}{m_\mathrm{l}^2} k_z^2 + (2\Xi_{u'})^2 {\varepsilon_{xy}}^2}\ . \tag{3.90}$$

Here, relations (3.75) and (3.77) are used to replace A_4 and A_1, and k_0 denotes the position of the unstrained conduction band minimum measured from the zone boundary X, $k_0 = 0.15 \cdot 2\pi/a_0$. Setting

$$\frac{\partial E_-}{\partial k_z} = 0\ , \tag{3.91}$$

the position of the conduction band minimum $k_{z,\min}$ can be obtained

$$k_{z,\min} = \begin{cases} k_0\sqrt{1 - \kappa^2 {\varepsilon_{xy}}^2} & , \quad |\varepsilon_{xy}| < 1/\kappa \\ 0 & , \quad |\varepsilon_{xy}| > 1/\kappa \end{cases}\ . \tag{3.92}$$

It can be seen that for strain smaller in magnitude than $1/\kappa$, the minimum position is shifted towards the X point. For $|\varepsilon_{xy}| = 1/\kappa$, the position of the minimum is located at the X point, thus $k_{z,\min} = 0$. The position is fixed, even when ε_{xy} is further increased. The changing position of $k_{z,\min}$ is visualized in Figure 3.12, where the impact of shear strain on the shape of the conduction bands Δ_1 and $\Delta_{2'}$ is plotted.

The strain dependent longitudinal mass $m_\mathrm{l}(\varepsilon_{xy})$ can be obtained from (3.73) by calculating

$$\frac{1}{m_\mathrm{l}(\varepsilon_{xy})} = \frac{1}{\hbar^2} \frac{\partial^2 E_-}{\partial k_z^2}\Big|_{\boldsymbol{k}=(0,0,k_{z,\min})}\ . \tag{3.93}$$

The following expressions can be derived after some algebraic manipulations

$$m_\mathrm{l}(\varepsilon_{xy}) = \begin{cases} m_\mathrm{l}\left(1 - \kappa^2 {\varepsilon_{xy}}^2\right)^{-1} & , \quad |\varepsilon_{xy}| < 1/\kappa \\ m_\mathrm{l}\left(1 - \dfrac{1}{\kappa|\varepsilon_{xy}|}\right)^{-1} & , \quad |\varepsilon_{xy}| > 1/\kappa \end{cases}\ . \tag{3.94}$$

Similar to (3.92), the dependence of the longitudinal mass on strain is described by two expressions, depending on whether the magnitude of strain is smaller or bigger than $1/\kappa$. *Note that* *in the approximation of Section 3.7.2 the effect of shear strain was modeled by the expansion (3.76) about the conduction band minimum of the unstrained lattice, such that the longitudinal mass* m_l *is not influenced by strain.*

To derive the transverse effective masses that include the dependence on $k_{z,\mathrm{min}}$, first (3.73) is transformed to the rotated coordinate system introduced in (3.85).

$$E_\pm(\hat{\varepsilon}, \boldsymbol{k}) = \lambda \pm \sqrt{A_4^2 k_z^2 + \left(2\Xi_{u'}\varepsilon_{xy} + \frac{A_3}{2}(k_x^2 - k_y^2)\right)^2} \,. \tag{3.95}$$

The effective mass in the $x' = [110]$ and $y' = [1\bar{1}0]$ direction can be obtained from

$$\frac{1}{m_{\mathrm{t},x'}(\varepsilon_{xy})} = \hbar^2 \frac{\partial^2 E_-}{\partial k_{x'}^2}\bigg|_{\boldsymbol{k}=(0,0,k_{z,\mathrm{min}})} \tag{3.96}$$

and

$$\frac{1}{m_{\mathrm{t},y'}(\varepsilon_{xy})} = \hbar^2 \frac{\partial^2 E_-}{\partial k_{y'}^2}\bigg|_{\boldsymbol{k}=(0,0,k_{z,\mathrm{min}})} \,. \tag{3.97}$$

The strain dependence of the transverse masses is given by

$$m_{\mathrm{t},x'}(\varepsilon_{xy}) = \begin{cases} m_\mathrm{t}\Big(1 + \eta\kappa\varepsilon_{xy}\Big)^{-1} & , \quad |\varepsilon_{xy}| < 1/\kappa \\ m_\mathrm{t}\Big(1 + \eta\,\mathrm{sgn}(\varepsilon_{xy})\Big)^{-1} & , \quad |\varepsilon_{xy}| > 1/\kappa \end{cases} \tag{3.98}$$

for the [110] direction and by

$$m_{\mathrm{t},y'}(\varepsilon_{xy}) = \begin{cases} m_\mathrm{t}\Big(1 - \eta\kappa\varepsilon_{xy}\Big)^{-1} & , \quad |\varepsilon_{xy}| < 1/\kappa \\ m_\mathrm{t}\Big(1 - \eta\,\mathrm{sgn}(\varepsilon_{xy})\Big)^{-1} & , \quad |\varepsilon_{xy}| > 1/\kappa \end{cases} \,. \tag{3.99}$$

for the $[1\bar{1}0]$ direction. Here, sgn denotes the signum function. For $|\varepsilon_{xy}| < 1/\kappa$ the effective masses derived in this way are consistent with equations (3.87) and (3.88). However, for $|\varepsilon_{xy}| > 1/\kappa$ the transverse masses are constant and depend on the sign of the strain only.

Finally, an analytical expression for the valley shift induced by shear strain ε_{xy}, which was given in Section 3.6.2, is calculated. According to (3.92), equation (3.90) has to be evaluated at $k_{z,\mathrm{min}} = k_0\sqrt{1 - \kappa^2{\varepsilon_{xy}}^2}$ for $|\varepsilon_{xy}| < 1/\kappa$. For $|\varepsilon_{xy}| > 1/\kappa$ the energy shift at the X point determines the overall valley shift. The shift between the valley pair along [001] and the valley pairs along [100] or [010] is obtained from

$$\delta E_\mathrm{shear} = E(\varepsilon_{xy}, \boldsymbol{k}_\mathrm{min}) - E(0, \boldsymbol{k}_\mathrm{min}) = \begin{cases} -\dfrac{\Delta}{4}\kappa^2{\varepsilon_{xy}}^2 & , \quad |\varepsilon_{xy}| < 1/\kappa \\ -\dfrac{\Delta}{4}\left(2\kappa|\varepsilon_{xy}| - 1\right) & , \quad |\varepsilon_{xy}| > 1/\kappa \end{cases} \,. \tag{3.100}$$

This derivation shows that the valley splitting cannot be obtained if the spectrum of the strained crystal is expanded around the conduction band minimum k_0 *of the unstrained crystal, as it was assumed in Section 3.7.2.*

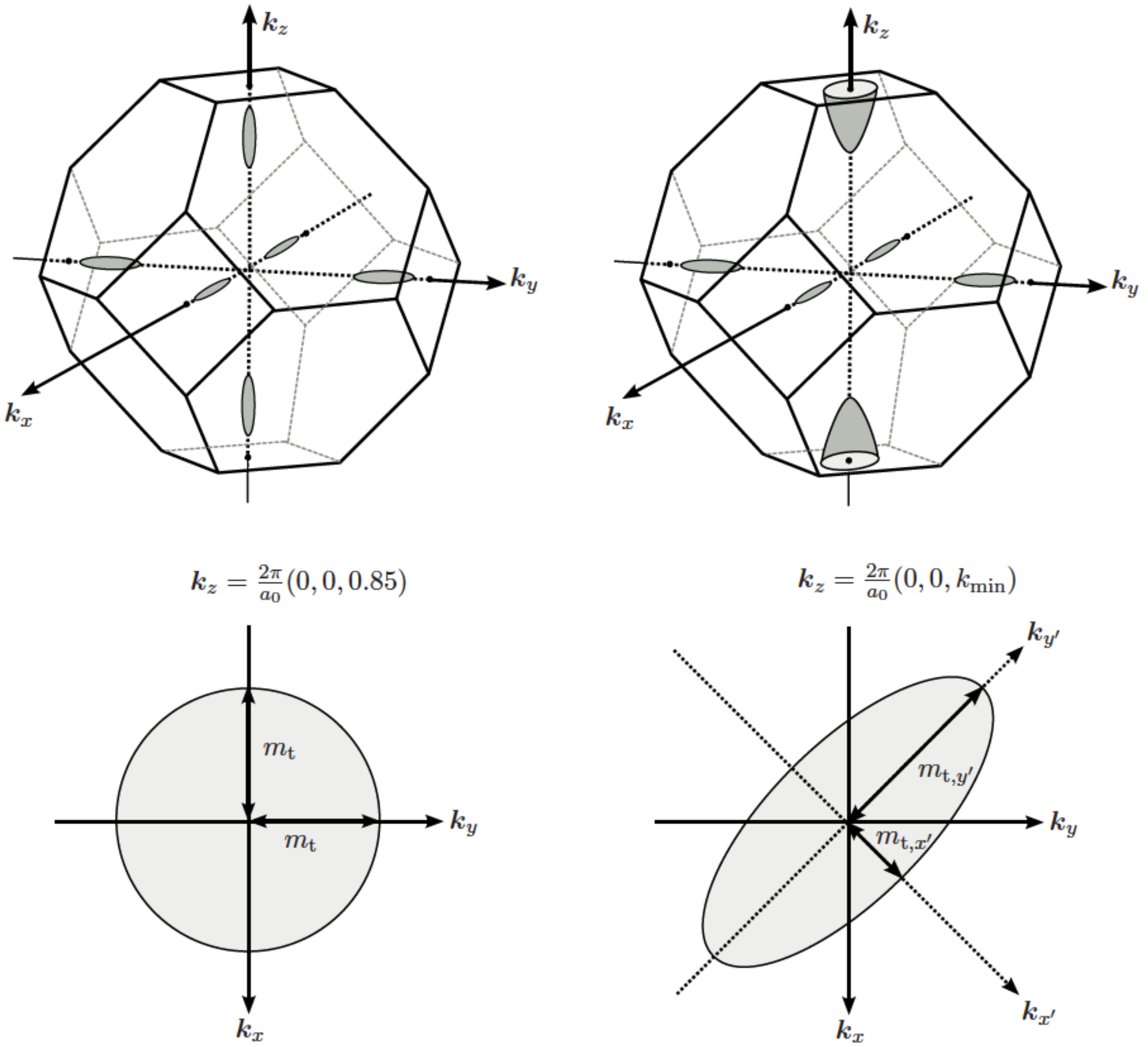


Figure 3.13: Constant energy ellipsoids for the Si conduction band with $\varepsilon_{xy} = 0$ (left side) and $\varepsilon_{xy} \neq 0$ (right side). In unstrained Si there are six equivalent valleys along the three principal axes. In the deformed lattice the valleys along $\boldsymbol{k}_z$ go down in energy, hence the surface of constant energy is bigger. The constant energy lines in the $\boldsymbol{k}_x\boldsymbol{k}_y$-plane are characterized by two masses, $m_{\mathrm{t},x'}$ and $m_{\mathrm{t},y'}$.

Discussion

In unstrained Si the constant energy surfaces of the conduction band valleys along [001] have a prolate ellipsoidal shape, where one of the three semiaxes is characterized by m_l and two semiaxes by m_t. At small shear strain ε_{xy} the constant energy surfaces take the form of scalene ellipsoids characterized by three masses, $m_{\mathrm{l},[001]}, m_{\mathrm{t},[110]}$, and $m_{\mathrm{t},[\bar{1}10]}$. These masses change under the influence of ε_{xy} and can be modeled using the equations (3.94), (3.98), and (3.99). From Figure 3.13 it can be seen that under shear strain ε_{xy}, the lines of constant energies in the $\boldsymbol{k}_x\boldsymbol{k}_y$-plane develop into ellipses with their semiaxes rotated 45° about the k_z axis. For levels of shear strain that significantly change the location of the conduction band edge, $\varepsilon_{xy} \approx 1/\kappa$, a large deformation of the shape of the conduction band takes place. In principle equations (3.94), (3.98), and (3.99) describe the change of the effective masses, but it will be shown in Section 6.1 that a (non-)parabolic approximation for the conduction band minimum is not valid in this case and full-band modeling is required for simulations of electron transport even at low electric fields.

As pointed out previously a key advantage of the $\boldsymbol{k}\cdot\boldsymbol{p}$ method is that it allows one to derive analytical expressions for the energy dispersion with the knowledge of only a small number of parameters. In the expressions derived in this section, the knowledge of only three parameters Δ, η, and κ is required to characterize the energy dispersion around the conduction band minimum under shear strain. These parameters can be calculated using the empirical pseudopotential method, which is briefly described in the next section.

3.8 Empirical Pseudopotential Theory for Arbitrary Strain

The empirical pseudopotential method in its nonlocal relativistic version was developed by Chelinkowsky and Cohen [Chelikowsky76, Cohen89]. It is frequently used to calculate the band structure of semiconductors. The method is efficient, and requires only a limited set of fitting parameters [Jungemann03, Fischetti96a, Rieger93, Bufler98]. These few parameters are usually calibrated in order to fit energy gaps, that can be determined from experiments, and are available for a large set of materials [Yu03].

3.8.1 The Empirical Pseudopotential Method

The pseudopotential theory is based on an ansatz which separates the total wave function into an oscillatory part and a smooth part, the so called pseudo wave function. The strong true potential of the ions is replaced by a weaker potential valid for the valence electrons, the pseudopotential $V(\boldsymbol{r})$, which approaches the unscreened Coulomb potential of the Si^{4+} ion at large values of r (see Figure 3.14). This replacement can be justified mathematically and shown to reproduce correctly the conduction and valence band states [Cohen89].

The one-electron Schrödinger equation is replaced by a pseudo-wave equation

$$\left(\frac{p^2}{2m} + V(\boldsymbol{r})\right)\phi_{\boldsymbol{k}}(\boldsymbol{r}) = E_{\boldsymbol{k}}\phi_{\boldsymbol{k}}(\boldsymbol{r}) \ , \tag{3.101}$$

with $\phi_{\boldsymbol{k}}(\boldsymbol{r})$ denoting the pseudo wave function and $V(\boldsymbol{r})$ the pseudopotential. This equation can

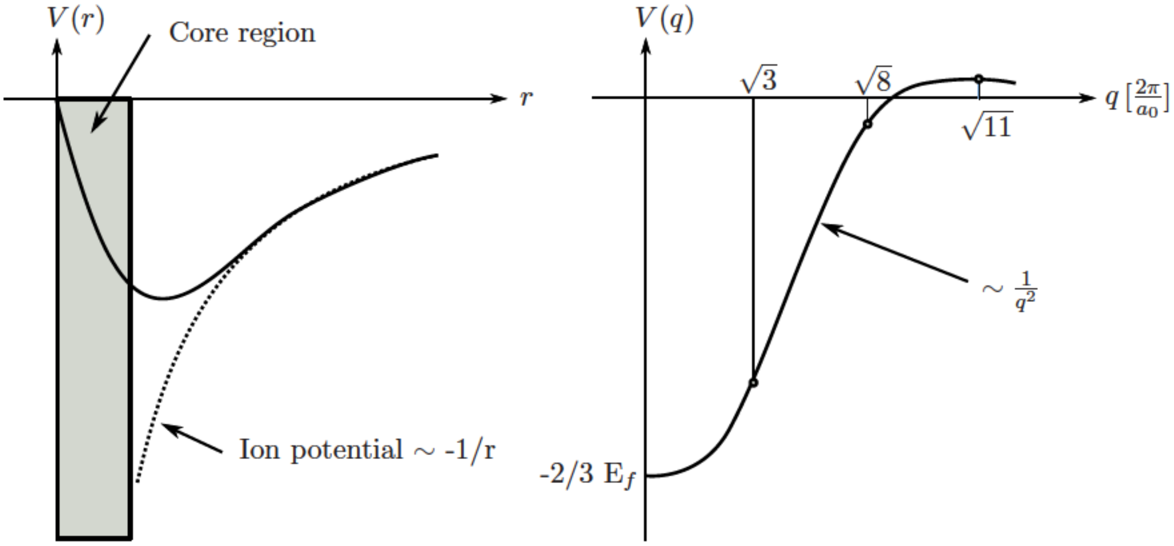


Figure 3.14: Schematic plot of a pseudopotential in real space (left) and reciprocal space (right).

be used to calculate physical properties of semiconductors which are dependent on the valence and conduction electrons only. Since pseudopotentials are only small perturbations, the energy bands are expected to be similar to those of nearly free electrons and an ansatz where the pseudo wave function is expanded into a sum of plane waves can be used

$$|\phi_{\boldsymbol{k}}\rangle = \sum_{\boldsymbol{G}} a_{\boldsymbol{G}} |\boldsymbol{k} + \boldsymbol{G}\rangle \tag{3.102}$$

to diagonalize (3.101). Here, $\boldsymbol{G}$ is a general reciprocal lattice vector as given in (3.24) and Dirac's notation $|\boldsymbol{k} + \boldsymbol{G}\rangle$ is applied to denote the plane wave with wave vector $\boldsymbol{k} + \boldsymbol{G}$. The coefficients $a_{\boldsymbol{G}}$ and the eigenvalues $E_{\boldsymbol{k}}$ can be determined from the solution of the secular equation

$$\det \left| \left[\frac{\hbar^2 k^2}{2m} - E_{\boldsymbol{k}} \right] \delta_{\boldsymbol{k},\boldsymbol{k}+\boldsymbol{G}} + \langle \boldsymbol{k} | V(\boldsymbol{r}) | \boldsymbol{k} + \boldsymbol{G} \rangle \right| = 0 \ . \tag{3.103}$$

The matrix elements of the pseudopotential are given by

$$\langle \boldsymbol{k} | V(\boldsymbol{r}) | \boldsymbol{k} + \boldsymbol{G} \rangle = \left[\frac{1}{N} \sum_{\boldsymbol{R}} e^{-i\boldsymbol{G}\cdot\boldsymbol{R}} \right] \frac{1}{\Omega} \int_{\Omega} V(\boldsymbol{r}) e^{-i\boldsymbol{G}\cdot\boldsymbol{r}} \mathrm{d}\boldsymbol{r} \ . \tag{3.104}$$

Here, $\boldsymbol{R}$ is a lattice vector in the real lattice, N is the number of atoms in the primitive unit cell[2] and Ω denotes its volume. The matrix elements are determined by the Fourier components $V_{\boldsymbol{G}}$ of the pseudopotential

$$V_{\boldsymbol{G}} = \frac{1}{\Omega} \int_{\Omega} V(\boldsymbol{r}) e^{-i\boldsymbol{G}\cdot\boldsymbol{r}} \mathrm{d}\boldsymbol{r} \ , \tag{3.105}$$

---

[2] By definition the primitive unit cell contains one lattice point.

which are frequently called form factors of the pseudopotential. If there is more than one atom in the primitive unit cell, a structure factor is introduced which depends on the relative position $\boldsymbol{r}_n$ of the respective atom in the primitive unit cell. The structure factor $S_{\boldsymbol{G}}$ is defined as

$$S_{\boldsymbol{G}} = \frac{1}{N}\sum_{n}^{N} e^{-i\boldsymbol{G}\cdot\boldsymbol{r}_n} \ , \tag{3.106}$$

where N denotes the number of atoms in the primitive unit cell. The pseudopotential $V(\boldsymbol{r})$ can be expressed in terms of the structure factor and the form factors by

$$V(\boldsymbol{r}) = \sum_{\boldsymbol{G}} V_{\boldsymbol{G}} S_{\boldsymbol{G}} \exp\left(i\boldsymbol{G}\cdot\boldsymbol{r}\right) \ . \tag{3.107}$$

In crystals with a diamond structure there are two atoms at the positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ in the primitive unit cell. By taking the midpoint between the two atoms in the unit cell as origin, the positions of the atoms are given by $\boldsymbol{r}_1 = \frac{a_0}{8}(1,1,1) = \tau$ and $\boldsymbol{r}_2 = -\frac{a_0}{8}(1,1,1) = -\tau$. Thus, the structure factor is given by

$$S_{\boldsymbol{G}} = \frac{1}{2}\big(\exp(-i\boldsymbol{G}\cdot\boldsymbol{\tau}) + \exp(i\boldsymbol{G}\cdot\boldsymbol{\tau})\big) = \cos(\boldsymbol{G}\cdot\boldsymbol{\tau}) \ . \tag{3.108}$$

In unstrained diamond structures the reciprocal lattice vectors in order of increasing magnitude are (in units of $\frac{2\pi}{a_0}$):

$$\begin{aligned}
\boldsymbol{G}_0 &= (0,0,0)\\
\boldsymbol{G}_3 &= (1,1,1),\ (\ \ 1,-1,1)\ ,\ldots\ ,(-1,-1,-1)\\
\boldsymbol{G}_4 &= (2,0,0),\ (-2,\ \ 0,0),\ \ldots\ ,(\ \ 0,\ \ 0,-2)\\
\boldsymbol{G}_8 &= (2,2,0),\ (\ \ 2,-2,0)\ ,\ldots\ ,(\ \ 0,-2,-2)\\
\boldsymbol{G}_{11} &= (3,1,1),\ (-3,\ \ 1,1)\ ,\ldots\ ,(-3,-1,-1)
\end{aligned}$$

Form factors with reciprocal lattice vectors larger than $G^2 > 11(\frac{2\pi}{a_0})^2$ are neglected, since typically $V_{\boldsymbol{G}}$ decreases as G^{-2} for large $\boldsymbol{G}$ (see Figure 3.14). Assuming that the atomic pseudopotentials are spherically symmetric $V(\boldsymbol{r}) = V(|\boldsymbol{r}|)$, the form factors only depend on the absolute value of the reciprocal lattice vector. The form factor belonging to $\boldsymbol{G}_0$ shifts the entire energy scale by a constant value, and can therefore be set to zero. The form factors belonging to the reciprocal lattice vectors $\boldsymbol{G}_3$ have an absolute value of $\sqrt{3}\cdot\frac{2\pi}{a_0}$ and are conventionally labeled V_3. Since the structure factor of the reciprocal lattice vectors $\boldsymbol{G}_4$ with magnitude $2\cdot\frac{2\pi}{a_0}$ vanishes,

$$\cos\left(\frac{2\pi}{a_0}\boldsymbol{\tau}\cdot(\pm2,0,0)\right) = \cos\left(\pm\frac{\pi}{2}\right) = 0 \ , \tag{3.109}$$

the respective form factor V_4 does not enter the pseudopotential (3.107). Thus, only three pseudopotentials form factors V_3, V_8 and V_{11} are required to calculate the band structure.

In Table 3.5 the parameters employed in the empirical pseudopotential calculations are listed. They consist of three local form factors $V_{\sqrt{3}}, V_{\sqrt{8}}, V_{\sqrt{11}}$, two parameters ($A_0$, R_0) to model the nonlocal correction, and two parameters (μ, ζ) entering the spin-orbit interaction term. The parameters coincide with the parameter set provided in [Rieger93] with the exception of $\mu = 0.00023$ Ry and $\zeta = 7.5589$ Å^{-1}, which have been adjusted in order to yield the desired split-off energy of 44 meV in the unstrained Si band structure. In the expansion of the pseudo wave function (3.102) plane waves with modulo $|\boldsymbol{G}_{lmn} - \boldsymbol{k}| < 5.7(2\pi/a_0)$ were included, which guarantees results converged to approximately 1 meV.

Table 3.5: Parameters employed in the band structure calculation of Si and Ge [Rieger93].

	Silicon	Germanium	Units
$V_{\sqrt{3}}$	-0.2241	-0.221	Rydberg
$V_{\sqrt{8}}$	-0.052	0.019	Rydberg
$V_{\sqrt{11}}$	-0.0724	0.056	Rydberg
A_l	0.03	0.275	Rydberg
R_l	1.06	1.22	Å
μ	0.00023	0.000965	Rydberg
ζ	7.5589	10.0911	Å^{-1}

3.8.2 Inclusion of Strain

To handle general strain conditions, four modifications in the band structure calculation have to be taken into account:

(i) The direct lattice vectors $\boldsymbol{a}_i{}'$ of the strained crystal are calculated by deforming the vectors $\boldsymbol{a}_i$ of the unstrained crystal according to (3.30). From the strained lattice basis vectors, the strained reciprocal lattice vectors $\boldsymbol{b}_i{}'$ can be obtained. These are used to calculate the strained lattice vectors of the reciprocal lattice which are used in the expansion of the pseudo wave function (3.102) and in the calculation of the normalizing volume of the strained unit cell Ω_0' as given in (3.31).

(ii) Since the local pseudopotential form factors enter the calculation at the strained reciprocal lattice vectors, an interpolation of the pseudopotential is required (see Figure 3.14). Different expressions have been proposed in [Friedel89, Rieger93]. In this work, the pseudopotential form factors of the strained lattice are obtained by performing a cubic spline interpolation through the pseudopotential form factors, V_0, V_3, V_8, V_{11}, and V_{3k_F}. Following [Rieger93], V_0 is set to $-2E_\text{F}/3$, and $V_{3k_\text{F}} = 0$, where k_F denotes the Fermi wave vector of the free electron gas.

(iii) In Figure 3.15 a schematic plot of the diamond structure with the primitive unit cell is plotted. The latter has a tetragonal shape. The vertex atoms of the tetrahedron and the central atom located at $(\frac{a_0}{4}, \frac{a_0}{4}, \frac{a_0}{4})$ belong to a different fcc-lattice. While the position of the vertex atoms of the tetrahedron (indicated in light-grey) can be calculated from macroscopic strain, the absolute position of the central atom in the bulk primitive unit cell (dark grey) remains undetermined. To obtain the exact position of the central atom an additional parameter for the displacement has to be taken into account.

A schematic plot showing the change of atomic positions in the primitive unit cell under strain is given in Figure 3.16. In the unstrained lattice, the central atom is positioned at the center of the tetrahedron, which is indicated by a white circle in Figure 3.16a. Under strain the vertex atoms change their positions and the central atom is displaced from the center of the distorted tetrahedron (see Figure 3.16b). In order to minimize the nearest neighbor central force energy of the system, the central atom moves towards the center of the four vertex atoms. However, opposing this reduction of energy is the increase of nearest neighbor non-central force energy

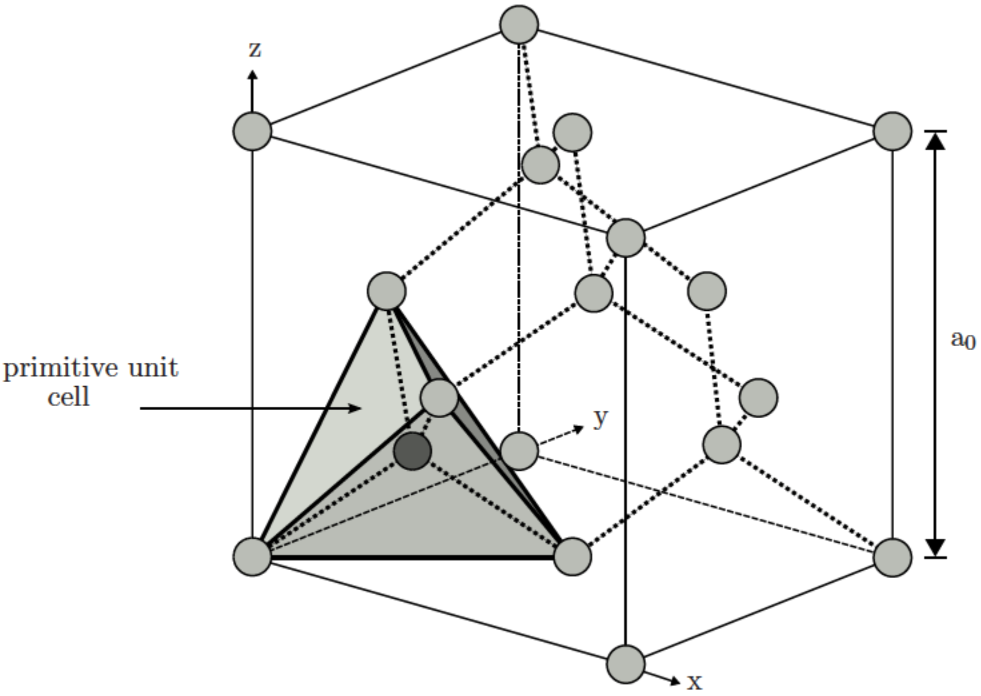
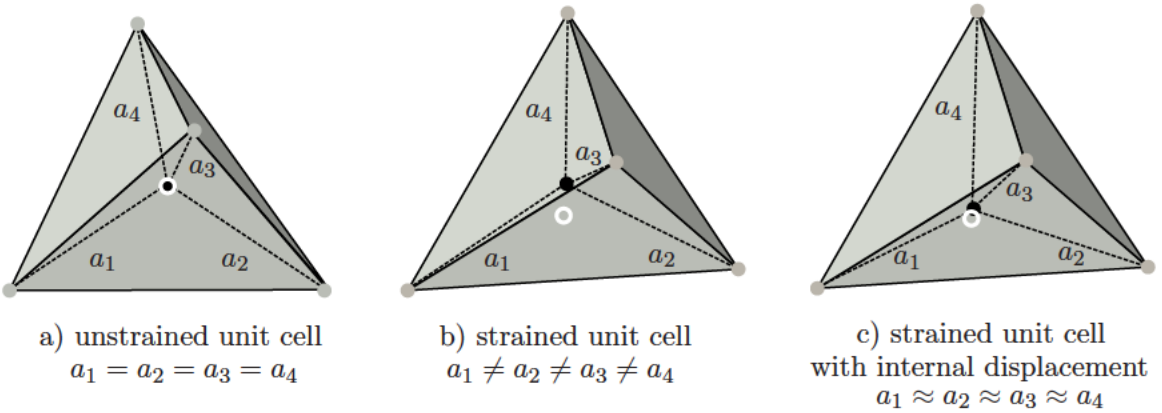


Figure 3.15: Schematic plot of the diamond structure and the primitive unit cell. The central atom at $(\frac{a_0}{4}, \frac{a_0}{4}, \frac{a_0}{4})$ is indicated by a dark sphere.



a) unstrained unit cell
$a_1 = a_2 = a_3 = a_4$

b) strained unit cell
$a_1 \neq a_2 \neq a_3 \neq a_4$

c) strained unit cell
with internal displacement
$a_1 \approx a_2 \approx a_3 \approx a_4$

Figure 3.16: Unit cell of diamond structure with center atom and its four nearest neighbors. a) In the unstrained crystal, all four bonds are of the same length. b) Under strain, atoms change their positions, leading to different bond lengths. c) The center atom is displaced to adjust a similar distance to the four neighbor atoms, whereas the positions of the vertex atoms is unchanged.

and far-neighbor energy [Kleinman62]. Thus, the central atom does not completely relax to the center of the strained tetrahedron as indicated in Figure 3.16c.

In the case of general strain, the additional displacement of the central atom with respect to the four vertex atoms of the unit tetrahedron in the diamond structure can be modeled in terms of an internal strain parameter (displacement factor) ξ: First, the positions of the vertex atoms and the central atom are derived from the macroscopic strain. Then the center of the four vertex atoms is determined. If the internal strain parameter ξ is set to zero, the central atom retains its position determined from the macroscopic strain only; if $\xi = 1$ the central atom moves to the center of the four vertex atoms, and all four bonds are of the same length. As previously discussed, neither of the two extrema occurs in a real crystal and a appropriate value $0 \leq \xi \leq 1$ for the internal strain parameter has to be used.

For the determination of the internal strain parameter of Si we performed calculations with the ab-initio total-energy and molecular-dynamics program VASP (Vienna ab-initio simulation program) [Kresse93, Kresse94, Kresse96a, Kresse96b, Kresse99]. A value of 0.5 was extracted, which is very close to previous theoretical calculations of Nielson [Nielsen85], who extracted a value of 0.53, and the experimental result 0.54 ± 0.04 [Cousins87].

If stress is applied along a fourfold axis $\langle 100 \rangle$ no internal displacement occurs. In this case the center of the deformed primitive unit cell coincides with the position of the central atom determined from macroscopic strain and all four bonds are of the same length. For the stress directions $\langle 110 \rangle$ and $\langle 111 \rangle$ analytical expressions for the internal displacement can be derived:

- For [110] stress the additional displacement along [001] is given by [Ungersboeck06a]

$$u_z = -\frac{\xi}{2} \frac{(1+\varepsilon_{xx})\varepsilon_{xy}}{1+\varepsilon_{zz}} a_0 \ . \tag{3.110}$$

- If the lattice is stressed along [111], the additional displacement is parallel to the stress direction [111]

$$\boldsymbol{u} = -\frac{\xi}{2} \varepsilon_{xy} a_0 (1,1,1) \ . \tag{3.111}$$

(iv) Finally, strain-induced loss of symmetry gives rise to a change in shape and volume of the irreducible wedge of the first BZ [Ungersboeck06a]. Irreducible wedges under various strain conditions were identified in Section 3.5. Only if the proper wedge is identified, redundancy in the band structure calculations can be avoided.

Chapter 4

Quantum Confinement and Subband Structure

CHARGE transport in the MOSFET channel is different from the bulk transport since the carriers interact with the Si-SiO$_2$ interface. Furthermore, a large electric field normal to the Si-SiO$_2$ interface causes the formation of a potential well, which confines charge carriers to a region close to the Si-SiO$_2$ interface. In the potential well a quasi two dimensional electron gas (2DEG) or hole gas (2DHG) is formed. Carriers are free to move parallel to the interface, but are tightly confined in the direction normal to the interface. This confinement leads to quantized energy levels, thus the conduction or valence bands are split into subbands.

The calculations of the energy levels of carriers confined in a quantum well are commonly based on three approximations:

- The Hartree approximation, stating that each electron moves in the average potential produced by all other electrons, thus neglecting many-body effects.
- In the semiconductor, the effective mass approximation (EMA) is applied [Bastard81].
- For the calculation of the energy levels it can also be assumed that the barrier between the insulator and the semiconductor is large enough that the envelope wave functions vanish at the semiconductor-insulator interface which is reasonable for the Si-SiO$_2$ interface [Ando82].

While the bulk band structure and its modification under the influence of mechanical strain was discussed in the previous chapter, this chapter is devoted to the subband structure of electrons in the MOS inversion layer. The influence of strain on the subband structure is shown for various orientations of the Si-SiO$_2$ interface.

4.1 Electron Confinement at the Semiconductor-Oxide Interface

In order to calculate the subband structure in the inversion layer the Schrödinger equation and the Poisson equation have to be considered as a coupled system of differential equations, which has to be solved self-consistently by numerical methods [Vasileska00]. The energy levels E and envelope wavefunctions ψ are determined by a solution of the effective Schrödinger equation

$$[T - e\Phi(z)]\psi = E\psi \ , \tag{4.1}$$

where $\Phi(z)$ is the electrostatic potential and T is the operator for the kinetic energy. The electrostatic potential determining the shape of the potential well is the solution of the Poisson equation

$$\nabla^2\Phi = -\frac{e}{\kappa_{\mathrm{si}}}[N_{\mathrm{dop}}(z) + p(z) - n(z)] \ . \tag{4.2}$$

Here, $N_{\mathrm{dop}}(z)$ is the doping profile in the semiconductor, and $p(z)$ and $n(z)$ denote the hole and electron concentration, respectively. The boundary conditions for the potential are $\lim_{z\to\infty}\Phi(z) = 0$ for the bulk interface and

$$\kappa_{\mathrm{diel}}\frac{\mathrm{d}\Phi}{\mathrm{d}z}\bigg|_{z=0^-} = \kappa_{\mathrm{sc}}\frac{\mathrm{d}\Phi}{\mathrm{d}z}\bigg|_{z=0^+} \ , \tag{4.3}$$

for the Si-SiO$_2$ interface. In (4.3) κ_{diel} denotes the dielectric permittivity of the insulating layer and κ_{sc} that of the semiconductor. Assuming the effective mass approximation, the kinetic energy operator T in (4.1) can be written as

$$T = \frac{1}{2}\sum_{i,j}\nu_{ij}p_ip_j \ , \qquad i,j = x,y,z \ . \tag{4.4}$$

where $p_j = -i\hbar(\partial/\partial x_j)$ denotes the momentum operator, and ν_{ij} is the reciprocal effective mass tensor. A coordinate system is chosen such that the z axis is normal to the semiconductor-insulator interface.

The conduction band valleys of semiconductors are typically oriented along high symmetry lines of the first Brillouin zone. To calculate the subband structure for any substrate orientation, it is necessary to introduce a unitary transformation from the crystallographic system x', y', z' to the interface coordinate system. The momentum operator and the reciprocal effective mass tensor in the interface coordinate system are transformed as

$$p_j = \sum_k u_{jk}{p_k}' \ , \tag{4.5}$$

$$\nu_{ij} = \sum_k u_{ik}u_{jk}{\nu_{kk}}' \ . \tag{4.6}$$

Here, u_{jk} are the elements of a unitary matrix, ${\nu_{kk}}' = 1/{m_k}'$, where ${m_k}'$ denote the principal effective masses of the constant-energy ellipsoid in the semiconductor. The unitary transformation matrix from the crystallographic coordinate system to the interface coordinate system is given by

$$\underline{U} = \begin{pmatrix} \cos\phi\cos\theta & -\sin\phi & \cos\phi\sin\theta \\ \sin\phi\cos\theta & \cos\phi & \sin\phi\sin\theta \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} , \tag{4.7}$$

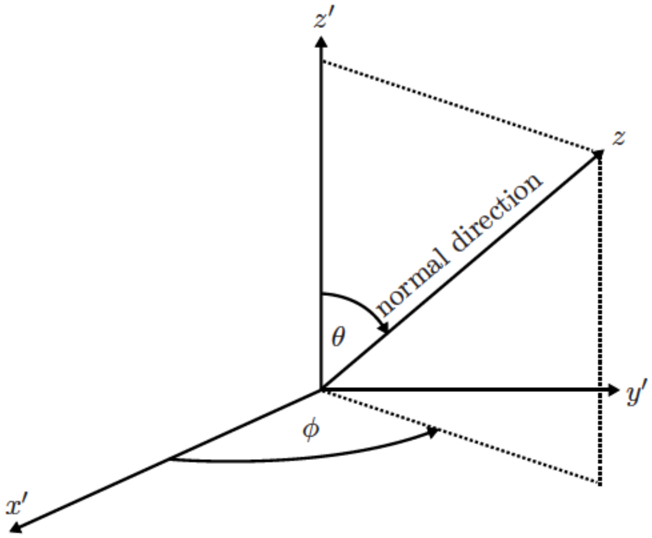


Figure 4.1: Angles ϕ and θ with respect to the coordinate system which diagonalizes the inverse effective mass tensor.

and involves a rotation of ϕ about the z' axis followed by a subsequent rotation of θ about the new y axis (compare Figure 4.1). The direction of the new z axis after the two rotations is thus given by the last column of $\underline{U}$,

$$\boldsymbol{e}_z = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta) \ . \tag{4.8}$$

Since the potential is assumed to be a function of z only, it is possible to the separate the trial solution of (4.1) into a z-dependant factor $\xi(z)$, and a plane wave factor representing free motion in the xy plane [Stern67]

$$\psi(x, y, z) = \xi(z)\exp(ik_1 x + ik_2 y) \ . \tag{4.9}$$

The functions ξ must satisfy the equation

$$\frac{\hbar^2}{2}\nu_{33}\frac{\mathrm{d}^2\xi}{\mathrm{d}z^2} + i\hbar^2(\nu_{13}k_1 + \nu_{23}k_2)\frac{\mathrm{d}\xi}{\mathrm{d}z} + (e\phi(z) + \hat{E})\xi(z) = 0 \ , \tag{4.10}$$

where

$$\hat{E} = E - \frac{\hbar^2}{2}(\nu_{11}k_1^2 + 2\nu_{12}k_1k_2 + \nu_{22}k_2^2) \ . \tag{4.11}$$

Following Stern and Howard [Stern67], the first derivative in the above equation can be eliminated using the substitution[1]

$$\xi(z) = \zeta(z)\exp\left(-\frac{iz}{\nu_{33}}(\nu_{13}k_1 + \nu_{23}k_2)\right) . \tag{4.12}$$

The differential equation for $\zeta(z)$ takes the form

$$\frac{\mathrm{d}^2\zeta_i(z)}{\mathrm{d}z^2} + \frac{2m_\perp}{\hbar^2}(E_i + e\phi(z))\zeta_i(z) = 0 \ , \quad \text{using} \quad m_\perp = 1/\nu_{33} \ . \tag{4.13}$$

---

[1] This separation of variables is not valid if the boundary condition $\zeta(0) = 0$ is relaxed [Stern67].

The eigenfunctions being subject to the boundary conditions $\lim_{z\to\infty}\zeta_i(z) = \lim_{z\to 0}\zeta_i(z) = 0$ and the eigenvalues of (4.13) are labeled by a subscript i. The energy spectrum is given by

$$E(k_1,k_2) = E_i + \frac{\hbar^2}{2}\left[\left(\nu_{11} - \frac{\nu_{13}^2}{\nu_{33}}\right)k_1^2 + 2\left(\nu_{12} - \frac{\nu_{13}\nu_{23}}{\nu_{33}}\right)k_1k_2 + \left(\nu_{22} - \frac{\nu_{23}^2}{\nu_{33}}\right)k_2^2\right] . \quad (4.14)$$

and represents constant-energy ellipses above the minimum energy E_i. The fact that E_i is independent of k_1 and k_2 is a result of the boundary condition $\zeta(0) = 0$. The energy levels E_i for a given value of $m_\perp$ generate a set of subband minima called subband ladder. Since the value of the quantization mass depends on the substrate orientation, so do the number and the degeneracy of the subband ladders. Obviously, if conduction band valleys have the same orientation with respect to the surface, these valleys belong to the same subband ladder. Because of the kinetic energy term (4.4) in the Schrödinger equation the valleys with the largest quantization mass $m_\perp$ have the lowest energy. Following a widely used convention, the subbands belonging to the ladder lowest in energy are labeled $0, 1, 2, \ldots$, those of the second ladder $0', 1', 2', \ldots$, the third ladder $0'', 1'', 2'', \ldots$, and so on [Ando82].

The principal effective masses $m_{\shortparallel,1}$ and $m_{\shortparallel,2}$, associated with motion parallel to the surface can be deduced from (4.14). This equation represents an ellipse whose principal axes are not parallel to k_1 and k_2. Introducing the matrix

$$\underline{M} = \begin{pmatrix} \nu_{11} - \nu_{13}^2/\nu_{33} & \nu_{12} - \nu_{13}\nu_{23}/\nu_{33} \\ \nu_{12} - \nu_{13}\nu_{23}/\nu_{33} & \nu_{22} - \nu_{23}^2/\nu_{33} \end{pmatrix} , \quad (4.15)$$

equation (4.14) can be written as

$$E(k_1,k_2) = E_i + \frac{\hbar^2}{2}\boldsymbol{k}^T\underline{M}\boldsymbol{k} \; . \quad (4.16)$$

The inverse effective masses $m_{\shortparallel,1}^{-1}$ and $m_{\shortparallel,2}^{-1}$ are the eigenvalues of $\underline{M}$ and can be calculated by solving the secular equation

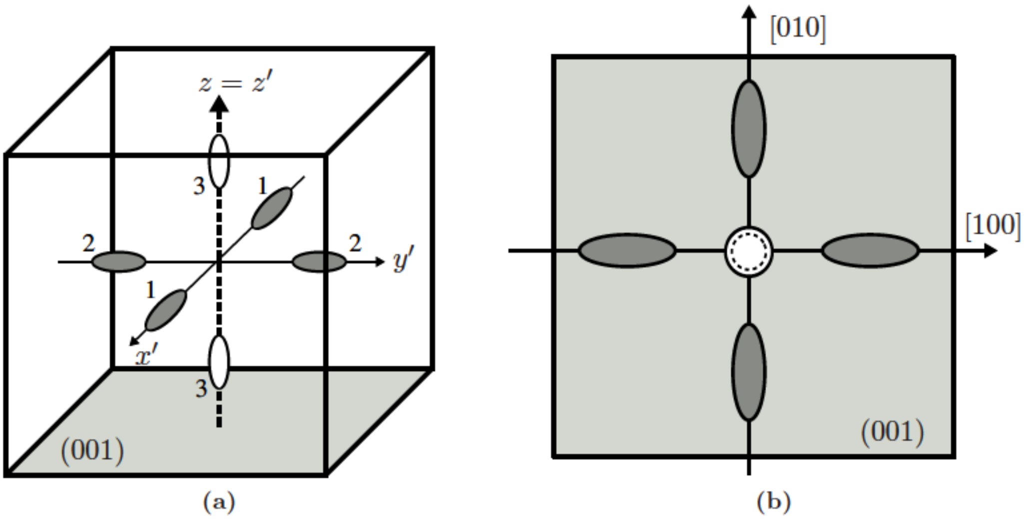
$$\det\left(\underline{M} - m_{\shortparallel}^{-1}\mathbb{1}\right) = 0 \; , \quad (4.17)$$

where $\mathbb{1}$ denotes the two dimensional unity matrix.

4.1.1 The Si-SiO$_2$ Interface

In the following section the subband structure of electrons confined near the Si-SiO$_2$ interface is discussed. As shown in Section 3.4.1 the conduction band edge of Si is located along the Δ direction and is located in the vicinity of the X symmetry point. Each valley is characterized by a longitudinal mass m_l and a transverse mass m_t. In the following only the three valley pairs along the equivalent Δ axes are considered, whereas the valleys at the L points are neglected. In the principal crystallographic system the inverse effective mass tensors describing the energy dispersion of the three valley pairs are given by

$$\underline{\nu}'^{(1)} = \begin{pmatrix} \frac{1}{m_\mathrm{l}} & 0 & 0 \\ 0 & \frac{1}{m_\mathrm{t}} & 0 \\ 0 & 0 & \frac{1}{m_\mathrm{t}} \end{pmatrix} , \quad \underline{\nu}'^{(2)} = \begin{pmatrix} \frac{1}{m_\mathrm{t}} & 0 & 0 \\ 0 & \frac{1}{m_\mathrm{l}} & 0 \\ 0 & 0 & \frac{1}{m_\mathrm{t}} \end{pmatrix} , \quad \text{and} \quad \underline{\nu}'^{(3)} = \begin{pmatrix} \frac{1}{m_\mathrm{t}} & 0 & 0 \\ 0 & \frac{1}{m_\mathrm{t}} & 0 \\ 0 & 0 & \frac{1}{m_\mathrm{l}} \end{pmatrix} . \quad (4.18)$$



(a) (b)

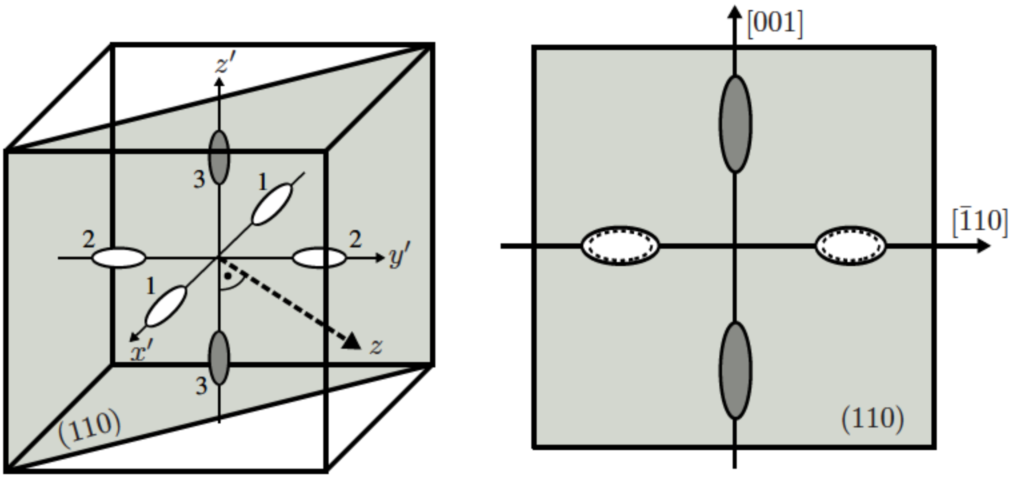
Figure 4.2: (a) Alignment of constant-energy surfaces of the Si conduction band with respect to the substrate surface (001). (b) Projection of constant-energy surfaces onto the (001) plane. Concentric spheres indicate the twofold degeneracy of the unprimed ladder. Constant-energy lines belonging to the primed ladder (filled in dark grey) have elliptic shape.

4.1.2 Substrate Orientation (001)

For cubic semiconductors the Miller indices of a plane $(hk\ell)$ are also the coordinates of the normal vector of this plane. For substrate orientation (001) the coordinate system in which the z axis is normal to the substrate surface coincides with the crystallographic system (see Figure 4.2a). Thus, no coordinate transformation needs to be involved and $\nu_{ij} = \nu_{ij}{}'$.

The quantization masses of the three different valleys v are $m_\perp^v = 1/\nu_{33}^v$. By comparison with (4.18) the quantization masses can be determined, yielding $m_\perp^{(1,2)} = m_\mathrm{t}$ for valleys labeled 1 and 2 in Figure 4.2a and $m_\perp^{(3)} = m_\mathrm{l}$ for the valley pair with the label 3. Since $m_\mathrm{l} > m_\mathrm{t}$, the two valleys with the large quantization mass $m_\perp^{(3)} = m_\mathrm{l}$ belong to the lowest (unprimed) subband ladder, whereas the four valleys with $m_\perp = m_\mathrm{t}$ constitute the primed subband ladder. The transport masses $m_{\shortparallel,1}$ and $m_{\shortparallel,2}$ for the three valleys can be found from (4.17). Since the matrix $\underline{M}^v$ contains only diagonal entries, $M_{11}^v = \nu_{11}^v$ and $M_{22}^v = \nu_{22}^v$, the transport masses are easily obtained from (4.18), $m_{\shortparallel,1}^v = 1/\nu_{11}^v$, and $m_{\shortparallel,2}^v = 1/\nu_{22}^v$.

For transport calculations not only the transport masses of a particular subband ladder are essential, but also the direction of the principal axes of the constant-energy ellipse with respect to the crystallographic axes of the wafer. In Figure 4.2b the projection of the constant-energy surfaces onto the substrate with orientation (001) is shown. The projection yields two subband ladders with spherical constant-energy lines (unprimed ladders) and four subbands with elliptic constant-energy lines (primed ladders). However, the transport properties of the four primed subband ladders are in general not equivalent since two of them are aligned along the [100] direction, whereas the other two are aligned along the [010] direction. Also, the principal axes of the constant-energy ellipsoids are interchanged.



(a) (b)

Figure 4.3: (a) Alignment of constant-energy surfaces of the Si conduction band with respect to the substrate surface (110). (b) Projection of constant-energy surfaces onto the (110) plane. Concentric ellipses (unfilled) indicate the fourfold degeneracy of the unprimed ladder. The constant-energy lines belonging to the primed ladder are filled in dark grey.

4.1.3 Substrate Orientation (110)

On (110) oriented substrate, the principal crystallographic system does not coincide with a system where the z axis is normal to the substrate surface and a coordinate transformation is necessary to obtain the quantization and transport masses. From the direction of the substrate normal

$$\mathbf{e}_n^{(110)} = \frac{1}{\sqrt{2}}(1,1,0) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta) \ , \tag{4.19}$$

the angles $\theta = 90°$ and $\phi = 45°$ of the coordinate transformation can be identified. The transformation matrix becomes

$$\underline{U} = \begin{pmatrix} 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ -1 & 0 & 0 \end{pmatrix} . \tag{4.20}$$

The inverse effective mass tensors of the three valley pairs can be calculated from (4.6)

$$\underline{\nu}^{(1,2)} = \begin{pmatrix} \frac{1}{m_\mathrm{t}} & 0 & 0 \\ 0 & \frac{m_\mathrm{t}+m_\mathrm{l}}{2m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}+m_\mathrm{l}}{2m_\mathrm{t}m_\mathrm{l}} \\ 0 & \frac{m_\mathrm{t}+m_\mathrm{l}}{2m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}+m_\mathrm{l}}{2m_\mathrm{t}m_\mathrm{l}} \end{pmatrix} , \underline{\nu}^{(3)} = \begin{pmatrix} \frac{1}{m_\mathrm{l}} & 0 & 0 \\ 0 & \frac{1}{m_\mathrm{t}} & 0 \\ 0 & 0 & \frac{1}{m_\mathrm{t}} \end{pmatrix} . \tag{4.21}$$

It can be seen that the quantization masses $m_\perp^{(1,2)} = 1/\nu_{33}^{(1,2)} = 2m_\mathrm{l}m_\mathrm{t}/(m_\mathrm{t}+m_\mathrm{l})$ for the two valley pairs labeled 1 and 2 in Figure 4.3a are equal. The quantization mass of the remaining two valleys (labeled 3 in Figure 4.3a) is $m_\perp^{(3)} = m_\mathrm{t}$. Since $2m_\mathrm{l}m_\mathrm{t}/(m_\mathrm{t}+m_\mathrm{l}) > m_\mathrm{t}$, the four valleys with the larger quantization mass $m_\perp^{(1,2)}$ belong to the lowest (unprimed) subband ladder, whereas the two valleys with $m_\perp^{(3)} = m_\mathrm{t}$ constitute the primed subband ladder. To calculate the transport masses the eigenvalue problem given in (4.17) for $\underline{M}^v$ has to be solved. On (110) oriented substrate $\underline{M}^v$ is given by

$$\underline{M}^{(1,2)} = \begin{pmatrix} \frac{1}{m_\mathrm{t}} & 0 \\ 0 & \frac{2}{m_\mathrm{t}+m_\mathrm{l}} \end{pmatrix} , \qquad \underline{M}^{(3)} = \begin{pmatrix} \frac{1}{m_\mathrm{l}} & 0 \\ 0 & \frac{1}{m_\mathrm{t}} \end{pmatrix} . \tag{4.22}$$

Thus, the transport masses of the unprimed subband ladder

$$m_{\parallel,1}^{(1,2)} = m_\mathrm{t} \ , \quad m_{\parallel,2}^{(1,2)} = (m_\mathrm{l}+m_\mathrm{t})/2 \ , \tag{4.23}$$

and the primed ladder

$$m_{\parallel,1}^{(3)} = m_\mathrm{l} \ , \quad m_{\parallel,2}^{(3)} = m_\mathrm{t} \ , \tag{4.24}$$

can be obtained. In Figure 4.3b the constant-energy lines of the subbands for a (110) oriented substrate are shown. The unprimed ladder is fourfold degenerate, whereas the primed ladder is twofold degenerate. This is opposite to the situation for (001) oriented substrate, where the unprimed ladder was twofold degenerate and the primed ladder was fourfold degenerate. The major principal axis of the unprimed subbands is $[\bar{1}10]$, whereas the major principal axis of the twofold degenerate primed subband ladder is [001].

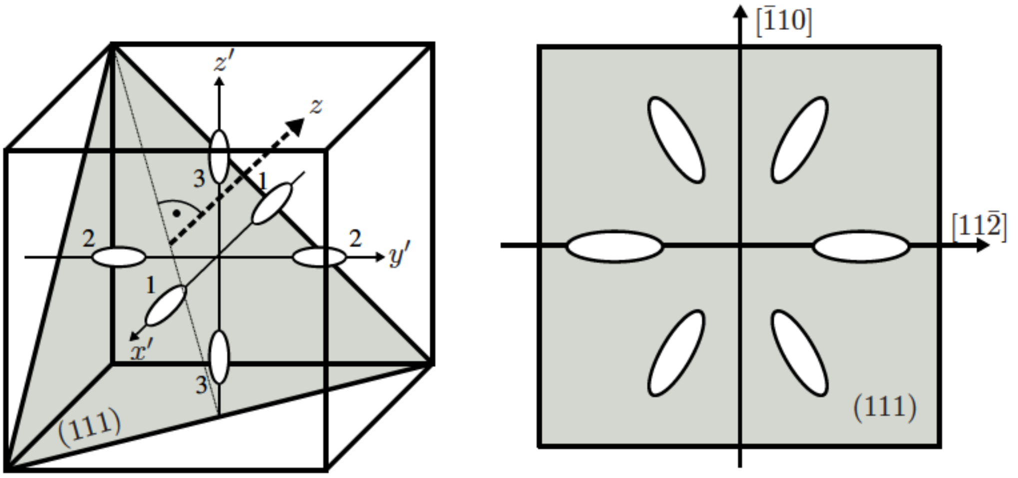
4.1.4 Substrate Orientation (111)

The substrate normal $\boldsymbol{e}_n^{(111)}$ for substrate orientation (111) is $\frac{1}{\sqrt{3}}(1,1,1)$. Thus, the axes of the principal crystallographic systems have to be rotated by the angles $\cos\theta = 1/\sqrt{3}$ ($\theta \approx 54.74°$) and $\phi = 45°$. The inverse effective mass tensor for the three valley pairs in the coordinate system with the z axis perpendicular to the substrate surface are

$$\underline{\nu}^{(1)} = \begin{pmatrix} \frac{m_\mathrm{t}+5m_\mathrm{l}}{6m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{l}-m_\mathrm{t}}{2\sqrt{3}m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}-m_\mathrm{l}}{3\sqrt{2}m_\mathrm{t}m_\mathrm{l}} \\ \frac{m_\mathrm{l}-m_\mathrm{t}}{2\sqrt{3}m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}+m_\mathrm{l}}{2m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{l}-m_\mathrm{t}}{\sqrt{6}m_\mathrm{t}m_\mathrm{l}} \\ \frac{m_\mathrm{t}-m_\mathrm{l}}{3\sqrt{2}m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{l}-m_\mathrm{t}}{\sqrt{6}m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}+2m_\mathrm{l}}{3m_\mathrm{t}m_\mathrm{l}} \end{pmatrix} , \tag{4.25}$$

$$\underline{\nu}^{(2)} = \begin{pmatrix} \frac{m_\mathrm{t}+5m_\mathrm{l}}{6m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}-m_\mathrm{l}}{2\sqrt{3}m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}-m_\mathrm{l}}{3\sqrt{2}m_\mathrm{t}m_\mathrm{l}} \\ \frac{m_\mathrm{t}-m_\mathrm{l}}{2\sqrt{3}m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}+m_\mathrm{l}}{2m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}-m_\mathrm{l}}{\sqrt{6}m_\mathrm{t}m_\mathrm{l}} \\ \frac{m_\mathrm{t}-m_\mathrm{l}}{3\sqrt{2}m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}-m_\mathrm{l}}{\sqrt{6}m_\mathrm{t}m_\mathrm{l}} & \frac{m_\mathrm{t}+2m_\mathrm{l}}{3m_\mathrm{t}m_\mathrm{l}} \end{pmatrix} , \tag{4.26}$$

$$\underline{\nu}^{(3)} = \begin{pmatrix} \frac{2m_\mathrm{t}+m_\mathrm{l}}{3m_\mathrm{t}m_\mathrm{l}} & 0 & \frac{\sqrt{2}(m_\mathrm{l}-m_\mathrm{t})}{3m_\mathrm{t}m_\mathrm{l}} \\ 0 & \frac{1}{m_\mathrm{t}} & 0 \\ \frac{\sqrt{2}(m_\mathrm{l}-m_\mathrm{t})}{3m_\mathrm{t}m_\mathrm{l}} & 0 & \frac{m_\mathrm{t}+2m_\mathrm{l}}{3m_\mathrm{t}m_\mathrm{l}} \end{pmatrix} . \tag{4.27}$$



(a) (b)

Figure 4.4: (a) Alignment of constant-energy surfaces of the Si conduction band with respect to the (111) substrate surface. (b) Projection of constant-energy surfaces onto the (111) plane. Since the quantization mass is equal for all valleys, only one sixfold degenerate subband ladder is formed.

It can be seen that all valleys have the same quantization mass $m_3^{(1,2,3)} = \frac{3m_\mathrm{t}m_\mathrm{l}}{m_\mathrm{t}+2m_\mathrm{l}}$ and therefore belong to the same subband ladder. To determine the transport masses, the eigenvalues of $\underline{M}^v$ have to be calculated

$$\underline{M}^{(1,2)} = \begin{pmatrix} \frac{3(m_\mathrm{l}+m_\mathrm{t})}{2m_\mathrm{t}(2m_\mathrm{l}+m_\mathrm{t})} & \frac{\sqrt{3}(m_\mathrm{l}-m_\mathrm{t})}{2m_\mathrm{t}(2m_\mathrm{l}+m_\mathrm{t})} \\ \frac{\sqrt{3}(m_\mathrm{l}-m_\mathrm{t})}{2m_\mathrm{t}(2m_\mathrm{l}+m_\mathrm{t})} & \frac{m_\mathrm{l}+5m_\mathrm{t}}{4m_\mathrm{l}m_\mathrm{t}+2m_\mathrm{t}^2} \end{pmatrix} , \quad \underline{M}^{(3)} = \begin{pmatrix} \frac{3}{2m_\mathrm{t}+m_\mathrm{t}} & 0 \\ 0 & \frac{1}{m_\mathrm{t}} \end{pmatrix} . \tag{4.28}$$

Solving the secular equation (4.15) the transport masses evaluate to

$$m_{\parallel,1}^{(1,2,3)} = m_\mathrm{t} , \quad \text{and} \quad m_{\parallel,2}^{(1,2,3)} = \frac{2m_\mathrm{l}+m_\mathrm{t}}{3} \tag{4.29}$$

for all three valley pairs.

The constant-energy lines of the subbands for a (111) substrate are shown in Figure 4.4b. The unprimed ladder is sixfold degenerate. Two ladders have their major principal axes parallel to the [11$\bar{2}$] direction. The principal axes of the other ladders is inclined by an angle of +30° and −30° from [$\bar{1}$10].

In Table 4.1 the principal effective masses of Si for the three discussed surface orientations are summarized. In unstrained Si the six conduction band valleys form a set of two subband ladders for substrate orientation (001) and (110), whereas for (111) oriented substrate only one subband ladder with sixfold degeneracy is formed. The energy alignments of the Si conduction subband ladders constituting on (001), (110), and (111) oriented substrate are shown Figure 4.5a and Figure 4.5b.

surface orientation	degeneracy	$m_{\parallel,1}$	$m_{\parallel,2}$	$m_{\perp}$	ladder
(001)	2	m_t	m_t	m_l	unprimed
	4	m_l	m_t	m_t	primed
(110)	4	m_t	$\frac{m_\mathrm{t}+m_\mathrm{l}}{2}$	$\frac{2m_\mathrm{t}m_\mathrm{l}}{m_\mathrm{t}+m_\mathrm{l}}$	unprimed
	2	m_l	m_t	m_t	primed
(111)	6	m_t	$\frac{2m_\mathrm{l}+m_\mathrm{t}}{3}$	$\frac{3m_\mathrm{t}m_\mathrm{l}}{m_\mathrm{t}+2m_\mathrm{l}}$	unprimed

Table 4.1: Principal effective masses of the six Si conduction band minima along Δ for three surface orientations. Here, $m_{\parallel,1}$, and $m_{\parallel,2}$ denote the transport masses, and $m_\perp$ is the quantization mass.

4.2 Strain Effects

In this section the effect of strain on the subband structure of Si inversion layers is analyzed. For this purpose we assume that the energy shifts of the individual subband ladders can be calculated using the deformation potential theory. Furthermore it is assumed that the effective mass change induced by shear strain in bulk Si is reflected in the subband structure as well.

Both strain and confinement breaks the cubic symmetry of bulk Si. As a result the degeneracy among different pairs of valleys is lifted, while each of these pairs retain their twofold valley degeneracy (Kramers degeneracy) [Ashcroft76].

The effect of strain on the subband structure is illustrated in Figure 4.5c. If strain is present, the energetic position of the subband ladders with respect to each other is not completely determined from their quantization masses. An additional shift of each subband ladder has to be considered according to (3.47). Since this shift depends on the valley orientation, it may be different for each valley and consequently the degeneracy between the subband ladders can be lifted.

Electron mobility in the inversion layer can be improved by preferentially populating the subband ladder with smallest transport mass. This can be achieved by shifting ladders with larger transport mass up in energy resulting not only in a smaller net effective transport mass, but also in a reduction of intervalley scattering. However, since the density of states is reduced by pushing subband ladders with large transport mass up in energy, the Fermi level accommodates at higher values and the electron gas experiences more pronounced degeneracy effects. These effects are crucial in the determination of the electron mobility and will be discussed in greater detail in Section 5.4.

4.2.1 Substrate Orientation (001)

On (001) substrate the energy difference between the lowest subband of the unprimed ladder and the lowest subband of the primed ladder stems from the different quantization masses.

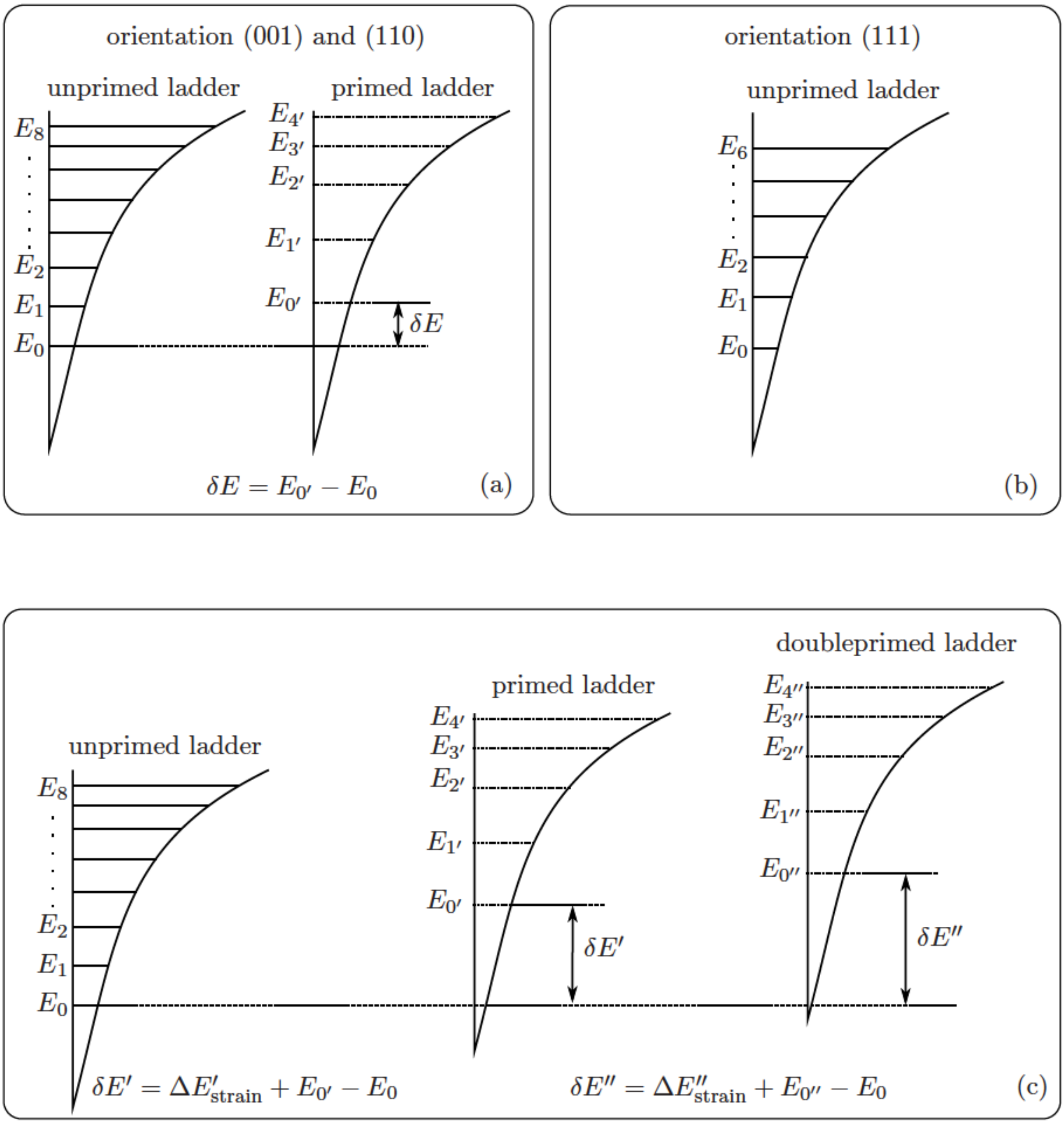


Figure 4.5: Energy lineups of the conduction subbands in unstrained Si for (a) substrate orientation (001), (110), and (b) substrate orientation (111). (c) Strain changes the lineup of the energies and thus is able to remove the degeneracy of subband ladders.

The larger quantization mass of the unprimed ladder (m_l=0.916 m$_0$) as compared to the small quantization mass of the primed ladder (m_t=0.19 m$_0$) yields a smaller splitting between the subband energies of the unprimed ladder than that of the unprimed ladder.

Electron transport in the inversion layer on (001) substrate can be improved by preferentially populating the twofold degenerate unprimed ladder and depopulating the primed ladder which is fourfold degenerate. A repopulation has two consequences for electron transport:

- A higher fraction of electrons has the smaller transport mass of the unprimed ladder.
- Additionally, by shifting the subbands of the primed ladder up in energy with respect to the unprimed ladder, the scattering rate to this subband ladder is decreased. This also contributes to the mobility increase.

The intrinsic splitting between unprimed and primed subband ladder can be increased by applying

- compressive stress along the interface normal,
- biaxial in-plane tensile strain (e.g. from epitaxially growing Si on Si$_{1-x}$Ge$_x$).
- tensile stress along the in-plane $\langle 110 \rangle$ direction.

Under the first two conditions the strain tensor only contains non-zero diagonal components ($\varepsilon_{xx} = \varepsilon_{yy} \neq \varepsilon_{zz}$). In the third case, where stress is applied along the $\langle 110 \rangle$ direction, additionally ε_{xy} is non-zero and an effective mass change occurs. Under the strain conditions listed above the degeneracy of the subband ladders is not changed. Thus, the unprimed ladder is still twofold degenerate, and the primed ladder remains fourfold degenerate.

Using linear deformation potential theory the splitting of the unprimed subband ladders can be calculated from (3.47). Under shear strain the additional shift between the primed and unprimed subband ladder given in (3.54) and the effective mass change (see equations (3.98), (3.99), and (3.94)) of electrons in the unprimed ladder has to be accounted for. As depicted in Figure 4.6, the energy dispersion of electrons in the unprimed subband ladder is not longer isotropic in the (001) plane and transport characteristics depends on the direction of transport.

Recent experiments suggest an enhancement of electron mobility if tensile stress is applied in the [100] direction [Irie04]. Under such stress conditions $\varepsilon_{xx} \neq \varepsilon_{yy}$, thus the fourfold degeneracy of the primed subbands is lifted and three sets of twofold degenerate ladders are formed (compare Figure 4.5).

4.2.2 Substrate Orientation (110)

On (110) substrate the energy difference between the two subband ladders formed in the unstrained case (compare Table 4.1) can be increased by applying tensile stress along [001]. Compressive stress along [$\bar{1}$10] or biaxial in-plane tensile stress will also increase the intrinsic splitting between the fourfold, unprimed and the twofold, primed subband ladder.

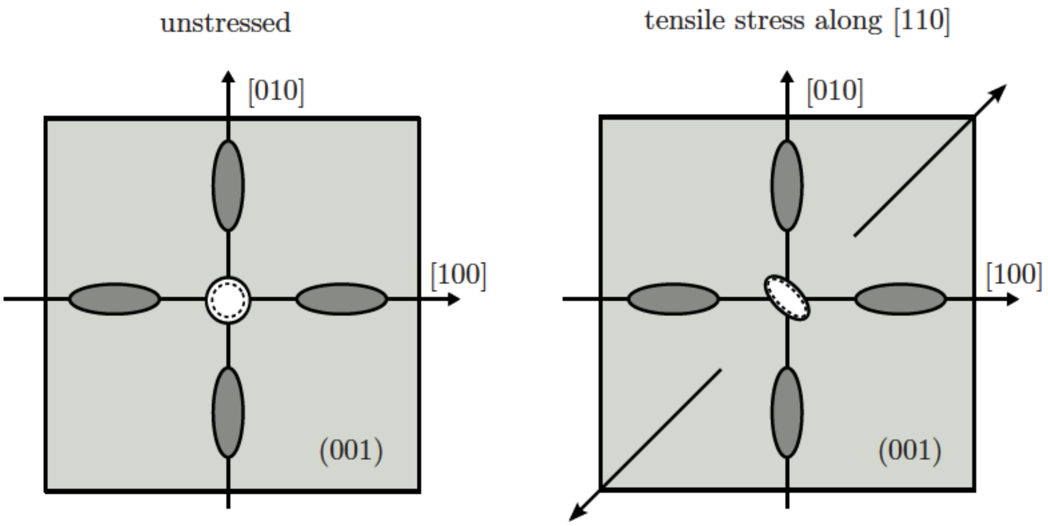


Figure 4.6: On the left the projection of constant-energy surfaces onto the (001) plane is plotted. The concentric circles indicate the twofold degeneracy of the unprimed ladder. Constant-energy lines of the primed ladder have ellipsoidal shape. On the right the effect of shear strain ε_{xy} on the subband dispersion is visualized: the effective masses of the unprimed subband ladder change.

4.2.3 Substrate Orientation (111)

On (111) substrate only one subband ladder forms which is sixfold degenerate. Biaxial in-plane strain does not split the Δ valleys, hence the subband ladders remain sixfold degenerate. The degeneracy can be lifted by applying stress along one out of the two in-plane directions, $[\bar{1}10]$ or $[11\bar{2}]$, since the resulting strain yields a different energy shift (3.47) for different subband ladders.

Chapter 5

Physical Mobility Modeling

Carrier mobility is a key parameter for the numerical simulation of the electrical characteristics of semiconductor devices. Many analytical models have been developed capturing accurately the dependence of mobility on temperature, doping, and electric field [Lombardi88, Selberherr89, Agostinelli91, Shirahata92, Darwish97, Kondo01, Reggiani02]. All these models were originally developed for unstrained Si. To describe the mobility in strained Si they have to be adapted accordingly [Dhar05].

In this work the electron mobility in bulk Si and Si inversion layers is analyzed by solving the semi-classical Boltzmann transport equation. The latter is solved numerically using a Monte Carlo (MC) method. For this purpose the Vienna Monte Carlo simulator VMC [VMC2.006] was developed, offering simulation algorithms for both bulk semiconductors and one-dimensional devices with models based on both analytical bands (ABMC) and the fullband structure (FBMC). VMC provides a comprehensive set of scattering models including phonon scattering, ionized impurity scattering, alloy scattering, and impact ionization.

The chapter is organized as follows: In Section 5.1, the basic features of the carrier mobility in semiconductors and the concept of the universality of the effective mobility are described. In Section 5.2 the validity of the Boltzmann transport equation as a fundamental equation for the description of carrier transport in semiconductor devices is discussed. The simulation method and an overview of the scattering models are given in Section 5.3. Finally, a new MC algorithm, which allows the inclusion of degeneracy effects, is presented in Section 5.4.

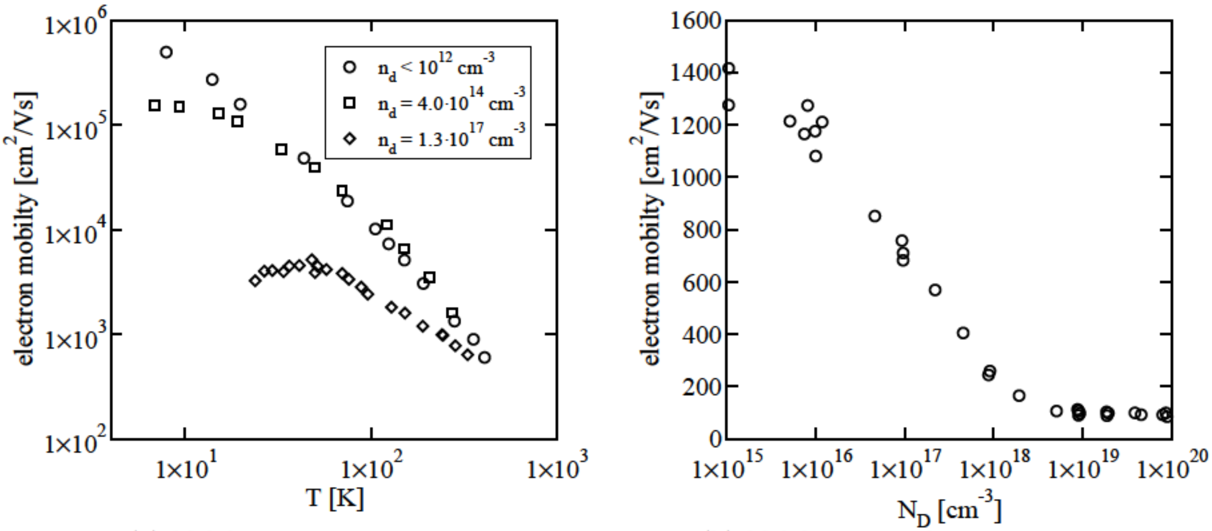
5.1 Mobility

Mobility describes the relation between drift velocity of electrons or holes and an applied electric field in a solid.

$$\boldsymbol{v}_\mathrm{d} = \hat{\mu}\boldsymbol{E} \tag{5.1}$$

Here, v_d is the drift velocity, $\boldsymbol{E}$ is the electric field, and $\hat{\mu}$ denotes the mobility tensor.

In a semiconductor the mobility of electrons is different from that of holes. The reason is the different band structure and scattering mechanisms of these two carrier types. When one charge



(a) Mobility versus temperature

(b) Mobility versus donor concentration

Figure 5.1: Experimental electron mobility data for Si [Jacoboni77, Tosic81].

carrier is dominant the conductivity of a semiconductor is directly proportional to the mobility of the dominant carrier. In Figure 5.1a and 5.1b the electron mobility of Si is plotted. The mobility in Si is a strong function of temperature and impurity concentration.

The device characteristics of MOSFETs is strongly influenced by transport in the inversion layer. Thus, the lattice mobility, representing a bulk quantity, cannot be directly used as a model parameter. In fully depleted silicon-on-insulator (FDSOI) and ultra-thin-body (UTB) MOSFETs all charge carriers reside in the inversion layer, thus quantum confinement and surface roughness scattering have to be taken into account when modeling the mobility.

The low-field mobility in inversion layers, when analyzed as a function of the confining electric field, is a function of doping, gate-voltage, back-bias voltage, and gate oxide thickness. Sabnis and Clemens found that the mobility data shows a *universal* behavior [Sabnis79], if it is plotted as a function of the effective field

$$E_{\text{eff}} = \frac{e}{\kappa_{\text{si}}}(N_{\text{depl}} + \eta N_{\text{s}}) \ . \tag{5.2}$$

Here, $\eta \sim 1/2$ for electrons, N_{depl} and N_{s} are the depletion and inversion charge densities, respectively, and κ_{si} is the dielectric constant of Si. This led to the conclusion that the MOSFET inversion layer mobility is a reproducible property associated with the Si-SiO$_2$ system, and not a parameter sensitive to nominal process variations.

On the basis of this work, investigations for different substrate orientations were performed. It was found that a similar universal behavior is achieved when the value of η [Takagi94] is properly adapted.

The first mobility studies were performed on MOSFETs with a spatially uniform doping profile in the channel. Later on, the validity of the universality of the effective mobility in channels with arbitrary doping profiles was investigated. It was discovered that the doping dependence of the effective mobility can be eliminated if plotted as a function of $E_{\text{eff}} \propto (\eta N_{\text{s}} + bN_{\text{depl}})$, where

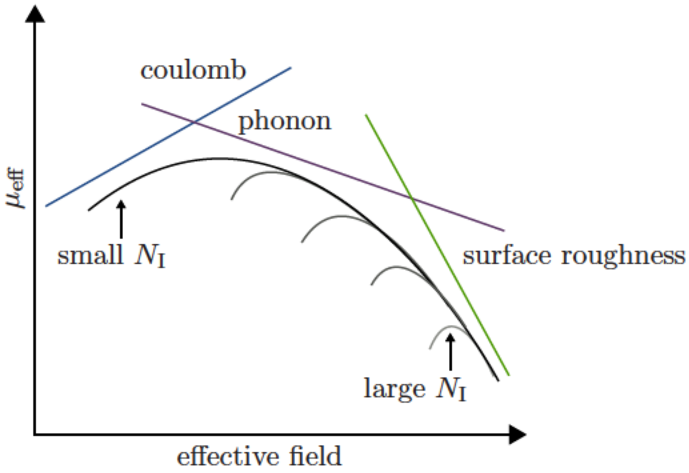


Figure 5.2: Universal mobility in Si inversion layer at $T = 300$ K.

the coefficient b is sensitive upon the shape of the doping profile [Vasileska97].

With the advent of SOI technologies, the physical basis and limitation of the universal nature of the effective mobility was examined for fully depleted SOI inversion layers. It was found that the universal mobility behavior does not hold if the electron density distribution reaches the lower surface of the Si layer. For such SOI structures with extremely thin Si film thickness T_{si} it has been predicted that there exists another kind of universal mobility behavior as a function of the inversion electron density N_{s}, independent of the impurity concentration N_{sub} and the buried oxide layer thickness T_{box} [Shoji97].

The effective mobility μ_{eff} in Si inversion layers can be experimentally determined from the drain conductance g_{d} in the linear region

$$\mu_{\mathrm{eff}} = \frac{L}{W}\frac{g_{\mathrm{d}}(V_{\mathrm{g}})}{N_{\mathrm{s}}(V_{\mathrm{g}})} \,, \tag{5.3}$$

where L and W are the transistor length and width, respectively, N_{s} is the surface carrier concentration, and V_{g} is the gate voltage. The value of E_{eff} can be calculated from (5.2) using

$$N_{\mathrm{depl}} = \sqrt{4\kappa_{\mathrm{si}}\phi_{\mathrm{B}}N_{\mathrm{sub}}/e} \,, \quad \text{where} \quad \phi_{\mathrm{B}} = \frac{\mathrm{k_B}T}{e}\ln(N_{\mathrm{sub}}/n_i) \,. \tag{5.4}$$

Here, N_{sub} is the substrate impurity concentration, $E_i - e\phi_{\mathrm{B}}$ is the bulk Fermi energy, and n_i is the intrinsic carrier concentration of Si ($n_i = 1.4 \times 10^{10}$ cm^{-3} at 300 K).

Figure 5.2 shows that the mobility characteristics of MOSFET inversion layers can be split into three distinctive regions. At low inversion charge densities (low vertical fields), mobility is limited by scattering with doping atoms and charges at the Si-SiO$_2$ interface (Coulomb scattering). Going to higher inversion densities, phonon scattering gains importance and dominates over Coulomb scattering. At large N_{s} scattering with surface roughness limits the total mobility.

To understand how the different scattering mechanisms affect the performance of a device on a circuit level, the device switching trajectory in the V_{DS}-V_{GS} space has to be considered. It has been found that for the delay of a ring oscillator the mobility at low and intermediate E_{eff} is important, because there the device spends comparatively more time than in the high gate field region [Mujtaba95].

5.2 Validity of the Boltzmann Transport Equation

As device dimensions in modern CMOS transistors are in the order of only several ten nanometers, the question of the validity of the Boltzmann transport equation,

$$\frac{\partial f_n(\boldsymbol{r},\boldsymbol{k},t)}{\partial t} + \boldsymbol{v}_n(\boldsymbol{k}) \cdot \nabla_{\boldsymbol{r}} f_n(\boldsymbol{r},\boldsymbol{k},t) + \frac{1}{\hbar}\boldsymbol{F}(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{k}} f_n(\boldsymbol{r},\boldsymbol{k},t) = \left(\frac{\partial f_n}{\partial t}\right)_{\mathrm{coll}} , \qquad (5.5)$$

as a fundamental description of carrier transport arises. This equation was originally derived for dilute gases. In the following some of the approximations of the Boltzmann transport equation and their implications are addressed.

The solution of the Boltzmann transport equation with an external force $\boldsymbol{F}(\boldsymbol{r})$ provides the distribution function $f_n(\boldsymbol{r},\boldsymbol{k},t)$ from which macroscopic quantities can be derived. The right-hand side of (5.5) describes the changes to the distribution function induced by scattering. The particle's group velocity is determined from the semiconductor band structure $E_n(\boldsymbol{k})$ as $\boldsymbol{v}_n(\boldsymbol{k}) = \hbar^{-1}\nabla_{\boldsymbol{k}} E_n(\boldsymbol{k})$. In the parabolic band approximation, $\hbar^{-1}\nabla_{\boldsymbol{k}} E_n(\boldsymbol{k}) = \hbar\boldsymbol{k}/\hat{m}^*$, and the particle's group velocity can be calculated from the effective mass tensor $\hat{m}^*$.

The distribution function $f_n(\boldsymbol{r},\boldsymbol{k},t)\mathrm{d}^3k\mathrm{d}^3r$ defines the probability density to find a particle in $\mathrm{d}^3k\mathrm{d}^3r$ at a given time t. Obviously, such a statistical description can only be appropriate when the number of carriers is large. Extremely down-scaled devices may contain too few carriers to justify this kind of statistical treatment.

Since carriers interact through their electric fields, the distribution function $f_n(\boldsymbol{r},\boldsymbol{k},t)$ at a particular point in the six dimensional position-momentum (phase) space at a given time can only be determined from the knowledge of f_n in all other points. This would involve a treatment using an $N-$particle system and an $N-$particle distribution function. However, if the carrier-carrier correlations are weak, the $N-$particle distribution function can be contracted to a one-particle distribution function [Harris04]. Alternatively, the influence of other carriers can be treated through the self-consistent electric field [Venturi89] and schemes where the Pauli exclusion principle is included [Bosi76, Lugli85, Yamakawa96, Ungersboeck06b].

A main assumption of the Boltzmann transport equation is that particles can be treated semiclassically, obeying Newton's law. Quantum mechanics enters the equation only through the band structure and the description of the collision term. Since both the position and the momentum of a particle are arguments of the distribution function, apparently the quantum mechanical uncertainty principle $\Delta p \Delta r \geq \hbar/2$ is violated. Assuming a spread in particle energy of $\mathrm{k_B}T$, one finds that the spread in position is

$$\Delta r \geq \hbar/(2\sqrt{2m^*\mathrm{k_B}T}) = \lambda_{\mathrm{B}}/2 \; . \qquad (5.6)$$

Here, λ_{B} denotes the particle's thermal average wavelength. Thus, one should not attempt to localize the particle's position exactly with respect to its thermal average wavelength. If the potential varies sharply on the scale of λ_{B}, which is typically in the order of 10 nm to 20 nm at room temperature, condition (5.6) is not satisfied, and instead of the Boltzmann equation a wave equation must be solved to study the propagation of a carrier wave through the device.

5.3 The Monte Carlo Method

Numerical techniques allow a solution of the Boltzmann transport equation with a rigorous treatment of the collision terms. Among them, the Monte Carlo (MC) method is well established [Kosina00b]. In the MC method, trajectories of charge carriers undergoing scattering events are calculated numerically both in real- and in momentum space under the influence of an external electric field. The method uses random numbers to select a scattering event and to determine the time of free flight. Because of its statistical nature, the calculation of a large number of trajectories is required to determine the average values of interest with a certain precision. In fact, this is the main disadvantage of the MC method as compared to deterministic solution methods of the Boltzmann transport equation. However, especially for the simulation of hot-carrier transport phenomena in two- or three dimensions MC is the method of choice, since complex physical models can be taken into account.

After the pioneering work of Kurosawa in 1966 [Kurosawa66], who was the first to apply the MC method to simulate carrier transport in semiconductors, a significantly improved MC method was successfully applied to transport calculations in a variety of semiconductors [Jacoboni83]. For electrons in Si, the most thoroughly investigated case, it is believed that a satisfactory understanding of the band structure and of the basic scattering mechanisms has been achieved giving rise to a “standard model” [Fischetti96b].

For predicting performance of modern CMOS bulk and SOI devices an accurate MC evaluation of carrier transport properties in inversion layers is of primary importance. Due to the strong confinement of carriers in the inversion layer of MOSFETs or due to the geometric confinement in multi-gate FETs the carrier motion is quantized in one or two confinement directions giving rise to the formation of subbands. The MC approach may incorporate the subband structure to describe the quantized carrier motion in the direction orthogonal to the current. The subbands are calculated by the self-consistent solution of the corresponding Schrödinger and Poisson equation. The free carrier motion within each subband may still be considered semiclassical and therefore can be well described by the corresponding Boltzmann equation for the subband distribution function $f_n(\boldsymbol{r}, \boldsymbol{k}, t)$, where $\boldsymbol{k}$ is a 2D wave vector. Because of possible carrier scattering between different subbands, the collision integrals on the right-hand-side of the Boltzmann equation have to include the terms responsible for the intersubband scattering processes. The transport in the inversion layer of a MOSFET is finally described by a set of Boltzmann equations for every subband, coupled by the intersubband scattering integrals. The set of the subband Boltzmann equations for $f_n(\mathbf{r}, \mathbf{k}, t)$ is conveniently solved by a MC method. Therefore, this approach combines the advantages of a quantum description in confinement direction with a semiclassical description in transport direction.

In the following, from the set of scattering models included in the simulator VMC only the electron phonon scattering mechanisms and their adaptation for the 2DEG are discussed in detail, whereas we refer to the documentation of VMC [VMC2.006] for a complete description of implemented scattering models. Additionally, the models for surface roughness scattering and dielectric screening of the 2DEG will be described.

5.3.1 Bulk Scattering Mechanisms

The collision integral at the right hand side of Boltzmann's transport equation is defined as the rate of change of the distribution function

$$\begin{aligned}\left(\frac{\partial f_n}{\partial t}\right)_{\mathrm{coll}} = \frac{\Omega}{(2\pi)^3}\sum_{n'}\int\limits_{\mathrm{BZ}}&(1-f_n(\boldsymbol{r},\boldsymbol{k},t))S_{nn'}(\boldsymbol{k},\boldsymbol{k}',\boldsymbol{r},t)f_{n'}(\boldsymbol{r},\boldsymbol{k}',t)\\ &-(1-f_{n'}(\boldsymbol{r},\boldsymbol{k}',t))S_{n'n}(\boldsymbol{k}',\boldsymbol{k},\boldsymbol{r},t)f_n(\boldsymbol{r},\boldsymbol{k},t)d^3k' \ . \end{aligned} \tag{5.7}$$

It describes the transition from an arbitrary state $(n', \boldsymbol{k}')$ into the state $(n, \boldsymbol{k})$ and the reverse process assuming that scattering does not change the particle's spin. The probability density rate for a transition from the initial state $(n', \boldsymbol{k}')$ to a final state $(n, \boldsymbol{k})$ depends on position and is proportional to the occupancy of the initial state $f_{n'}(\boldsymbol{r}, \boldsymbol{k}')$, proportional to the transition rate $S_{nn'}(\boldsymbol{k}, \boldsymbol{k}', \boldsymbol{r}, t)$, and to the probability that the final state is not occupied $(1 - f_n(\boldsymbol{r}, \boldsymbol{k}))$. The latter factor is a result of the Pauli Principle and will be discussed in more detail in Section 5.4.

Omitting time and position dependence for the sake of brevity, the scattering rates are defined as

$$S_n(\boldsymbol{k}) = \frac{\Omega}{(2\pi)^3}\int\limits_{\mathrm{BZ}}\sum_{n'} S_{n'n}(\boldsymbol{k}',\boldsymbol{k})\mathrm{d}^3k' \ , \tag{5.8}$$

where the integration can be performed over the first BZ or any primitive cell in the reciprocal space. The scattering rate represents the rate at which particles are scattered out of the initial state $(n, \boldsymbol{k})$. From Fermi's Golden Rule it follows that the scattering rates are proportional to the final density of states per spin, $g_n\left(E\right)$, which is given by

$$g_n\left(E\right) = \frac{1}{(2\pi)^3}\int\limits_{\mathrm{cell}} \delta(E - E_n(\boldsymbol{k}))\, d^3k \ , \tag{5.9}$$

In fullband MC simulations this integral is evaluated numerically. Using an analytical description for the conduction bands, the minima of the conduction bands (valleys) are approximated using the bandform function

$$\begin{aligned}\gamma^v(E) = \frac{\hbar^2}{2}\sum_{i,j=1}^{3} k_i\frac{1}{m^{*v}_{ij}}k_j &= \left(E^v_{\mathrm{nonpar}}(\boldsymbol{k}) - E^v_0\right)\left(1+\alpha^v(E^v_{\mathrm{nonpar}}(\boldsymbol{k}) - E^v_0)\right) & (5.10)\\ &= E\left(1+\alpha^v E\right) & (5.11)\end{aligned}$$

Note that the band index n was replaced by the valley index v, and E denotes the energy with respect to the valley offset E_0. m^{*v}_{ij} denotes the effective mass tensor, and α^v the non-parabolicity coefficient of the valley with index v. A parabolic band dispersion is obtained if the nonparabolicity coefficient α^v is zero.

The density of states of the analytical band structure evaluates to

$$g^v\left(E\right) = \frac{1}{\sqrt{2}}\frac{\{m^v_{\mathrm{dos}}\}^{3/2}}{\pi^2\hbar^3}\sqrt{\gamma^v\left(E\right)}(1+2\alpha^v E) \ , \tag{5.12}$$

where m^v_{dos} denotes the density of states mass of the v-th valley

$$m^v_{\mathrm{dos}} = \sqrt[3]{\{m^*_{11}\}^v\{m^*_{22}\}^v\{m^*_{33}\}^v} \ , \tag{5.13}$$

which can be calculated from the effective mass tensor.

The transition rates from state $(v, \boldsymbol{k})$ to state $(v', \boldsymbol{k}')$ for phonon scattering in a non-polar semiconductor can be written as [Jacoboni83]

$$\{S^{\text{abs}}_{\text{emi}}\}^{v'v}(\boldsymbol{k}', \boldsymbol{k}) = \frac{\pi}{\rho\Omega\omega_q}\left(N_q + \frac{1}{2} \mp \frac{1}{2}\right)\mathscr{O}|\mathcal{D}_{ij}q_j\xi_i|^2\delta[E^{v'}(\boldsymbol{k}') - E^v(\boldsymbol{k}) \mp \hbar\omega_q] \, . \tag{5.14}$$

Here, the upper and lower symbols refer to phonon absorption and emission, respectively. The rate depends on the momentum transfer $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}'$, the phonon number N_q, the deformation potential tensor $\mathcal{D}_{ij}$, the mass density of the crystall ρ, the phonon angular frequency ω_q , its polarization ξ_i, and the overlap integral $\mathscr{O}$,

$$\mathscr{O} = \left|\int_{BZ} u^*_{\boldsymbol{k}'}(\boldsymbol{r})u_{\boldsymbol{k}}(\boldsymbol{r})\exp(i\boldsymbol{G}\cdot\boldsymbol{r})d^3r\right|^2 \, . \tag{5.15}$$

The overlap factors depend on the type of transition. For intravalley transitions of electrons, $\mathscr{O}$ is frequently set to unity, even though this is true only for exact plane waves or for wave functions formed with pure s states [Jacoboni83]. Because the lowest conduction band of cubic semiconductors is a mixture of a s and p-type state, an overlap factor less than unity is obtained. Both for intra- and intervalley transitions $\mathscr{O}$ was found to be almost constant for each type of scattering process [Reggiani73], thus the values for $\mathscr{O}$ may be included in the coupling constants.

Acoustic Intravalley Scattering

The expression for the scattering probability for electron intravalley scattering from acoustic phonons can be simplified by using the elastic and equipartition approximation. Within this approximation, the energy transfer in a scattering process is neglected, and the phonon population given by the Bose-Einstein statistics

$$N_q = \frac{1}{\exp\left(\frac{\hbar\omega_q}{\mathrm{k_B}T_\mathrm{L}}\right) - 1} \tag{5.16}$$

is represented by the equipartition expression $N_q \simeq \dfrac{\mathrm{k_B}T_\mathrm{L}}{\hbar q u_\mathrm{s}} - \dfrac{1}{2}$. Thus, (5.14) becomes

$$\{S^{\text{abs}}_{\text{emi}\,\text{ac}}\}^{v}(\boldsymbol{k}', \boldsymbol{k}) = \frac{\pi q\{\Xi^v_{\text{adp}}\}^2}{\Omega u_\mathrm{s}\rho}\left(\frac{\mathrm{k_B}T_\mathrm{L}}{\hbar q u_\mathrm{s}} \mp \frac{1}{2}\right)\delta[E^v(\boldsymbol{k}') - E^v(\boldsymbol{k})] \, , \tag{5.17}$$

where v denotes the valley index, T_L is the lattice temperature, Ξ^v_{adp} is the acoustic deformation potential of the v-th valley, $\mathrm{k_B}$ is Boltzmann's constant, u_s denotes the average sound velocity, and ρ is the mass density of the crystal.

Since in the elastic approximation no distinction is made between absorption or emission processes, both transition probabilities can be added. In this approximation acoustic scattering is isotropic: any state $\boldsymbol{k}'$ belonging to the equi-energy surface has the same probability of occurrence, independent of the angle of the initial state $\boldsymbol{k}$. Thus, the rate for acoustic scattering is a function of energy only

$$\{S_{\text{ac}}\}^v(E) = \frac{2\pi\mathrm{k_B}T_\mathrm{L}\{\Xi^v_{\text{adp}}\}^2}{\hbar u_\mathrm{s}^2\rho}g^v(E) \, , \tag{5.18}$$

Table 5.1: Parameters for phonon scattering. The intervalley scattering parameters for the Δ-valleys are taken from [Jacoboni83].

		Silicon	Units
Intra	ρ	2.33^{a}	$\mathrm{g/cm^3}$
	v_{s}	9.05^{b}	cm/sec
	$\Xi_{\mathrm{adp}}^{\Delta}$	9.0^{b}	eV
	$\hbar\omega_{\mathrm{op}}$	61.2^{c}	meV
	D_tK_L	1.75^{c}	$10^8\mathrm{eV/cm}$
$\mathrm{Inter}_{\Delta\Delta}$	$\hbar\omega_{\mathrm{g1}}$	12.06	meV
	D_tK_{g1}	0.5	$10^8\mathrm{eV/cm}$
	$\hbar\omega_{\mathrm{g2}}$	18.53	meV
	D_tK_{g2}	0.8	$10^8\mathrm{eV/cm}$
	$\hbar\omega_{\mathrm{g3}}$	62.04	meV
	D_tK_{g3}	11.0	$10^8\mathrm{eV/cm}$
	$\hbar\omega_{\mathrm{f1}}$	18.86	meV
	D_tK_{f1}	0.3	$10^8\mathrm{eV/cm}$
	$\hbar\omega_{\mathrm{f2}}$	47.39	meV
	D_tK_{f2}	2.0	$10^8\mathrm{eV/cm}$
	$\hbar\omega_{\mathrm{f3}}$	59.03	meV
	D_tK_{f3}	2.	$10^8\mathrm{eV/cm}$

[a] [Jacoboni83], [b] [Jungemann03], [c] [Fischetti96b]

where $g^v(E)$ is the density of states per spin.

Optical Intravalley Scattering

The scattering probability can be written, starting from (5.14), by replacing $\Xi^2q^2\xi^2$ with a squared optical coupling constant $\{D_tK^v\}^2$ [Jacoboni83]. This constant can also include an overlap factor $\mathscr{O}$. The energy associated with the optical phonon $\hbar\omega_{\mathrm{op}}$ and the phonon number $N_q = N_{\mathrm{op}}$ can be assumed to be constant. Hence, the resulting scattering probability is

$$\{S_{\mathrm{op}}^{\substack{\mathrm{abs}\\\mathrm{emi}}}\}^v(\boldsymbol{k}',\boldsymbol{k}) = \frac{\pi}{\rho\Omega\omega_{\mathrm{op}}}\left(N_{\mathrm{op}} + \frac{1}{2} \mp \frac{1}{2}\right)\{D_tK^v\}^2\delta[E^v(\boldsymbol{k}') - E^v(\boldsymbol{k}) \mp \hbar\omega_{\mathrm{op}}]\ , \tag{5.19}$$

The scattering rate for optical phonons is a function of the final energy $E \pm \hbar\omega_{\mathrm{op}}$

$$\{S_{\mathrm{op}}^{\substack{\mathrm{abs}\\\mathrm{emi}}}\}^v(E) = \frac{\pi\{D_tK^v\}^2}{\rho\omega_{\mathrm{op}}}\left(N_{\mathrm{op}} + \frac{1}{2} \mp \frac{1}{2}\right)g^v(E \pm \hbar\omega_{\mathrm{op}})\ . \tag{5.20}$$

From the matrix element theorem one can derive that this type of scattering occurs only in the conduction band valleys along the ⟨111⟩ directions [Harrison56].

Intervalley Scattering

Both acoustic and optical phonons can cause electron transitions between states in different conduction band valleys. This scattering can be formally treated in the same way as intravalley scattering by optical phonons[Harrison56, Conwell67].

The scattering rate for intervalley scattering out of a valley v for a phonon mode η is given by

$$\{S_\eta^{\substack{\mathrm{abs}\\\mathrm{emi}}}\}^v(E) = \sum_{v' \neq v} \frac{\pi \{D_t K_\eta^{v'v}\}^2 Z^{v'}}{\rho \omega_\eta} \left(N_\eta + \frac{1}{2} \mp \frac{1}{2} \right) g^{v'} \left(E^{v'} \pm \hbar\omega_\eta - \Delta E^{v'v} \right) , \tag{5.21}$$

where a summation over the final valleys is performed. The allowed final valleys v' are determined by the selection rules for a phonon mode η: g-type phonons can induce transitions between valleys which are located on the same axis in the three dimensional $\boldsymbol{k}$ space, whereas transitions among orthogonal axes are labeled f-type. The squared coupling constants $\{D_t K_\eta^{v'v}\}^2$ depend upon the initial and final valley and the phonon branch η involved in the transition. $Z^{v'}$ denotes the number of possible final equivalent valleys, N_η is the phonon number, and $\Delta E^{v'v}$ is the difference between the energies of the minima of the final and initial valley. Since strain is able to change $\Delta E^{v'v}$, the intervalley scattering rate can be efficiently reduced through strain.

The numerical values entering the bulk phonon scattering rates are summarized in Table 5.1.

5.3.2 Scattering Mechanisms in the 2DEG

In analogy with the 3D case, the scattering rates for electrons residing in subband n of valley v are obtained from an integration over all possible two-dimensional $\boldsymbol{k}'$ states after scattering

$$S_n^v(\boldsymbol{k}) = \frac{A}{4\pi^2} \int\limits_{\mathrm{BZ}} \mathrm{d}^2k' \sum_{n',v'} S_{n',n}^{v',v}(\boldsymbol{k}', \boldsymbol{k}) . \tag{5.22}$$

Here, A denotes the area, and a summation over all subbands n' and valleys v' after scattering is performed. The density of states per spin for the 2DEG is obtained as

$$g_n(E) = \frac{1}{(2\pi)^2} \int \delta[E - E_n(\boldsymbol{k})] \, \mathrm{d}^2k . \tag{5.23}$$

As in the three dimensional case the energy subband dispersion $E_n^v(\boldsymbol{k})$ of subband n in a valley v can be approximated close to its minimum using a nonparabolic relation [Ando82, Laux88]

$$E(1 + \alpha_n^v E) = \frac{\hbar^2}{2} \left(\frac{k_1^2}{m_{\parallel,1}^v} + \frac{k_2^2}{m_{\parallel,2}^v} \right) , \tag{5.24}$$

where $E = E_n^v(\boldsymbol{k}) - E_{n,0}^v$ denotes the energy with respect to the subband minimum $E_{n,0}^v$, and α_n^v is the nonparabolicity coefficient. The integration over the energy in (5.23) can be performed analytically yielding

$$g_n^v(E) = \frac{\{m_{\mathrm{dos}}\}_n^v}{2\pi\hbar^2} (1 + 2\alpha_n^v E) \Theta[E] , \tag{5.25}$$

for the density of states of the subband n of valley v. Here, Θ denotes the Heaviside step function, and $\{m_\mathrm{dos}\}^v = \sqrt{m^v_{\parallel,1} m^v_{\parallel,2}}$ is the density of states mass. In Si the subband dispersion is frequently assumed to be independent of the subband index n [Jungemann93, Fischetti02, Roldan96]. The masses $m^v_{\parallel,1}$ and $m^v_{\parallel,2}$ for three substrate orientations of Si are listed in Table 4.1.

The bulk phonon scattering model for the six valleys along Δ has been adapted for the 2DEG following the treatment of Price [Price81]. The well accepted model for the bulk phonon spectrum of Jacoboni [Jacoboni83] can be used for the Si inversion layer.

Acoustic Intravalley Scattering

In the 2DEG scattering with long-wavelength acoustic phonons causes intra- and intersubband transitions within the same valley. The transition rate of electrons in the inversion layer due to interaction with acoustic phonons is given by [Price81]

$$\{S_\mathrm{ac}\}^v_{n'n}(\boldsymbol{k}',\boldsymbol{k}) = \frac{2\pi}{\hbar}\frac{\mathrm{k_B}T_\mathrm{L}\{\Xi^v_\mathrm{adp}\}^2}{A\rho u_l^2}\frac{1}{b^{vv}_{n'n}}\delta[E^v_{n'}(\boldsymbol{k}') - E^v_n(\boldsymbol{k})] \ . \tag{5.26}$$

Here, n and n' denote the subband indices, v the valley index, and u_l is the longitudinal sound velocity. In (5.26) an overlap integral occurs,

$$\frac{1}{b^{v'v}_{n'n}} = \int\limits_0^\infty |\zeta^{v'}_{n'}(z)|^2 |\zeta^v_n(z)|^2 \mathrm{d}z \ , \tag{5.27}$$

where $\zeta^{v'}_{n'}(z)$ denote the one dimensional envelope wave functions being solutions of the Schrödinger equation (4.13). The parameters $b^{v'v}_{n'n}$ have the unit of length and are usually referred to as effective widths. It can be understood as the effective extent of the interaction region between electrons in different subbands and valleys in the z direction.

Using the nonparabolic density of states (5.25) the scattering rate can be written as

$$\{S_\mathrm{ac}\}^v_n(E) = \frac{2\pi}{\hbar}\frac{\mathrm{k_B}T\{\Xi^v_\mathrm{adp}\}^2}{\rho u_l^2}\sum_{n'}\frac{1}{b^{vv}_{n'n}} g^v_{n'}(E - E^v_{n'}) \ . \tag{5.28}$$

Intervalley Scattering

In the 2DEG intervalley scattering causes transitions to valleys belonging to either the same or to different subband ladders. Following Price's treatment for intravalley optical phonon scattering the transition rate for transitions between different valleys becomes

$$\{S_\eta^{\substack{\mathrm{abs}\\ \mathrm{emi}}}\}^{v'v}_{n'n}(\boldsymbol{k}',\boldsymbol{k}) = (1-\delta_{v'v})\frac{\pi(D_tK_\eta)^2_{v'v}}{A\rho\omega_\eta}\frac{1}{b^{v'v}_{n'n}}\left(N_\eta + \frac{1}{2} \mp \frac{1}{2}\right)\delta\left[E^{v'}_{n'}(\boldsymbol{k}') - E^v_n(\boldsymbol{k}) \mp \hbar\omega_\eta\right] \ , \tag{5.29}$$

where v and v' denote the valley index before and after the transition, respectively [Price81, Jungemann93]. A summation is performed over all final valleys v' which are determined by the selection rules for phonon mode η.

Using the effective mass approximation, the scattering rate for intervalley transitions from an electron in valley v and subband n can be written as (5.22)

$$\{S_\eta^{\overset{\text{abs}}{\text{emi}}}\}_n^v(E) = \sum_{n'}\sum_{v'\neq v} \frac{\hbar\{D_t K_\eta^{v'v}\}^2}{\rho\omega_\eta}\frac{1}{b_{n'n}^{v'v}}\left(N_\eta+\frac{1}{2}\mp\frac{1}{2}\right) g_{n'}^{v'}(E\pm\hbar\omega_\eta - E_{n'}^{v'}) \; . \tag{5.30}$$

Surface Roughness Scattering

The deviations of the interface from an ideal flat plane can be described by a two-dimensional roughness fluctuation, $\Delta(\boldsymbol{r})$, where $\boldsymbol{r}$ is the two-dimensional position vector in the plane of the interface [Ferry97]. The potential associated with the roughness $\Delta(\boldsymbol{r})$ can be viewed as a combination of two effects:

- The boundary perturbation causes the envelope functions to be displaced from their unperturbed positions.
- The imposed fluctuation of the electric field and the charge density at the rough interface give rise to an electrostatic contribution to the potential.

The original formulation of Prange and Nee [Prange68] of the unscreened matrix elements for surface roughness scattering has been adopted. It can be applied for scattering at two interfaces [Esseni04]

$$|\{M_{\text{sr}}^{\text{unscr}}\}_{n'n}^{v}(\boldsymbol{q})|^2 = \frac{C(\boldsymbol{q})}{A}\{F_{n'n}^{v}\}^2 \; , \quad \text{where} \quad F_{n'n}^{v} = \frac{\hbar^2}{2m_q^v}\frac{\mathrm{d}\zeta_{n'}^{v}(z_{\text{int}})}{\mathrm{d}z}\frac{\mathrm{d}\zeta_{n}^{v}(z_{\text{int}})}{\mathrm{d}z} \; . \tag{5.31}$$

Here, $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}'$ is the momentum transfer, m_q^v is the quantization mass of electrons in valley v, and $\mathrm{d}\zeta_n^v(z_{\text{int}})/\mathrm{d}z$ denotes the derivative of the envelope function with respect to z at the position of the interface (for instance, $z_{\text{int}} = 0$, and $z_{\text{int}} = T_{\text{si}}$ for the front and back-interface of a thin Si film). The spectral density $C(\boldsymbol{q})$ is the 2D Fourier transform of the autocovariance function

$$C(\boldsymbol{r}) = \langle\Delta(\boldsymbol{r}')\Delta(\boldsymbol{r}' - \boldsymbol{r})\rangle \; , \tag{5.32}$$

where the brackets denote the ensemble average of the roughness fluctuation $\Delta(\boldsymbol{r})$. The roughness spectrum is frequently assumed to be Gaussian [Jungemann93, Esseni03, Esseni04]

$$C(\boldsymbol{q}) = \pi\Delta_{\text{rms}}^2 L_{\text{c}}^2 \exp(-(qL_{\text{c}}/2)^2) \; , \tag{5.33}$$

or of exponential shape [Goodnick85, Ferry97]

$$C(\boldsymbol{q}) = \frac{\pi\Delta_{\text{rms}}^2 L_{\text{c}}^2}{(1+(q^2L_{\text{c}}^2/2))^{3/2}} \; . \tag{5.34}$$

Here, Δ_{rms} is the root mean square value of the roughness fluctuations and L_{c} is the autocovariance length.

The transition rate for surface roughness scattering is

$$S_{n'n}^{v}(\boldsymbol{k},\boldsymbol{k}') = \frac{2\pi}{\hbar}\frac{C(\boldsymbol{q})}{A}\{F_{n'n}^{v}\}^2\delta[E_{n'}^{v}(\boldsymbol{k}') - E_{n}^{v}(\boldsymbol{k})] \; , \tag{5.35}$$

where intersubband transitions due to surface roughness are restricted to the same valley [Esseni03, Cheng71]. In the nonparabolic band approximation the scattering rate for a Gaussian spectrum is given by

$$\begin{aligned}&\frac{A}{4\pi^2}\sum_{n'}\int \mathrm{d}^2k'\, S^v_{n'n}(\boldsymbol{k},\boldsymbol{k}') = \\ &\frac{\Delta^2_{\mathrm{rms}}L^2_{\mathrm{c}}}{2\hbar}\sum_{n'}\{F^v_{n'n}\}^2\int \mathrm{d}^2k'\exp(-(qL_{\mathrm{c}}/2)^2)\delta[E^v_{n'}(\boldsymbol{k}')-E^v_n(\boldsymbol{k})] = \\ &\frac{\pi\Delta^2_{\mathrm{rms}}L^2_{\mathrm{c}}}{\hbar}\sum_{n'}g^v_{n'}(E-E^v_{n'})\{F^v_{n'n}\}^2\int_0^{2\pi}\mathrm{d}\phi'\exp(-(qL_{\mathrm{c}}/2)^2). \qquad (5.36)\end{aligned}$$

Assuming isotropic bands $m^v = m^v_{\shortparallel,1} \approx m^v_{\shortparallel,2} \approx \sqrt{m_{\shortparallel,1}m_{\shortparallel,2}}$, the integral over the angle can be written as

$$\begin{aligned}\int_0^{2\pi}\mathrm{d}\phi'\exp(-(qL_{\mathrm{c}}/2)^2) &= \exp\left(\frac{-L^2_{\mathrm{c}}(k'^2+k^2)}{4}\right)\int_0^{2\pi}\exp\left(\frac{-L^2_{\mathrm{c}}(k'k)\cos\phi'}{2}\right)\mathrm{d}\phi' \\ &= \exp\left(\frac{-L^2_{\mathrm{c}}(k'^2+k^2)}{4}\right)2\pi\mathcal{I}_0(y)\ , \qquad (5.37)\end{aligned}$$

where $y = -L^2_{\mathrm{c}}k'k/2$, and $\mathcal{I}_0$ denotes the modified Bessel function of the first kind.

Since the electrons in the inversion layer screen the scattering potential, the transition rate for surface roughness scattering is reduced. The dielectric function relates the unscreened and screened matrix elements of the scattering potential through the dielectric function ϵ_{D}

$$\{M^{\mathrm{scr}}_{\mathrm{sr}}\}^v_{n'n}(\boldsymbol{q}) = \sum_w\sum_{m',m}\frac{\{M^{\mathrm{unscr}}_{\mathrm{sr}}\}^w_{n'n}(\boldsymbol{q})}{\epsilon_{\mathrm{D}}{}^{vw}_{(n'n)(m'm)}(\boldsymbol{q})}\ . \qquad (5.38)$$

Because surface roughness represents a static potential the dependence on the frequency ω can be dropped. Since the number of relevant subbands can be of order 100 [Jungemann93], further simplifications are required to numerically evaluate the impact of screening.

In the long-wavelength limit, $\boldsymbol{q} \to 0$, intersubband transitions are completely unscreened [Ferry97], thus $\{M^{\mathrm{scr}}_{\mathrm{sr}}\}^v_{n'n}(\boldsymbol{q}) = \{M^{\mathrm{unscr}}_{\mathrm{sr}}\}^v_{n'n}(\boldsymbol{q})$. Furthermore, the multisubband dielectric function reduces to a scalar function when neglecting the intersubband polarizabilities and the correction terms due to the intrasubband polarizabilities of the other subbands [Ferry97]. This approximation is frequently applied for transport simulations [Esseni03, Esseni04]. The scalar dielectric function for intrasubband transitions can be given in terms of the polarization function $L_n(\boldsymbol{q})$ and the form factor $F_n(\boldsymbol{q})$

$$\epsilon_{\mathrm{D}}(\boldsymbol{q}) = 1 + \sum_v\sum_n\frac{e^2}{2\kappa_{\mathrm{si}}q}F^v_n(\boldsymbol{q})L^v_n(\boldsymbol{q})\ , \qquad (5.39)$$

where κ_{si} is the dielectric constant of Si [Ferry97]. The polarization function can be expressed in terms of the Fermi-Dirac distribution function f_{FD}

$$L^v_n(\boldsymbol{q}) = 2\sum_{\boldsymbol{k}}\frac{f_{\mathrm{FD}}[E^v_n(\boldsymbol{k})] - f_{\mathrm{FD}}[E^v_n(\boldsymbol{k}+\boldsymbol{q})]}{E^v_n(\boldsymbol{k}+\boldsymbol{q}) - E^v_n(\boldsymbol{k})}\ . \qquad (5.40)$$

Table 5.2: Parameters for scattering in the 2DEG for {001} and {110} substrate orientation. For intervalley scattering the bulk parameters of Table 5.1 are used.

	{001}	{110}	Units
$\Xi^{\Delta}_{\mathrm{adp}}$	14.8	13.0	eV
L_{c}	1.3	1.5	nm
Δ_{rms}	0.4	0.55	nm

The form factor can be calculated from

$$F^v_n(\boldsymbol{q}) = \int dz \int dz' |\zeta^v_n(z')|^2 |\zeta^v_n(z)|^2 G(z,z',q) \ . \tag{5.41}$$

Here, $G(z,z',q)$ denotes the Green's function. For a semi-infinite Si layer the Green's function evaluates to [Ando82]

$$G(z,z',q) = \frac{1}{2}\left[\left(1-\frac{\kappa_{\mathrm{si}}}{\kappa_{\mathrm{sio_2}}}\right)e^{-q|z-z'|} + \left(1+\frac{\kappa_{\mathrm{si}}}{\kappa_{\mathrm{sio_2}}}\right)e^{q|z-z'|}\right] \ . \tag{5.42}$$

For a Si layer sandwiched between two semi-infinite SiO$_2$ films (from $z=0$ to $z=T_{\mathrm{si}}$) it is given by [Fischetti03]

$$G(z,z',q) = \frac{1}{(1-\tilde{\kappa}^2 e^{-2qT_{\mathrm{si}}})}\left[e^{-q|z-z'|} + \tilde{\kappa}e^{-q|z+z'|} + \tilde{\kappa}e^{-2qT_{\mathrm{si}}}\left(e^{q|z-z'|} + \tilde{\kappa}e^{q|z+z'|}\right)\right] \ , \tag{5.43}$$

where $\tilde{\kappa} = (\kappa_{\mathrm{si}} - \kappa_{\mathrm{sio_2}})/(\kappa_{\mathrm{si}} + \kappa_{\mathrm{sio_2}})$.

Both the form factors and the polarization function are evaluated numerically from the wave functions and are used to calculate the screened surface roughness scattering rate.

5.3.3 Coupling to the Schrödinger Poisson Solver

In this work the Schrödinger-Poisson solver SCHRED 2.0 [Vasileska00] was extended to allow general strain conditions and arbitrary substrate orientation. The dependence of the subband spectrum on the Si substrate orientation has been implemented by taking the proper masses $m_{\shortparallel,1}, m_{\shortparallel,2}$, and $m_\perp$ according to Table 4.1. Recently, a coupling with VSP [Karner06] has been established.

The impact of strain is implemented by taking into account the strain-induced effective mass change using the analytic expressions (3.94),(3.98), and (3.99). Furthermore, the energy shifts of the subbands induced by strain are superimposed to the subband energies during the self-consistent solution of the Schrödinger-Poisson system.

The tool flow for a transport simulation in Si inversion layers is depicted in Figure 5.3. First, values for the input parameters of the Schrödinger-Poisson solver, such as N_{dop}, V_g or T_{si}, have to be specified in an according input file. These values, together with the strain and the substrate orientation are passed to the Schrödinger-Poisson solver. In a postprocessing step the eigenenergies and wave functions of the subbands and the Fermi level are used to calculate the

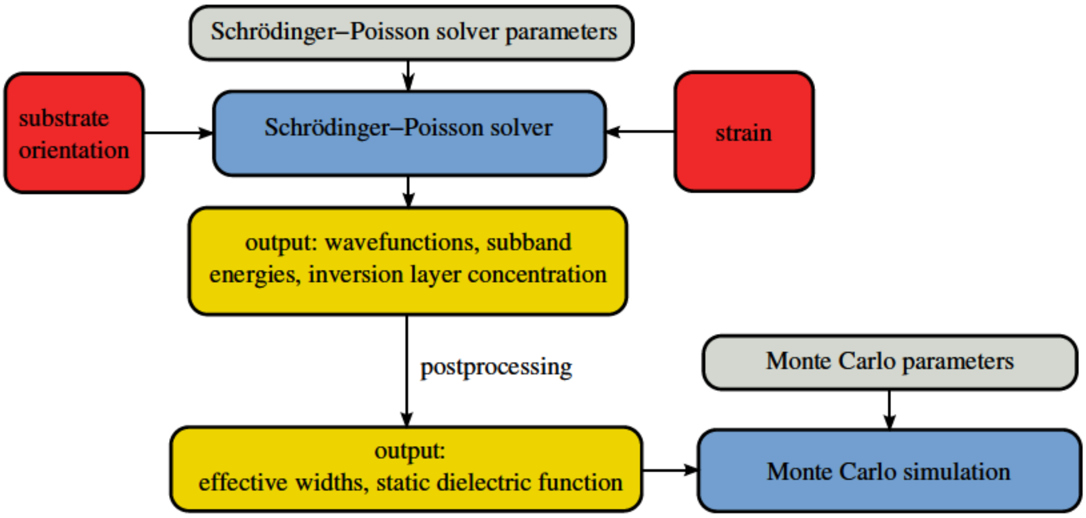


Figure 5.3: Flow chart for coupled Schrödinger-Poisson solver and Monte Carlo simulation.

- effective widths (5.27) which enter the phonon scattering rates,
- surface roughness scattering matrix elements (5.31),
- polarization function (5.40) and form factors (5.41) entering the static dielectric function.

Hereafter, the Monte Carlo parameters have to be initialized and are passed to the MC simulator together with the postprocessed output of the Schrödinger-Poisson solver. If the quantity of interest is the effective electron mobility, a special Monte Carlo algorithm can be used [Kosina00a, Smirnov03]. It was originally designed for small signal analysis of the three dimensional electron gas, and in this work generalized for the simulation of the zero field mobility in inversion layers [Ungersboeck06b, Sverdlov06].

5.4 MC Algorithm Including Degeneracy Effects

In transport calculations of the 2DEG forming in the channel of a MOSFET the inclusion of the Pauli principle is expected to be important since the lowest subband may lie well below the Fermi level even in the regime of moderate and high effective fields (high inversion layer concentrations). This leads to modified subband populations and an elevated mean kinetic energy of electrons as compared to the nondegenerate case. A change in the mobility is therefore to be expected.

Surprisingly, there is a discordance in the literature whether and how degeneracy effects should be included in transport calculations of inversion layers. Therefore, the ways to include the Pauli principle in a MC algorithm are revised and critically compared to each other. The usual method, where the Pauli blocking factor $[1 - f(\boldsymbol{k})]$ is approximated using the equilibrium distribution function $f_{\mathrm{FD}}(\boldsymbol{k})$, can be shown to lead to unphysical subband populations, kinetic energies, and mobilities. The reason being that at high degeneracy the error $\delta(\boldsymbol{k}) = f(\boldsymbol{k}) - f_{\mathrm{FD}}(\boldsymbol{k})$ is dominant.

A new MC algorithm accounting for the Pauli exclusion principle is proposed which is less sensitive to the error $\delta(\boldsymbol{k})$. The proposed algorithm is based on the following reformulation of the collision operator

$$\begin{aligned}\mathcal{Q}[f]_{\boldsymbol{k}} &= \int f(\boldsymbol{k}')[1-f(\boldsymbol{k})]S(\boldsymbol{k},\boldsymbol{k}')\,\mathrm{d}\boldsymbol{k}' - \int f(\boldsymbol{k})[1-f(\boldsymbol{k}')]S(\boldsymbol{k}',\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}' \\ &= \int f(\boldsymbol{k}')S(\boldsymbol{k},\boldsymbol{k}') - f(\boldsymbol{k})S(\boldsymbol{k}',\boldsymbol{k}) + \underbrace{f(\boldsymbol{k})f(\boldsymbol{k}')[S(\boldsymbol{k}',\boldsymbol{k}) - S(\boldsymbol{k},\boldsymbol{k}')]}_{\text{additional term}}\,\mathrm{d}\boldsymbol{k}'\ , \qquad (5.44)\end{aligned}$$

where the band index n and the time dependence of the distribution are dropped. The last term represents a nonlinear correction to the nondegenerate collision operator. To linearize the scattering operator it is common to keep one factor of the product $f(\boldsymbol{k})f(\boldsymbol{k}')$ constant and to treat the other as the unknown.

Near thermodynamic equilibrium, the distribution function f can be approximated by the Fermi-Dirac distribution function f_{FD}. The key point of the new method is that a symmetric approximation with respect to $\boldsymbol{k}$ and $\boldsymbol{k}'$ is employed

$$f(\boldsymbol{k})f(\boldsymbol{k}') \approx \frac{1}{2}\big(f(\boldsymbol{k})f_{FD}(\boldsymbol{k}') + f_{FD}(\boldsymbol{k})f(\boldsymbol{k}')\big)\ . \qquad (5.45)$$

Using this approximation the scattering operator can be expressed in terms of a modified transition rate $\widehat{S}(\boldsymbol{k},\boldsymbol{k}')$ and scattering rate $\widehat{\lambda}_{\boldsymbol{k}}$ as

$$\mathcal{Q}[f]_{\boldsymbol{k}} = \int f(\boldsymbol{k}')\widehat{S}(\boldsymbol{k},\boldsymbol{k}')\,\mathrm{d}\boldsymbol{k}' - f(\boldsymbol{k})\widehat{\lambda}_{\boldsymbol{k}}\ .$$

with

$$\widehat{S}(\boldsymbol{k}',\boldsymbol{k}) = S(\boldsymbol{k}',\boldsymbol{k})\left(1-\frac{1}{2}f_{\mathrm{FD}}(\boldsymbol{k}')\right) + S(\boldsymbol{k},\boldsymbol{k}')\frac{1}{2}f_{\mathrm{FD}}(\boldsymbol{k}')\ , \quad \text{and} \qquad (5.46)$$

$$\widehat{\lambda}_{\boldsymbol{k}} = \int \widehat{S}(\boldsymbol{k}',\boldsymbol{k})\,\mathrm{d}\boldsymbol{k}'\ . \qquad (5.47)$$

A simple error analysis shows the advantage of this formulation. Consider a highly degenerate state $\boldsymbol{k}$, characterized by $f(\boldsymbol{k}) \approx 1$. A direct approximation of the blocking factor $[1-f(\boldsymbol{k})]$ can give completely wrong results, because the approximation of the blocking factor is determined by the error, $1-(f_{FD}+\delta) \approx \delta$. In the formulation (5.46), however, because of $\delta \ll 1$ the effect of the error will be negligible, $1-(f_{FD}+\delta)/2 \approx 1/2$.

The modified transition rate (5.46) is given by a linear combination of the forward rate $S(\boldsymbol{k},\boldsymbol{k}')$ and backward rate $S(\boldsymbol{k}',\boldsymbol{k})$. The latter can be expressed in terms of the forward rate by means of the principle of detailed balance [Ashcroft76]. The modified scattering rates for phonon emission and absorption become,

$$\widehat{\lambda}_{\mathrm{em}} = \lambda_{\mathrm{em}}\left(1-\frac{1}{2}\frac{f_{\mathrm{FD}}(E_{\mathrm{f}})}{N_0+1}\right)\ , \qquad (5.48)$$

$$\widehat{\lambda}_{\mathrm{ab}} = \lambda_{\mathrm{ab}}\left(1+\frac{1}{2}\frac{f_{\mathrm{FD}}(E_{\mathrm{f}})}{N_0}\right)\ , \qquad (5.49)$$

where E_f denotes the final energy and N_0 the equilibrium phonon distribution function,

$$N_0 = \frac{1}{\exp\left(\frac{\hbar\omega_0}{\mathrm{k_B}T}\right) - 1} \ .$$

To implement the Pauli principle in a conventional MC program for nondegenerate statistics the only modifications necessary are the replacement of the classical scattering rates by the modified ones.

For elastic scattering mechanisms the modified scattering rates do not change from the classical ones, $\widehat{\lambda}_{\boldsymbol{k}} = \lambda_{\boldsymbol{k}}$. For the simulation of the 2DEG one can assume scattering with surface roughness, impurities, and acoustic phonons to be elastic.

In Section 6.3.2 simulation results using the new MC method including the Pauli exclusion principle are discussed. It is shown that in the low field limit the proposed algorithm yields the same mobility as the Kubo-Greenwood formula, while other algorithms do not. We use the new method to extract velocity profiles and to illustrate the large effect of degeneracy on the electron system.

Chapter 6

Simulation Results

In this chapter the results of band structure calculations and of electron mobility simulations for Si under general strain conditions are presented. Special emphasis is put on shear strain arising from stress along $\langle 110 \rangle$ because of its high relevance from a technological point of view.

The chapter is organized as follows: In Section 6.1 results from band structure calculations using the empirical pseudopotential method are presented. Especially the impact of strain on the Δ-valleys of the lowest conduction band is analyzed as the electron mobility is mainly determined by electrons residing in these valleys. Band structure calculations are compared to the analytical expressions derived using a degenerate $\mathbf{k}\cdot\mathbf{p}$-theory at the zone boundary X point. The effective mass change of electrons predicted by the calculations is verified by comparison with experimental data. Next, the effect of strain on the subband structure of Si inversion layers as obtained from the Schrödinger-Poisson solver is examined.

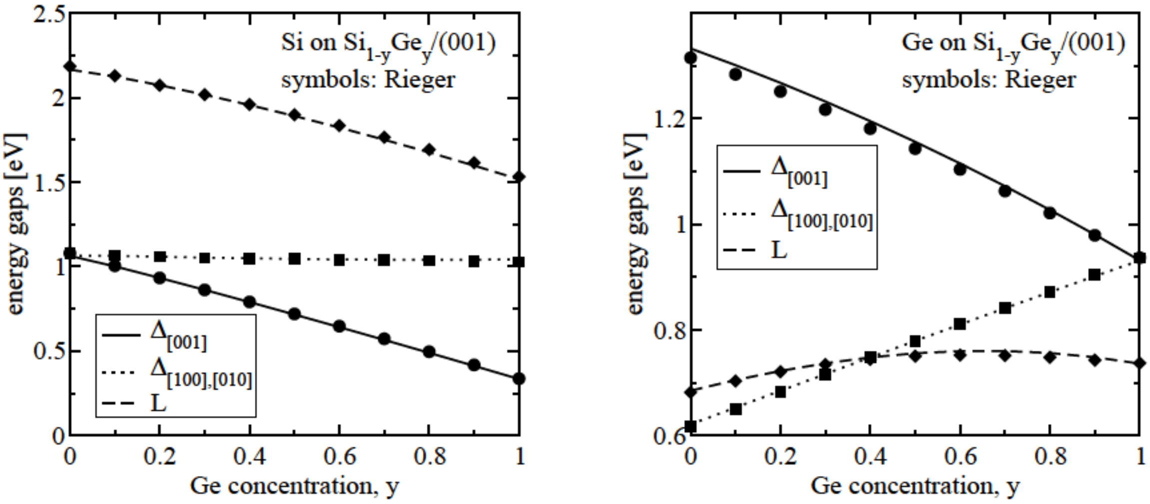
In Section 6.2 the Si bulk electron mobility is analyzed for various strain conditions. The impact of the strain-induced valley shifts and the strain-induced effective mass change on the mobility is evaluated. Section 6.3 contains the simulation results of the effective mobility in Si inversion layers. Again, focus is put on substrate orientations and strain configurations that are of technological relevance. Finally, the influence of degeneracy effects on the inversion layer mobility is analyzed and simulation results are compared to experimental data.

6.1 Bandstructure Calculations

The parameters relevant for the modeling of strain effects on the conduction bands of Si were extracted from EPM calculations and are given in Table 6.1. The value for the uniaxial deformation potential Ξ_u^Δ coincides with the value extracted by Rieger [Rieger93]. The value for the shear deformation potential $\Xi_{u'} = 7.0$ eV compares well with the values extracted from experimental data: a value of 5.7±1 eV [Hensel65] has been predicted from cyclotron resonance experiments. From more recent measurements of the indirect exciton spectrum of Si the value 7.5±2 eV has been obtained [Laude71]. By calculations based on a self-consistent perturbation theory [Goroff63] the value 7.8 eV was obtained for the shear deformation potential. The effective masses and the band separation $\Delta = 0.53$ eV between the two lowest conduction bands at

Table 6.1: Bandstructure parameters extracted from numerical band structure calculations.

parameter	EPM value	Units
Ξ_u^Δ	9.29	eV
$\Xi_{u'}$	7.0	eV
Δ	0.53	eV
m_l	0.914	m_0
m_t	0.196	m_0
κ	52.83	1
η	-0.809	1



(a)

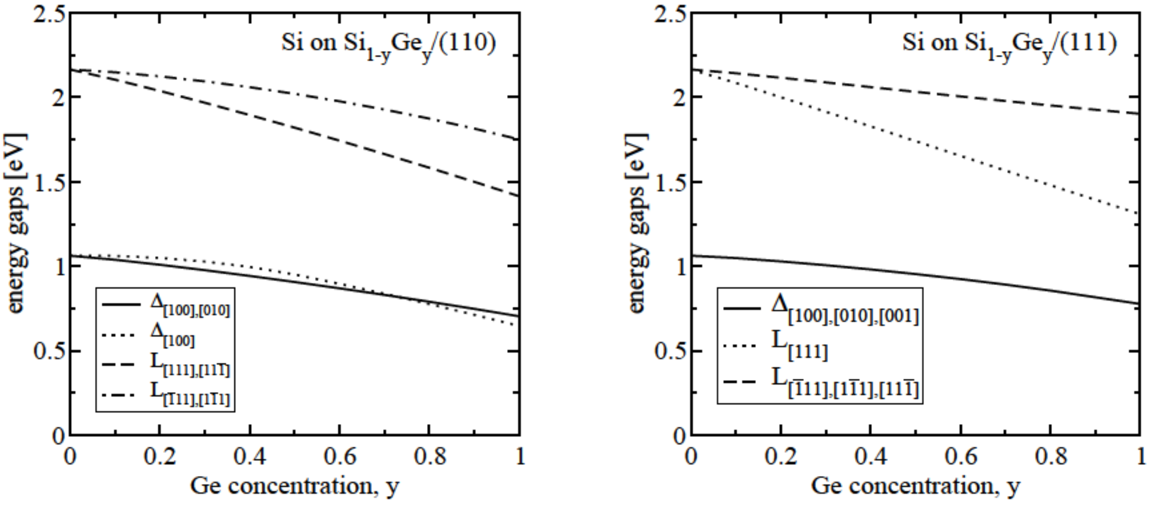
(b)

Figure 6.1: Calculated band edge energies of the Δ-valleys and L-valleys of (a) strained Si and (b) strained Ge on $\text{Si}_{1-y}\text{Ge}_y$ buffer with respect to valence band edge. Calculations (lines) are compared to results obtained by Rieger [Rieger93] (symbols).

the conduction band minimum are in good agreement to experimental values [Hensel65]. Finally, the two parameters defined in (3.89), $\kappa = 4\Xi_{u'}/\Delta = 52.83$ and $\eta = -0.809$ were extracted from EPM calculations. These two parameters are used in the analytical expressions for the effective mass change (3.94), (3.98), and (3.99), and the valley shift (3.100) induced by shear strain.

6.1.1 Strain-Induced Shift of the Conduction Band Minimum

In Figure 6.1 the calculated band edge energies of biaxially strained Si and Ge are compared to those of Rieger and Vogl [Rieger93]. Good agreement can be observed for both strained Si for (001)-oriented $\text{Si}_{1-y}\text{Ge}_y$ buffers and strained Ge for (001)-oriented $\text{Si}_{1-y}\text{Ge}_y$ buffers for the



(a)

(b)

Figure 6.2: Calculated band edge energies of the Δ-valleys and L-valleys of strained Si on (a) (110)-oriented and (b) (111)-oriented Si$_{1-y}$Ge$_y$ relaxed buffer with respect to valence band edge.

whole range of substrate mole-fractions y.

In Figure 6.2 the band edge energies of the conduction band valleys of Si grown on Si$_{1-y}$Ge$_y$ buffers with orientation (110) and (111) are presented. For (111)-oriented buffers the Δ-valleys do not split in agreement to the theoretical model (3.47). For (110)-oriented buffer the splitting between the Δ-valleys of strained Si is very small. From a deformation potential theory which neglects the lifting of the degeneracy of the two lowest conduction bands at the X points the splitting cannot be reproduced. Using the deformation potential theory the valley shift between the $\Delta_{[001]}$-valley pair and the $\Delta_{[010]}$-valley pair increases linearly with strain [Singh93]

$$\delta E_0^{\Delta[001]} - \delta E_0^{\Delta[100]} = \Xi_u^{\Delta} \varepsilon_{||} \tilde{c} \;, \tag{6.1}$$

where $\varepsilon_{||} = (a_{\mathrm{sige}} - a_{\mathrm{si}})/a_{\mathrm{si}}$, $\tilde{c} = (c_{11} + 2c_{12})/(c_{11} + c_{12} + 2c_{44})$, and a_{sige} is the lattice constant of the Si$_{1-y}$Ge$_y$ buffer.

When taking into account the additional valley shift of the $\Delta_{[001]}$-valley due to shear strain, the total valley shift is obtained from

$$\delta E_0^{\Delta[001]} + \delta E_1^{\Delta[001]} - \delta E_0^{\Delta[100]} = \Xi_u^{\Delta} \varepsilon_{||} \tilde{c} \left(1 - \varepsilon_{||} \tilde{c} \frac{\Delta}{4} \kappa^2\right) \;. \tag{6.2}$$

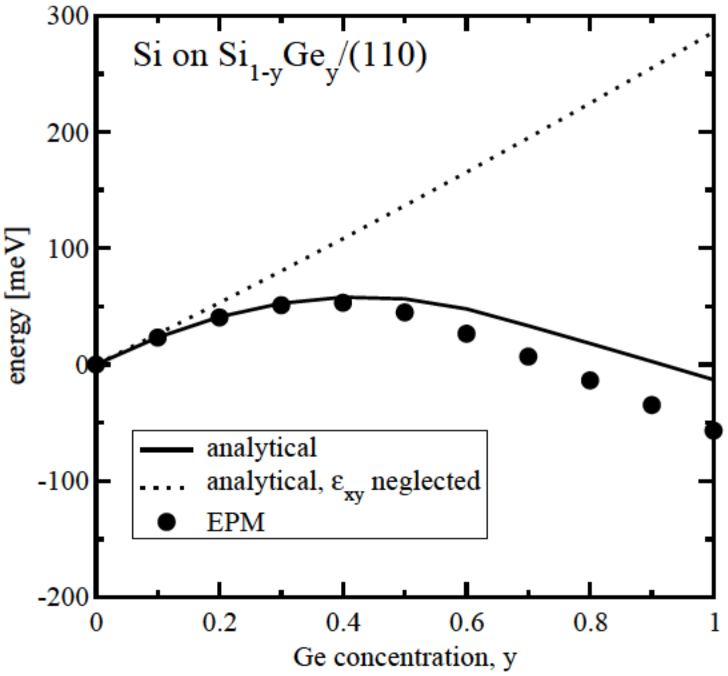


Figure 6.3: Valley splitting of the $\Delta_{[001]}$ and $\Delta_{[100]}$ valleys of biaxially strained Si on a (110) $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ buffer.

Figure 6.3 shows that the model (6.2) agrees much better with the results of EPM calculations.

The shifts of the band edge of the Δ-valleys and the L-valleys are calculated for uniaxial stress along four different directions. Figure 6.4 shows that the valley shifts are linear for stress up to 2 GPa, and that the largest splitting among the Δ-valleys is obtained for the stress direction $\langle 100 \rangle$. Since the conduction band edges are given with respect to edge of the top valence band, which is either the heavy-hole or the light-hole band depending on the sign of stress, the slope of the valley splitting at 0 GPa changes.

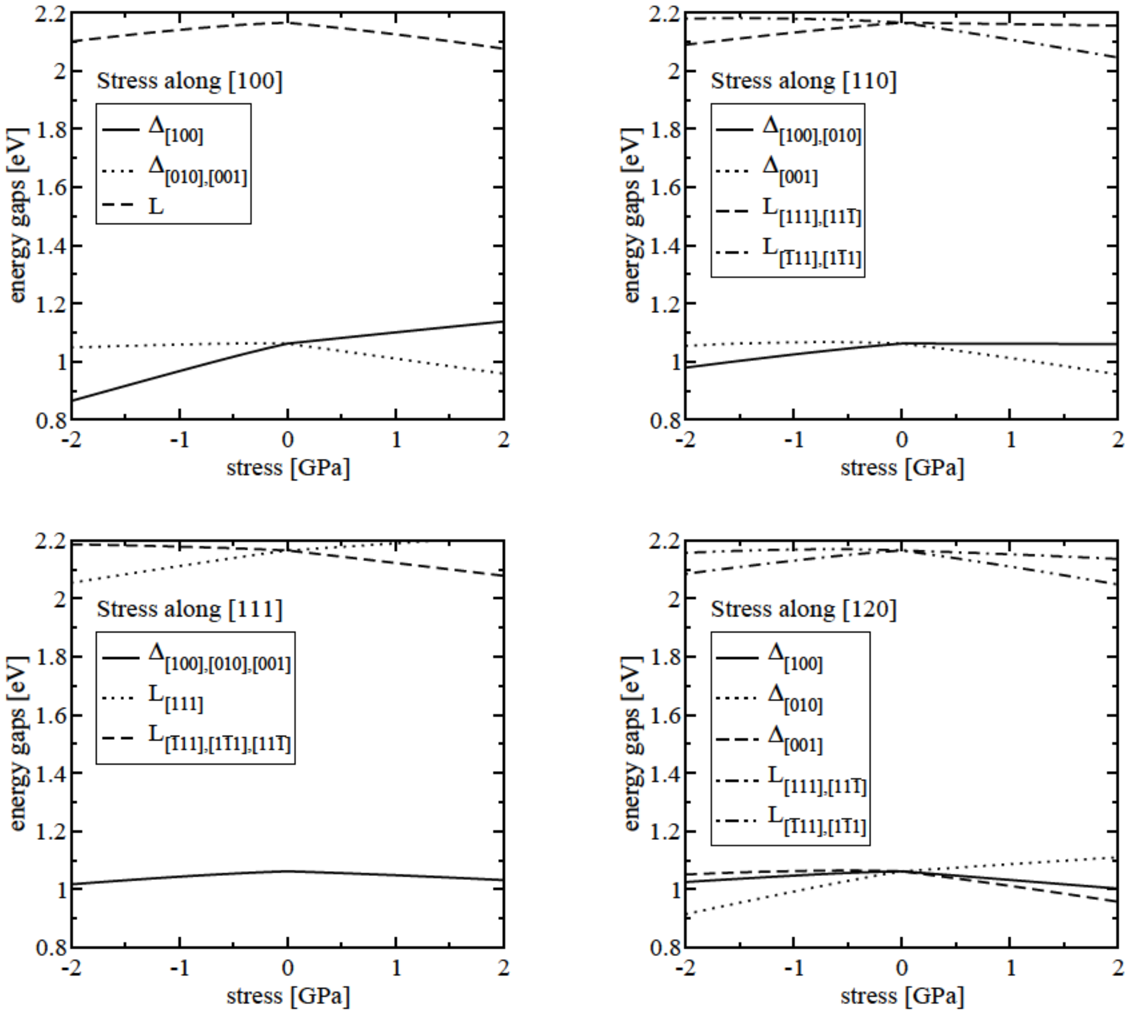


Figure 6.4: Calculated energies of the Δ-valleys and L-valleys of Si with respect to the valence band edge as a function of uniaxial stress for various stress directions.

6.1.2 Strain-Induced Change of the Shape of the Conduction Band Edge

In Figure 6.5 the energy dispersion of the two lowest conduction bands along the Δ symmetry lines are given for 2 GPa uniaxial stress along four stress directions. The different splitting can be interpreted from the strain tensor resulting from stress as given in (3.18):

- Since the off-diagonal elements of the strain tensor vanish if stress is applied along [100], the degeneracy of the conduction bands at the X-points is not lifted. Stress shifts the $\Delta_{[001]}$-valleys up in energy with respect to the other four valleys.
- For stress direction [110] the non-zero ε_{xy} component lifts the degeneracy of the conduction bands at the [00±1] X-points. Thus, a noticeable effect of strain on the effective masses and the mobility is expected.
- Stress along [111] is not expected to give rise to electron mobility enhancement. The non-zero off-diagonal components equally lift the degeneracy of the conduction bands in all six X-points, however, the valley degeneracy remains.
- For [120] stress all three valley-pairs are shifted with respect to each other. Additionally, the $\Delta_{[001]}$-valley is affected by the splitting of the two conduction bands along [001] due to the non-zero ε_{xy} component.

In Figure 6.6 the lifting of the degeneracy of the two lowest conduction bands is analyzed as a function of ε_{xy}. The splitting is linear with ε_{xy} and can be approximated using the analytic expression $\delta E = 4\Xi_{u'}\varepsilon_{xy}$ using 7.0 eV for the shear deformation potential.

The change of position of the minima of the conduction band edge and the Δ-valley shifts resulting from shear strain ε_{xy} are plotted in Figure 6.7, where results from EPM calculations are compared to the analytical expressions (3.92) and (3.100).

The shear strain-induced effective mass change of the transverse and longitudinal mass characterizing the $\Delta_{[001]}$-valley pair of Si is plotted in Figure 6.8. Again, results from the EPM are compared to the analytical expressions, (3.94), (3.98), and (3.99). Good agreement can be observed for $\varepsilon_{xy} < 1.5\%$. For larger values of shear strain, the change of the effective masses as obtained from $\boldsymbol{k}\cdot\boldsymbol{p}$ theory is smaller than that from the empirical pseudopotential method.

The calculated change of the effective mass induced by shear strain has been compared to values extracted from cyclotron resonance measurements. Good agreement is achieved as can be seen in Figure 6.9.

Finally, in Figure 6.10 the constant-energy lines in the plane $k_z = k_{\min}$ are shown. For increasing ε_{xy} the evolving ellipsoid is characterized by two different transverse masses given in (3.98) and (3.99). The principal axes of the ellipses are [110] and [1$\bar{1}$0].

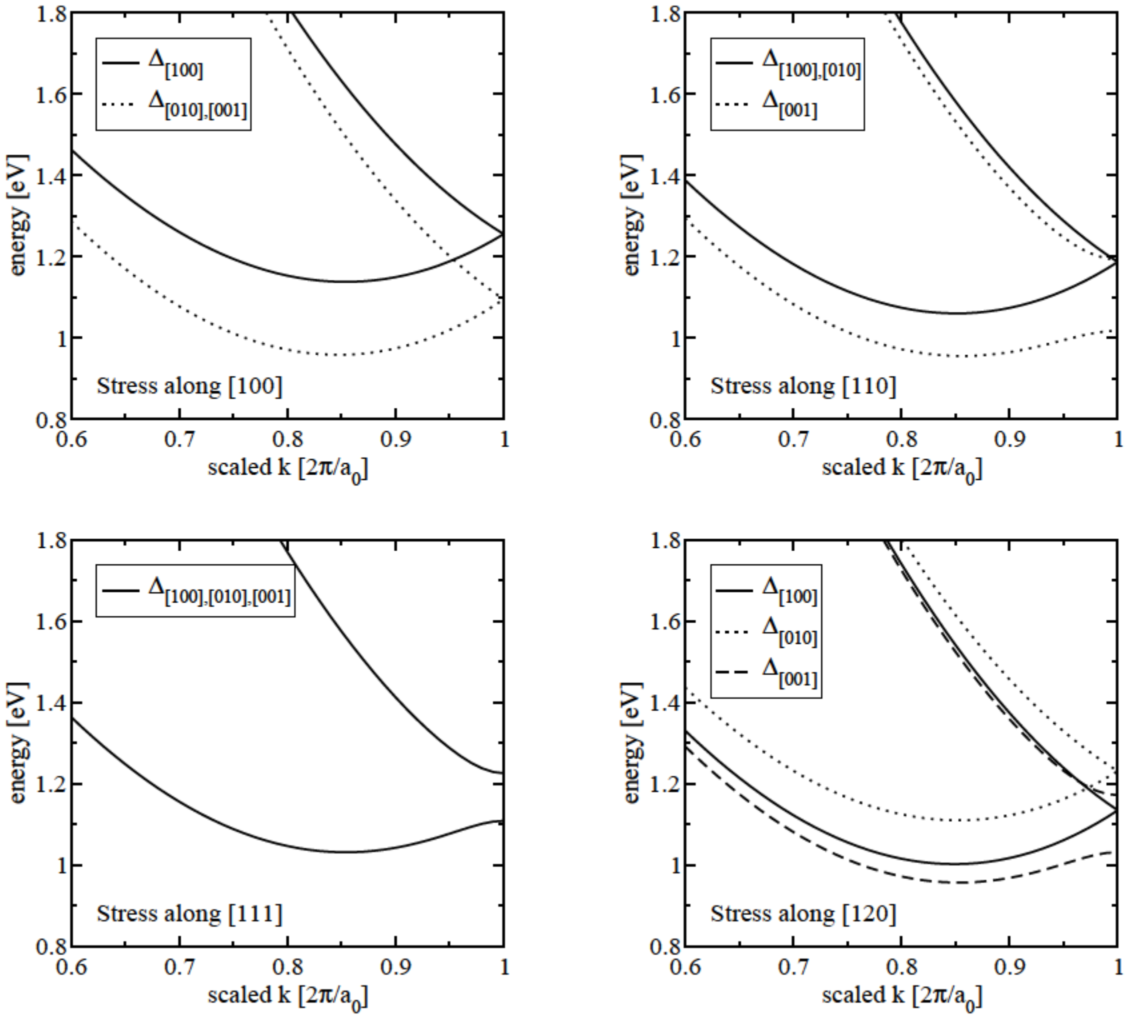
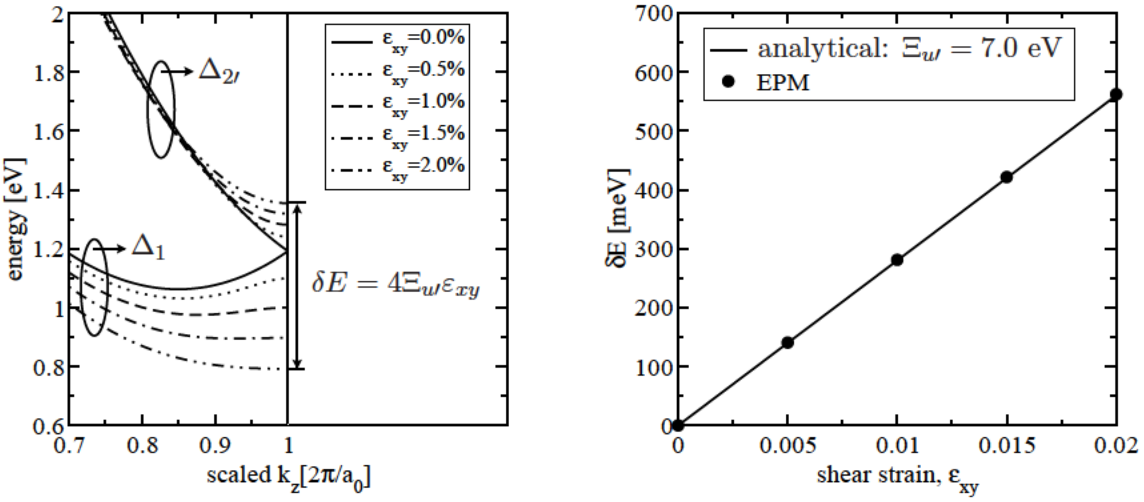
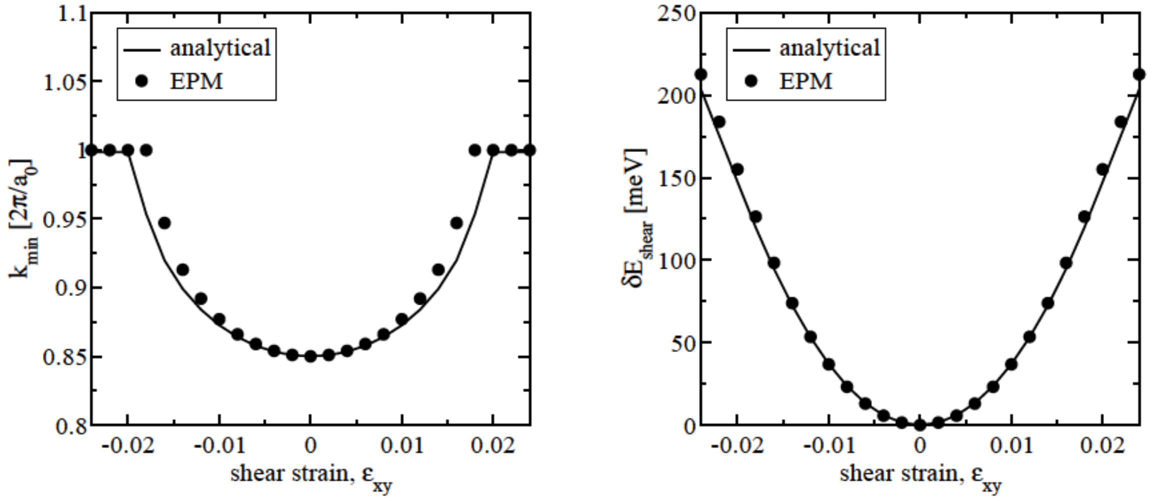


Figure 6.5: The two lowest conduction bands along the Δ-symmetry lines with respect to valence band edge at 2 GPa uniaxial stress for various stress directions.



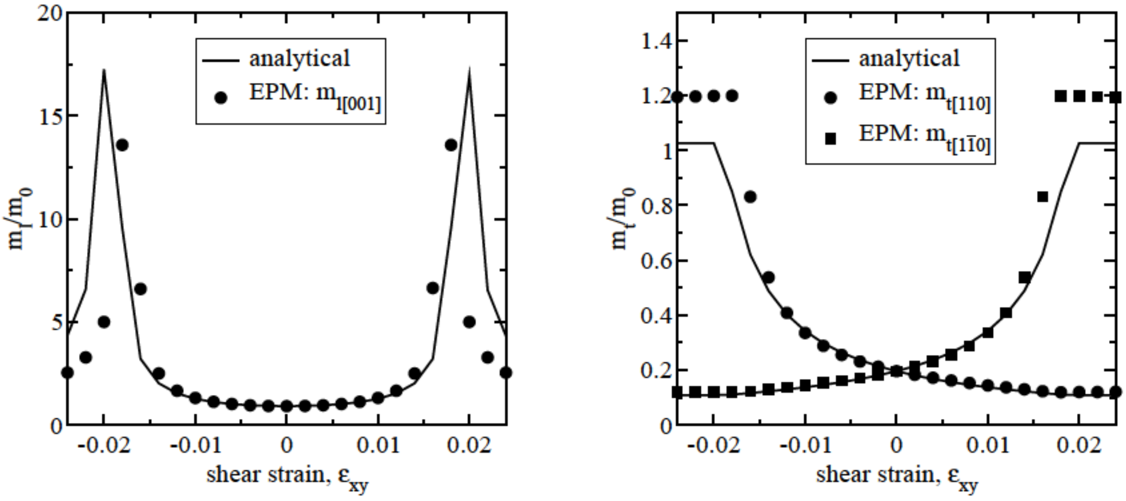
(a) (b)

Figure 6.6: (a) Calculated energy dispersion of the Δ_1 and $\Delta_{2'}$ conduction bands of Si along the [001] direction in the vicinity of the zone boundary at various levels of shear strain ε_{xy}. (b) Calculated energetic split between the two lowest conduction bands at zone boundary from EPM calculations (symbols) and from analytic expression (3.53) using 7.0 eV for the deformation potential $\Xi_{u'}$ (line).



(a) (b)

Figure 6.7: (a) Position of the conduction band minimum as a function of shear strain ε_{xy} evaluated numerically by EPM calculations (symbols) and analytically (lines) from (3.92). (b) Shift of the minima of the $\Delta_{[100]}$ and $\Delta_{[010]}$-valley pairs with respect to the $\Delta_{[001]}$-valleys induced by shear strain ε_{xy}. Comparison with analytical result from (3.100).



(a) (b)

Figure 6.8: Comparison of EPM calculations for the effective masses (a) along the directions [001], and (b) along [110] and [$\bar{1}$10] with analytical expressions (3.94),(3.98), and (3.99).

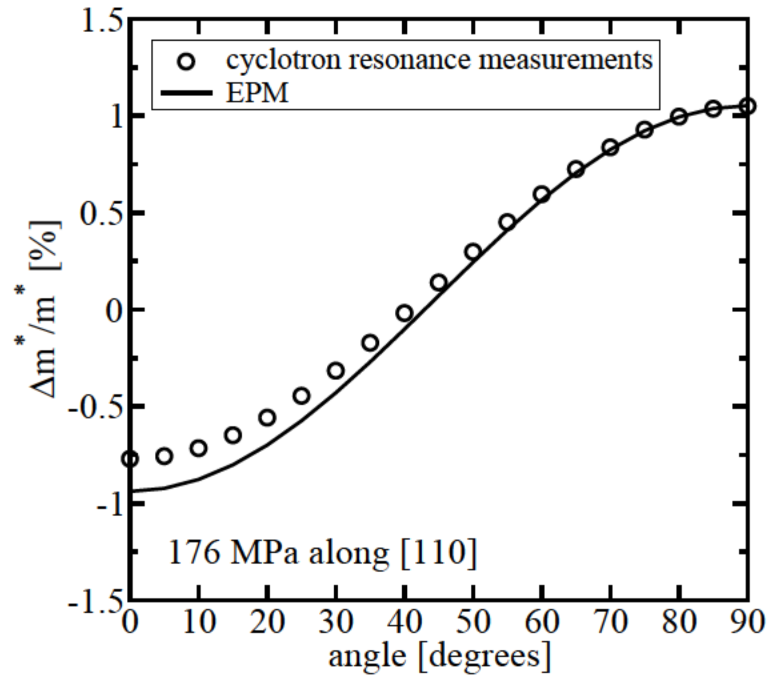
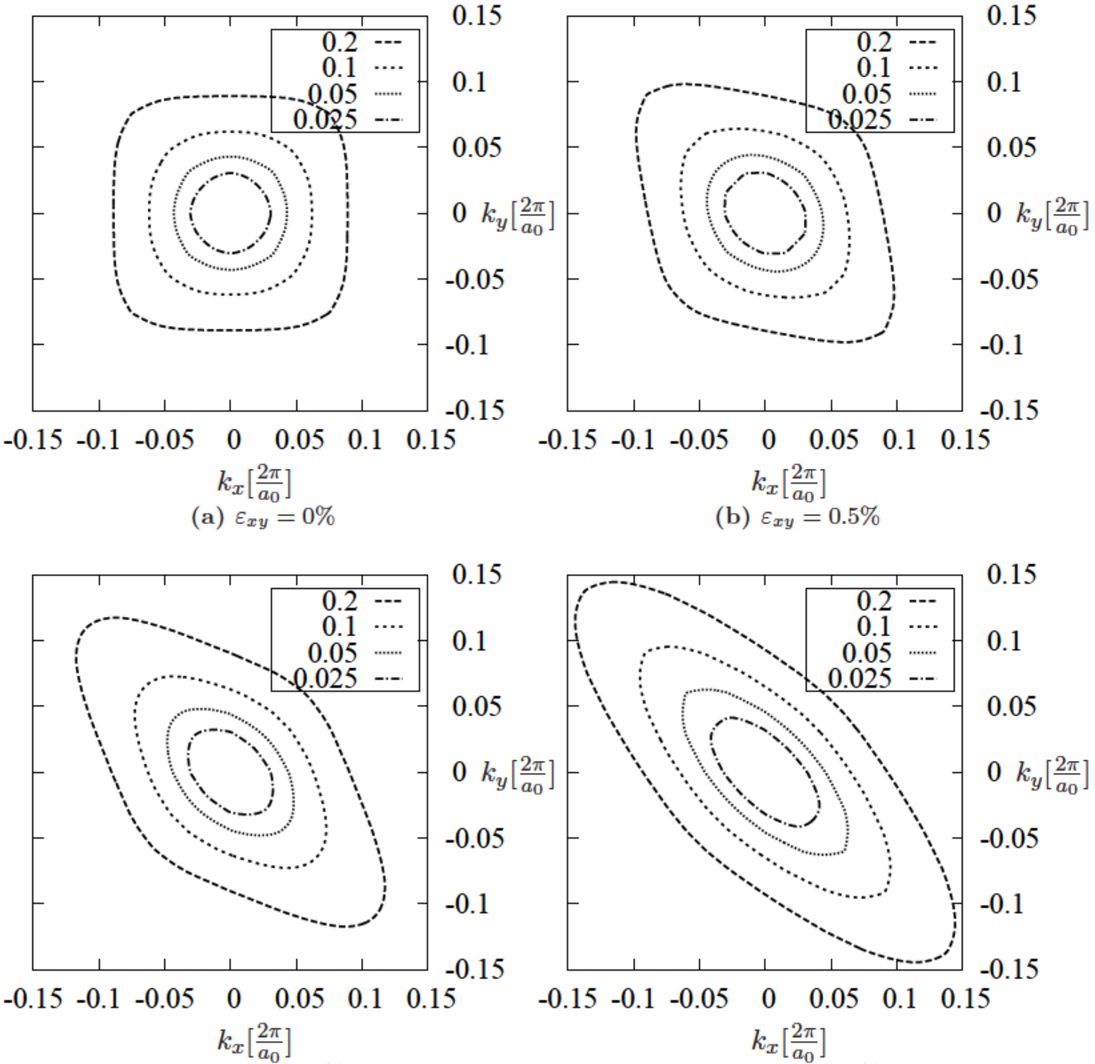


Figure 6.9: Calculated anisotropy of the cyclotron resonance effective mass, $m^* = \sqrt{m_t m_l}$, of the (001) ellipsoid for 176 MPa tensile stress along [110] compared to measurements [Hensel65]. The effective mass change in the (001) plane is consistent with experimental data.



(a) $\varepsilon_{xy} = 0\%$ (b) $\varepsilon_{xy} = 0.5\%$

(c) $\varepsilon_{xy} = 1.0\%$ (d) $\varepsilon_{xy} = 1.5\%$

Figure 6.10: Constant-energy lines (units of eV) of the lowest conduction band valley along [001] in the $k_z = k_{\text{min}}$ plane at four levels of shear strain.

6.1.3 Subband Structure

Figure 6.11 shows the three lowest energy levels, $E_n^v(\boldsymbol{k}=0)$, and the populations of the subband ladders of Si inversion layers for substrate orientation (001), (110), and (111) as a function of the inversion layer concentration at a substrate doping concentration of 1×10^{16} cm^{-3}. The energy levels are plotted relative to the lowest subband energy level. Two subband ladders are formed in (001)- and (110)-oriented substrate, whereas in (111)-oriented substrate the six Δ-valleys form only one subband ladder. It can be seen that the subband splitting increases with the inversion layer concentration. Furthermore, the Fermi level crosses the lowest subband level. This happens at an inversion layer concentration of 5×10^{12} cm^{-2} for (001) oriented substrate and above 1×10^{13} cm^{-2} for (110) and (111)-oriented substrate. Therefore, the 2DEG is degenerate and the Pauli exclusion principle has to be included in simulations.

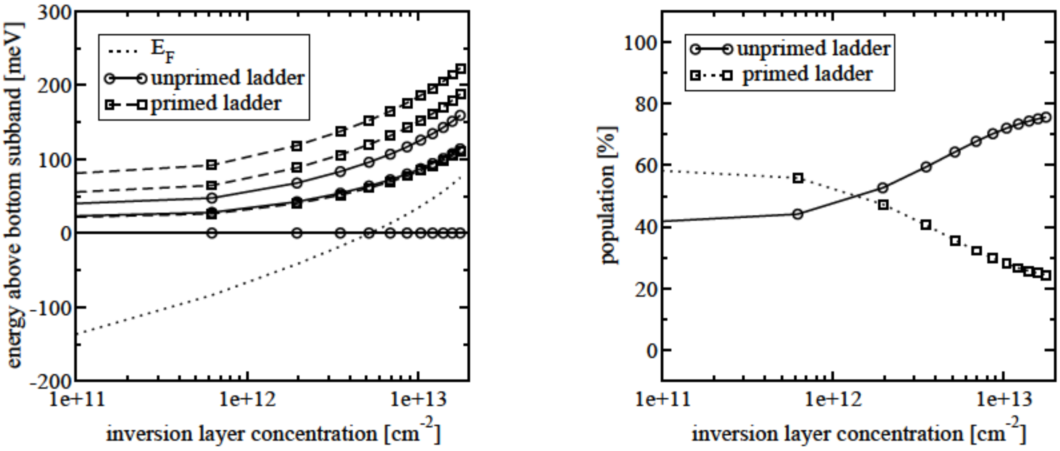
The subband populations for the three substrate orientations are compared in Figure 6.11b. The population of the primed ladder is the largest for (001)-oriented substrate. Keeping in mind that for this substrate orientation the transport masses of the primed ladder are larger than for the unprimed ladder, one can anticipate the large potential of strain to reduce the primed ladder population.

In Figure 6.12 the three lowest subband levels and subband populations for a substrate with (001)-orientation are given at three stress configurations. Figure 6.12a and 6.12b show that a tensile stress along [100] causes the primed ladder, being fourfold degenerate in the unstrained case, to split into two ladders with twofold degeneracy. A stress of 1 GPa is able to shift up in energy one of these ladders (the unprimed one) such that the population of this ladder is below 2% at all inversion layer concentrations.

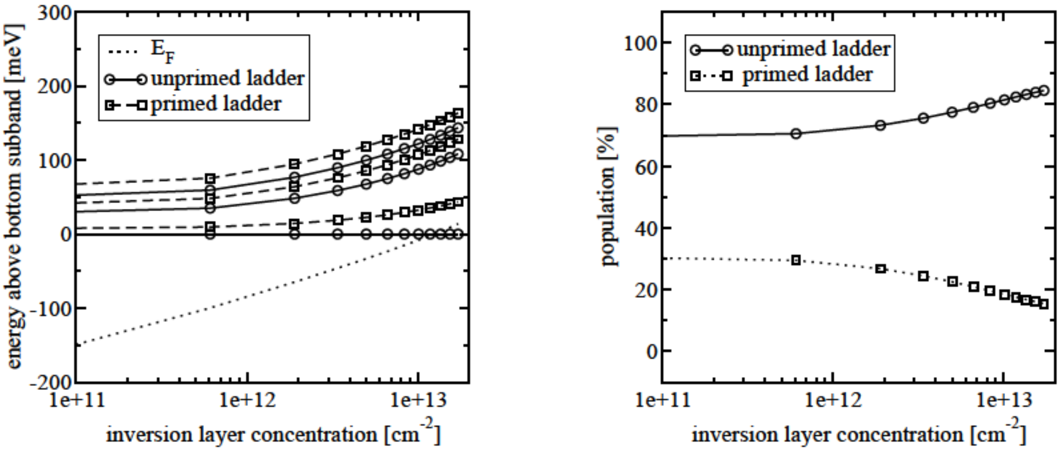
Stress along [110] does not change the degeneracy of the subband ladders. Figure 6.12d shows that the population of the primed ladder decreases below 20% for this type of stress. Additionally, the effective mass change induced by stress along the [110]-direction does not significantly change the subband levels and subband populations at 1 GPa stress. The small effect of stress on the unprimed ladder is not surprising since the subband energies and populations are determined by the quantization mass $\mathrm{m}_{[001]}$ and the density of states mass $\sqrt{m_{[110]}m_{[1\bar{1}0]}}$, respectively, which are not largely affected by the shear component $\varepsilon_{xy}=0.314\%$ resulting from 1 GPa stress.

Biaxial tensile stress in the (001)-plane can be seen to be the most efficient configuration to reduce the population of the primed subband ladder below 5% at 1 GPa (Figure 6.12f).

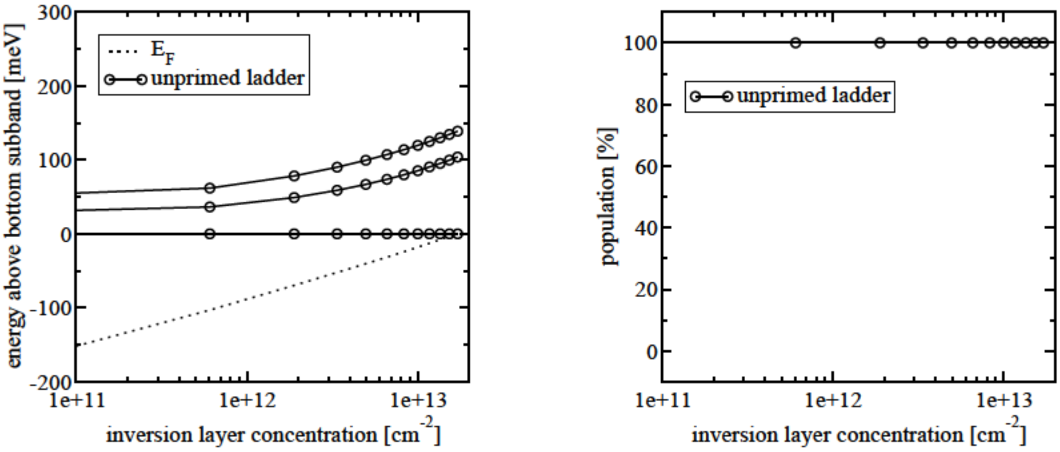
In Figure 6.13 the lowest subband levels and subband populations for a substrate with (110)-orientation and 1 GPa tensile stress along [001] are shown. The stress moves the primed subband ladder up in energy, thereby reducing its occupation below 2%.



(a) substrate orientation (001) (b) substrate orientation (001)

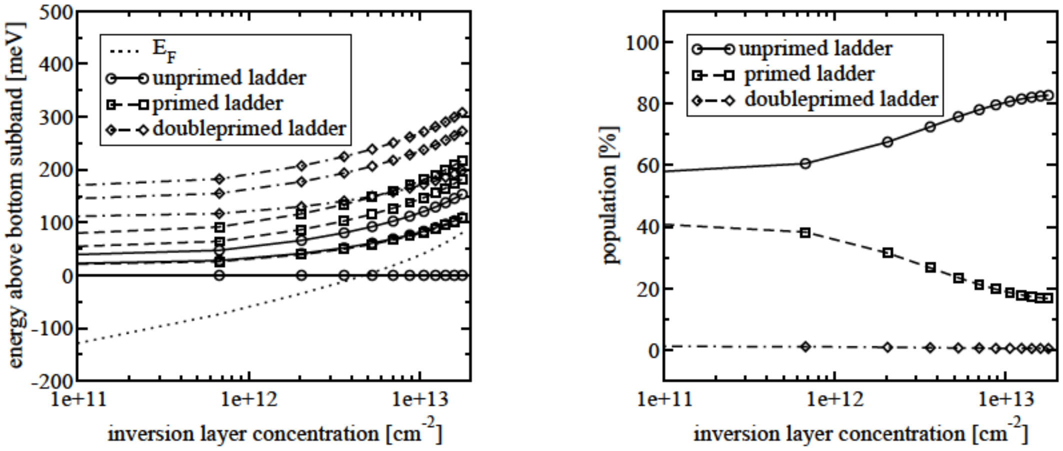


(c) substrate orientation (110) (d) substrate orientation (110)

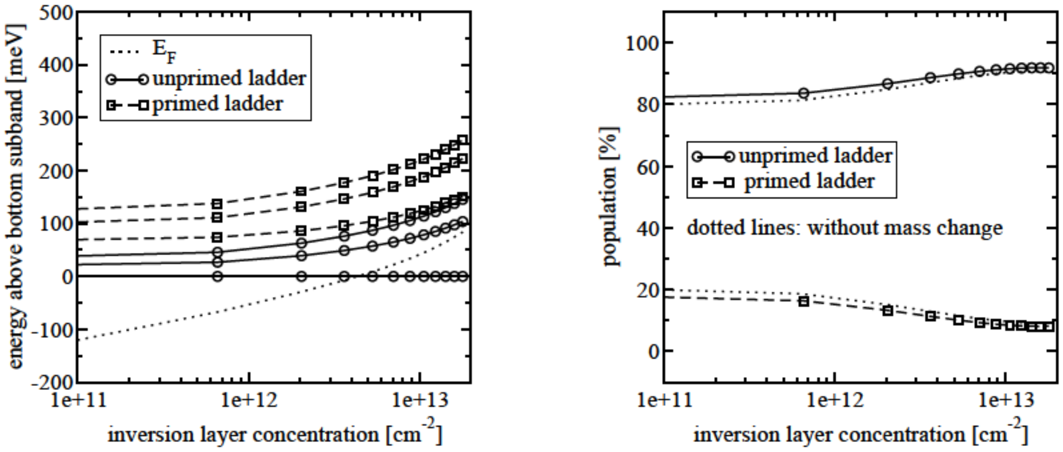


(e) substrate orientation (111) (f) substrate orientation (111)

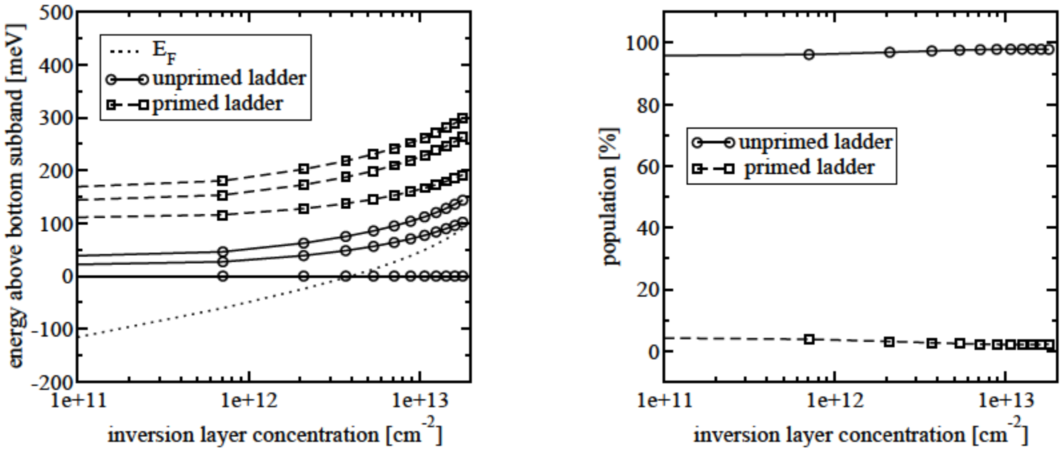
Figure 6.11: Three lowest subband energies and subband ladder populations for substrate orientation (001), (110), and (111). $N_{\mathrm{sub}} = 1 \times 10^{16}\ \mathrm{cm}^{-3}$ and $T = 300$ K.



(a) Stress along [100] (b) Stress along [100]



(c) Stress along [110] (d) Stress along [110]



(e) Biaxial stress (f) Biaxial stress

Figure 6.12: Three lowest subband energies and subband populations for (001)-oriented substrate at 1 GPa tensile stress and three stress directions. $N_{\text{sub}} = 1 \times 10^{16}$ cm^{-3} and $T = 300$ K.

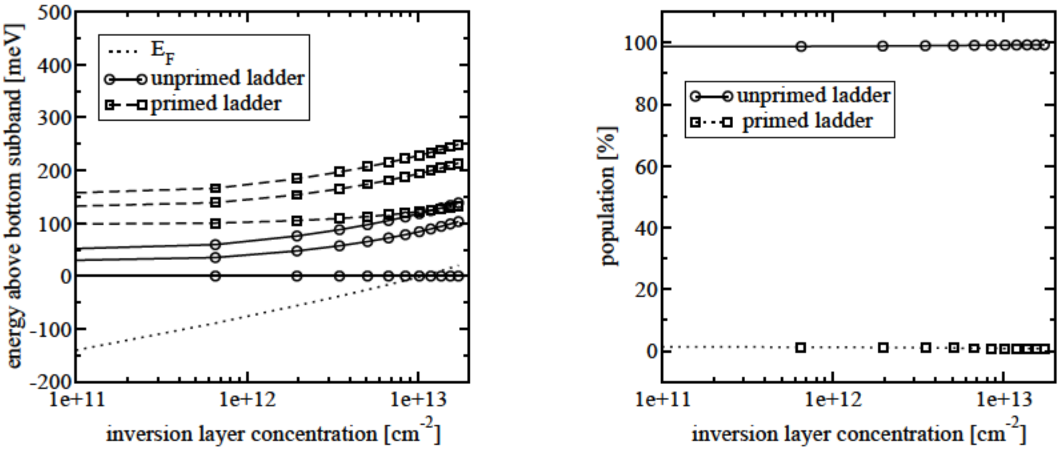
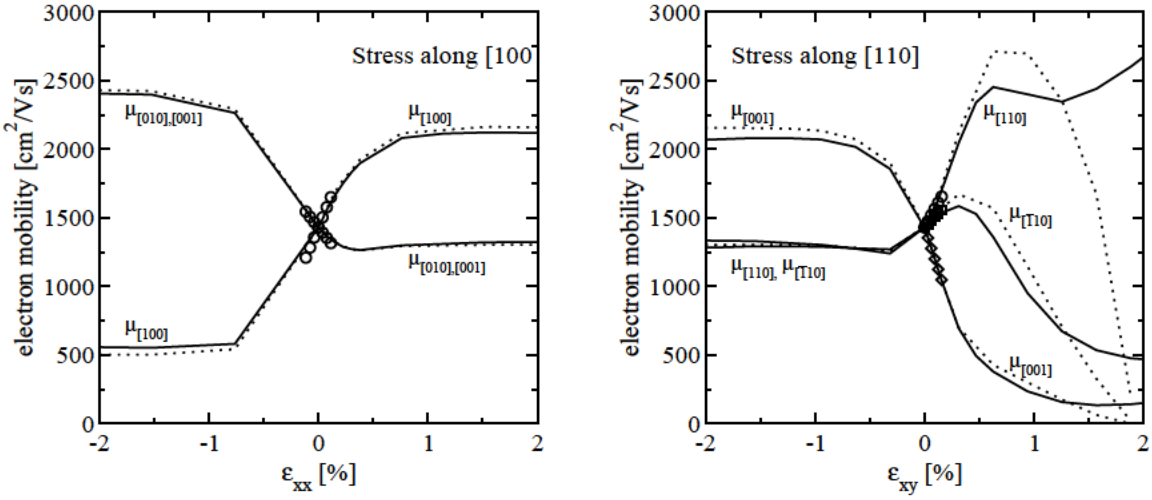


Figure 6.13: Three lowest subband energies and subband populations for (110)-oriented substrate at 1 GPa tensile stress along [001]. $N_{\mathrm{sub}} = 1 \times 10^{16}$ cm^{-3} and $T = 300$ K.



(a) (b)

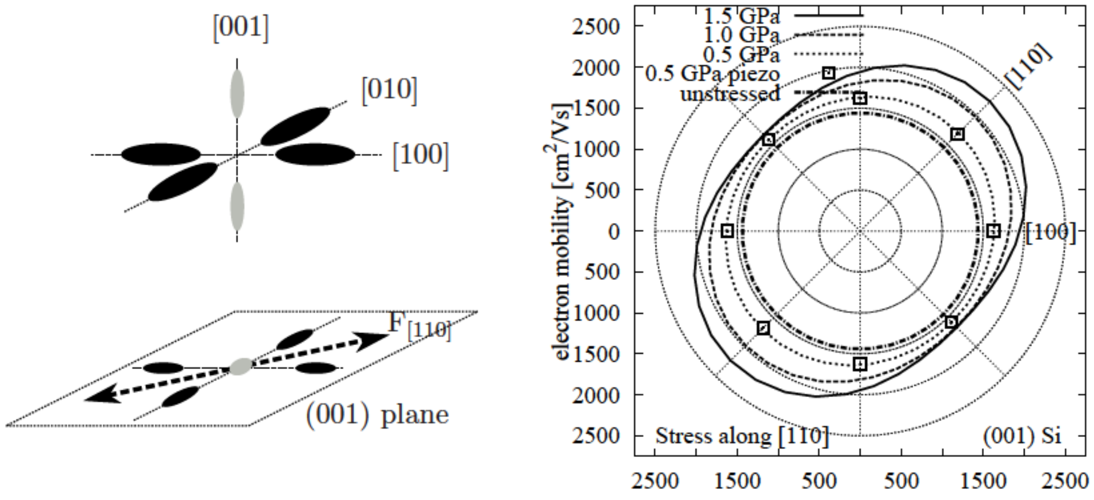
Figure 6.14: Simulated bulk electron mobility of intrinsic Si as a function of strain for stress direction [100] (a) and [110] (b). Mobility from FBMC (solid lines) and ABMC (dotted lines) simulations is plotted along the stress direction and two orthogonal directions. Symbols indicate the change of mobility calculated using the piezoresistance coefficients [Smith54].

6.2 Bulk Electron Mobility of Strained Si

In Figure 6.14 we compare simulation results for the electron mobility of strained Si obtained with FBMC and ABMC for the stress directions [100] and [110]. Mobility is plotted in three orthogonal directions, one being parallel and two being perpendicular to stress. Comparing simulation results from ABMC with FBMC allows: (i) a direct interpretation of the FBMC results, and (ii) the extraction of the limits of validity of the analytical band model.

In Figure 6.14a the simulation results from ABMC and FBMC for stress along [100] are compared and good agreement is observed. The resulting mobility is anisotropic in the (001) plane ($\mu_{[100]} \neq \mu_{[010]}$), which is a result of the strain-induced Δ-valley shifts. Mobility saturates at approximately 1% strain, regardless of the sign of strain. The saturated mobility values are larger for compressive strain, since for this type of strain four Δ-valleys are being depopulated, while for tensile strain only two valleys are being depopulated. Thus, the larger amount of intervalley scattering and the larger transport mass of the two populated valley-pairs reduces the mobility enhancement for tensile stress.

In Figure 6.14b simulation results are shown for stress along [110]. At compressive stress (negative ε_{xy}), four valleys move down in energy, yielding a decreased mobility in the (001) plane and a mobility enhancement along [001]. However, if tensile stress is applied along [110], the mobilities along the three directions [110], [1$\bar{1}$0], and [001] are different from each other with the largest mobility enhancement in [110] direction. Furthermore, no clear mobility saturation is observed as ε_{xy} increases. The reason for the different mobility enhancement under [110] tensile stress compared to [100] stress is that for this particular stress direction the valley pair along



(a)

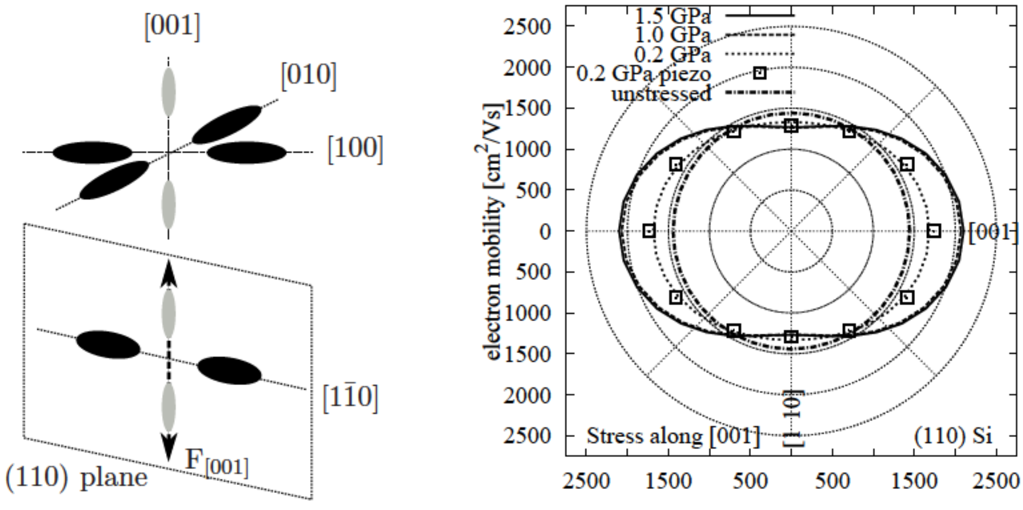
(b)

Figure 6.15: (a) Constant-energy surfaces of the lowest Si conduction band under uniaxial tensile stress along [110] and projection on the (001) plane. (b) Simulated in-plane electron mobility for various stress levels. At 0.5 GPa simulation results are compared to the mobility calculated from the piezoresistance coefficients (open squares).

[001] is primarily populated. As shown in Figure 6.10, this valley pair experiences a pronounced deformation as a function of shear strain ε_{xy}. In ABMC simulations this deformation was accounted for by using expressions (3.100), (3.94), and (3.99). It can be seen that the simulation results from ABMC qualitatively agree with those from FBMC up to 0.5% shear strain. At larger strain levels the band deformation is so pronounced that an energy band description in terms of an effective mass is no longer accurate, and FBMC have to be used even in the case of low-field transport to calculate the mobility. It is anticipated that a somewhat better agreement between ABMC and FBMC could be obtained, if the strain effect on the nonparabolicity coefficient α were included in the analytical band model.

The simulated mobility enhancement for stress along [100] and [110] was compared with predictions from a model based on the linear piezoresistance coefficients [Smith54]. Good agreement is found for both stress directions at small stress (< 0.2 GPa) where the model is valid (see Figure 6.14). Models solely based on strain-induced intervalley electron transfer [Herring56] fail to explain the origin of the non-vanishing shear piezoresistance coefficient of $\pi_{44} = -13.6 \times 10^{-11}$ Pa^{-1} [Smith54]. Hence, these models are not capable of reproducing the anisotropy of electron mobility in uniaxially stressed channels with [110] channel direction [Maruyama90, Kanda91].

The anisotropy of the electron mobility in the (001)-plane arising from stress along the [110] direction is shown in Figure 6.15. It is a result of the stress-induced effective mass change only. In Figure 6.16 the anisotropic electron mobility resulting from stress along the [001] direction in the (110)-plane is plotted. The anisotropy of the electron mobility originates from a repopulation of the subband ladders induced by stress.



(a)

(b)

Figure 6.16: (a) Constant energy surfaces of the lowest Si conduction band under uniaxial tensile stress along [001] and projection on the (110) plane. (b) Simulated in-plane electron mobility for various stress levels. At 0.2 GPa simulation results are compared to the mobility calculated from the piezoresistance coefficients (open squares).

The remainder of this section is temporarily not shown.

6.3 Electron Inversion Layer Mobility of Strained Si

The inversion layer mobility was calculated by means of a Monte Carlo method taking into account the subband structure calculated from a Schrödinger-Poisson solver (see Chapter 4). Phonon scattering and screened surface roughness scattering were taken into account. Depending on the substrate orientation, the sixfold degeneracy of the Δ-valley is lifted and up to three different subband ladders are formed. While for (001) substrate the subbands of the lowest subband ladder (unprimed ladder) are circular and the ladder is two-fold degenerate, for (110) substrate the subbands are elliptical and the unprimed ladder is fourfold degenerate. The higher density of states and larger transport masses for (110) substrate yield a lower inversion layer mobility as compared to (001) substrate. In Figure 6.17 simulation results are compared to experimental data [Uchida03, Tsutsui05].

In the following the origin of the electron mobility enhancement induced by uniaxial stress of the 2DEG is discussed. To demonstrate the influence of the substrate orientation, subband MC simulations are performed for (001) and (110) oriented substrates. It is shown that uniaxial stress leads to a pronounced anisotropy of the in-plane mobility for both substrate orientations. While for (110) substrate this effect stems from the anisotropic transport mass of the lowest subband ladder, the effective mass change induced by stress in [110] direction is responsible for the anisotropic electron mobility on (001) oriented substrate.

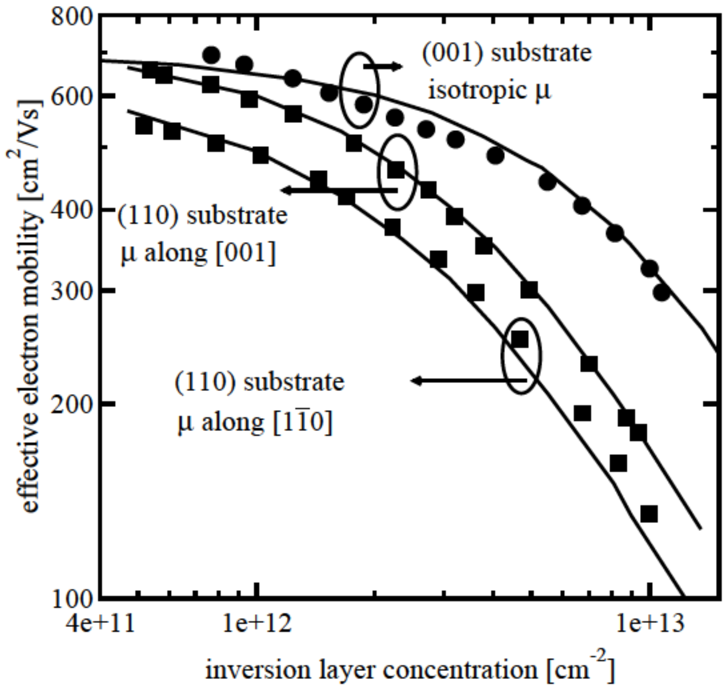


Figure 6.17: Simulated effective electron mobility μ_{eff} for substrate orientation (001) and (110) compared to measurements [Uchida03, Tsutsui05] (symbols). The anisotropic mobility for the (110) substrate is given along [001] and [1$\bar{1}$0].

The stress-induced change of the effective mass of electrons is incorporated in the subband MC simulations. In Figure 6.18 the mobility components parallel and perpendicular to stress direction [110] are compared to the unstrained mobility for two tensile stress levels. Tensile stress along [110] has two beneficial effects on the parallel mobility component: the splitting between the unprimed and primed ladders is increased, and the transport mass in direction of stress is reduced with respect to the unstrained case. From these two effects one can understand the mobility enhancement parallel to the stress direction for all inversion layer densities. Perpendicular to stress, the effective mass is increased, which leads to a smaller mobility enhancement in this direction. strained Si,

To enhance the mobility for (110) substrate a uniaxial tensile stress along [001] is applied, as this stress condition increases the splitting between the primed and unprimed ladders. From EPM calculations only a negligible change of the effective transport masses in the subband ladders is observed. Stress increases the component of μ_{eff} parallel to the stress direction, whereas the perpendicular mobility is smaller as compared to the unstressed case as shown in Figure 6.19. Since the effective masses are not changed, the mobility change is expected to saturate at larger stress (~ 1 GPa), as soon as the primed ladder becomes depopulated. These results are in good agreement with experimental data for the anisotropic mobility enhancement for (001) and (110) oriented substrates [Irie04].

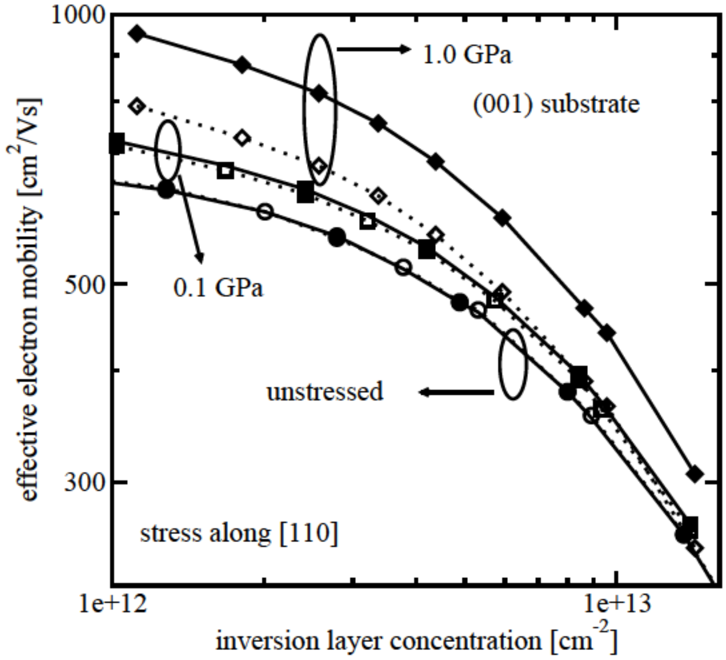


Figure 6.18: Mobility components parallel (closed symbols) and perpendicular (open symbols) to stress direction [110] without stress (circles), 0.1 GPa (squares), and 1 GPa stress (diamonds).

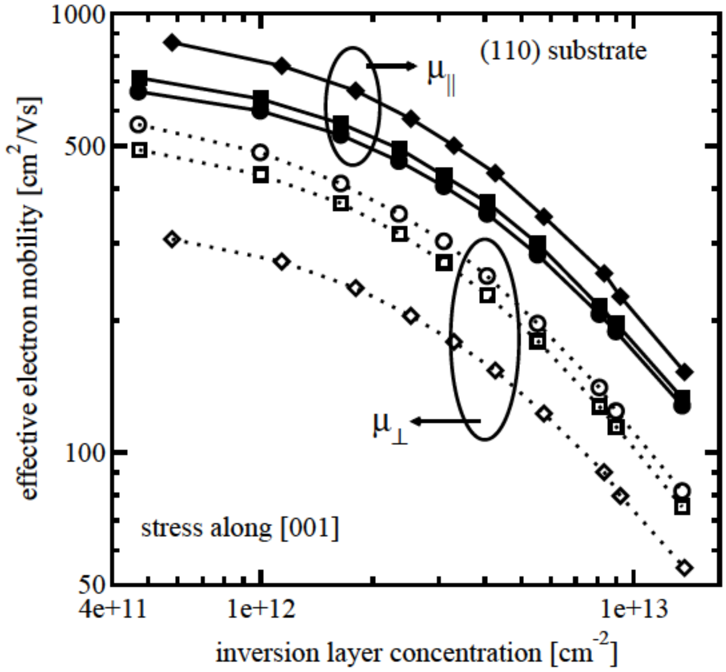
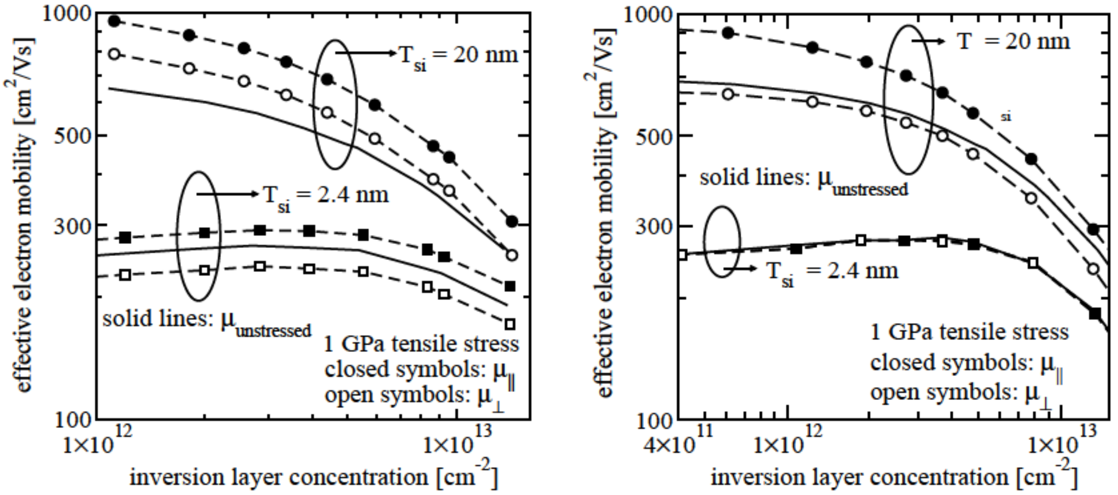


Figure 6.19: Mobility components parallel (closed symbols) and perpendicular (open symbols) to stress direction [001] without stress (circles), 0.1 GPa (squares), and 1 GPa stress (diamonds).



(a) stress along [110] (b) stress along [100]

Figure 6.20: Simulated effective mobility for substrate orientation (001), two channel orientations and two body thicknesses of unstressed Si (solid lines) and with 1 GPa stress along (a) [110] and (b) along [100] (dashed lines). The mobilities are plotted parallel (closed symbols) and perpendicular (open symbols) to the stress direction.

6.3.1 Channel Mobility in UTB MOSFETs for (001) and (110) Substrates

The influence of stress on the effective electron mobility enhancement in UTB MOSFETs is analyzed. For (001) oriented substrate it was shown before that a stress in the [110] direction induces two beneficial effects for the electron mobility in the [110] direction, namely (i) stress-induced valley splitting and (ii) stress-induced effective mass change. At relatively large body thicknesses like $T_{\mathrm{si}} = 20$ nm, the mobility enhancement $\Delta\mu_{\mathrm{eff}}$ can be understood from a combination of these two effects yielding an anisotropic μ_{eff} as compared to the unstressed system (see Figure 6.20). In ultra-thin Si bodies, however, the strong geometrical quantum confinement induces a large intrinsic valley splitting, thus the additional stress-induced valley shifts have a negligible effect on the mobility. At T_{si}=2.4 nm, the larger component $\mu_{\mathrm{eff}||}$ parallel and the smaller component $\mu_{\mathrm{eff}\perp}$ perpendicular to the stress direction result from the effective mass change only, which is in good agreement with experimental data [Uchida05].

In Figure 6.20 the effect of uniaxial stress along channels in [100] direction is shown. Stress in direction [100] lifts the degeneracy of the fourfold degenerate, primed ladder. Since no effective mass change occurs, $\Delta\mu_{\mathrm{eff}}$ occurring at large T_{si} is a result of subband ladder repopulation only. As the body thickness is decreased, the population of the higher subband ladders inevitably decreases, and strain cannot further decrease the population. Therefore, in Si films with a thickness of $T_{\mathrm{si}} = 2.4$ nm even stress at a level of 1 GPa does not induce a mobility enhancement.

Stress-induced $\Delta\mu_{\mathrm{eff}}$ for (110) oriented substrates can be understood from similar arguments. Tensile stress along [001], which increases the bulk- and inversion layer mobility parallel to stress (see Figure 6.16 and Figure 6.19), does not alter the mobilities at small T_{si}, because it does not change the effective masses, but merely increases the splitting between the primed and unprimed

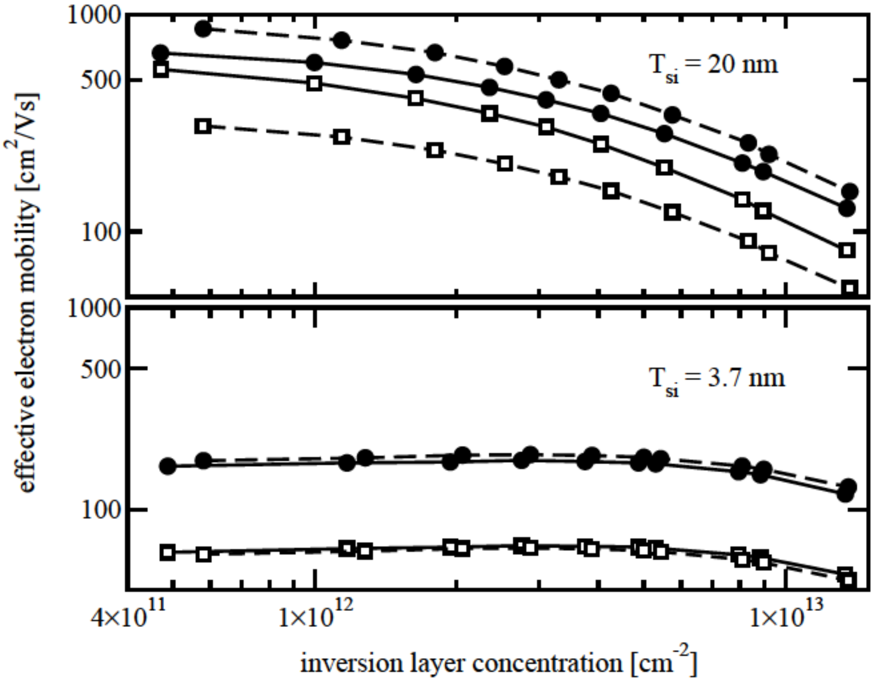


Figure 6.21: Effective mobility for substrate orientation (110) of unstressed (solid lines) and 1 GPa stressed (dashed lines) Si for two body thicknesses. The mobility components are plotted parallel (closed symbols) and perpendicular (open symbols) to stress direction [001].

subband ladders. Figure 6.21 shows how the stress-induced mobility enhancement, that can be observed at $T_{\text{si}} = 20$ nm, vanishes at $T_{\text{si}} = 3.7$ nm.

6.3.2 Degeneracy Effects on Inversion Layer Mobility

The MC algorithm presented in Section 5.4 taking into account the Pauli principle has been compared to two other methods found in the literature [Bosi76, Yamakawa96]. The first algorithm to include the Pauli principle in the MC technique [Bosi76] is based on a self-consistent iterative algorithm that uses a rejection technique to account for the occupation probability of the final state after each scattering event. Since this auxiliary self-scattering mechanism is proportional to the occupation of the final states, the algorithm prevents a large number of classically allowed transitions.

A different approach to include degeneracy in MC simulations was given in [Yamakawa96]. Inelastic scattering rates are multiplied with a factor of $(1 - f_{\text{FD}}(\epsilon_f))/(1 - f_{\text{FD}}(\epsilon_i))$, where ϵ_i and ϵ_f denote the initial and final electron energy, respectively. This additional factor stems from the use of Fermi-Dirac statistics within the relaxation time approximation [Roychoudhury80].

In the limit of vanishing field the mobility can also be calculated based on the relaxation time approximation using the Kubo-Greenwood formula [Fischetti03]. Figure 6.22 shows that the new method yields the closest agreement, whereas a non-self-consistent implementation of the algorithm proposed in [Bosi76], where $f(\mathbf{k})$ has been approximated by the equilibrium distribu-

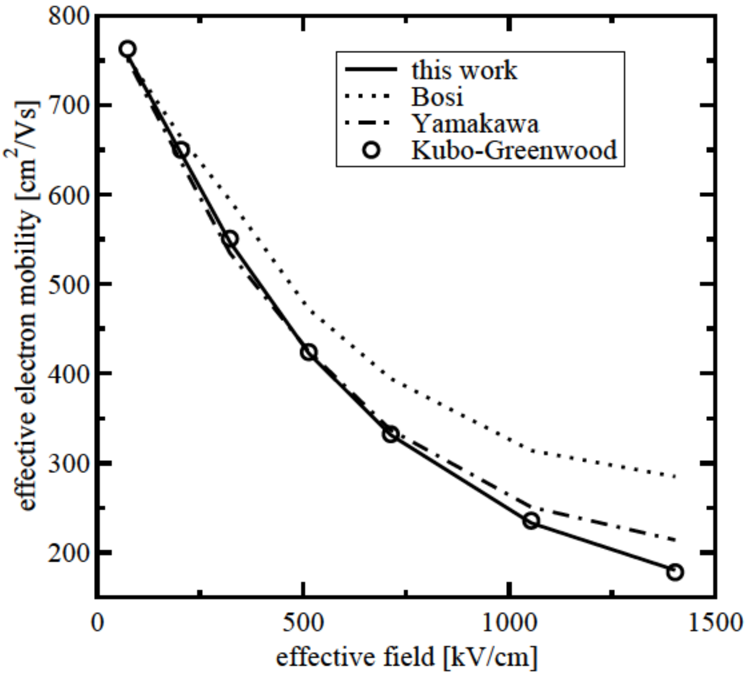


Figure 6.22: The simulated effective mobility using a new algorithm (solid line) is compared to results of a non-self-consistent version of Bosi's algorithm [Bosi76] (dotted line), the algorithm proposed in [Yamakawa96], and to the mobility calculated with the Kubo-Greenwood formalism (open circles).

tion function $f_{\mathrm{FD}}(\mathbf{k})$, and the algorithm proposed by [Yamakawa96] overestimate the effective mobility.

Using the proposed MC algorithm the mean electron velocity as a function of total electron energy was extracted in the small driving field limit. For this purpose the particle's energy domain was divided into a set of intervals ΔE. The mean velocity component of an electron in direction of the driving field in a particular interval $E \in [E_0, E_0 + \Delta E]$ can be obtained during a MC simulation from a history of duration T

$$\overline{v}(E_0) = \frac{1}{T} \int_0^T v[\mathbf{k}(t)] \big(\theta[E(t) - E_0] - \theta[E(t) - E_0 - \Delta E]\big) \mathrm{d}t \;, \tag{6.3}$$

where $\theta(E)$ denotes the unit step function and $\mathbf{k}(t)$ represents the trajectory of the electron wave vector. Note that the overall mobility is proportional to the sum of the mean velocities of all intervals.

A interesting behavior can be observed when comparing the mean velocities resulting from simulations with classical and Fermi-Dirac statistics from Figure 6.23 and 6.24. As shown in Figure 6.23 the mean velocities coincide for both simulation modes in the nondegenerate regime (small E_{eff}) when only phonon scattering is considered in the MC simulation. At high E_{eff}, where the 2DEG is highly degenerate, a shift of the mean velocity distribution toward

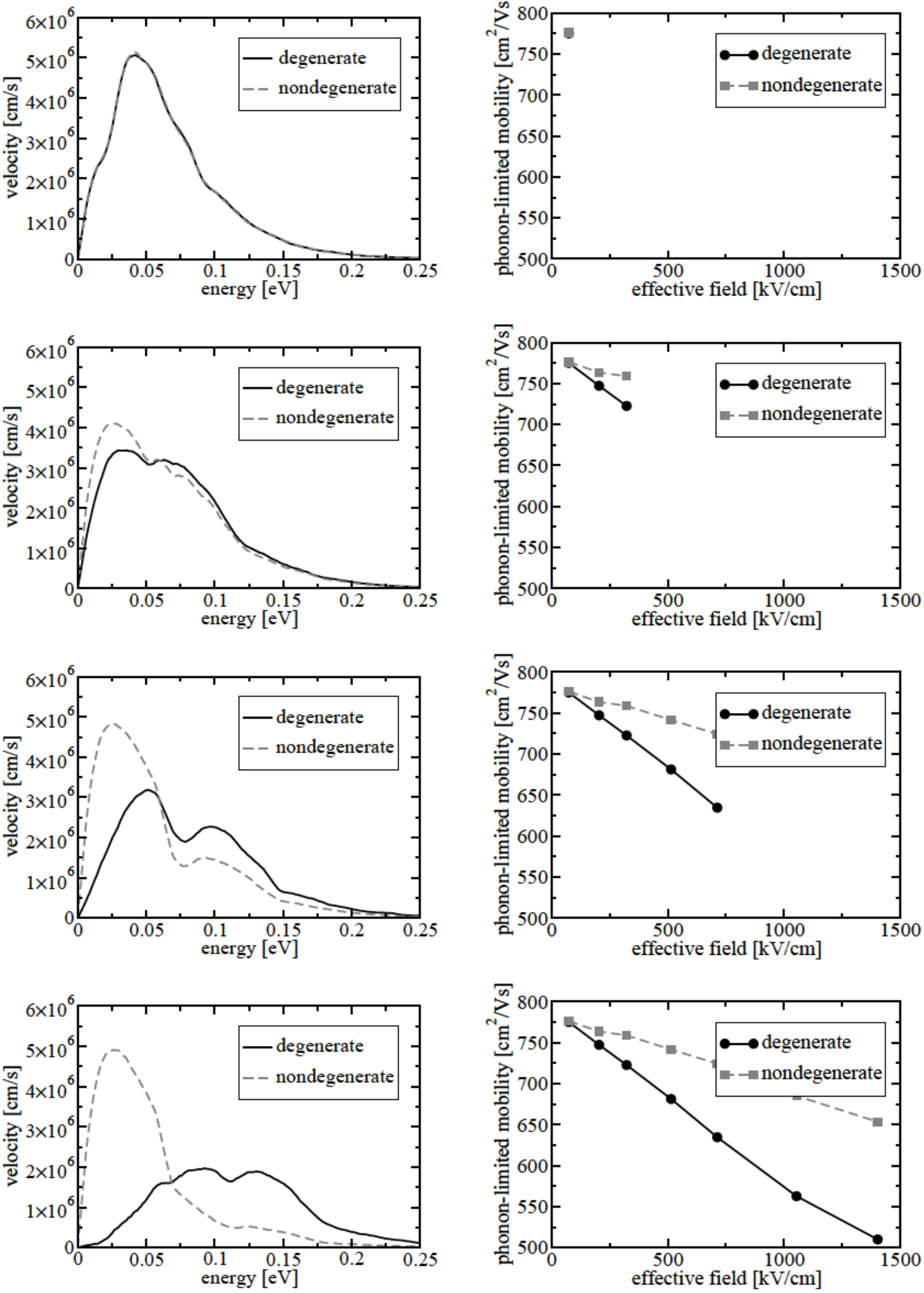


Figure 6.23: Velocity distribution functions (left) and the corresponding electron mobilities limited by phonon scattering (right) for various effective fields.

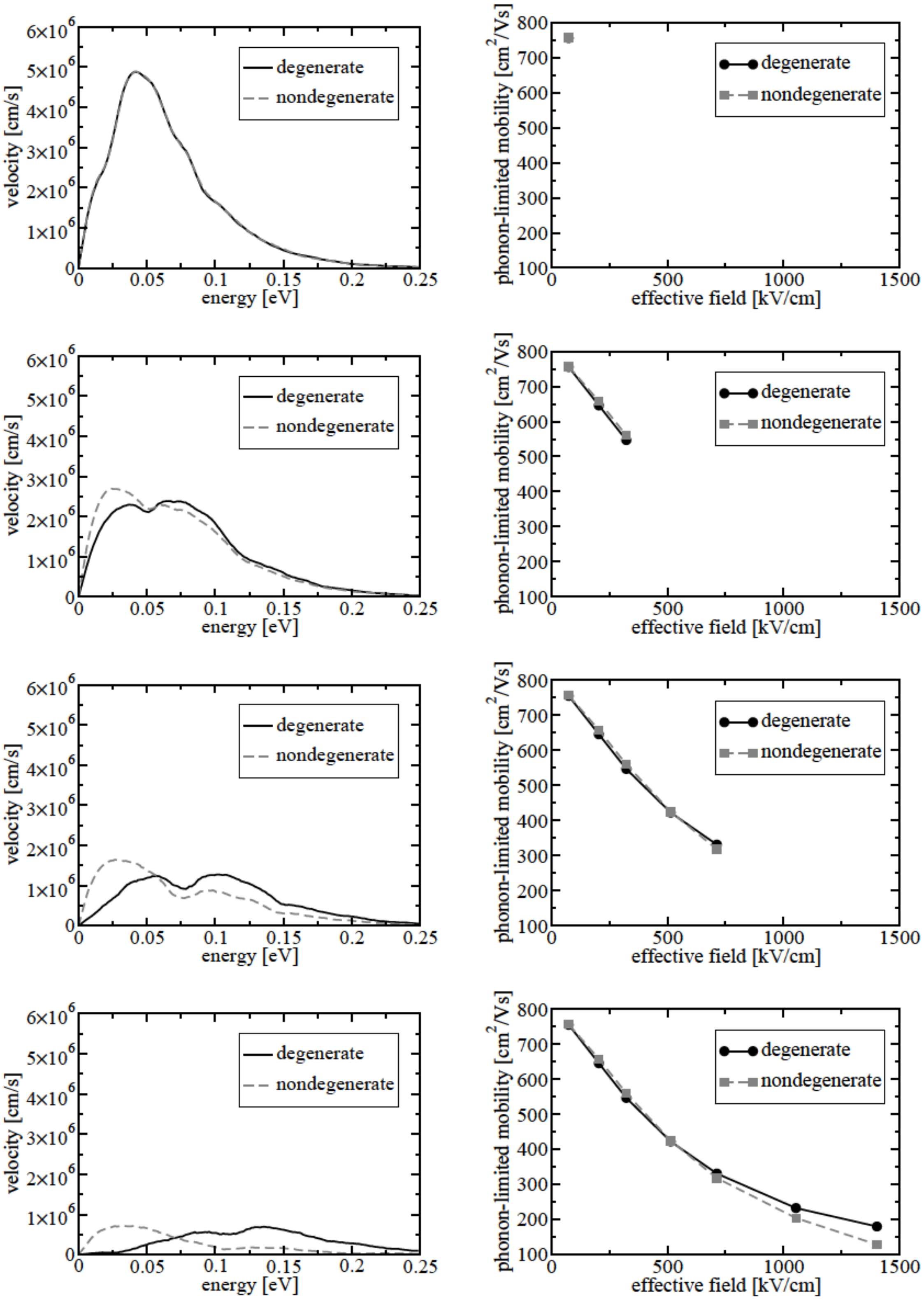


Figure 6.24: Velocity distribution functions (left) and the corresponding electron mobilities limited by phonon- and surface scattering (right) for various effective fields.

higher energies and a decrease of its peak can be observed as compared to the mean velocity resulting from nondegenerate simulations without the Pauli principle. The coincidence of the mean velocities in the nondegenerate regime is a check of consistency that the algorithm with the Pauli principle included converges to the classical algorithm for the nondegenerate 2DEG. At high E_{eff} the different mean velocities can be interpreted as follows: In simulations neglecting the Pauli principle electrons have an equilibrium energy of $\mathrm{k_B}T$ whereas the mean energy resulting from simulations with the Pauli principle can be more than twice as much. Since phonon scattering is proportional to the density of states, which is an increasing step-like function for the 2DEG, electrons at higher energies – as it is the case in simulations with the Pauli principle – experience more scattering and thus the phonon-limited mobility is strongly decreased (see right plots of Figure 6.23).

The plots of Figure 6.24 show the mean velocities and the effective mobilities at various effective fields when surface roughness scattering is included in MC simulations. At low E_{eff} surface roughness scattering does not play an important role, and the mean velocities compare well with the simulation results for the phonon-limited mobility in Figure 6.23. However, now even at high E_{eff} the mean velocities stemming from simulations with and without the Pauli principle do not differ as much. The large peak that was observed in the nondegenerate phonon-limited mean velocity at small energies is suppressed. This is a direct consequence of surface roughness scattering, which is at small energies more effective than phonon scattering. Thus, the effective mobility resulting from simulations with degenerate statistics are incidentally in close agreement to those using classical statistics even though the phonon-limited mobility experiences a noticeable reduction when using degenerate statistics. As previously discussed, this close agreement can only be understood from the cancellation of two effects: Degeneracy leads to an increase of the mean kinetic energy. This leads to an increase in phonon scattering and a decrease in the mobility. At the same time electrons with larger kinetic energies experience less effective surface-roughness scattering, thus the surface roughness limited mobility is increased. By coincidence, in Si inversion layers at room temperature these two effects cancel each other at all effective fields, and the difference between a simulation with nondegenerate and degenerate statistics is very small.

As a final observation from Figure 6.23 and 6.24 one can see that due to degeneracy effects electrons at energies below the Fermi level have smaller mean velocities, which corresponds well to the general picture that highly occupied states have little contribution to transport.

Chapter 7

Summary and Conclusions

STRAIN techniques are used by many prominent microprocessor manufacturers primarily for their sub-130 nm CMOS technologies. The most prominent strain technologies were outlined in this work. A revision of current strain technologies showed that shear strain induced by uniaxial stress in $\langle 110 \rangle$ direction is widely used to enhance the electron mobility for $\{001\}$ Si substrate. Hence, in this work special emphasis was put on this strain configuration.

The effect of strain on the band structure of Si was investigated systematically using the linear deformation potential theory and the $\boldsymbol{k}\cdot\boldsymbol{p}$ method. Shear strain lifts the degeneracy of the two lowest conduction bands Δ_1 and $\Delta_{2'}$ at the zone boundary X points. The change of the shape of the lowest conduction band was quantified in terms of (i) an effective mass change, (ii) a shear-strain-induced valley splitting, and (iii) a change in position of the valley minimum in $\boldsymbol{k}$-space. Additionally, the empirical pseudopotential method was adapted to incorporate strain effects. The results from numerical band structure calculations were compared to the analytical expressions derived using the $\boldsymbol{k}\cdot\boldsymbol{p}$ theory and good agreement was observed.

Furthermore, the effect of strain on the subband structure of Si inversion layers formed at the surface between one or two Si-SiO$_2$ interfaces with $\{001\}$, $\{110\}$, and $\{111\}$ substrate orientation was shown. The transport masses and the degeneracy of the subband ladders depend on the substrate orientation and can be modified by strain. The strain configurations that enhance carrier mobility were identified for each substrate orientation.

Fullband MC simulations were performed using VMC to analyze the effect of strain on the electron mobility. MC simulations using an analytical description of the electron bands were shown to be valid in a limited range of shear strain ($< \pm 0.5\%$). At larger shear strain the band deformation is so pronounced, that fullband modeling is required. Hence, for modeling of transport in strained Si a fullband description is of particular importance. MC simulations and a rigorous modeling of the strain effect on the electron band structure reproduce experimentally observed mobility data for bulk Si and Si inversion layers on different substrate orientations.

A method for the inclusion of the Pauli principle in a Monte Carlo algorithm is presented to study the effect of degeneracy both on the phonon-limited mobility and the effective mobility including surface-roughness scattering in Si inversion layers. It is shown that at room temperature and for $\{001\}$ substrate orientation incidentally degeneracy has a minor effect on the effective mobility, despite non-degenerate statistics yields unphysical subband populations and an underestimation of the mean electron energy. In general a correct treatment of the degenerate carrier statistics of the 2DEG is important.

Bibliography

[Agostinelli91] V. M. Jr. Agostinelli, H. Shin, and A. F. Jr. Tasch. *A comprehensive model for inversion layer hole mobility for simulation of submicrometer MOSFET's.* IEEE Trans.Electron Devices, vol. 38, pages 151–159, 1991.

[Ando82] T. Ando, A.B. Fowler, and F. Stern. *Electronic properties of two-dimensional systems.* Review of Modern Physics, vol. 54, no. 2, pages 437–672, 1982.

[Andrieu06] F. Andrieu, T. Ernst, O. Faynot, O. Rozeau, Y. Bogumilowicz, J. M. Hartmann, L. Brevard, A. Toffoli, D. Lafond, and B. Ghyselen. *Performance and physics of sub-50 nm strained Si on* $Si_{1-x}Ge_x$ *on insulator (SGOI) nMOSFETs.* Solid-State Electron., vol. 50, pages 566–572, Apr 2006.

[Ang04] Kah-Wee Ang, King Jien Chui, V. Bliznetsov, Anyan Du, N. Balasubramanian, Ming Fu Li, Ganesh Samudra, and Yee-Chia Yeo. *Enhanced performance in 50 nm n-MOSFETs with silicon-carbon source/drain regions.* In Proc. Intl. Electron Devices Meeting, pages 1069–1071, 2004.

[Ang05] Kah-Wee Ang, King-Jien Chui, Vladimir Bliznetsov, Chih-Hang Tung, Anyan Du, Narayanan Balasubramanian, Ganesh Samudra, Ming Fu Li, and Yee-Chia Yeo. *Lattice strain analysis of transistor structures with silicon–germanium and silicon–carbon source/drain stressors.* Appl.Phys.Lett., vol. 86, no. 9, page 093102, 2005.

[Arghavani06] R. Arghavani, L. Xia, H.M. Saad, M. Balseanu, G. Karunasiri, A. Mascarenhas, , and S. E. Thompson. *A reliable and manufacturable method to induce a stress of* $> 1 GPa$ *on a p-channel MOSFET in high volume manufacturing.* IEEE Electron Device Lett., vol. 27, no. 2, pages 114–116, 2006.

[Ashcroft76] N. W. Ashcroft, and N. D. Mermin. “Solid state physics,”. W. B. Saunders, Philadelphia, 1976.

[Bai04] P. Bai, C. Auth, S. Balakrishnan, M. Bost, R. Brain, V. Chikarmane, R. Heussner, M. Hussein, J. Hwang, D. Ingerly et al. *A 65 nm logic technology featuring 35 nm gate lengths, enhanced channel strain, 8 Cu interconnect* *layers, low-k ILD and 0.57 μm^2 SRAM cell.* In Proc. Intl. Electron Devices Meeting, pages 657–660, 2004.

[Balslev66] I. Balslev. *Influence of uniaxial stress on the indirect absorption edge in silicon and germanium.* Physical Review, vol. 143, pages 636 – 647, 1966.

[Bardeen50] J. Bardeen, and W. Shockley. *Deformation potentials and mobilities in non-polar crystals.* Physical Review, vol. 80, no. 1, pages 72–80, 1950.

[Bastard81] G. Bastard. *Superlattice band structure in the envelope-function approximation.* Physical Review B, vol. 24, no. 10, pages 5693–5697, Nov 1981.

[Belford01] R.E. Belford. *Uniaxial, tensile-strained Si devices.* J.Electron.Mater., vol. 30, no. 7, pages 807–811, 2001.

[Bera06] L.K. Bera, W.Y. Loh, , L.H. Guo, X.W. Zhang, G.Q. Lo, N. Balasubramanian, and D.-L. Kwong. *The impact of uniform strain applied via bonding onto plastic substrate on MOSFET performance.* IEEE Electron Device Lett., vol. 27, no. 1, pages 58–61, 2006.

[Bir74] G. L. Bir, and G. E. Pikus. “Symmetry and strain induced effects in semiconductors,”. Wiley, New York, 1974.

[Boeuf01] F. Boeuf, T. Skotnicki, S. Monfray, C. Julien, D. Dutartre, J. Martins, P. Mazoyer, R. Palla, B. Tavel, P. Ribot, E. Sondergard, and A. Sanquer. *16 nm planar NMOSFET manufacturable within state-of-the-art CMOS process thanks to specific design and optimisation.* In Proc. Intl. Electron Devices Meeting, pages 29.5.1–29.5.4, 2-5 Dec. 2001.

[Bosi76] S. Bosi, and C. Jacoboni. *Monte Carlo high-field transport in degenerate GaAs.* J.Physics C, vol. 9, pages 315–319, 1976.

[Bufler98] F.M. Bufler. “Full-band monte carlo simulation of electrons and holes in strained Si and SiGe,”. Herbert Utz Verlag, 1998.

[Cardona66] Manuel Cardona, and Fred H. Pollak. *Energy-band structure of germanium and silicon: The k·p method.* Physical Review, vol. 142, no. 2, pages 530–543, Feb 1966.

[Chan03] V. Chan, R. Rengarajan, N. Rovedo, Wei Jin, T. Hook, P. Nguyen, Jia Chen, E. Nowak, Xiang-Dong Chen, D. Lea, A. Chakravarti, V. Ku, S. Yang, A. Steegen, C. Baiocco, P. Shafer, Hung Ng, Shih-Fen Huang, and C. Wann. *High speed 45nm gate length CMOSFETs integrated into a 90nm bulk technology incorporating strain engineering.* In Proc. Intl. Electron Devices Meeting, pages 77–80, 2003.

[Chan05] V. Chan, Ken Rim, Meikei Ieong, Sam Yang, Rajeev Malik, Young Way Teh, Min Yang, and Qiqing. *Strain for CMOS performance improvement.* In Custom Integrated Circuits Conference, 2005. Proceedings of the IEEE 2005, pages 667–674, 2005.

[Chang03] L. Chang, Yang-kyu Choi, D. Ha, P. Ranade, Shiying Xiong, J. Bokor, Chenming Hu, and T. J. King. *Extremely scaled silicon nano-CMOS devices.* Proc.IEEE, vol. 91, pages 1860–1873, 2003.

[Chelikowsky76] James R. Chelikowsky, and Marvin L. Cohen. *Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors.* Physical Review B, vol. 14, no. 2, pages 556–582, Jul 1976.

[Chen04] Chien-Hao Chen, T. L. Lee, T. H. Hou, C. L. Chen, C. C. Chen, J. W. Hsu, K. L. Cheng, Y. H. Chiu, H. J. Tao, Y. Jin, C. H. Diaz, S. C. Chen, and M. S. Liang. *Stress memorization technique (SMT) by selectively strained-nitride capping for sub-65nm high-performance strained-Si device application.* In VLSI Symp. Tech. Dig., pages 56–57, 2004.

[Cheng71] Y. C. Cheng. *On the scattering of electrons in magnetic and electric surface states by surface roughness.* Surface Science, vol. 27, pages 663–666, Oct 1971.

[Cheng01] Zhi-Yuan Cheng, M. T. Currie, C. W. Leitz, G. Taraschi, E. A. Fitzgerald, J. L. Hoyt, and D. A. Antoniadas. *Electron mobility enhancement in strained-Si n-MOSFETs fabricated on SiGe-on-insulator (SGOI) substrates.* IEEE Electron Device Lett., vol. 22, pages 321–323, 2001.

[Chidambaram06] P. R. Chidambaram, C. Bowen, S. Chakravarthi, C. Machala, and R. Wise. *Fundamentals of silicon material properties for successful exploitation of strain engineering in modern CMOS manufacturing.* IEEE Trans.Electron Devices, vol. 53, no. 5, pages 944–964, 2006.

[Chui07] K. J. Chui, K. W. Ang, N. Balasubramanian, M. F. Li, G. S. Samudra, and Y. C. Yeo. *n-MOSFET with silicon–carbon source/drain for enhancement of carrier transport.* IEEE Trans.Electron Devices, vol. 54, pages 249–256, 2007.

[Cohen89] M. L. Cohen, and J. Chelikowsky. "Electronic structure and optical properties of semiconductors,", volume 75 of *Springer Ser. Solid-State Sci.* Springer, Berlin, Heidelberg, 1989.

[Connelly03] D. Connelly, C. Faulkner, and D.E. Grupp. *Performance advantage of Schottky source/drain in ultrathin-body silicon-on-insulator and dual-gate CMOS.* IEEE Trans.Electron Devices, vol. 50, no. 5, pages 1340–1345, May 2003.

[Conwell67] E.M. Conwell. "High field transport in semiconductors,", volume 9 of *Solid State Physics.* Academic Press, New York, London, 1967.

[Cousins87] C. S. G. Cousins, L. Gerward, J. S. Olsen, B. Selsmark, and B. J. Sheldon. *Surface effects in uniaxially stressed crystals: the internal-strain parameters of silicon and germanium revised.* Journal of Physics C Solid State Physics, vol. 20, pages 29–37, January 1987.

[Darwish97] M. N. Darwish, J. L. Lentz, M. R. Pinto, P. M. Zeitzoff, T. J. Krutsick, and Hong Ha Vuong. *An improved electron and hole mobility model for general* *purpose device simulation.* IEEE Trans.Electron Devices, vol. 44, pages 1529–1538, 1997.

[Dhar05] S. Dhar, H. Kosina, V. Palankovski, E. Ungersboeck, and S. Selberherr. *Electron mobility model for strained-Si devices.* IEEE Trans.Electron Devices, vol. 52, no. 4, pages 527–533, 2005.

[Eneman05] S. Eneman, P. Verheyen, R. Rooyackers, F. Nouri, L. Washington, R. Degraeve, B. Kaczer, V. Moroz, A. De Keersgieter, R. Schreutelkamp, M. Kawaguchi, Y. Kim, A. Samoilov, L. Smith, P. P. Absil, K. De Meyer, M. Jurczak, and S. Biesemans. *Layout impact on the performance of a locally strained p-MOSFET.* In VLSI Symp. Tech.Dig., pages 22–23, 2005.

[Esseni03] D. Esseni, and A. Abramo. *Modeling of electron mobility degradation by remote coulomb scattering in ultrathin oxide MOSFETs.* IEEE Trans.Electron Devices, vol. 50, no. 7, pages 1665–1674, 2003.

[Esseni04] D. Esseni. *On the modeling of surface roughness limited mobility in SOI MOSFETs and its correlation to the transistor effective field.* IEEE Trans.Electron Devices, vol. 51, no. 3, pages 394–401, 2004.

[Ferry97] D. K. Ferry, and S. M. Goodnick. “Transport in nanostructures,”. Cambridge University Press, 1997.

[Fischetti96a] M. V. Fischetti, and S. E. Laux. *Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys.* J.Appl.Phys., vol. 80, no. 4, pages 2234–2252, 1996.

[Fischetti96b] M.V. Fischetti, and S.E. Laux. *Monte Carlo simulation of electron transport in Si: The first 20 years.* In G. Baccarani, and M. Rudan, editeurs, 26th European Solid State Device Research Conference, pages 813–820, Bologna, Italy, 1996. Editions Frontiers.

[Fischetti02] M. V. Fischetti, F. Gamiz, and W. Hänsch. *On the enhanced electron mobility in strained-silicon inversion layers.* J.Appl.Phys., vol. 92, no. 12, pages 7320–7324, 2002.

[Fischetti03] M. V. Fischetti, and Z. Ren. *Six-band k·p calculation of the hole mobility in silicon inversion layers: Dependence on surface orientation, strain, and silicon thickness.* J.Appl.Phys., vol. 94, no. 2, pages 1079–1095, 2003.

[Fitzgerald91] E.A. Fitzgerald, Y.H Xie, M.L. Green, D. Brasen, A.R. Kortan, J. Michel, Y.J. Mii, and B. Weir. *Totally relaxed* Ge_xSi_{1-x} *layers with low threading dislocation densities grown on Si substrates.* J.Appl.Phys., vol. 59, no. 7, pages 811–813, 1991.

[Friedel89] P. Friedel, M. S. Hybertsen, and M. Schlüter. *Local empirical pseudopotential approach to the optical properties of Si/Ge superlattices.* Physical Review B, vol. 39, no. 11, pages 7974–7977, Apr 1989.

[Fritze04] M. Fritze, C.L. Chen, S. Calawa, D. Yost, B. Wheeler, P. Wyatt, C.L. Keast, J. Snyder, and J. Larson. *High-speed Schottky-barrier pMOSFET with* $f_T =$ *280 GHz*. IEEE Electron Device Lett., vol. 25, no. 4, pages 220–222, April 2004.

[Ghyselen04] B. Ghyselen, J. M. Hartmann, T. Ernst, C. Aulnette, B. Osternaud, Y. Bogumilowicz, A. Abbadie, P. Besson, O. Rayssac, and A. Tiberj. *Engineering strained silicon on insulator wafers with the Smart CutTM technology*. Solid-State Electron., vol. 48, pages 1285–1296, Aug 2004.

[Goodnick85] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek. *Surface roughness at the Si(100)-*SiO_2 *interface*. Physical Review B, vol. 32, no. 12, pages 8171–8186, Dec 1985.

[Goroff63] Iza Goroff, and Leonard Kleinman. *Deformation potentials in silicon. III. Effects of a general strain on conduction and valence levels*. Physical Review, vol. 132, no. 3, pages 1080–1084, Nov 1963.

[Haensch06] W. Haensch, E. J. Nowak, and Dennard R. H. *Silicon CMOS devices beyond scaling*. IBM J.Res.Dev., vol. 50, no. 4/5, pages 339–361, 2006.

[Hall51] Harry H. Hall, J. Bardeen, and G. L. Pearson. *The effects of pressure and temperature on the resistance of* $p-n$ *junctions in germanium*. Physical Review, vol. 84, no. 1, pages 129–132, Oct 1951.

[Harris04] S. Harris. “An Introduction to the Theory of the Boltzmann Equation,”. Courier Dover Publications, 2004.

[Harrison56] W. A. Harrison. *Scattering of electrons by lattice vibrations in nonpolar crystals*. Physical Review, vol. 104, no. 5, pages 1281–1290, 1956.

[Haugerud03] B. M. Haugerud, L. A. Bosworth, and R. E. Belford. *Mechanically induced strain enhancement of metal–oxide–semiconductor field effect transistors*. J.Appl.Phys., vol. 94, no. 6, pages 4102–4107, 2003.

[Hensel65] J. C. Hensel, H. Hasegawa, and M. Nakayama. *Cyclotron resonance in uniaxially stressed silicon. II. Nature of the covalent bond*. Physical Review, vol. 138, no. 1A, pages A225–A238, Apr 1965.

[Herring56] Conyers Herring, and Erich Vogt. *Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering*. Physical Review, vol. 101, no. 3, pages 944–961, Feb 1956.

[Hinckley90] J.M. Hinckley, and J. Singh. *Influence of substrate composition and crystallographic orientation on the band structure of pseudomorphic Si-Ge alloy films*. Physical Review B, vol. 42, pages 3546–3566, 1990.

[Hoeneisen72] B. Hoeneisen, and C. A. Mead. *Fundamental limitations in microelectronics–I.MOS technology*. Solid-State Electron., vol. 15, pages 819–829, July 1972.

[Horstmann05] M. Horstmann, A. Wei, T. Kammler, J. Höntschel, H. Bierstedt, T. Feudel, K. Frohberg, M. Gerhardt, A. Hellmich, K. Hempel, J. Hohage, P. Javorka, J. Klais, G. Koerner, M. Lenski, A. Neu, R. Otterbach, P. Press, C. Reichel, M.Trentsch, B. Trui, H. Salz, M. Schaller, H.-J. Engelmann, O. Herzog, H. Ruelke, P. Hübler, R. Stephan, D. Greenlaw, M. Raab, and N. Kepler. *Integration and optimization of embedded-SiGe, compressive and tensile stressed liner films, and stress memorization in advanced SOI CMOS technologies.* In Proc. Intl. Electron Devices Meeting, pages 233–236, 2005.

[Hu96] Chenming Hu. *Gate oxide scaling limits and projection.* In Proc. Intl. Electron Devices Meeting, pages 319–322, 1996.

[Huang01] L.-J. Huang, J. Chu, S. Goma, C. Emic, S. Koester, D. Canaperi, P. Mooney, S. Cordes, J. Speidell, R. Anderson, and H. Wong. *Carrier mobility enhancement in strained Si-on-Insulator fabricated by wafer bonding.* In VLSI Symp. Tech.Dig., pages 57–58, 2001.

[Intel07] Intel. *45nm technology*, 2007. http://www.intel.com/technology/silicon-/45nm_technology.htm.

[Irie04] H. Irie, K. Kita, K. Kyuno, and A. Toriumi. *In-plane mobility anisotropy and universality under uni-axial strains in n- and p-MOS inversion layers on (100), (110), and (111) Si.* In Proc. Intl. Electron Devices Meeting, pages 225–228, 2004.

[Ito00] S. Ito, H. Namba, K. Yamaguchi, T. Hirata, K. Ando, S. Koyama, S. Kuroki, N. Ikezawa, T. Suzuki, T. Saitoh, , and T. Horiuchi. *Mechanical stress effect of etch-stop nitride and its impact on deep submicron transistor design.* In Proc. Intl. Electron Devices Meeting, pages 247–251, 2000.

[Jacoboni77] C. Jacoboni, C. Canali, G. Ottaviani, and A. Alberigi Quaranta. *A review of some charge transport properties of silicon.* Solid-State Electron., vol. 20, pages 77–89, Feb 1977.

[Jacoboni83] C. Jacoboni, and L. Reggiani. *The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials.* Reviews of Modern Physics, vol. 55, no. 3, pages 645–705, 1983.

[Jan05] C.-H. Jan, P. Bai, J. Choi, G. Curello, S. Jacobs, J. Jeong, K. Johnson, D. Jones, S. Klopcic, J. Lin, N. Lindert, A. Lio, S. Natarajan, J. Neirynck, P. Packan, J. Park, I. Post et al. *A 65nm ultra low power logic platform technology using uni-axial strained silicon transistors.* In Proc. Intl. Electron Devices Meeting, pages 60–63, 2005.

[Jungemann93] C. Jungemann, A. Edmunds, and W.L. Engl. *Simulation of linear and non-linear electron transport in homogeneous silicon inversion layers.* Solid-State Electron., vol. 36, no. 11, pages 1529–1540, 1993.

[Jungemann03] C. Jungemann, and B. Meinerzhagen. “Hierarchical Device Simulation: The Monte Carlo Perspective,”. Computational Microelectronics. Springer, 2003.

[Kanda91] Yozo Kanda, and Katuhisa Suzuki. *Origin of the shear piezoresistance coefficient* π_{44} *of n-type silicon.* Physical Review B, vol. 43, no. 8, pages 6754–6756, Mar 1991.

[Kane56] E.O. Kane. *Energy band structure in p-type germanium and silicon.* J.Phys.Chem.Solids, vol. 1, no. 1-2, pages 82–99, 1956.

[Karner06] M. Karner, A. Gehring, S. Holzer, M. Pourfath, M. Wagner, H. Kosina, T. Grasser, and S. Selberherr. *VSP-a multi-purpose Schrödinger-Poisson solver for TCAD applications.* In 11th International Workshop on Computational Electronics Book of Abstracts, pages 255–256, 2006.

[Khamankar04] R. Khamankar, H. Bu, C. Bowen, S. Chakravarthi, P. R. Chidambaram, M. Bevan, A. Krishnan, H. Niimi, B. Smith, J. Blatchford, B. Hornung, J. P. Lu, P. Nicollian, B. Kirkpatrick, D. Miles, M. Hewson, D. Farber, L. Hall, H. Alshareef, A. Varghese, A. Gurba, V. Ukraintsev, B. Rathsack, J. DeLoach, J. Tran, C. Kaneshige, M. Somervell, S. Aur, C. Machala, and T. Grider. *An enhanced 90nm high performance technology with strong performance improvements from stress and mobility increase through simple process changes.* In VLSI Symp. Tech. Dig., pages 162–163, 2004.

[Kittel96] C. Kittel. “Introduction to solid state physics,”. John Wiley & Sons, 7 edition, 1996.

[Kleinman62] Leonard Kleinman. *Deformation potentials in silicon. I. Uniaxial strain.* Physical Review, vol. 128, no. 6, pages 2614–2621, Dec 1962.

[Kondo01] M. Kondo, and H. Tanimoto. *An accurate Coulomb mobility model for MOS inversion layer and its application to NO-oxynitride devices.* IEEE Trans.Electron Devices, vol. 48, pages 265–270, 2001.

[Kosina00a] H. Kosina, M. Nedjalkov, and S. Selberherr. *A Monte Carlo Method for Small Signal Analysis of the Boltzmann Equation.* J.Appl.Phys., vol. 87, pages 4308–4314, 2000.

[Kosina00b] H. Kosina, M. Nedjalkov, and S. Selberherr. *Theory of the Monte Carlo method for semiconductor device simulation.* IEEE Trans.Electron Devices, vol. 47, no. 10, pages 1899–1908, 2000.

[Kresse93] G. Kresse, and J. Hafner. *Ab initio molecular dynamics for liquid metals.* Physical Review B, vol. 47, no. 1, pages 558–561, Jan 1993.

[Kresse94] G. Kresse, and J. Hafner. *Ab initio molecular-dynamics simulation of the liquid-metal amorphous-semiconductor transition in germanium.* Physical Review B, vol. 49, no. 20, pages 14251–14269, May 1994.

[Kresse96a] G. Kresse, and J. Furthmüller. *Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set.* Comput. Mat. Sci., vol. 6, pages 15–50, 1996.

[Kresse96b] G. Kresse, and J. Furthmüller. *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set.* Physical Review B, vol. 54, no. 16, pages 11169–11186, Oct 1996.

[Kresse99] G. Kresse, and D. Joubert. *From ultrasoft pseudopotentials to the projector augmented-wave method.* Physical Review B, vol. 59, no. 3, pages 1758–1775, Jan 1999.

[Kurosawa66] T. Kurosawa. *Monte Carlo calculation of hot electron problems.* J.Phys.Soc.Jap, vol. 21, pages 424–426, 1966.

[Laude71] Lucien D. Laude, Fred H. Pollak, and Manuel Cardona. *Effects of uniaxial stress on the indirect exciton spectrum of silicon.* Physical Review B, vol. 3, no. 8, pages 2623–2636, Apr 1971.

[Laux88] S. E. Laux, and M. V. Fischetti. *Monte-Carlo simulation of submicrometer Si n-MOSFETs at 77 and 300 K.* IEEE Electron Device Lett., vol. 9, pages 467–469, 1988.

[Lee02] Hyunjin Lee, Jongho Lee, and Hyungcheol Shin. *DC and AC characteristics of sub-50-nm MOSFETs with source/drain-to-gate nonoverlapped structure.* IEEE Trans.Nanotechnology, vol. 1, no. 4, pages 219–225, Dec. 2002.

[Levinshtein99] M. Levinshtein, S. Rumyantsev, and M. Shur, editeurs. “Handbook Series on Semiconductor Parameters,”, volume 1,2. World Scientific, London, 1999.

[Lim04] Ji-Song Lim, S. E. Thompson, and J. G. Fossum. *Comparison of threshold-voltage shifts for uniaxial and biaxial tensile-stressed n-MOSFETs.* IEEE Electron Device Lett., vol. 25, pages 731–733, 2004.

[Lombardi88] C. Lombardi, S. Manzini, A. Saporito, and M. Vanzi. *A physically based mobility model for numerical simulation of nonplanar devices.* IEEE Trans.Computer-Aided Design, vol. 7, pages 1164–1171, 1988.

[Lugli85] P. Lugli, and D. K. Ferry. *Degeneracy in the ensemble Monte Carlo method for high-field transport in semiconductors.* IEEE Trans.Electron Devices, vol. 32, pages 2431–2437, 1985.

[Luttinger55] J. M. Luttinger, and W. Kohn. *Motion of electrons and holes in perturbed periodic fields.* Physical Review, vol. 97, no. 4, pages 869–883, Feb 1955.

[Luttinger56] J. M. Luttinger. *Quantum theory of cyclotron resonance in semiconductors: General theory.* Phys. Rev., vol. 102, no. 4, pages 1030–1041, May 1956.

[Maruyama90] T. Maruyama, S. Zaima, Y. Koide, Y. Kanda, and Y. Yasuda. *Anisotropy of piezoresistance in n-channel inversion layers of metal-oxide-semiconductor transistors on (001)Si.* J.Appl.Phys., vol. 68, no. 11, pages 5687–5691, 1990.

[Matsumoto02] T. Matsumoto, S. Maeda, H. Dang, T. Uchida, K. Ota, Y. Hirano, H. Sayama, T. Iwamatsu, T. Ipposhi, H. Oda, S. Maegawa, Y. Inoue, , and T. Nishmura. *Novel SOI wafer engineering using low stress and high mobility CMOSFET with* $\langle 100 \rangle$ *channel for embedded RF/analog applications.* In Proc. Intl. Electron Devices Meeting, pages 663–666, 2002.

[Mohta05] N. Mohta, and S. E. Thompson. *Mobility enhancement.* IEEE Circuits and Devices Magazine, vol. 21, pages 18–23, 2005.

[Mujtaba95] S. A. Mujtaba. *Advanced mobility models for design and simulation of deep submicron MOSFETs.* PhD thesis, Stanford University, CA, 1995.

[Nielsen85] O.H. Nielsen, and R.M. Martin. *Stresses in semiconductors: Ab initio calculations on Si, Ge, and GaAs.* Physical Review B, vol. 32, no. 6, pages 3792–3805, 1985.

[Nowotny98] Helmut Nowotny. *Theoretische Festkörperphysik II*, 1998. Vorlesungsskript Technische Universität Wien.

[Ota02] K. Ota, K. Sugihara, H. Sayama, T. Uchida, H. Oda, T. Eimori, H. Morimoto, and Y. Inoue. *Novel locally strained channel technique for high performance 55nm CMOS.* In Proc. Intl. Electron Devices Meeting, pages 27–30, 2002.

[Ouyang05] Q. Ouyang, M. Yang, J. Holt, S. Panda, H. Chen, H. Utomo, M. Fischetti, N. Rovedo, J. Li, N. Klymko, H. Wildman, T. Kanarsky, G. Costrini, D.M. Fried, A. Bryant, J.A. Ott, M. Ieong, and C.Y. Sung. *Investigation of CMOS devices with embedded SiGe source/drain on hybrid orientation substrates.* In VLSI Symp. Tech. Dig., pages 28–29, 2005.

[Pidin04] S. Pidin, T. Mori, K. Inoue, S. Fukuta, N. Itoh, E. Mutoh, K. Ohkoshi, R. Nakamura, K. Kobayashi, K. Kawamura, T. Saiki, S. Fukuyama, S. Satoh, M. Kase, and K. Hashimoto. *A novel strain enhanced CMOS architecture using selectively deposited high tensile and high compressive silicon nitride films.* In Proc. Intl. Electron Devices Meeting, pages 213–216, 2004.

[Prange68] R.E. Prange, and T.W Nee. *Quantum spectroscopy of the low-field oscillations in the surface impedance.* Physical Review, vol. 168, pages 779–786, 1968.

[Price81] P. J. Price. *Two-dimensional electron transport in semiconductor layers. I. Phonon scattering.* Annals of Physics, vol. 133, pages 217–239, May 1981.

[Reggiani73] L. Reggiani, and C. Calandra. *Bloch states mixing in Si conduction band.* Physics Letters A, vol. 43, pages 339–340, Mar 1973.

[Reggiani02] S. Reggiani, A. Valdinoci, L. Colalongo, M. Rudan, G. Baccarani, A. Stricker, F. Illien, N. Felber, W. Fichtner, S. Mettler, S. Lindenkreuz, and L. Zullino. *Surface mobility in silicon at large operating temperature.* In Simulation of Semiconductor Processes and Devices, 2002. SISPAD 2002. International Conference on, pages 15–20, 2002.

[Rieger93] Martin M. Rieger, and Peter Vogl. *Electronic-band parameters in strained* $Si_{1-x}Ge_x$ *alloys on* $Si_{1-y}Ge_y$ *substrates.* Physical Review B, vol. 48, no. 19, pages 14276–14287, Nov 1993.

[Rim00] K. Rim, J. L. Hoyt, and J. F. Gibbons. *Fabrication and analysis of deep submicron strained-Si n-MOSFET's.* IEEE Trans.Electron Devices, vol. 47, pages 1406–1415, 2000.

[Rim02] K. Rim, J. Chu, H. Chen, K. Jenkins, T. Kanarsky, K. Lee, A. Mocuta, H. Zhu, R. Roy, J. Newbury, J. Ott, K. Petrarca, P. Mooney, D. Lacey, S. Koester, K. Chan, D. Boyd, M. Ieong, and H. Wong. *Characteristics and device design of sub-100 nm strained Si n- and p-MOSFETs.* In VLSI Symp. Tech.Dig., pages 98–99, 2002.

[Rim03] K. Rim, K. Chan, L. Shi, D. Boyd, J. Ott, N. Klymko, F. Cardone, L. Tai, S. Koester, M. Cobb, D. Canaperi, B. To, E. Duch, I. Babich, R. Carruthers, P. Saunders, G. Walker, Y. Zhang, M. Steen, and M. Ieong. *Fabrication and mobility characteristics of ultra-thin strained Si directly on insulator (SSDOI) MOSFETs.* In Proc. Intl. Electron Devices Meeting, pages 49–52, 2003.

[Roldan96] J. B. Roldan, F. Gamiz, J. A. Lopez-Villanueva, and J. E. Carceller. *A Monte Carlo study on the electron-transport properties of high-performance strained-Si on relaxed* $Si_{1-x}Ge_x$ *channel MOSFETs.* J.Appl.Phys., vol. 80, no. 9, pages 5121–5128, 1996.

[Roychoudhury80] D. Roychoudhury, and P. K. Basu. *Mobility of electrons in a quantized inversion layer due to phonon scattering.* Physical Review B, vol. 22, no. 12, pages 6325–6329, 1980.

[Sabnis79] A. G. Sabnis, and J. T. Clemens. *Characterization of the electron mobility in the inverted* $\langle 100 \rangle$ *Si surface.* In Proc. Intl. Electron Devices Meeting, volume 25, pages 18–21, 1979.

[Sadaka04] M. Sadaka, AV.Y. Thean, A. Barr, D. Tekleab, S. Kalpat, and T. White. *Fabrication and operation of sub-50 nm strained-Si on* $Si_{1-x}Ge_x$ *on Insulator (SGOI) CMOSFETs.* In Proceedings IEEE International SOI Conference, pages 209–211, 2004.

[Scott99] G. Scott, J. Lutze, M. Rubin, F. Nouri, and M. Manley. *NMOS drive current reduction caused by transistor layout and trench isolation induced stress.* In Proc. Intl. Electron Devices Meeting, pages 827–830, 1999.

[Seitz35] Frederick Seitz. *The theoretical constitution of metallic lithium.* Physical Review, vol. 47, no. 5, pages 400–412, Mar 1935.

[Selberherr89] S. Selberherr. *MOS device modeling at 77 K.* IEEE Trans.Electron Devices, vol. 36, pages 1464–1474, 1989.

[Sheraw05] C.D. Sheraw, M. Yang, D.M. Fried, G. Costrini, T. Kanarsky, W.H. Lee, V. Chan, M.V. Fischetti, J. Holt et al. *Dual stress liner enhancement in hybrid orientation technology.* In VLSI Symp. Tech.Dig., pages 12–13, 2005.

[Shimizu01] A. Shimizu, K. Hachimine, N. Ohki, H. Ohta, M. Koguchi, Y. Nonaka, H. Sato, , and F. Ootsuka. *Local mechanical-stress control (LMC): A new technique for CMOS-performance enhancement.* In Proc. Intl. Electron Devices Meeting, pages 433–436, 2001.

[Shirahata92] M. Shirahata, H. Kusano, N. Kotani, S. Kusanoki, and Y. Akasaka. *A mobility model including the screening effect in MOS inversion layer.* IEEE Trans.Computer-Aided Design, vol. 11, pages 1114–1119, 1992.

[Shoji97] M. Shoji, Y. Omura, and M. Tomizawa. *Physical basis and limitation of universal mobility behavior in fully depleted silicon-on-insulator Si inversion layers.* J.Appl.Phys., vol. 81, pages 786–794, January 1997.

[SIA06] Semiconductor Industry Association SIA. *International Technology Roadmap for Semiconductors - 2006 Update*, 2006. http://www.itrs.net.

[Singh93] Jasprit Singh. "Physics of semiconductors and their heterostructures,". McGraw-Hill, 1993.

[Skotnicki05] T. Skotnicki, J. A. Hutchby, Tsu-Jae King, H. S. P. Wong, and F. Boeuf. *The end of CMOS scaling: toward the introduction of new materials and structural changes to improve MOSFET performance.* IEEE Circuits & Devices, vol. 21, pages 16–26, 2005.

[Smirnov03] S. Smirnov, H. Kosina, M. Nedjalkov, and S. Selberherr. *Monte Carlo method for modeling of small signal response including the Pauli exclusion principle.* J.Appl.Phys., vol. 94, no. 9, pages 5791–5799, 2003.

[Smith54] Charles S. Smith. *Piezoresistance effect in germanium and silicon.* Physical Review, vol. 94, no. 1, pages 42–49, Apr 1954.

[Stanley98] John Stanley, and Neil Goldsman. *New irreducible wedge for scattering rate calculations in full-zone Monte Carlo simulations.* VLSI Design, vol. 8, no. 1–4, pages 413–417, 1998.

[Stathis98] J. H. Stathis, and D. J. DiMaria. *Reliability projection for ultra-thin oxides at low voltage.* In Proc. Intl. Electron Devices Meeting, pages 167–170, 1998.

[Steegen99] A. Steegen, M. Stucchi, A. Lauwers, and K. Maex. *Silicide induced pattern density and orientation dependent transconductance in MOS transistors.* In Proc. Intl. Electron Devices Meeting, pages 497–500, 1999.

[Stern67] Frank Stern, and W. E. Howard. *Properties of semiconductor surface inversion layers in the electric quantum limit.* Physical Review, vol. 163, no. 3, pages 816–835, Nov 1967.

[Sverdlov06] V. Sverdlov, E. Ungersboeck, H. Kosina, and S. Selberherr. *Orientation dependence of the low field mobility in double- and single-gate SOI FETs.* Proceedings ESSDERC 2006, pages 178–181, 2006.

[Takagi94] S. Takagi, A. Toriumi, M. Iwase, and H. Tango. *On the Universality of Inversion Layer Mobility in Si MOSFET's: Part II-effects of surface orientation.* IEEE Trans.Electron Devices, vol. 41, no. 12, pages 2363–2368, 1994.

[Thompson04] S.-E. Thompson, M. Armstrong, C. Auth, M. Alavi, and M. Buehler. *A 90-nm logic technology featuring strained-silicon.* IEEE Trans.Electron Devices, vol. 51, no. 11, pages 1790–1797, 2004.

[Thompson06] S.-E. Thompson, G. Sun, Y.S. Choi, and T. Nishida. *Uniaxial-process-induced strained-Si: Extending the CMOS Roadmap.* IEEE Trans.Electron Devices, vol. 53, no. 5, pages 1010–1020, 2006.

[Tinkham64] M. Tinkham. “Group theory and quantum mechanics,”. McGraw-Hill, New York, 1964.

[Tosic81] T. I. Tosic, D. A. Tjapkin, and M. M. Jevtic. *Mobility of majority carriers in doped noncompensated silicon.* Solid-State Electron., vol. 24, pages 577–582, Jun 1981.

[Tsutsui05] G. Tsutsui, M. Saitoh, T. Saraya, T. Nagumo, and T. Hiramotoy. *Mobility enhancement due to volume inversion in (110)-oriented ultra-thin body double-gate nMOSFETs with body thickness less than 5 nm.* In Proc. Intl. Electron Devices Meeting, pages 747–750, 2005.

[Uchida03] K. Uchida, J. Koga, and S. Takagi. *Experimental study on carrier transport mechanisms in double- and single-gate ultrathin-body MOSFETs - Coulomb scattering, volume inversion, and δTSOI-induced scattering.* In Proc. Intl. Electron Devices Meeting, pages 805–808, 2003.

[Uchida05] K. Uchida, T. Krishnamohan, K.C. Saraswat, and Y.Nishi. *Physical mechanisms of electron mobility enhancement in uniaxial stressed MOSFETs and impact of uniaxial stress engineering in ballistic regime.* In Proc. Intl. Electron Devices Meeting, pages 135–138, 2005.

[Ungersboeck06a] E. Ungersboeck, S. Dhar, G. Karlowatz, H. Kosina, and S. Selberherr. *Physical modeling of electron mobility enhancement for arbitrarily strained silicon.* In 11th International Workshop on Computational Electronics Book of Abstracts, pages 141–142, 2006.

[Ungersboeck06b] E. Ungersboeck, and H. Kosina. *Monte Carlo study of electron transport in strained silicon inversion layers.* J.Comput.Electronics, vol. 5, no. 2-3, pages 79–83, July 2006.

[Uppal04] S. Uppal, M. Bollani, A.F.W. Willoughby, J.M. Bonar, R.J.H. Morris, and M.G. Dowsett. *Diffusion of ion-implanted boron in high Ge content SiGe alloys.* In Electrocemical Society Proceedings, volume 07, pages 159–165, 2004.

[Van de Walle86] Chris G. Van de Walle, and Richard M. Martin. *Theoretical calculations of heterojunction discontinuities in the Si/Ge system.* Physical Review B, vol. 34, no. 8, pages 5621–5634, Oct 1986.

[Vasileska97] D. Vasileska, and D. K. Ferry. *Scaled silicon MOSFET's: universal mobility behavior.* IEEE Electron Device Lett., vol. 44, pages 577–583, 1997.

[Vasileska00] D. Vasileska, and Z. Ren. *SCHRED 2.0 User's Manual.* http://www.nanohub.org, 2000.

[Venturi89] F. Venturi, R.K. Smith, E.C. Sangiorgi, M.R. Pinto, and B. Riccó. *A general purpose device simulator coupling poisson and Monte Carlo transport with applications to deep submicron MOSFET's.* IEEE Trans.Computer-Aided Design, vol. 8, no. 4, pages 360–369, 1989.

[VMC2.006] VMC2.0. *Vienna Monte Carlo 2.0 User's Guide.* Institut für Mikroelektronik, http://www.iue.tuwien.ac.at/software, Technische Universität Wien, Austria, 2006.

[Wallmark75] J. T. Wallmark. *Fundamental physical limitations in integrated electronic circuits.* Inst. Phys. Conf. Ser., vol. 25, pages 133–167, 1975.

[Wang03] Y. G. Wang, D. B. Scott, J. Wu, J. L. Waller, J. Hu, K. Liu, and V. Ukraintsev. *Effects of uniaxial mechanical stress on drive current of 0.13 μm MOSFETs.* IEEE Trans.Electron Devices, vol. 50, pages 529–531, 2003.

[Welser92] J. Welser, J.L. Hoyt, and J.F. Gibbons. *NMOS and PMOS transistors fabricated in strained silicon/relaxed silicon-germanium structures.* In Proc. Intl. Electron Devices Meeting, pages 1000–1002, 1992.

[Welser94] J. Welser, J.L. Hoyt, and J.F. Gibbons. *Electron Mobility Enhancement in Strained-Si N-Type Metal-Oxide-Semiconductor Field-Effect Transistors.* IEEE Electron Device Lett., vol. 15, no. 3, pages 100–102, 1994.

[Yamakawa96] S. Yamakawa, H. Ueno, K. Taniguchi, C. Hamaguchi, K. Miyatsuji, K. Masaki, and U. Ravaioli. *Study of interface roughness dependence of electron mobility in Si inversion layers using the Monte Carlo method.* J.Appl.Phys., vol. 79, no. 2, pages 911–916, 1996.

[Yang04] H. S. Yang, R. Malik, S. Narasimha, Y. Li, R. Divakaruni, P. Agnello, S. Allen, A. Antreasyan, J. C. Arnold, K. Bandy et al. *Dual stress liner for high performance sub-45nm gate length SOI CMOS manufacturing.* In Proc. Intl. Electron Devices Meeting, pages 1075–1077, 2004.

[Yang06] M. Yang, V.W.C. Chan, K.K Chan, L. Shi, D.M. Fried, J.H. Stathis et al. *Hybrid-orientation technology (HOT): Opportunities and challenges.* IEEE Trans.Electron Devices, vol. 53, pages 965–978, 2006.

[Yu03] P. Yu, and M. Cardona. "Fundamentals of Semiconductors,". Springer, 2003.

[Zhang05] D. Zhang, B.Y. Nguyen, T. White, B. Goolsby et al. *Embedded SiGe S/D PMOS on thin body SOI substrate with drive current enhancement.* In VLSI Symp. Tech. Dig., pages 26–27, 2005.

[Zhu04] Shiyang Zhu, H.Y. Yu, S.J. Whang, J.H. Chen, Chen Shen, Chunxiang Zhu, S.J. Lee, M.F. Li, D.S.H. Chan, W.J. Yoo, A. Du, C.H. Tung, J. Singh, A. Chin, and D.L. Kwong. *Schottky-barrier S/D MOSFETs with high-k gate dielectrics and metal-gate electrode.* IEEE Electron Device Lett., vol. 25, no. 5, pages 268–270, May 2004.

Own Publications

[1] S. Dhar, E. Ungersbock, H. Kosina, T. Grasser, and S. Selberherr, "Electron Mobility Model for ⟨110⟩ Stressed Silicon Including Strain-Dependent Mass," *IEEE Trans.Nanotechnology*, vol. 6, no. 1, pp. 97–100, 2007.

[2] V. Sverdlov, E. Ungersboeck, H. Kosina, and S. Selberherr, "Volume inversion mobility in SOI MOSFETs for different thin body orientations," *Solid-State Electron.*, vol. 51, pp. 299–305, 2007.

[3] S. Dhar, H. Kosina, G. Karlowatz, S. E. Ungersboeck, T. Grasser, and S. Selberherr, "High-Field Electron Mobility Model for Strained-Silicon Devices," *IEEE Trans.Electron Devices*, vol. 53, pp. 3054–3062, Dec. 2006.

[4] S. Dhar, H. Kosina, G. Karlowatz, E. Ungersboeck, T. Grasser, and S. Selberherr, "A tensorial high-field electron mobility model for strained silicon," *SiGe Technology and Device Meeting, ISTDM 2006*, pp. 72–73, 2006.

[5] S. Dhar, E. Ungersboeck, H. Kosina, T. Grasser, and S. Selberherr, "Analytical Modeling of Electron Mobility in Strained Germanium," in *Proc. Simulation of Semiconductor Processes and Devices*, (Monterey, USA), pp. 39–42, Sept. 2006.

[6] S. Dhar, E. Ungersboeck, H. Kosina, T. Grasser, and S. Selberherr, "Electron Mobility Model for ⟨110⟩ Stressed Si Including Strain-Dependent Mass," in *Abstracts IEEE 2006 Silicon Nanoelectronics Workshop*, (Honolulu, USA), pp. 153–154, June 2006.

[7] S. Dhar, E. Ungersboeck, M. Nedjalkov, and V. Palankovski, "Monte Carlo Simulation of the Electron Mobility in Strained Silicon," in *The Fifteenth International Scientific and Applied Science Conference Book 2*, (Sozopol, Bulgaria), pp. 169–173, 2006.

[8] G. Karlowatz, E. Ungersboeck, W. Wessner, and H. Kosina, "Full-Band Monte Carlo Analysis of Electron Transport in Arbitrarily Strained Silicon," in *Proc. Simulation of Semiconductor Processes and Devices*, (Monterey, USA), pp. 63–66, Sept. 2006.

[9] G. Karlowatz, E. Ungersboeck, W. Wessner, H. Kosina, and S. Selberherr, *Analysis of Hole Transport in Arbitrarily Strained Germanium*, vol. 3 of *ECS Transactions*. The Electrochemical Society, 2006.

[10] M. Karner, E. Ungersboeck, A. Gehring, S. Holzer, H. Kosina, and S. Selberherr, "Strain Effects on Quasi-Bound State Tunneling in Advanced SOI CMOS Technologies," in *Proc. Simulation of Semiconductor Processes and Devices*, (Monterey, USA), pp. 314–317, Sept. 2006.

[11] T. Krishnamohan, C. Jungemann, D. Kim, E. Ungersboeck, S. Selberherr, P. Wong, Y. Nishi, and K. Saraswat, "Theoretical Investigation Of Performance In Uniaxially- and Biaxially-Strained Si, SiGe and Ge Double-Gate p-MOSFETs," in *Proc. Intl. Electron Devices Meeting*, pp. 36.2.1–36.2.4, 11-13 Dec. 2006.

[12] V. Palankovski, A. Marchlewski, E. Ungersboeck, and S. Selberherr, "Identification of Tranport Parameters for Gallium Nitride Based Semiconductor Devices," in *Proc. of 5th International Symposium on Mathematical Modeling*, (Vienna, Austria), pp. 14–1–14–9, 2006.

[13] V. Sverdlov, E. Ungersboeck, H. Kosina, and S. Selberherr, "Orientation Dependence of the Low Field Mobility in Double-and Single-gate SOI FETs," in *Proc. 36th European Solid-State Device Research Conf.*, (Montreux, Switzerland), pp. 178–181, 2006.

[14] V. Sverdlov, E. Ungersboeck, and H. Kosina, "Mobility for High Effective Field in Double-Gate and Single-Gate SOI for Different Substrate Orientations," in *EUROSOI 2006 Second Workshop of the Thematic Network on Silicon On Insulator Technology, Devices and Circuits*, (Grenoble, France), pp. 133–134, 2006.

[15] V. Sverdlov, E. Ungersboeck, H. Kosina, and S. Selberherr, "Comparative Study of Low-Field Mobilities in Double- and Single-Gate Ultra-Thin Body SOI for Different Substrate Orientations," in *Abstracts IEEE 2006 Silicon Nanoelectronics Workshop*, (Honolulu, USA), pp. 17–18, June 2006.

[16] V. Sverdlov, E. Ungersboeck, and H. Kosina, "Mobility Modeling in SOI FETs for Different Substrate Orientations and Strain Conditions," in *NATO Advanced Research Workshop Conference Abstracts*, (Sudak, Ukraine), pp. 77–78, 2006.

[17] E. Ungersboeck and H. Kosina, "Monte Carlo Study of Electron Transport in Strained Silicon Inversion Layers," *J.Comput.Electronics*, vol. 5, pp. 79–83, July 2006.

[18] E. Ungersboeck, S. Dhar, G. Karlowatz, H. Kosina, and S. Selberherr, "Physical Modeling of Electron Mobility Enhancement for Arbitrarily Strained Silicon," in *11th International Workshop on Computational Electronics Book of Abstracts*, pp. 141–142, 2006.

[19] E. Ungersboeck, V. Sverdlov, H. Kosina, and S. Selberherr, "Electron Inversion Layer Mobility Enhancement by Uniaxial Stress on (001) and (110) Oriented MOSFETs," in *Proc. Simulation of Semiconductor Processes and Devices*, (Monterey, USA), pp. 43–46, Sept. 2006.

[20] E. Ungersboeck, V. Sverdlov, H. Kosina, and S. Selberherr, "Strain Engineering for CMOS Devices," in *2006 Eight International Conference on Solid-State and Integrated Cicuit Technology Proceedings (Part 1 of 3)*, pp. 124–127, 2006.

[21] E. Ungersboeck, V. Sverdlov, H. Kosina, and S. Selberherr, *Low-Field Electron Mobility in Stressed UTB SOI MOSFETs for Different Substrate Orientations*, vol. 3 of *ECS Transactions*. The Electrochemical Society, 2006.

[22] E. Ungersboeck, V. Sverdlov, H. Kosina, and S. Selberherr, *Modeling of Advanced Semiconductor Devices*, vol. 4 of *ECS Transactions*. The Electrochemical Society, 2006.

[23] E. Ungersboeck, V. Sverdlov, H. Kosina, and S. Selberherr, "Low-Field Electron Mobility in Stressed UTB SOI MOSFETs for Different Substrate Orientations," in *Meeting Abstracts: 210th Meeting of The Electrochemical Society*, p. 1397, 2006.

[24] E. Ungersboeck, H. Kosina, and S. Selberherr, "The Influence of Stress on Inversion Layer Mobility," in *Abstracts Advanced Heterostructure Workshop*, (Kona, USA), pp. TH–2, 2006.

[25] S. Dhar, G. Karlowatz, E. Ungersboeck, and H. Kosina, "Numerical and Analytical Modeling of the High-Field Electron Mobility in Strained Silicon," in *Proc. Simulation of Semiconductor Processes and Devices*, (Tokio, Japan), pp. 223–226, 01-03 Sept. 2005.

[26] S. Dhar, H. Kosina, V. Palankovski, E. Ungersboeck, and S. Selberherr, "Electron Mobility Model for Strained-Si Devices," *IEEE Trans.Electron Devices*, vol. 52, no. 4, pp. 527–533, 2005.

[27] S. Dhar, G. Karlowatz, E. Ungersboeck, H. Kosina, and S. Selberherr, "Modeling of Velocity Field Characteristics in Strained Si," in *Proceedings of the Fourteenth International Workshop on the Physics of Semiconductor Devices*, pp. 1060–1063, 2005.

[28] S. Dhar, H. Kosina, V. Palankovski, E. Ungersboeck, and S. Selberherr, "A Physically-Based Electron Mobility Model for Strained Si Devices," in *Proc. NSTI Nanotech*, pp. 13–16, 2005.

[29] M. Pourfath, A. Gehring, E. Ungersboeck, H. Kosina, S. Selberherr, B. Cheong, and W. Park, "Separated carrier injection control in carbon nanotube field-effect transistors," *J.Appl.Phys.*, vol. 97, p. 106103, 2005.

[30] M. Pourfath, E. Ungersboeck, A. Gehring, B. Cheong, W. Park, H. Kosina, and S. Selberherr, "Optimization of Schottky barrier carbon nanotube field effect transistors," *Microelectronic Engineering*, vol. 81, no. 2-4, pp. 428–433, 2005.

[31] M. Pourfath, E. Ungersboeck, A. Gehring, H. Kosina, S. Selberherr, W. PARK, and B. Cheong, "Numerical Analysis of Coaxial Double Gate Schottky Barrier Carbon Nanotube Field Effect Transistors," *J.Comput.Electronics*, vol. 4, no. 1, pp. 75–78, 2005.

[32] E. Ungersboeck, M. Pourfath, H. Kosina, A. Gehring, B. Cheong, W. Park, and S. Selberherr, "Optimization of Single-Gate Carbon-Nanotube Field-Effect Transistors," *IEEE Trans.Nanotechnology*, vol. 4, no. 5, p. 533, 2005.

[33] E. Ungersboeck and H. Kosina, "The Effect of Degeneracy on Electron Transport in Strained Silicon Inversion Layers," in *Proc. Simulation of Semiconductor Processes and Devices*, (Tokio, Japan), pp. 311–314, 01-03 Sept. 2005.

[34] E. Ungersboeck and H. Kosina, "Monte Carlo study of electron transport in strained silicon inversion layers," in *15th Workshop on Modeling and Simulation of Electron Devices*, (Pisa, Italy), pp. 10–11, July 2005.

[35] S. Dhar, H. Kosina, V. Palankovski, E. Ungersboeck, and S. Selberherr, "Modeling of Electron Mobility in Strained Si Devices," in *Proc. SAFE*, pp. 793–796, 2004.

[36] M. Nedjalkov, H. Kosina, E. Ungersboeck, and S. Selberherr, "A quasi-particle model of the electron-Wigner potential interaction," *Semicond.Sci.Technol.*, vol. 19, no. 4, pp. S226–S228, 2004.

[37] M. Pourfath, E. Ungersboeck, A. Gehring, B.-H. Cheong, H. Kosina, and S. Selberherr, "Three-Dimensional Analysis of Schottky Barrier Carbon Nanotube Field Effect Transistors," in *Proc. Simulation of Semiconductor Processes and Devices*, (Munich, Germany), pp. 149–152, Sept. 2004.

[38] M. Pourfath, E. Ungersboeck, A. Gehring, W. Park, B.-H. Cheong, H. Kosina, and S. Selberherr, "Numerical Analysis of Coaxial Double Gate Schottky Barrier Carbon Nanotube Field Effect Transistors," in *10th International Workshop on Computational Electronics Book Abstracts*, (West Lafayette, USA), pp. 237–238, 2004.

[39] M. Pourfath, E. Ungersboeck, A. Gehring, B. Cheong, W. Park, H. Kosina, and S. Selberherr, "Improving the ambipolar Behavior of Schottky Barrier Carbon Nanotube Field Effect Transistors," in *Proc. 34th European Solid-State Device Research Conf.*, (Leuven, Belgium), pp. 429–432, 2004.

[40] M. Pourfath, E. Ungersboeck, A. Gehring, B.-H. Cheong, W. Park, H. Kosina, and S. Selberherr, "Optimization of Schottky Barrier Carbon Nanotube Field Effect Transistors," in *Nano and Giga Challenges in Microelectronics Book of Abstracts*, (Krakow, Poland), p. 201, 2004.

[41] E. Ungersboeck, M. Pourfath, A. Gehring, H. Kosina, B.-H. Cheong, and S. Selberherr, "Optimization of Carbon Nanotube Field Effect Transistors," in *Proceedings of the Symposium on Nano Device Technology, SNDT*, pp. 43–46, 2004.

[42] E. Ungersboeck, A. Gehring, H. Kosina, S. Selberherr, B.-H. Cheong, and W. Choi, "Analysis of Carrier Transport in Carbon Nanotube FET Devices," in *Proceedings of the Twelfth International Workshop on the Physics of Semiconductor Devices*, pp. 1060–1063, 2003.

[43] E. Ungersboeck, A. Gehring, H. Kosina, S. Selberherr, B.-H. Cheong, and W. B. Choi, "Simulation of Carrier Transport in Carbon Nanotube Field Effect Transistors," in *Proc. 33rd European Solid-State Device Research Conf.*, (Estoril, Portugal), pp. 411–414, 2003.

	Author	Co-Author	Total
Journals & Contributions to books	5	8	13
Conferences	9	21	30
Total	14	29	43

Table 1: Publication statistics.

Curriculum Vitae

[redacted]

Born in Vienna, Austria.

May 1995

High school graduation (*Matura*)
at the Akademisches Gymnasium Wien (with honors).

October 1995 – June 1996

Compulsory military service.

October 1996

Enrolled in Technical Physics at the Vienna University of Technology, Austria.

March 2002

Received degree of *Diplom-Ingenieur* (M.Sc.) in Technical Physics from the Vienna University of Technology.

June 2002

Entered the position of teaching assistant and doctoral program at the Institute for Microelectronics, Vienna University of Technology, under the supervision of Prof. Hans Kosina.

October 2003 – November 2003

Held a position as a visiting researcher at the Samsung Advanced Institute of Technology (Seoul, South Korea).

May 2004

First prize of 11th Symposium on Nano Device Technology (SNDT 2004) student paper contest held in Hsinchu, Taiwan.

January 2004 – December 2006

Coordinator of Sinano project funded by the European Commission under the 6th EU Framework Programme for Research and Technological Development at the Institute for Microelectronics.

October 2006 – April 2007

Research and software development for project with the Semiconductor Technology Development Division, Semiconductor Business Group, Sony Coorporation.