Stretching Silicon:
A Uniaxial and Biaxial Strain Generation Process and the Resulting Mobility Enhancement in Silicon-on-Insulator MOSFETs

Rebecca Lorenz Peterson

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
ELECTRICAL ENGINEERING

SEPTEMBER 2006

© Copyright by Rebecca Lorenz Peterson, 2006.
All rights reserved.

To Derek

Abstract

Strained silicon is of technological interest for its ability to increase charge carrier mobilities in MOSFETs and thereby improve circuit performance without requiring device scaling. At high vertical electric fields, biaxial tensile strain enhances electron mobility while uniaxial compressive strain enhances hole mobility, for example.

In the present work, a process is developed to integrate upon a single wafer, for the first time, uniaxially-strained, biaxially-strained and unstrained silicon islands. The ultra-thin (< 30 nm) strained silicon-on-insulator (SOI) layer does not require relaxed silicon germanium buffers. The tensile strain is uniform over large areas (hundreds of μm^2). The magnitude of the uniaxial strain depends on its crystal direction and is greater than the biaxial strain, due to the Poisson effect. A maximum uniaxial silicon strain of +1.0% is achieved in the <100> crystal direction.

The strain generation method works by lateral expansion of a silicon / silicon germanium bi-layer island on a viscous borophosphorosilicate glass (BPSG) layer during high temperature (> 700°C) anneal. By manipulating the island geometry from squares to narrow rectangles, biaxial and uniaxial silicon strains are achieved. Numerical simulations of the lateral expansion process for islands of various aspect ratios identify process windows and island geometry requirements for maximum uniaxial strain. The modeled and measured strain results agree well.

An undesired buckling process can roughen the film surface during anneals; the buckling and lateral expansion processes compete to determine the final state of the film.

The amplitude of buckling can be reduced and its time scale lengthened by thinning the compliant BPSG layer. Differences in 1-D and 2-D buckling on uniaxially- and biaxially-stressed SiGe islands, respectively, are measured, and by comparison to models, attributed to the different magnitudes of 1-D and 2-D stress present in the films.

Fully-depleted strained-SOI MOSFETs are fabricated using this strain generation process. By adding a silicon nitride barrier layer below the BPSG, device performance is greatly improved. Biaxial tensile strain of 0.38-0.52% enhances electron and hole mobilities by 30-55%. Comparable uniaxial tension parallel to a <100> or <110> channel improves electron mobilities by up to 72% and hole mobilities by up to 54%.

Acknowledgements

This thesis represents the culmination of many years of work. I am indebted to many people who have enabled this research and supported me during my time at Princeton.

First, I thank Professor Jim Sturm for his constant enthusiasm and support for the research. His intellectual curiosity is infectious and his insightful questions have indelibly and productively shaped my work. I also thank Professors Sigurd Wagner and Claire Gmachl for reading this thesis. I am particularly grateful to Professor Wagner for the encouragement he has provided throughout my PhD program.

Special thanks are due to Dr. Karl Hobart and Dr. Fritz Kub at the Naval Research Lab for lending their critical wafer bonding expertise, and for their interest in this project. With gratitude I acknowledge the deposition of BPSG done by Tony Margarella at Northrop Grumman, without which this work would not be possible. I thank Professor Rui Huang at University of Texas at Austin for providing the numerical calculation script for square island expansion, and thank both him and Professor Zhigang Suo of Harvard University for helpful discussions about buckling.

At Princeton, I would like to thank Professor Thomas Duffy and his group for use of the Raman spectroscopy apparatus, Dr. Nan Yao and Garret Lau for assistance with x-ray diffraction, and Joe Palmer and Dr. Helena Gleskova for assistance in the cleanroom. I am grateful for EBSD measurements made by Matt Nowell at EDAX/TSL and SIMS done by Dr. Temel Büyüklimanli at Evans East. Thanks also to the American Association of University Women, who financially supported me in my final year.

A special thanks goes to past and present members of the RTCVD crowd: Malcolm Carroll, Eric Stewart, Haizhou Yin, Xiang-Zheng Bo, Keith Chung, Kun Yao, Weiwei Zheng and Mike Pasqual, for the excellent camaraderie and technical co-operation. Thanks also to the other members of Sturm Group for their kind support. I have appreciated collaboration and friendship with Stéphanie P. Lacour, I-Chun Chen, Ke Long, and Iris Hsu, as well as many others whose paths crossed mine on campus.

A warm thanks to the Jakubowski, LeMorvan, and Wong families, who have seen me through good days and bad, and have brought me joy throughout. I thank my parents for being continual sources of inspiration, and the rest of my family and in-laws for their up-beat encouragement. Finally, no words can express my deepest thanks to my husband Derek, who has walked alongside me through the past six years. His steadfast care, patience and earnest interest have made the work possible.

Contents

Abstract iv

Acknowledgements vi

1 Introduction 1

1.1. Thesis Overview ... 1

1.2. Thesis Outline ... 3

2 Toward Uniaxially-Strained Silicon MOSFETs 5

2.1. Limits of CMOS Scaling ... 5

2.2. Strain-Induced Enhancement of Electron Mobility ... 10

2.2.1. Effect of Biaxial and Uniaxial Strain on the Silicon Conduction Band ... 11

2.2.2. Electron Transport in Strained-Silicon Layers ... 16

2.3. Strain-Induced Enhancement of Hole Mobility ... 24

2.3.1. Effect of Biaxial and Uniaxial Strain on the Silicon Valence Band ... 24

2.3.2. Hole Effective Mass and Transport in Strained-Silicon Layers ... 27

2.4. Fabrication Approaches for Strained Channel MOSFETs ... 35

2.4.1. Substrate-Induced Strain ... 35

2.4.2. Process-Induced Strain ... 39

2.4.3. Post-Process Mechanically-Induced Strain ... 42

2.4.4. Strained Silicon-on-Insulator by Wafer Bonding ... 42

2.5. Summary ... 46

3 Generation of Biaxially- and Uniaxially-Stressed SiGe on BPSG 47

3.1. Introduction to Strain Engineering Using Compliant Substrates ... 47

3.2. Fabrication of Strained SiGe on Compliant BPSG ... 50

3.2.1. BPSG Requirements ... 50

3.2.2. Wafer Bonding and Layer Transfer Technique ... 53

3.3.Characterization of Film Strain and Crystal Orientation 61
3.3.1. X-ray Diffraction 63
3.3.2. Micro-Raman Spectroscopy....................65
3.3.3. Electron Back-Scattering Diffraction 72
3.4.Lateral relaxation of SiGe films on BPSG....................74
3.4.1. Lateral Relaxation Used to Achieve Unstrained SiGe....................74
3.4.2. Defect Density of Relaxed SiGe Films on BPSG....................78
3.5.Generation of Uniaxial Stress in SiGe 80
3.5.1. Stripe Geometry for Uniaxial Stress 80
3.5.2. Contribution of the Poisson Effect and its Crystal-Direction Dependence81
3.5.3. Experimental Results for Uniaxial Stress 83
3.6.Summary 90

4 Dynamic Model of Lateral Relaxation and Process Window for Maximum Asymmetry of Strain 91

4.1.Introduction to the Lateral Relaxation Model....................91
4.2.Lateral Expansion of Square Islands....................95
4.3.Lateral Expansion of Rectangular Islands 104
4.4.Process Windows for Maximum Strain Asymmetry 109
4.5.Comparison of Model and Experimental Data for Rectangular Islands....................111
4.6.Summary 113

5 Generation of Uniaxial Strain in Silicon 114

5.1.Stress Balance of Si/SiGe Bilayers....................114
5.1.1. Introduction to Stress Balance 115
5.1.2. Derivation of the Stress Balance Equation 118
5.1.3. Observation of Biaxial Stress Balance in Square Islands 122
5.2.Uniaxially Strained Silicon 124
5.2.1. Stress Balance of Rectangular Islands 124
5.2.2. Experimental Measurements of Uniaxial Silicon Strain....................125
5.3.Application of Lateral Expansion Model to Uniaxially-Strained Silicon....................129
5.4.Scalability of the Lateral Expansion Process to Smaller Islands....................131

5.5.Summary 131

6 Studies of Buckling of Strained SiGe films on BPSG 133

6.1.Two-Dimensional Buckling in Biaxially-Stressed Films 134

6.1.1. Measurement of 2-D Buckling..134

6.1.2. Modeling 2D buckling: Exponential Growth Rate 139

6.1.3. Modeling 2D buckling: Saturated Buckling Amplitude 142

6.2.Compliant Layer Thinning to Reduce Buckling 148

6.2.1. Benefits of Thin BPSG: Model Predictions...148

6.2.2. Buckling Measurements on Thin BPSG 152

6.2.3. Buckling on Ultra-Thin BPSG 157

6.3.One-dimensional Buckling in Uniaxially-Stressed SiGe Films....................................163

6.3.1. Comparing 1-D and 2-D Buckling: τ_B and λ 164

6.3.2. Comparing 1-D and 2-D Buckling: a_{max} 168

6.3.3. Conclusions on 1-D vs 2-D Buckling 169

6.4.Summary 171

7 FETs on Biaxially and Uniaxially Strained Silicon 172

7.1.A Strained-Channel FET Process 172

7.1.1. Preparing the Strained-Silicon Channel Layer 173

7.1.2. FET Process Flow...185

7.2.Bonded SOI Transistor Characteristics...193

7.2.1. Minimizing Source/Drain Resistance 195

7.2.2. Choosing a Substrate Bias for Full Channel Depletion 203

7.2.3. Optimizing the BPSG/Silicon Interface...206

7.3.Strain-Induced Mobility Enhancement...215

7.3.1. Strain-Induced Mobility Enhancement Predicted by Bulk Piezoresistance Theory 215

7.3.2. Measuring Mobility in Thin SOI Devices 223

7.3.3. Electron and Hole Mobility Enhancement due to Biaxial Strain.....................227

7.3.4. Uniaxial-strain Induced Electron Mobility Enhancement 231

7.3.5. Uniaxial-strain Induced Hole Mobility Enhancement 241

7.4.Summary 246

8 Conclusions **248**

8.1.Summary 248

8.2.Future Work 250

Appendix A Wafer Bonding and Smart-Cut Layer Transfer Process **251**

Appendix B Hooke's Law Applied to $Si_{1-x}Ge_x$ with Various Crystal Directions **259**

Appendix C Matlab Scripts Used In This Work **267**

Appendix D RTCVD Growth Sequences Used In This Work **293**

Appendix E Optimizing RTCVD Temperature Control **300**

Appendix F Bonding Sample List **316**

Appendix G Publications and Presentations Resulting from this Thesis **320**

References **324**

Chapter 1

Introduction

1.1 Thesis Overview

In an era of aggressive CMOS device scaling, the use of enhanced mobility channels has been proposed to improve circuit speed without further gate length reductions. Heteroepitaxially-generated biaxial tensile silicon strain has long been known to increase electron mobility, but the strain-induced hole mobility increase disappears at high vertical electric field [1-3]. For this reason, recent attention has shifted to process-induced uniaxial strain, as uniaxial compressive strain along a <110> channel enhances hole mobility even at high vertical fields [4,5]. However, such strain degrades electron mobility and even with a different, optimized process applied locally to nFETs the electron mobility enhancement is less than that achieved by biaxial tensile strain [6,7].

Most present techniques for generating strain in silicon are limited to a single type of strain (uniaxial or biaxial) within a given sample, which is clearly not optimum for simultaneous enhancement of hole and electron mobilities. A key achievement of this thesis is the development of a new technology to generate biaxial and uniaxial silicon strain side-by-side on a single silicon-on-insulator (SOI) wafer.

Initially, the process is developed for silicon germanium (SiGe) layers, yielding islands of uniaxially-stressed, biaxially-stressed, and unstrained SiGe, all on the same sample [8]. A biaxially-strained SiGe layer is transferred onto a host silicon wafer coated with borophosphorosilicate glass (BPSG) by wafer bonding and Smart-Cut™ layer transfer. At high temperature, the BPSG flows and SiGe islands patterned by reactive ion etching (RIE) can relax their strain by lateral expansion. By varying the island geometry from squares to narrow rectangular stripes, biaxially-symmetric and uniaxial stress, respectively, are obtained. The magnitude of uniaxial stress is dependent on the crystal-direction of the rectangle, due to the variation of Poisson's ratio in the surface plane of

the film. The Poisson effect also causes the uniaxial stress to be less than biaxial stress for the same starting film.

The lateral expansion process is numerically modeled to identify processing windows and island geometry constraints for maximum strain asymmetry. Rectangular islands with length to width aspect ratios of greater than four are needed, and as the aspect ratio increases, the process window for the lateral expansion anneal widens significantly. A fit to the measured SiGe biaxial and uniaxial strain data confirms the validity of the model. The modeling also points the way toward nanometer scale strained-silicon islands, which have not been attempted here, but are of great technological interest for application in advanced CMOS technologies.

When annealing to control SiGe strain, the strained SiGe film can also release its strain by buckling vertically, out of the plane of the film. The buckling and lateral expansion processes compete to determine the final state of the film. Two ways to mitigate this undesired buckling are demonstrated. First, when the compliant BPSG layer is thinned the buckling amplitude (surface roughness) is reduced and the buckling process slowed. Second, films with uniaxial and biaxial stress buckle in one- and two-dimensions (1-D) and (2-D), respectively. Here, reduced buckling in 1-D is observed, and by comparison to buckling models is attributed to the lower stress in the 1-D vs 2-D stressed SiGe islands made in this work.

The strain-generation process described above is then extended to yield uniaxially- and biaxially-strained silicon, by using a silicon/SiGe bi-layer. Upon high temperature anneal, the BPSG flows, and a Si/SiGe island laterally expands to reach a force balance. The SiGe becomes less compressively strained while the silicon becomes tensile, with the ratio of the two layer thicknesses determining the final strains. By varying island size, biaxially-strained, uniaxially-strained, and unstrained silicon films are fabricated all on the same sample. Like the single SiGe layer, the uniaxial silicon strain level is found to vary depending on the crystal direction of the strain, and reaches a maximum of 1.0% in the <100> direction.

Finally, MOSFETs are fabricated using these strained-silicon layers for the transistor channels. The devices are fully-depleted SOI transistors. By adding silicon nitride and silicon dioxide barrier layers between the back silicon and BPSG interface,

the transistor performance is greatly improved, with more rapid device turn-on and higher mobilities. Biaxial tensile strain of ~0.4% is shown to increase both hole and electron mobilities by ~40%. Uniaxial silicon strain also increases hole and electron mobilities by up to ~70%, although the presence and amount of mobility enhancement is shown to be dependent on the crystal-direction of the strain, and on the strain direction relative to the channel, as expected by bulk piezoresistance theory.

1.2 Thesis Outline

Chapter 2 presents the motivation for this research to address the challenge of keeping up with Moore's law in an era of aggressive device scaling. The effect of biaxial and uniaxial strain on conduction and valence bands of silicon is predicted from the "model-solid" theory, and strain-induced valence band warping is modeled using perturbation theory. The resulting changes in electron and hole transport in strained-silicon layers are considered, and a survey is made of existing fabrication approaches to achieve biaxial or uniaxial strained silicon.

Chapter 3 presents our experimental approach for generating biaxial and uniaxial stress in thin silicon germanium (SiGe) layers on insulator. Details of the wafer bonding and layer transfer technique are given and the SiGe/Si layer characterization techniques used in this work are described. A lateral expansion technique used to achieve biaxial and uniaxial SiGe stress is demonstrated, and the observed crystal-direction dependence of the uniaxial stress is explained.

In Chapter 4 a numerical model of strain generation to determine process windows and island geometry requirements for maximum asymmetry of strain is studied. The time scale of lateral expansion is calculated for biaxial and uniaxial SiGe stress on rectangular islands of various aspect ratios. The model results are compared to experimental data.

Chapter 5 presents an extension of the lateral expansion method described above to SiGe/Si bi-layers, in order to achieve ultra-thin (< 30nm) biaxial and uniaxial-strained silicon on insulator. A dependence of the uniaxial tensile strain on crystal direction is

observed, and the experimental results are compared to the numerical model of Ch. 4. The scalability of the strain generation process down to the nanometer scale is discussed.

Chapter 6 addresses an undesired buckling process affecting these strained layers. A method for reducing the buckling amplitude and slowing its temporal progress is demonstrated. Differences in 1-D and 2-D buckling in uniaxially and biaxially-stressed SiGe films, respectively, are observed and explained through modeling.

In Chapter 7, fully-depleted n- and p-channel SOI MOSFETs are fabricated with uniaxially-strained, biaxially-strained and unstrained channels. The strain-induced changes in electron and hole mobility are measured. The use of silicon dioxide and silicon nitride barrier layers to optimize the back BPSG/silicon interface is demonstrated.

Chapter 8 provides a summary of this work, and some suggestions for possible future research directions. In Appendix A, details of the wafer bonding process are supplied. Appendix B discusses Hooke's law as it applies to crystalline $Si_{1-x}Ge_x$. Presented in Appendix C are various *matlab* scripts used for this thesis, and Appendix D provides epitaxial growth sequences for the MOSFET channel and poly-Si layers. Appendix E details the control of temperature for our silicon epitaxy apparatus, as pertains to the work here, Appendix F lists the samples available for future wafer bonding, and Appendix G lists publications and conference presentations resulting from this research.

Chapter 2

Toward Uniaxially-Strained Silicon MOSFETs

This thesis presents a novel technique for the generation of biaxially- and uniaxially-strained thin silicon films integrated upon the same wafer, and measures the effect of strain on metal-oxide-semiconductor field-effect transistor (MOSFET) transistor parameters. Strained silicon is of technological interest because strain can increase the electron and hole mobility. Using a strained-silicon MOSFET channel thus has the potential to increase circuit speed without requiring the usual reduction in channel length. In this chapter, the technological motivation toward strained silicon is discussed, and the physical mechanisms of strain-induced electron and hole mobility enhancement (*i.e.*, conduction and valence band shifting and bending) are examined. Finally, existing methods for obtaining strained silicon are compared.

2.1 Limit of CMOS Scaling

In 1965, Gordon Moore, who went on to co-found Intel Corporation (Santa Clara, CA), predicted in a trade magazine article [9] that the number of transistors on an integrated circuit would double every year. He later amended this statement, saying that doubling should be expected at a slower pace of every two years [10]. A quick glance at the historic progression of Intel microprocessors shows that this pace has been kept for more than 30 years. In the top graph of Fig. 2.1, the number of transistors per chip from 1971 to the present (solid symbols) adheres closely to Moore's prediction (solid line) [11,12].

This progress has been largely enabled by the constant scaling down of transistor dimensions over time, as shown in the bottom plot of Fig. 2.1. In 1972 the circuit line width was 10 μm; by 2002 it was only 0.18 μm. By putting more transistors in the same area, the cost per transistor (middle plot of Fig. 2.1) has exponentially decreased and circuit speed has exponentially increased with time (top plot of Fig. 2.1, open symbols).

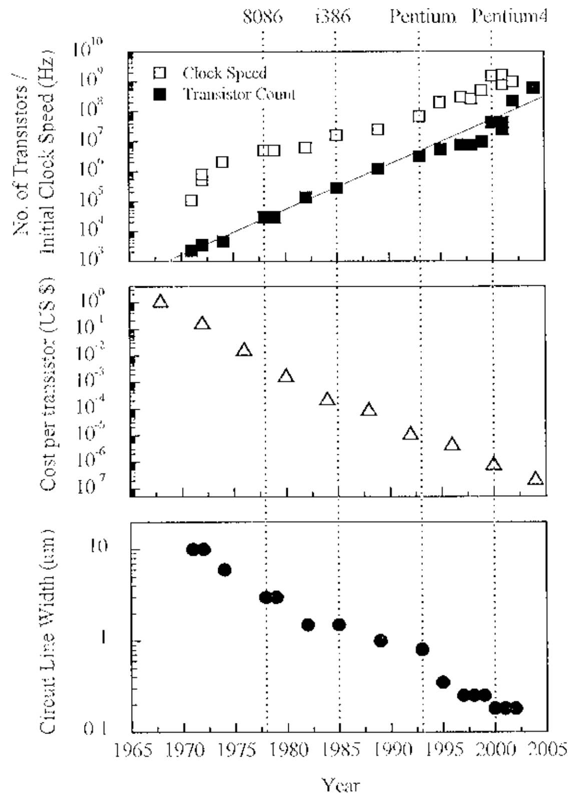


Figure 2.1: Historic changes in Intel microprocessors: number of transistors per chip, initial clock speed, cost per transistor, and circuit line width, from 1968 to 2004. The solid line in the top graph indicates the Moore's law expectation that the transistor count doubles every two years. The vertical lines highlight several common microprocessors. Data from Refs. [11, 12, 13].

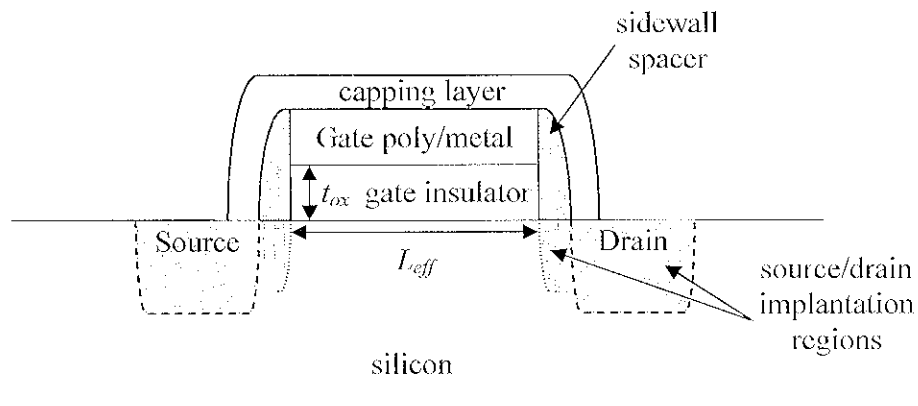


Figure 2.2: Schematic of a short-channel MOSFET with an effective channel length of L_{eff} and gate oxide thickness of t_{ox}. The sidewall spacers and multiple source/drain implants are optimized to provide ultra-shallow junctions with low series resistance. Capping layers can be used to transfer strain into the silicon channel.

The new technologies these advances have enabled have come to define the present "information age."

But more than just an astute observation, Moore's law, as his prediction is known, has become a guiding principle for the integrated circuit industry. Since 1992, the Semiconductor Industry Association has published National Technology Roadmaps [14], now called the International Technology Roadmap for Semiconductors [15]. In order to keep performance gains on an exponential rise, the industry uses these roadmaps to ensure that the needed tools, materials and processes are simultaneously available.

For the 65-nm technology node currently in production at Intel, the effective gate length of logic transistors is L_{eff} = 35 nm, and the effective gate insulator (oxide/nitride) thickness is t_{ox}=1.2 nm [6]. These two quantities are illustrated in Fig. 2.2 for a typical short-channel transistor. Constant-field scaling [16] requires that these dimensions scale down according to Moore's law; the industry roadmap calls for an effective gate length of L_{eff} = 9 nm in 2016. At such small dimensions not only will patterning and dimensional control be extremely difficult, but as the number of atoms in the channel becomes so low as to be countable, the resulting statistical fluctuation of doping may make it difficult to control the variation of transistor parameters such as the threshold voltage across an entire wafer. Clearly, a scaling limit will eventually be reached, possibly within the next 10 to 20 years.

One way to improve transistor performance that does not require scaling is to increase the charge carrier mobility in the transistor channel. The drain current in a long-channel transistor is given for small drain-source voltages by

$$I_{DS\ linear} = \mu \frac{W}{L_{eff}} \frac{\varepsilon_{ox}\varepsilon_o}{t_{ox}} \left(V_{GS} - V_T\right) V_{DS} \tag{2.1}$$

where μ is the carrier mobility, W and L_{eff} are the transistor width and effective length, ε_{ox} and t_{ox} are the gate insulator relative dielectric constant and thickness, V_{DS} is the drain-source voltage and V_T is the threshold voltage. By increasing the mobility, μ, one can obtain the desired higher drain current without needing to scale down L_{eff}.

Because silicon is a piezoresistive material, its resistivity changes with applied stress or strain. Therefore one technique that has been proposed to increase FET mobility is the introduction of strain into the silicon channel. In this work two types of strain,

biaxial and uniaxial, will be discussed. These strain types are schematically drawn in Fig. 2.3, with the solid arrows indicating directions in which the cubic crystal is being pulled into tensile strain. For the biaxial strain case of Fig. 2.3a, the level of tension is symmetric in the two in-plane (x- and y-) directions. In Fig. 2.3b, the uniaxial strain case is depicted: tension exists in only one in-plane direction (x- or y-); the other in-plane strain is equal to zero. Here, in the schematic drawing of Fig. 2.3b, the uniaxial strain is in the y-direction: $\varepsilon_{xx}=0$ while $\varepsilon_{yy}\neq 0$. Regardless of whether the in-plane strain is biaxial or uniaxial, in the z-direction the silicon layer has zero stress ($\sigma_{zz}=0$). The z-direction strain is therefore non-zero, and can be calculated from Hooke's law as described in App. B. For tensile in-plane strain, the crystal will be compressively strained in the out-of-plane z-direction.

Biaxial tensile silicon strain has been investigated extensively over the past decade to enhance the electron mobility in strained-silicon MOSFETs [1,2,17-20]. The strain-induced electron mobility enhancement occurs even at high vertical electric fields such as are present in inverted FETs [2,17,18,20]. Hole mobility enhancement has also been seen under biaxially-tensile strained silicon, but for holes the mobility enhancement disappears at high vertical fields [3,19,21]. Therefore, despite the electron mobility enhancement, the net increase in circuit speed is small since at high vertical field biaxial tensile strain only exacerbates the existing hole/electron mobility imbalance in unstrained silicon (μ_n/μ_p=2.4-2.9) [22].

Recently, hole mobility enhancement that persists at high fields has been obtained using uniaxially-compressive strained silicon. Here the strain is induced by process effects in very short-channel transistors (L_{eff} < 60 nm) [4,5,23,24]. The hole mobility enhancement of uniaxial compressive strain is, in fact, anticipated by bulk piezoresistance theory, as will be discussed in depth in Ch. 7.

However the same uniaxial strain which so positively affects hole mobility does not increase electron mobility. In order to simultaneously optimize both hole and electron mobility, a hybrid strain approach is needed. Intel has used different processing methods on nFETs and pFETs to obtain uniaxially-tensile channels for electrons and uniaxially-compressive channels for holes [4,5,23,24]. These strain generation methods will be discussed more fully below, in Sec. 2.4. IBM has developed a hybrid-crystal orientation

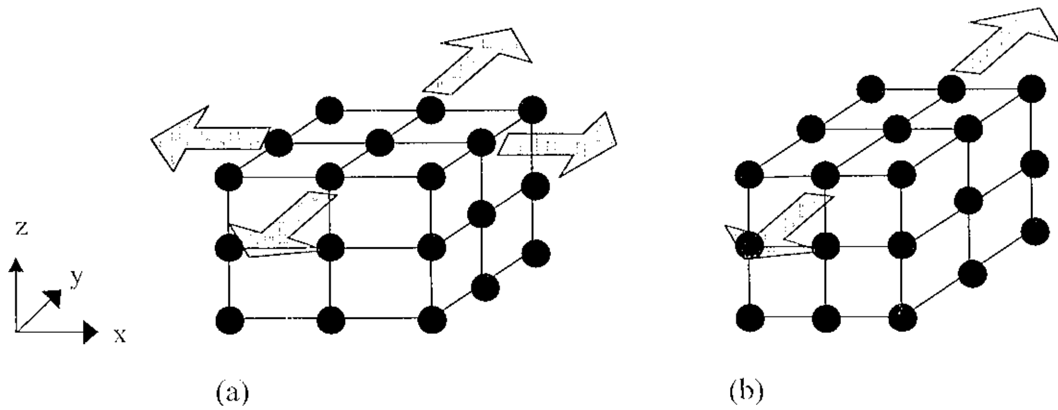


Figure 2.3: Schematic of (a) biaxial tensile strain and (b) uniaxial tensile strain in a cubic crystal. The arrows indicate directions of tension. Biaxial strain is symmetric in the x- and y-directions, whereas uniaxial strain is characterized by strain being present in either x or y (here, y), with zero strain in the opposite direction. In both cases, there is no stress in the z-direction, and the amount of z-axis strain is calculated from Hooke's law (see App. B).

process with nFETs on silicon with an (001) surface orientation and pFETs on (110) surface silicon to enhance both mobilities on the same wafer [25]. Recently they have combined this hybrid orientation method with process-induced strain similar to that of Intel, in order to further maximize the mobility enhancement [26]. In both IBM and Intel's methods, the amount of strain is dependent on the precise process techniques, is limited to very short channel transistors (as L_{eff} increases the strain decreases), can only be applied longitudinally along the channel, and the strain is rather anisotropic [27].

A key achievement of the present work is the development of a novel process to integrate biaxial and uniaxial silicon strain on the same wafer. This offers the possibility of enhancing both electron and hole mobility by using different types of strain for nFET and pFET transistors. Moreover, the restrictions listed above for process induced strain are lifted: our technology generates strain that is precisely uniaxial in the plane of the silicon film, with well-controlled strain levels of up to 1.0%, and uniform strain over large areas (hundreds of μm^2). In the remainder of this chapter, the physical mechanisms by which biaxial and uniaxial strain affect electron and hole mobility will be examined, and then a detailed look will be taken at existing methods that can be used to generate uniaxial strain in silicon.

2.2 Strain-Induced Enhancement of Electron Mobility

This thesis focuses on the use of a novel process to obtain biaxially- and uniaxially-strained silicon in order to enhance electron and hole mobility in MOSFETs. Strain-induced mobility enhancement is possible because silicon is a piezoresistant material: its conductivity changes with applied strain. In this section, the physical mechanisms by which uniaxial and biaxial strain change electron transport in silicon will be examined. First, the "model-solid" energy band theory for biaxial and uniaxial strain will be presented, and then the impact of energy band shifts and splitting on electron transport will be discussed.

2.2.1 Effect of Biaxial and Uniaxial Strain on the Silicon Conduction Band

The "model-solid" energy band theory has been used previously [28] to describe the energy shifting and splitting of the conduction and valence bands in biaxially-strained $Si_{1-x-y}Ge_xC_y$ with germanium contents of less than 40% and carbon fractions of less than 3%. Here, starting with the original theory [29-33] the model is extended to the case of uniaxial strain in silicon.

The "model-solid" theory makes two assumptions. First, assume that the conduction band position, $E^j_{c,\alpha}$, can be modeled as linear sum of the "intrinsic" (unstrained) band position, $E^{\text{int}}_{c,\alpha}$ and the strain-induced band shift $\Delta E^{j,strain}_{c,\alpha}$:

$$E^j_{c,\alpha} = E^{\text{int}}_{c,\alpha} + \Delta E^{j,strain}_{c,\alpha}, \tag{2.2}$$

where α is a given position in k-space (*e.g.*, the Δ, Γ or L locations), and j is a given k-space direction (*e.g.*, for the Δ bands, [100], [010] or [001]).

The "intrinsic" conduction band position at various points in k-space can be calculated from the known valence band position and band-gap [33]:

$$E^{\text{int}}_{c,\alpha} = E^{\text{int}}_v + \frac{1}{3}\Delta_o + E^{\text{int}}_{g,\alpha} \tag{2.3}$$

For bulk silicon, the maximum energy of unstrained valence band is $E^{\text{int}}_v + \frac{1}{3}\Delta_o$, where $E^{\text{int}}_v = -7.03eV$ is the average position of the three valence band maximums (of the heavy hole, light hole and split-off bands) in intrinsic, unstrained silicon and $\Delta_o = 0.04eV$ [33]. (The "zero" of energy is the vacuum level.) The bandgap to the minimum energy Δ conduction bands is $E^{\text{int}}_{g,\Delta} = 1.17eV$ at 300 K [33]. Therefore the "intrinsic" conduction band position for silicon is $E^{\text{int}}_{c,\Delta} = -5.85eV$.

Second, assume that the strain-induced conduction band shift, $\Delta E^{j,strain}_{c,\alpha}$, is linearly related to the strain in the film according to [33]:

$$\Delta E^{j,strain}_{c,\alpha} = \left[\Xi^{\alpha}_d \overline{\overline{1}} + \Xi^{\alpha}_u \left\{\hat{a}_{j,\alpha} \cdot \hat{a}_{j,\alpha}\right\}\right] : \varepsilon, \tag{2.4}$$

where $\tilde{1}$ is the unity tensor

$$\tilde{1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.5}$$

ε is the strain matrix

$$\varepsilon = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix}, \tag{2.6}$$

and where $\{\hat{a}_{j,\alpha} \cdot \hat{a}_{j,\alpha}\}$ represents dyadic product of unit vectors in the j direction. Note that this general formulation allows for arbitrary strain states. In the case of biaxial in-plane strain, it reduces to the equations given by Yang [28].

The strain matrix, ε, can be determined by transforming an arbitrary biaxial or uniaxial strain into the cubic crystal axes components ([100], [010] and [001]) as described in Appendix B. Here, there are three cases of interest. For biaxial strain in the (001) surface plane of a film, the strain matrix is simply

$$\varepsilon = \begin{bmatrix} \varepsilon_{biax} & 0 & 0 \\ 0 & \varepsilon_{biax} & 0 \\ 0 & 0 & -2\dfrac{c_{12}}{c_{11}}\varepsilon_{biax} \end{bmatrix}, \tag{2.7a}$$

whereas for uniaxial strain in the <100> direction of a (001) suface, ε is given by

$$\varepsilon = \begin{bmatrix} \varepsilon_{uni} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\dfrac{c_{12}}{c_{11}}\varepsilon_{uni} \end{bmatrix} \tag{2.7b}$$

and for uniaxial strain in the <110> direction of a (001) surface, the matrix is

$$\varepsilon = \begin{bmatrix} \dfrac{1}{2}\varepsilon_{uni} & -\dfrac{1}{2}\varepsilon_{uni} & 0 \\ -\dfrac{1}{2}\varepsilon_{uni} & \dfrac{1}{2}\varepsilon_{uni} & 0 \\ 0 & 0 & -\dfrac{c_{12}}{c_{11}}\varepsilon_{uni} \end{bmatrix} \tag{2.7c}$$

The elastic stiffness constant values, c_{11} and c_{12}, for silicon are 164.8 and 63.5 GPa, respectively [34].

For silicon, the Δ conduction bands will always have the lowest energy, and thus other local energy minima (*e.g.*, at *L* points) can be ignored. The silicon conduction band has six-constant energy ellipsoids along the [100], [010] and [001] directions, with minima at the Δ points. These ellipsoids are drawn schematically in Fig. 2.4a.

To calculate the change in the energy levels of these six ellipsoids with strain, one can use Eqn. 2.4 with the unit vectors $\hat{a}_\Delta = [1 \quad 0 \quad 0], [0 \quad 1 \quad 0]$, and $[0 \quad 0 \quad 1]$ parallel to the ellipsoid axes. The result is:

$$\Delta E_{c,\Delta}^{100,\bar{1}00} = \Xi_d^\Delta \cdot (\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) + \Xi_u^\Delta \cdot \varepsilon_{xx} \qquad (2.8a)$$

$$\Delta E_{c,\Delta}^{010,0\bar{1}0} = \Xi_d^\Delta \cdot (\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) + \Xi_u^\Delta \cdot \varepsilon_{yy} \qquad (2.8b)$$

$$\Delta E_{c,\Delta}^{001,00\bar{1}} = \Xi_d^\Delta \cdot (\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) + \Xi_u^\Delta \cdot \varepsilon_{zz} \qquad (2.8c)$$

For silicon, the Δ band deformation potentials, Ξ_d^Δ and Ξ_u^Δ, are equal to 1.127 eV and 9.16 eV [33], respectively. Because each direction is governed by a separate equation, the application of strain can shift the conduction band and/or split it into two or three distinct sub-bands.

First the trivial case of hydrostatic strain (for which $\varepsilon_{xx}=\varepsilon_{yy}=\varepsilon_{zz}$) is examined, for comparison with the biaxial and uniaxial strain cases of interest. When a hydrostatic strain of ε_h is applied such that $\varepsilon_{xx}=\varepsilon_{yy}=\varepsilon_{zz}=\varepsilon_h$, the conduction band will remain 6-fold degenerate (no band-splitting occurs), but according to Eqn. 2.8 the band will be shifted by a constant equal to

$$\Delta E_{c,\Delta 6} = 3 \cdot a_c^\Delta \cdot \varepsilon_h = +12.54 \cdot \varepsilon_h \ (eV), \qquad (2.9)$$

where the deformation potential for hydrostatic strain, a_c^Δ, is defined as

$$a_c^\Delta = \Xi_d^\Delta + \frac{1}{3} \cdot \Xi_u^\Delta \qquad (2.10)$$

and is equal to 4.18 eV for silicon.

The biaxial strain case can be solved by substituting the strains of Eqn. 2.7a into Eqn. 2.8. The hydrostatic strain component will shift the bands uniformly up or down in energy, while the non-hydrostatic component of the strain will split the bands. Thus, a

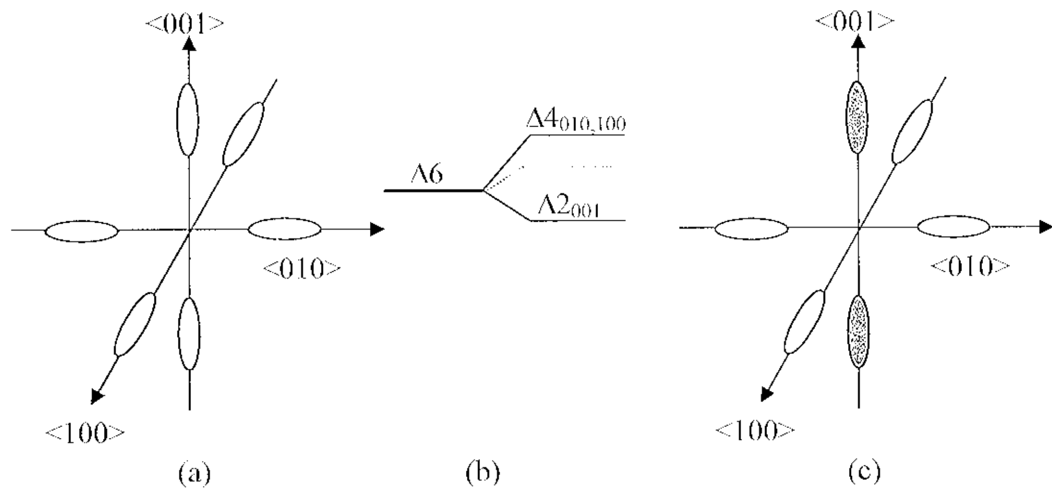


Figure 2.4: Constant energy ellipsoids in reciprocal (k-) space for the conduction band of silicon. In unstrained silicon (a), the six ellipsoids are identical, and the conduction band is 6-fold degenerate. In silicon with biaxial tensile strain or uniaxial tensile strain aligned to <110>, the bands split as shown in (b), where the dotted line indicates the shift due to the hydrostatic strain component and the solid lines represent the total effect of the strain. The lowest energy sub-band, shaded in (c), is two-fold degenerate.

biaxial strain lifts the 6-fold degeneracy of the unstrained or hydrostatically-strained silicon. The conduction band splits into a two-fold band, $\Delta 2_{001}$, with a energy shift of

$$\Delta E_{c,\Delta 2}^{001,00\bar{1}} = \left[2 \cdot a_c^{\Delta} \cdot \left(1 - \frac{c_{12}}{c_{11}}\right) - \frac{2}{3} \cdot \Xi_u^{\Delta} \cdot \left(1 + 2 \cdot \frac{c_{12}}{c_{11}}\right)\right] \cdot \varepsilon_{biax} = -5.67 \cdot \varepsilon_{biax}\ (eV) \quad (2.11a)$$

and a four-fold band, $\Delta 4_{010,100}$,

$$\Delta E_{c,\Delta 4}^{010,0\bar{1}0,100,\bar{1}00} = \left[2 \cdot a_c^{\Delta} \cdot \left(1 - \frac{c_{12}}{c_{11}}\right) + \frac{1}{3} \cdot \Xi_u^{\Delta} \cdot \left(1 + 2 \cdot \frac{c_{12}}{c_{11}}\right)\right] \cdot \varepsilon_{biax} = +10.54 \cdot \varepsilon_{biax}\ (eV).$$

(2.11b)

The first term in Eqns. 2.11a and 2.11b represents the hydrostatic shift of both sub-bands, while the second term corresponds to the band splitting, as illustrated in Fig. 2.4b. For tensile strain, the $\Delta 2_{001}$ band (shaded in Fig. 2.4c) will have the lowest energy. For a biaxial strain of 1.0%, the $\Delta 2$ and $\Delta 4$ sub-bands will be separated by 160 meV. This separation is equal to ~$6kT$ at room temperature, so most of the electrons will be confined to the lowest band.

The case for uniaxial strain in the <110> direction, Eqn. 2.7c, is very similar. Just as in the biaxial strain case, the conduction band splits into two sub-bands:

$$\Delta E_{c,\Delta 2}^{001,00\bar{1}} = \left[2 \cdot a_c^{\Delta} \cdot \left(1 - \frac{c_{12}}{c_{11}}\right) - \frac{2}{3} \cdot \Xi_u^{\Delta} \cdot \left(1 + 2 \cdot \frac{c_{12}}{c_{11}}\right)\right] \cdot \frac{\varepsilon_{uni}}{2} = -2.83 \cdot \varepsilon_{uni}\ (eV) \quad (2.12a)$$

$$\Delta E_{c,\Delta 4}^{010,0\bar{1}0,100,\bar{1}00} = \left[2 \cdot a_c^{\Delta} \cdot \left(1 - \frac{c_{12}}{c_{11}}\right) + \frac{1}{3} \cdot \Xi_u^{\Delta} \cdot \left(1 + 2 \cdot \frac{c_{12}}{c_{11}}\right)\right] \cdot \frac{\varepsilon_{uni}}{2} = +5.27 \cdot \varepsilon_{uni}\ (eV)$$

(2.12b)

The band shift and splitting are qualitatively identical to the case of biaxial strain, as shown in Fig. 2.4b,c. However, the magnitudes of the band shift and splitting relative to the strain are reduced by a factor of two, with $\varepsilon_{uni}/2$ replacing ε_{biax} in Eqn. 2.11. If the uniaxial <110> strain is 1.0%, the two sub-bands will be separated by 80 meV, or ~$3kT$ at room temperature, still a significant separation.

For uniaxial strain in the <100> crystal direction, the bands will split quite differently, into three separate sub-bands, each two-fold degenerate, as shown in Fig. 2.5. The sub-band positions are given by Eqn. 2.7b and 2.8:

$$\Delta E_{c,\Delta 2}^{001,00\bar{1}} = \left[2 \cdot a_c^{\Delta} \cdot \left(1 - \frac{c_{12}}{c_{11}}\right) - \frac{2}{3} \cdot \Xi_u^{\Delta} \cdot \left(1 + 2 \cdot \frac{c_{12}}{c_{11}}\right)\right] \cdot \frac{\varepsilon_{uni}}{2} = -2.83 \cdot \varepsilon_{uni}\ (eV) \quad (2.13a)$$

$$\Delta E_{c,\Delta 2}^{010,0\bar{1}0} = \left[2 \cdot a_c^{\Delta} \cdot \left(1 - \frac{c_{12}}{c_{11}}\right) + \frac{2}{3} \cdot \Xi_u^{\Delta} \cdot \left(-1 + \frac{c_{12}}{c_{11}}\right)\right] \cdot \frac{\varepsilon_{uni}}{2} = +0.69 \cdot \varepsilon_{uni}\ (eV) \quad (2.13b)$$

$$\Delta E_{c,\Delta 2}^{100,\bar{1}00} = \left[2 \cdot a_c^{\Delta} \cdot \left(1 - \frac{c_{12}}{c_{11}}\right) + \frac{2}{3} \cdot \Xi_u^{\Delta} \cdot \left(2 + \frac{c_{12}}{c_{11}}\right)\right] \cdot \frac{\varepsilon_{uni}}{2} = +9.85\ \varepsilon_{uni}\ (eV)\,. \quad (2.13c)$$

For tensile strain again the $\Delta 2_{001}$ sub-band will be lowest, but the amount of band-splitting to the next sub-band ($\Delta 2_{010}$), is reduced. For 1.0% tensile uniaxial strain, the band-splitting between $\Delta 2_{001}$ and $\Delta 2_{010}$ will be only 35 meV, or ~1.3kT at room temperature. It is therefore likely that conduction band electrons will shared between these two sub-bands.

The energy bands are plotted vs biaxial and uniaxial strain in Fig. 2.6a. The middle of the plot, indicated by the dashed line, indicates zero strain. Tensile strain is to the right and compressive strain to the left. As described by Eqns. 2.11-2.12, for biaxial strain and uniaxial strain aligned to the <110> crystal direction, the conduction band splits into two sub-bands, with the lowest energy band a two-fold degenerate band aligned to the <001> direction as drawn in Fig. 2.4b,c. For uniaxial strain aligned to the <100> crystal direction, three sub-bands are formed with much closer-spaced splitting, but with the $\Delta 2_{001}$ band still at the lowest energy, as shown in Fig. 2.5.

2.2.2 Electron Transport in Strained-Silicon Layers

The strain-induced band-splitting described in Sec. 2.2.1 modifies electron transport in several ways. First, when the unstrained 6-fold degenerate conduction band is split into two or three sub-bands in strained silicon, the conduction band electrons will be distributed differently in k-space. Assuming for simplicity that the Boltzmann distribution of electron population density holds in the sub-bands, plotted in Fig. 2.6b is the sub-band population as a function of strain. The Boltzmann distribution states that the fractional population of each sub-band, n_i, is proportional to the exponential of the conduction band position:

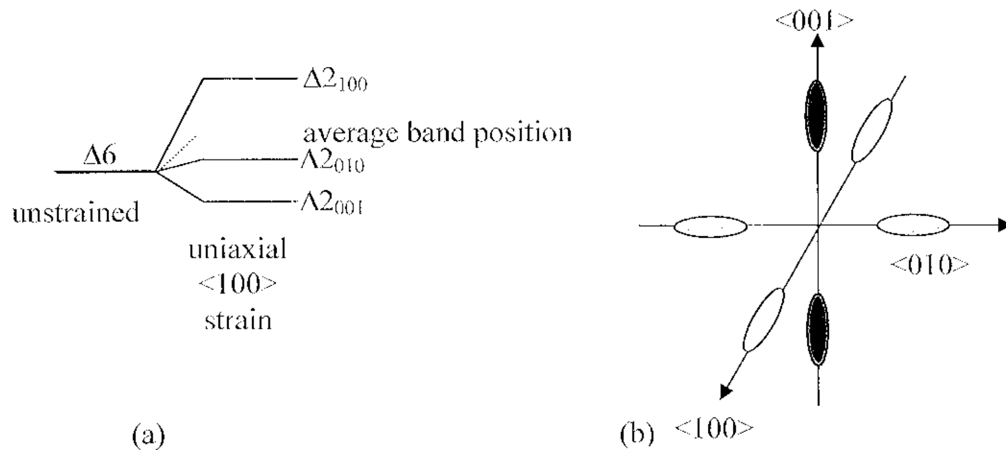


Figure 2.5: (a) Energy band splitting and (b) constant energy ellipsoids in reciprocal (k-) space for the conduction band of silicon with uniaxial strain in the <100> direction. For this type of strain, the conduction band splits into three sub-bands, each of which are two-fold degenerate. The lowest energy $\Delta 2_{001}$ ellipsoids are shown solid in (b), while the next sub-band, $\Delta 2_{010}$, are gray-shaded. The band shift due to the hydrostatic component of the strain is shown by the dotted line in (a).

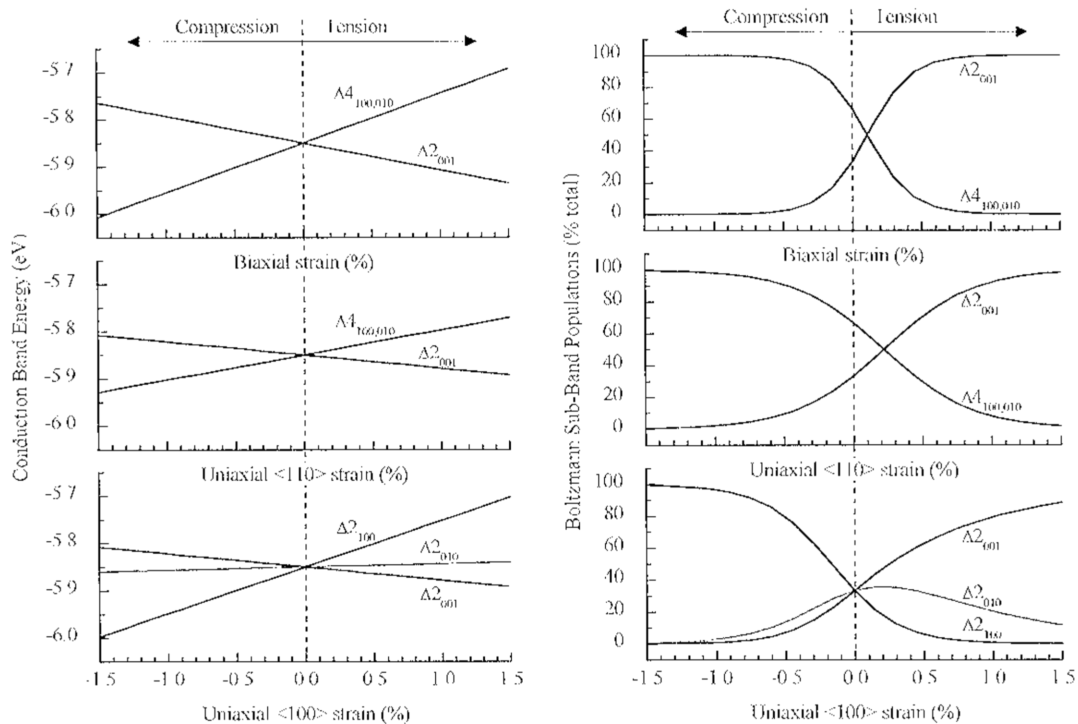


Figure 2.6: Sub-band energy positions (a) and populations (b) for the conduction band of silicon at 300 K according to the "model-solid" theory of Sec. 2.2.1 for the cases of biaxial strain, uniaxial strain parallel to <110>, and uniaxial strain parallel to <100>.

$$n_j = \frac{N_j e^{-E_j/kT}}{\sum_j N_j e^{-E_j/kT}}, \quad (2.14)$$

where N_j is the degeneracy of the sub-band and E_j is the sub-band energy. As tensile strain increases, electrons shift into the lowest sub-band. For biaxial or uniaxial strain greater than ~0.5%, more than two-thirds of the electrons in the conduction band are located in the $\Delta 2_{001}$ sub-band. This change in band population has two primary effects which enhance the electron mobility: (1) the effective electron mass is reduced and (2) intervalley phonon scattering is reduced [35].

Because of the ellipsoidal shape of the constant energy surfaces for the delta conduction bands, the effective mass of electrons is different in the transverse direction (m^*_T=0.19) and the longitudinal direction (m^*_L=0.92) [36]. In unstrained silicon, the conductivity effective mass is a composite of these two values [22,37,38]:

$$m^*_{n\ unstrained} = 3\cdot\left(\frac{1}{m^*_l} + \frac{2}{m^*_t}\right)^{-1} = 0.26. \quad (2.15)$$

If the biaxial or uniaxial <110> strain is large enough, almost all of the electrons will be $\Delta 2_{001}$ sub-band, and transport in the (001) plane will be dominated by the transverse effective mass, m^*_T, which is lower than the $m^*_{n\ unstrained}$ value of Eqn. 2.15 for unstrained silicon. (The value of m^*_T is independent of biaxial strain [39,40].) Thus, strain significantly decreases the average conductivity mass, which increases the mobility, μ, according to [41]

$$\mu = \frac{q\tau}{m^* m_e} \quad (2.16)$$

where q is the electron charge, τ is the scattering time constant, and m^* is the effective mass (unitless) and m_e is the electron mass (9.11×10^{-31} kg).

A second effect is that the band-splitting, if of sufficient magnitude (*i.e.*, several kT or more), by confining most of the electrons to the $\Delta 2_{001}$ sub-band, reduces intervalley phonon scattering. The scattering rates from various scattering events sum, as per Matthiessen's rule, to determine the total scattering rate:

$$\frac{1}{\tau_{total}} = \frac{1}{\tau_{intervalley\ phonon}} + \frac{1}{\tau_{intravalley\ phonon}} + \frac{1}{\tau_{surface\ roughness}} + \frac{1}{\tau_{ionized\ impurity}} + \ldots \quad (2.17)$$

For unstrained silicon, intervalley scattering usually dominates over intravalley scattering, so a reduction in its rate reduces the total scattering rate, increasing the mobility. So, the strain-induced band splitting and resulting electron population redistribution reduce the effective electron mass and reduce intervalley phonon scattering. These two effects are thought to dominate the strain-induced increase of mobility observed in silicon where the electrons exist in a 3-D electron gas.

In an inverted n-MOSFET, a 2-D triangular potential well is formed by the FET gate bias. This potential well confines charge carriers to the surface channel, quantizing the conduction band energies and forming a 2-D electron gas. As the gate bias is increased, the electron sheet density, n_s, and the effective electric field in the channel both increase, deepening the potential well. Because of a different effective mass in the vertical direction, this quantum confinement splits the 6-fold degenerate conduction band into a two sub-bands for 2-fold and 4-fold degeneracy, in much the same way as described above for biaxial tensile strain [42]. The situation is sketched in Fig. 2.7, where the $\Delta 2$ band has discrete energy levels of E_0, E_1, E_2, and so on, and the $\Delta 4$ band has levels of E'_0, E'_1, E'_2, etc. Fischetti [43], *et. al.*, have calculated that in relaxed silicon with a carrier concentration, n_s, of 10^{13} cm^{-2}, that the difference between the lowest energy levels in the $\Delta 2_{001}$ and $\Delta 4_{100,010}$ sub-bands ($\Delta E_0 = E'_0 - E_0$) at 300 K is ~95meV, which implies that 75% of the electrons are located in the lower $\Delta 2_{001}$ sub-band. So, the situation in an *unstrained* silicon MOSFET channel in strong inversion is exactly analogous to that of a silicon layer with a moderate biaxial tensile strain, as described above: the band splitting will reduce intervalley scattering, and the high fractional electron population in the $\Delta 2_{001}$ implies a lower effective electron mass, both increasing the electron mobility [42-44].

Fischetti [43] has also examined the superposition of biaxial silicon strain on this 2-D potential well structure, and has calculated that for biaxial tensile strain of 1.0%, the energy level splitting significantly increases to $\Delta E_0 = 263$ meV, increasing the electron population fraction in the $\Delta 2_{001}$ band to 90%. This shift is shown in Fig. 2.8. The addition of strain thus further increases the electron mobility. At high levels of strain, the intervalley scattering becomes quite weak as the band splitting is larger than the required phonon energy. Then, the maximum strain-induced mobility enhancement is limited

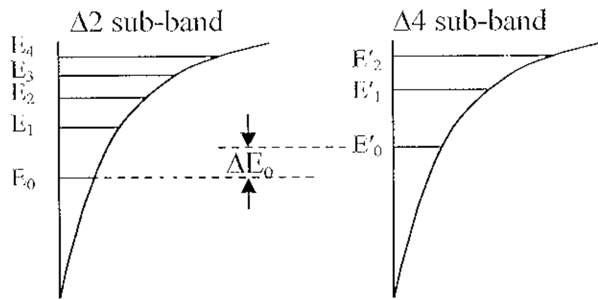


Figure 2.7: Discrete energy levels formed within the 2-D potential of an unstrained silicon surface-channel FET for the (a) $\Delta 2_{001}$ sub-band perpendicular to the surface and (b) the $\Delta 4_{100,010}$ sub-band in the plane of the surface. Note that the lowest energy levels in the two sub-bands are offset by ΔE_0 because of the different vertical effective masses in the two sub-bands. After Ref. [43,45].

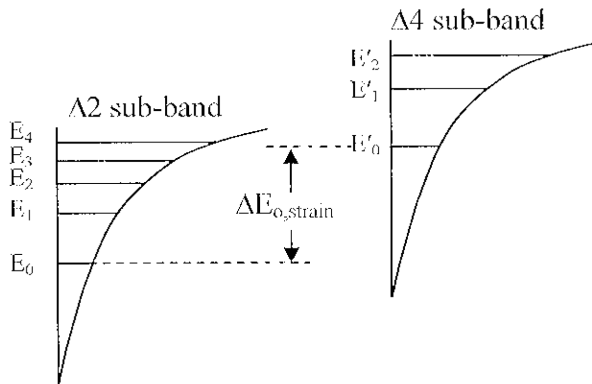


Figure 2.8: Discrete energy levels formed within the 2-D potential of a biaxially-strained silicon surface-channel FET for the (a) $\Delta 2_{001}$ sub-band perpendicular to the surface and (b) the $\Delta 4_{100,010}$ sub-band in the plane of the surface. Note that the energy splitting between the two bands, ΔE_0, has increased greatly from the case of zero strain depicted in Fig. 2.7. After Ref. [43,45].

according to Eqn. 2.17 by the *intra*valley phonon scattering rate, which is independent of strain [42,43]. The predicted electron mobility enhancement vs biaxial tensile silicon strain of Ref. [42] is plotted in Fig. 2.9, and has been well-matched by experimental work. (See for example, Ref. [7].)

The observed electron mobility enhancement with biaxial tensile strain has been shown experimentally to persist at high vertical electric field, under which most logic transistors operate [2,17,18,20]. At low fields the reduction in intervalley scattering and reduced effective mass due to band splitting dominate electron transport. At high fields, the larger band bending splits the bands sufficiently so that strain provides a trivial additional benefit to the electron mobility. Therefore, the electron mobility enhancement due to biaxial tension should decrease with increasing electric field, but this is not what is observed. Several authors [43,46,47] have concluded that the high field electron mobility enhancement can be accounted for by assuming that the strained silicon has a reduced surface roughness at the SiO_2/Si interface, compared to unstrained silicon. At high field, the carriers are more tightly confined at the interface, and thus role of surface roughness in determining the mobility (via Eqns. 2.16 and 2.17) is amplified, compared to the low field case. The decreased surface roughness in strained silicon required to explain the observed high field mobility enhancement has been experimentally measured [48-50]. It is the high-field enhancement of electron mobility due to biaxial strain (and, presumably, a coincident reduced surface roughness) that makes the use of biaxial tension so attractive for high performance FET applications.

So far, the implications of biaxial tensile silicon strain on electron transport have been discussed. The uniaxial strain cases may, of course, differ, and very little theoretical work has been done on transport in layers with uniaxial stress or uniaxial strain [5]. As seen in Sec. 2.2.1, for uniaxial strain aligned to <110> the band splitting will be identical to that generated by biaxial strain, but will be of lower magnitude. Thus, one expects that the effect of uniaxial <110> strain on mobility will similar to biaxial strain -- tension increases the mobility via reduced intervalley scattering and reduced effective mass – but with less mobility enhancement per percentage strain. At high uniaxial <110> strains, one expects that intravalley scattering will again limit the maximum mobility enhancement. Based on these speculations, a qualitative sketch of mobility enhancement vs uniaxial

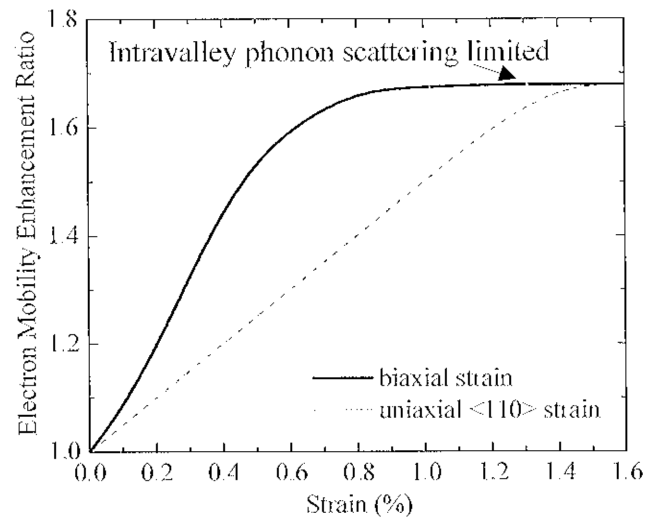


Figure 2.9: Enhancement of electron mobility in silicon at 300 K due to biaxial tensile strain, as calculated in Ref. [42], and due to uniaxial strain aligned to <110>, as qualitatively predicted here. The maximum mobility enhancement in both cases is expected to be limited by intravalley phonon scattering.

<110> strain is given in Fig. 2.9. Such electron mobility enhancement for <110> strain parallel to a <110> channel has been observed by Thompson, *et. al.* [4,5,23,24], and has been shown to persist at high vertical effective fields [5], presumably due to the same physical effects that govern biaxial tensile strain: increased band splitting and reduced surface roughness scattering in strained-silicon layers.

The case of uniaxial strain aligned to the <100> crystal direction is complicated by the existence of three sub-bands, with much smaller energy separation between the two lowest bands, each of which are two-fold degenerate. We are not aware of any published data on <100> n-channel transistors. For [100] uniaxial strain, the electron population will be shared between the [001] and [010] bands as shown in Fig. 2.6b. The band-splitting induced by the strain largely excludes carriers from the high energy $\Delta 2_{100}$ band. Comparing in Fig. 2.6b the fraction of electrons in the upper bands, it is clear that of the three strain cases presented, uniaxial <100> strain has the lowest electron population in its highest band:

$$n_{j,\,\Delta 2\ 100\ \text{uniaxial}\ <100>}(\varepsilon) < n_{j,\,\Delta 4\ 100,010,\ \text{biaxial or uniaxial}\ <110>}(\varepsilon) \qquad (2.18)$$

Consider the case of [100] uniaxial strain applied parallel to the flow of carriers in the channel. For electron flow in the [100] direction, the [001] and [010] bands will be characterized by effective masses of m^*_T, while the [100] band will be characterized by the much larger m^*_L. The large energy splitting between the $\Delta 2_{100}$ and other bands caused by uniaxial [100] strain confines electrons in the $\Delta 2_{001}$ and $\Delta 2_{010}$ sub-bands (Fig. 2.6b), which have low effective electron mass. Therefore, like biaxial-strain, it is predicted that the total effective mass for uniaxial <100> strain will be low, $m^*_{n,strained} \approx m^*_T = 0.19$. A strong mobility enhancement from [100] uniaxial tensile strain applied parallel to the channel is expected. Alternate strain configurations are also possible: consider the case of [100] uniaxial strain applied perpendicular to a [010] flow of electrons in the channel. Now the effective mass will be a combination of m^*_L from the [010] ellipsoids and m^*_T from the [001] ellipsoids. Thus for tensile strain the effective mass may actually increase, decreasing the mobility compared to the unstrained silicon case.

In summary, the electron mobility in strained-silicon layers is determined by strain-induced band splitting, which shifts electron populations into specific sub-bands, modifying the intervalley phonon scattering rates, and changing the effective electron

mass. The addition of biaxial and uniaxial <110> tensile silicon strain is predicted to increase the electron mobility, with the uniaxial strain showing a weaker mobility-strain relationship than biaxial strain. In both cases, the electron mobility enhancement is predicted to saturate at about +68% due to intravalley phonon scattering, as shown in Fig. 2.9. The mobility enhancement has been shown to persist at high vertical electric fields for biaxial strain, and is expected to likewise persist for uniaxial <110> strain. For the case of uniaxial <100> tensile strain, it is proposed that the mobility will increase sharply with uniaxial strain when the strain and channel are parallel, due to the low effective electron mass. However, when the channel is perpendicular to a uniaxial <100> tensile strain, the mobility is expected to remain unchanged or perhaps even decrease. In Ch. 7, these predictions will be compared to the phenomenological piezoresistance theory, and to our biaxially- and uniaxially-strained nFET devices results. The analysis above has been limited to a quantitative analysis of strain-induced band splitting, followed by qualitative speculation on its effects on electron mobility. A more thorough analysis, beyond the scope of this work, is needed to more robustly determine the effect of uniaxial strain on electron scattering rates, and to include the impact of ellipsoidal band deformation on effective mass.

2.3 Strain-Induced Enhancement of Hole Mobility

In the previous section the effect of biaxial and uniaxial strain on electron mobility was analyzed and discussed. Next, hole mobility will be considered. First, the impact of strain on the valence band energy will be explored, and then the effect of these distortions and shifts of the energy bands on hole transport in strained-silicon layers is analyzed.

2.3.1 Effect of Biaxial and Uniaxial Strain on the Silicon Valence Band

The valence band structure of unstrained silicon is schematically drawn in Fig. 2.10a. It consists of three bands with energy maxima at the Γ point. The spin-orbit (SO) interaction lifts the three-fold degeneracy and splits the bands by $\Delta_o = 0.04eV$ [33]. The

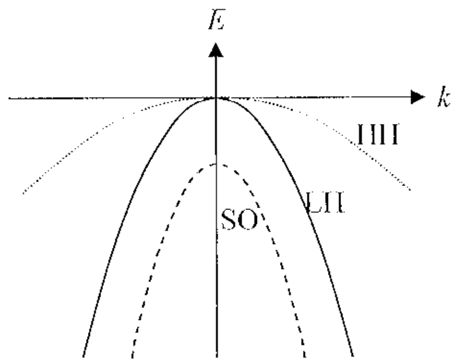


Figure 2.10: Schematic of the valence bands *E*-*k* diagrams in unstrained silicon: the heavy hole (HH), light hole (LH) and split-off (SO) bands. The band curvature determines the effective mass according to [36]: $\frac{1}{m_{ij}^{*}m_{e}} = \frac{1}{\hbar^{2}}\frac{\partial^{2}E}{\partial k_{i}\partial k_{j}}$.

other two bands, which are degenerate in unstrained silicon, are called the light hole (LH) and heavy hole (HH) bands, because the bands have different curvature and thus different effective hole mass.

As with the analysis of electron transport in strained layers, one can begin by applying the "model-solid" theory to determine how the valence band structure is altered by the application of biaxial or uniaxial silicon strain. Again, the theory assumes that the position of valence band j, E_v^j, is equal to the sum of the unstrained "intrinsic" band position, E_v^{int}, and the spin-orbit and strain-induced band shift, $\Delta E_v^{j,spin/strain}$:

$$E_v^j = E_v^{int} + \Delta E_v^{j,spin/strain}, \tag{2.19}$$

where the "intrinsic" band position (the average valence band position in unstrained silicon) was given earlier as $E_v^{int} = -7.03eV$ [33], and where j refers to a specific valence band (LH, HH or SO). The band edge shift is sum of a hydrostatic strain component (h) and a non-hydrostatic (nh) spin/strain component [33]:

$$\Delta E_v^{j,spin/strain} = \Delta E_v^h + \Delta E_{v,j}^{nh}. \tag{2.20}$$

The valence bands shift up or down in energy with hydrostatic strain according to [33]

$$\Delta E_v^h = a_v^\Gamma \left(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}\right), \tag{2.21}$$

where $a_v^\Gamma = 2.46eV$ [33].

The non-hydrostatic strain shift, $\Delta E_{v,j}^{nh}$, includes the split-off contribution and the strain-induced energy band shifts. For uniaxial strain in the <100> direction, the non-hydrostatic valence band shifts are given [28,32,33,51] by solving for the eigenvalues of the Hamiltonian matrix, such that:

$$\begin{bmatrix} \frac{\Delta_0}{3} - \frac{1}{2}\delta E_{001} & 0 & 0 \\ 0 & \frac{\Delta_0}{3} + \frac{1}{2}\delta E_{001} & \frac{\sqrt{2}}{2}\delta E_{001} \\ 0 & \frac{\sqrt{2}}{2}\delta E_{001} & -\frac{2}{3}\Delta_0 \end{bmatrix} - \begin{bmatrix} \Delta E_{v,HH}^{nh} & 0 & 0 \\ 0 & \Delta E_{v,LH}^{nh} & 0 \\ 0 & 0 & \Delta E_{v,SO}^{nh} \end{bmatrix} = 0 \tag{2.22a}$$

which can be simplified to

$$\Delta E_{v,HH}^{nh} = +\frac{1}{3}\Delta_0 - \frac{1}{2}\delta E_{001} \tag{2.23a}$$

$$\Delta E^{nh}_{v,LH} = -\frac{1}{6}\Delta_0 + \frac{1}{4}\delta E_{001} + \frac{1}{2}\sqrt{\Delta_0^2 + \Delta_0 \cdot \delta E_{001} + \frac{9}{4}\left(\delta E_{001}\right)^2} \tag{2.23b}$$

$$\Delta E^{nh}_{v,SO} = -\frac{1}{6}\Delta_0 + \frac{1}{4}\delta E_{001} - \frac{1}{2}\sqrt{\Delta_0^2 + \Delta_0 \cdot \delta E_{001} + \frac{9}{4}\left(\delta E_{001}\right)^2} \tag{2.23c}$$

where $\delta E_{001} = 2 \cdot b \cdot \left(\varepsilon_{zz} - \varepsilon_{xx}\right)$ and $b = -2.35eV$ for silicon. For the case of zero strain, the light hole and heavy hole bands are degenerate with $\Delta E^{nh}_{v,LH/HH} = +1/3\Delta_0$ while the SO band is separated by Δ_0: $\Delta E^{nh}_{v,SO} = -2/3\Delta_0$. Eqn. 2.23 can also be used to describe the effect of biaxial silicon strain.

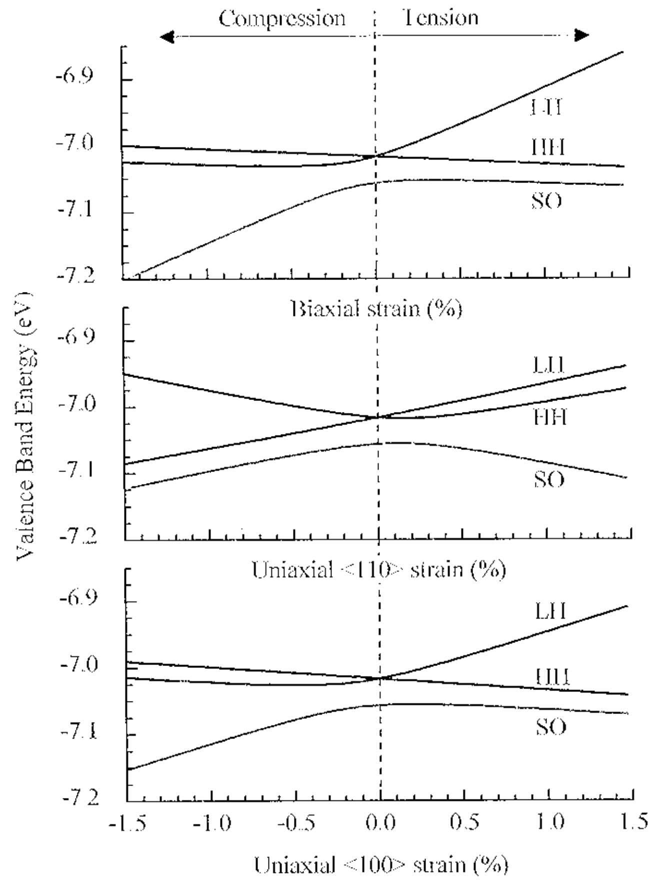
For uniaxial strain in the <110> direction, the non-hydrostatic valence band shifts are found [33,51] by solving

$$\left|\begin{bmatrix} \frac{\Delta_0}{3} - \frac{1}{8}\left(\delta E_{001} + 3 \cdot \delta E_{111}\right) & -\frac{\sqrt{3}}{8}\left(\delta E_{001} - \delta E_{111}\right) & \frac{\sqrt{6}}{8}\left(\delta E_{001} - \delta E_{111}\right) \\ -\frac{\sqrt{3}}{8}\left(\delta E_{001} - \delta E_{111}\right) & \frac{\Delta_0}{3} + \frac{1}{8}\left(\delta E_{001} + 3 \cdot \delta E_{111}\right) & \frac{\sqrt{2}}{8}\left(\delta E_{001} + 3 \cdot \delta E_{111}\right) \\ \frac{\sqrt{6}}{8}\left(\delta E_{001} - \delta E_{111}\right) & \frac{\sqrt{2}}{8}\left(\delta E_{001} + 3 \cdot \delta E_{111}\right) & -\frac{2}{3}\Delta_0 \end{bmatrix} - \begin{bmatrix} \Delta E^{nh}_{v,HH} & 0 & 0 \\ 0 & \Delta E^{nh}_{v,LH} & 0 \\ 0 & 0 & \Delta E^{nh}_{v,SO} \end{bmatrix}\right| = 0 \tag{2.24}$$

where $\delta E_{001} = 4 \cdot b \cdot \left(\varepsilon_{xx} - \varepsilon_{zz}\right)$, $\delta E_{111} = \left(4/\sqrt{3}\right) \cdot d \cdot \varepsilon_{xy}$, and $d = -5.32eV$. Equation 2.24 is best solved numerically.

2.3.2 Hole Effective Mass and Transport in Strained-Silicon Layers

The strain-induced band shifts of Eqns. 2.23-2.24 are plotted in Fig. 2.11. The application of any non-hydrostatic strain splits the light hole/heavy hole degeneracy and maintains or expands the separation of the SO band from the LH/HH bands. As the strain level increases, the band splitting increases. For biaxial tensile strains greater than 0.54%, shown in the top plot on the right, the LH/HH band splitting becomes larger than the optical phonon energy in silicon, 58 meV, significantly reducing interband phonon scattering [52]. A reduction in the interband hole scattering rate will increase the hole mobility via Eqns. 2.16 and 2.17. From Fig. 2.11, the largest LH/HH band splitting occurs for biaxial tensile strain, uniaxial tensile strain aligned to <100>, and uniaxial



Figures 2.11: Valence band shifts of the light hole (LH), heavy hole (HH) and split-off (SO) bands induced by biaxial and uniaxial strain, from Eqns. 2.23 and 2.24.

compressive strain aligned to <110>, and therefore one expects these types of strain to yield significant mobility enhancements.

The application of strain changes not only the valence band edge position, but also the valence band shape. This is in contrast to the conduction band in silicon, where it was assumed that the parabolicity of the unstrained silicon constant energy surface was unaffected by strain. Even for unstrained silicon, the LH/HH degeneracy will cause the constant energy surfaces around the Γ point to be slightly non-parabolic [53,54]. In bulk silicon, the effective conductivity hole mass can be approximated as a combination of the LH and HH masses, m^*_{LH} and m^*_{HH}:

$$m^*_{p,unstrained} \cong \frac{3}{\frac{1}{m^*_{LH}} + \frac{1}{m^*_{HH}}} = 0.35, \qquad (2.25)$$

where $m^*_{LH} = 0.15$ and $m^*_{HH}=0.54$ [22,37,38,55]. When strain is applied, the bands become more warped and non-parabolic. This strain-induced band warping changes the effective hole mass so that the meaning of the band labels ("light hole" and "heavy hole") is essentially lost. Bir and Pikus provide a thorough treatment of band shape in diamond-lattice semiconductors [53] using strain as a perturbation to the unstrained crystal Hamiltonian. Their model provides generalized *E-k* relations for the LH and HH bands, but ignores interactions between SO band and the LH/HH bands. Also, the *E-k* curves have been calculated without adding any corrections to keep the average valence band position constant. Therefore, the model of Sec. 2.3.1 should be used to describe band shifts, but this model provides a helpful glimpse of the strain-induced LH and HH band warping. Since these bands are consistently of higher energy than the SO band, as shown in Fig. 2.11, they dominate hole transport, and their shape will determine m^*_p. The *E-k* relationship is defined by Bir and Pikus [53] as:

$$E_{LH,HH}(k) = Ak^2 + a \cdot tr(\varepsilon) \pm \left[\varepsilon_k(k) + \varepsilon_{\varepsilon k}(k) + \varepsilon_\varepsilon\right]^{1/2} \qquad (2.26a)$$

$$\varepsilon_k(k) = B^2k^4 + C^2\left(k_x^2k_y^2 + k_x^2k_z^2 + k_y^2k_z^2\right) \qquad (2.26b)$$

$$\varepsilon_{\varepsilon k}(k) = Bb\left[3\left(k_x^2\varepsilon_{xx} + k_y^2\varepsilon_{yy} + k_z^2\varepsilon_{zz}\right) - k^2 \cdot tr(\varepsilon)\right] + 2Dd\left(k_xk_y\varepsilon_{xy} + k_xk_z\varepsilon_{xz} + k_yk_z\varepsilon_{yz}\right) \qquad (2.26c)$$

$$\varepsilon_{\varepsilon} = \frac{b^2}{2}\left[\left(\varepsilon_{xx} - \varepsilon_{yy}\right)^2 + \left(\varepsilon_{yy} - \varepsilon_{zz}\right)^2 + \left(\varepsilon_{zz} - \varepsilon_{xx}\right)^2\right] + d^2\left(\varepsilon_{xy}^2 + \varepsilon_{xz}^2 + \varepsilon_{yz}^2\right), \tag{2.26d}$$

where tr(ε) is the trace of the strain matrix. The constants are given [56,57] by

$$|A| = 4.1\frac{\hbar^2}{2m_e} = 1.56\times10^{-15}cm^2\cdot eV\,, \tag{2.27a}$$

$$|B| = 1.6\frac{\hbar^2}{2m_e}\,, \tag{2.27b}$$

$$|C| = 3.3\frac{\hbar^2}{2m_e}\,,\ \text{and} \tag{2.27c}$$

$$D^2 = C^2 + 3B^2 \tag{2.27d}$$

which were calculated to match the measured effective hole masses. The other values of the other constants ($a = a_v^1$, b, and d), are given in the previous section.

The resulting *E-k* curves obtained from Eqns. 2.26 and 2.27 are plotted in Fig. 2.12 for the cases of unstrained, compressively strained and tensile strained silicon. Three different types of strain are analyzed. In Fig. 2.12, in order from top to bottom are plotted biaxial strain in the {001} plane, uniaxial strain aligned to [100] and uniaxial strain aligned to [110]. For each type of strain, two plots are given. On the left, the *E-k* relationship is plotted for zero strain in the z-axis [001] direction (ε_{zz}=0). These plots are provided for comparison with other published results [5,53]. However for the silicon films considered in this work, the z-axis is under zero stress (σ_{zz}=0), and has a non-zero strain described by Eqn. 2.7; the *E-k* diagrams for these cases (given on the right-hand-side), were calculated here using the above expressions as they were not available in the literature. Finally, within each graph, as specified on the horizontal axis, on the left-hand-side *E-k* is plotted for the crystal direction parallel (in-plane) to the applied strain, while the right-hand-side shows *E-k* in the direction perpendicular to the uniaxial strain but in the (001) plane of the strained layer.

The application of biaxial strain splits the bands and changes the band shape slightly, as drawn in Fig. 2.12a,b. For biaxial tensile strains of ~0.8%, the holes will be completely confined to the upper band, and its effective mass, slightly reduced from the unstrained silicon HH band value, will dominate hole transport. The relationship of in-

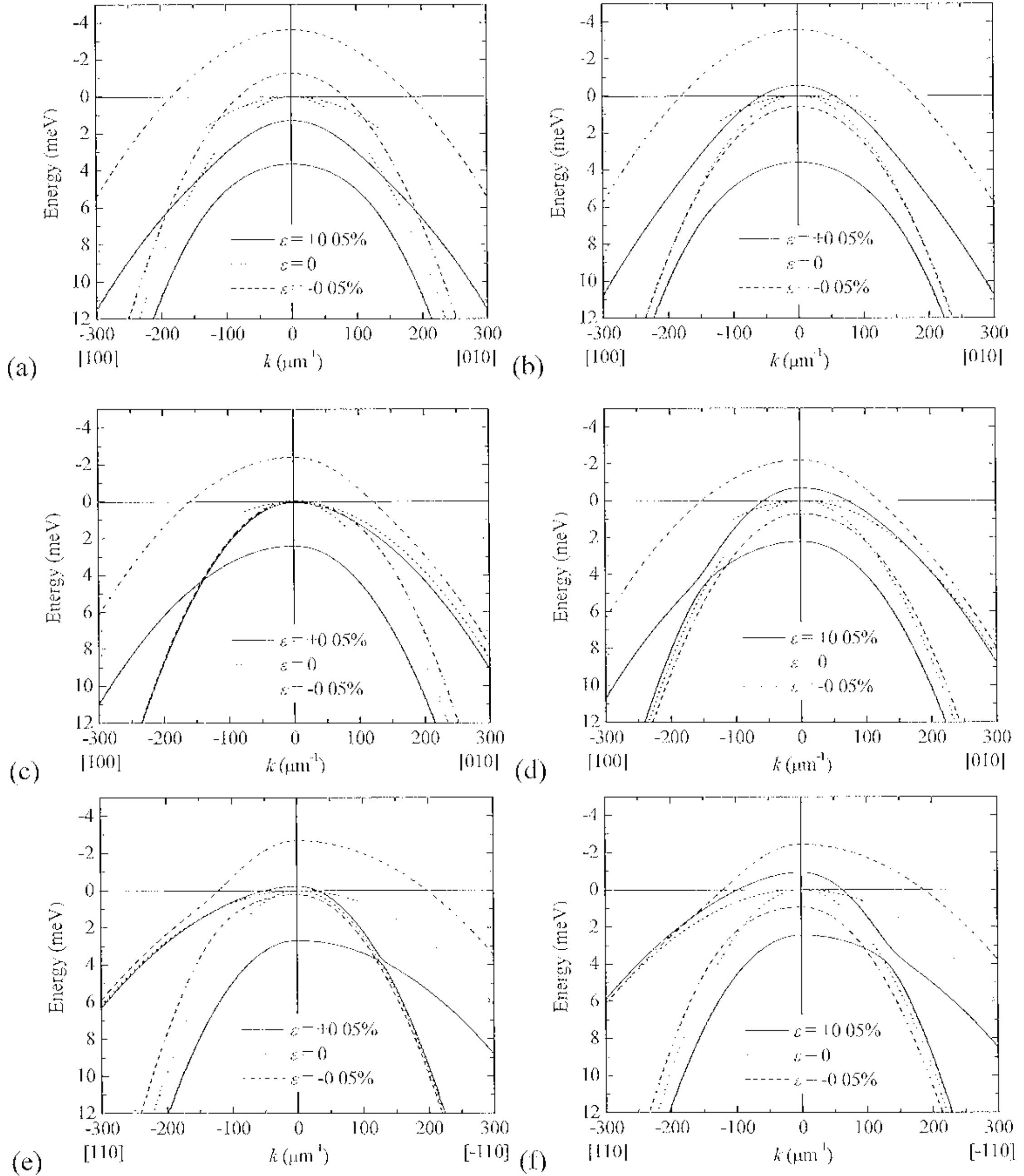


Figure 2.12: Schematic of LH and HH valence band bending for (a), (b) biaxial strain in the {001} plane, (c), (d) uniaxial strain aligned to [100]; and (e), (f) uniaxial strain aligned to [110]. For each type of strain, the plot on the left is for ε_{zz}=0 (adapted from Ref. [53]), while that on the right is for σ_{zz}=0 (*i.e.*, the matrices of Eqn. 2.7). Within each graph, the dotted lines indicate bands at zero strain, solid lines indicate tensile strain and dashed lines are for compressive strain. Bands parallel and perpendicular to the strain (but in the {001} plane of the strained film) are shown on the left-hand and right-hand side of each plot, respectively.

plane mobility to biaxial tensile strain in silicon at 300 K has been calculated using the six k·p model by Oberhuber, *et. al*., [58]; the result is plotted in Fig. 2.13. The hole mobility is expected to increase monotonically with biaxial tensile strain for low level strain due to strain-induced band splitting, sub-band population re-distribution and band warping, which reduce inter-band scattering and lower the in-plane effective hole mass. For strains of greater than 1.6%, the in-plane hole mobility enhancement is predicted to plateau at a factor of about 2.5 [58]. At these large strain levels, the transfer of holes into the upper valence band is complete and the maximum mobility enhancement is determined by its effective mass and intraband scattering rates. Measurements of p-channels FETs show that, unfortunately, the biaxial strain-induced hole mobility enhancement in FETs disappears at high vertical electric field [3,19,21]. The loss of mobility enhancement at high field has been attributed to a narrowing of the LH/HH band splitting [54], which mixes the hole population between two bands and thus increases the effective hole mass.

Compared to the case of biaxial tensile strain, the effect of uniaxial strain on the silicon valence band and hole mobility is poorly understood. Valence band warping resulting from uniaxial tensile strain in germanium (with ε_{zz}=0) has been calculated by Bir and Pikus [53], and their results agree qualitatively with the band shapes calculated in Fig. 2.12c,e for silicon under the same strain conditions. For uniaxial compressive strain along [110] (dashed lines in Fig. 2.12e), the LH/HH band degeneracy is lifted. The in-plane [110] *E-k* curve has a "dimple" [5,23,59] near k=0 that is characterized by an light-hole-like effective mass of ~m^*_{LH}. This reduced hole effective mass (in addition to the reduced phonon scattering due to band splitting) is likely responsible for the large hole mobility increases observed experimentally by a group at Intel [4,5,23,24,59]. They have noted that, significantly, the uniaxial strain-induced hole mobility enhancement persists at high vertical fields, presumably because the LH/HH band splitting is well maintained, regardless of the applied electric field. It is this high vertical field hole mobility enhancement which provides a strong impetus to pursue uniaxial strain over biaxial strain (where no high field μ_p enhancement is observed) in order to better equalize the hole and electron mobilities in CMOS circuitry.

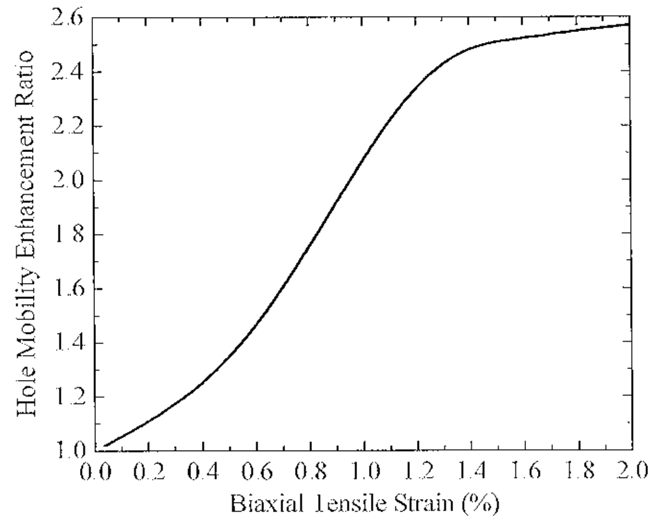


Figure 2.13: Enhancement of in-plane hole mobility in silicon at 300 K due to biaxial tensile strain, as calculated in Ref. [58].

As predicted elsewhere [3,52,60,61], under uniaxial strain the model of Eqn. 2.26 predicts that the effective masses for the directions parallel and perpendicular to the uniaxial strain (in the plane of the strained film) are not equal. Perpendicular to the uniaxial [110] compressive strain, the effective mass is heavy-hole-like, $\sim m^*_{HH}$. The inclusion of the z-axis strain (Fig. 2.12f) in the E-k equations amplifies these results, with a slightly more pronounced "dimple" and larger separation between the two bands across the observed k-space.

For uniaxial strain aligned to the [100] crystal direction, the strain-induced band warping is shown in Fig. 2.12c,d. The trends are qualitatively similar to those observed for [110] strain, but now tension is required instead of compression. For silicon tension applied uniaxially along the [100] direction, the [100] in-plane direction shows a dimpled E-k surface, with an effective mass equal to m^*_{LH} near k=0. In the perpendicular direction, the effective mass is much larger, $\sim m^*_{HH}$. Again, if the z-axis strain is included as shown in Fig. 2.12d, in the solution, the two bands are separated in energy across k-space.

In summary, the "model-solid" theory has been used to predict valence band splitting and perturbation theory used to predict valence band warping due to biaxial and uniaxial strain. Strain can induce large valence band splitting and band warping, lifting the degeneracy and repopulating the unstrained silicon light hole/heavy hole bands, and lowering the effective hole mass. To achieve low effective mass in the plane of the strained film, it is possible to use biaxial tension, uniaxial [110] compression (with carrier transport parallel to the strain), uniaxial [110] tension (with carrier transport perpendicular to the strain), or uniaxial [100] tension (with carrier transport parallel to the strain). Biaxial strain has been shown experimentally to increase FET hole mobility, but the effect is lost at high vertical electric field. In contrast, observed uniaxial <110> compressive strain-induced hole mobility enhancement has been shown to persist even at high fields. By orienting the channel in-plane but perpendicular to the uniaxial strain it may be possible to obtain high hole mobilities for other types of strain. For example, a $[\bar{1}10]$ channel might be used in conjunction with uniaxial tension along [110] to obtain a low m^*_p value. The strain-generation technique presented in this thesis allows for unprecedented control of strain orientation relative to the channel, in order to better investigate and exploit such possibilities. In Ch. 7, the phenomenological piezoresistance

theory will be presented which allows quantitative prediction of mobility enhancement, and will be compared to measured results on our biaxially- and uniaxially-strained pFETs.

2.4 Fabrication Approaches for Strained Channel MOSFETs

In the previous sections it has been shown that biaxial or uniaxial silicon strain can improve the electron and hole mobilities. Here, existing methods for generating the desired silicon strain are considered. Four approaches to strain generation are discussed: substrate-induced strain, where a heteroepitaxial mismatch between the substrate and the silicon layer creates a biaxial silicon strain; FET-process induced strain; mechanically-induced strain, introduced after FET processing; and wafer bonding methods that result in strained silicon-on-insulator.

2.4.1. Substrate-Induced Strain

Heteroepitaxy, the psuedomorphic growth of one crystal film on another, can be used to control film strain. For example, a SiGe film grown on a bulk silicon substrate will be compressively strained, as shown in Fig. 2.14. The biaxial compressive strain arises due to the mismatch between the relaxed SiGe lattice constant, a^r_{SiGe}, and the smaller silicon lattice constant, a^r_{Si}. The larger SiGe lattice is compressed to match the relaxed silicon lattice spacing. For a $Si_{1-x}Ge_x$ film grown on a silicon substrate, the resulting compressive strain in the plane of the film is

$$\varepsilon_{//}(x) = \frac{a^r_{Si} - a^r_{SiGe}(x)}{a^r_{SiGe}(x)} = -0.417x\ [\%] \qquad (2.28)$$

Conversely, if a silicon layer is heteroepitaxially grown on a relaxed SiGe layer, it will be under biaxial tensile strain.

For thin epitaxial layers, the full lattice mismatch strain is present. As the epitaxial layer thickens, the energetic cost of the strain becomes too great, and misfit dislocations are generated at the Si/SiGe interface, as shown in Fig. 2.15. The

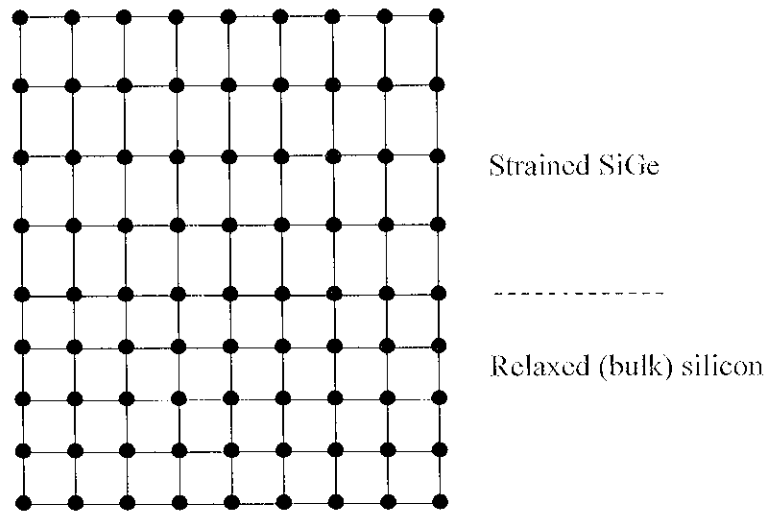


Figure 2.14: Lattice schematic of heteroepitaxially-grown $Si_{1-x}Ge_x$ on a bulk silicon substrate. The SiGe layer is under biaxial compressive strain, and the silicon substrate is assumed to remain fully relaxed even at the Si/SiGe interface.

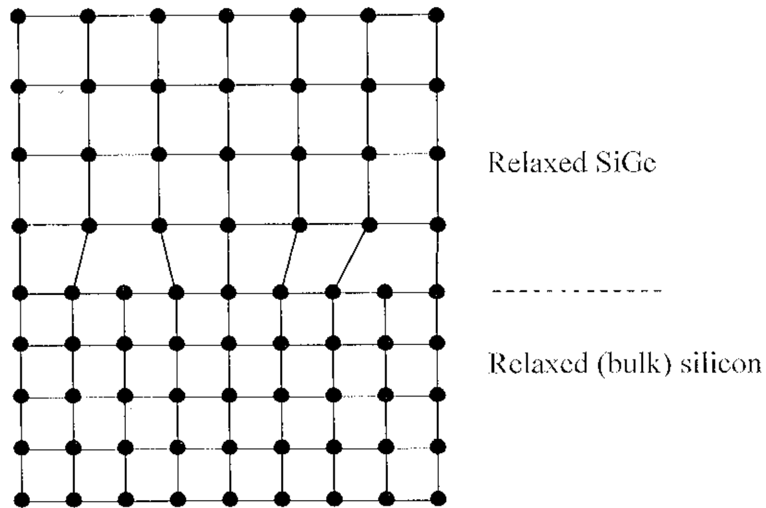


Figure 2.15: Lattice schematic of a thick $Si_{1-x}Ge_x$ layer grown on a bulk silicon substrate. The SiGe relaxes by the generation of misfit dislocations.

equilibrium, or Matthews-Blakeslee critical thickness for the epilayer, $h_{c,eq}$, below which dislocations are energetically unfavorable is defined [62-64] as

$$h_{c,eq} = \frac{b}{8\pi} \cdot \frac{1}{\varepsilon_{//}} \cdot \frac{\left(1 - \nu \cos^2\theta\right)}{\left(1 + \nu\right)\cos\lambda} \cdot \ln\left(4h_{c,eq}/b\right), \qquad (2.29)$$

where for a misfit dislocation in $Si_{1-x}Ge_x$, ν is the epilayer Poisson ratio, b is the Burgers vector equal to $a^r_{SiGe}/\sqrt{2}$, θ=60° is the angle between the dislocation line and the Burgers vector, and λ=60° is the angle between the slip direction and the direction in the interface plane that is perpendicular to the intersection of the interface with the slip plane. Epitaxial films with thickness less than $h_{c,eq}$ should, in theory, be free of dislocations. If epitaxy conditions are carefully controlled, in practice thicker layers can be obtained, up to the metastable thickness, $h_{c,MS}$. The metastable critical thickness was defined by People and Bean [65,66] as the film thickness at which its strain energy density becomes greater than the self-energy of an isolated screw dislocation:

$$h_{c,MS} = \frac{b^2}{16\pi\sqrt{2}} \cdot \frac{1}{a^r_{SiGe} \cdot \varepsilon^2_{//}} \cdot \frac{\left(1 - \nu\right)}{\left(1 + \nu\right)} \cdot \ln\left(h_{c,MS}/b\right) \qquad (2.30)$$

In Fig. 2.16 the equilibrium and meta-stable values of critical thickness are plotted versus the germanium content, x, of a $Si_{1-x}Ge_x$ epilayer grown on a silicon substrate. As germanium content increases, the strain in the SiGe layer increases according to Eqn. 2.28 and thus the critical thickness decreases.

As described above, due to the lattice mismatch between SiGe and silicon, biaxially-tensile strained silicon can be obtained by heteroepitaxially growing a thin layer of silicon on top of relaxed SiGe. To obtain a relaxed $Si_{1-x}Ge_x$ layer for the starting material, typically a SiGe buffer layer is grown on a silicon wafer, with the germanium content, x, graded from zero to the desired level (20-30%), as shown in Fig. 2.17. The graded SiGe layer is capped with a relaxed (unstrained) buffer layer of SiGe with constant germanium content. Such substrates are called "strain-relaxed substrates" or "SiGe virtual substrates." In the graded layer, the SiGe is grown above its critical thickness, so that the SiGe strain is relaxed through misfit dislocations, as drawn in Fig. 2.15. The incorporation of dislocations can be used to obtain relaxed SiGe epilayers, but at the cost of high crystal defect densities. Much work has been done to reduce the resulting surface level defect density to approximately 10^5 cm^{-2} and to obtain a smooth

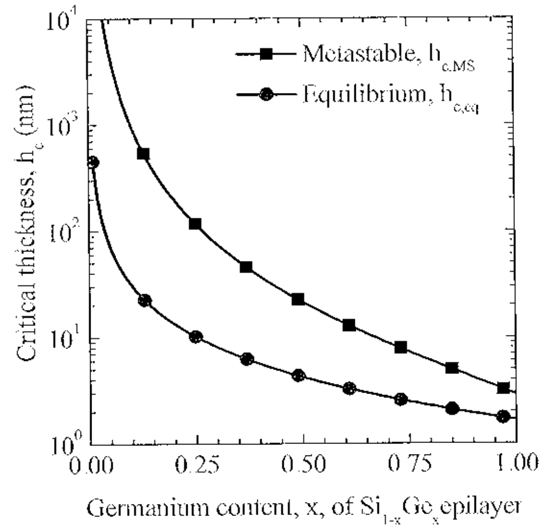


Figure 2.16: The equilibrium and metastable critical thickness vs Germanium content, x, of a $Si_{1-x}Ge_x$ epilayer grown on a silicon substrate, from Eqns. 2.29 and 2.30.

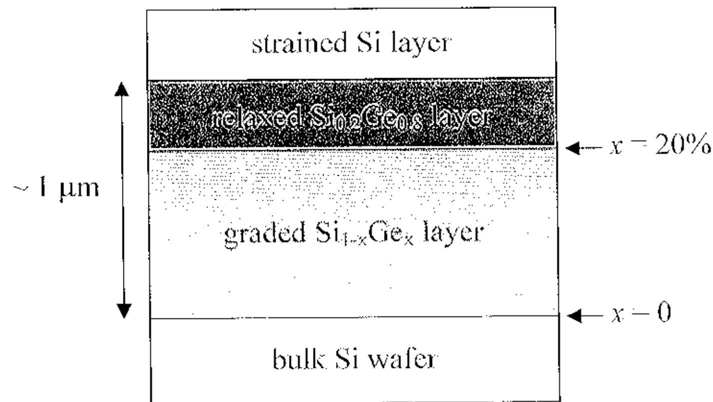


Figure 2.17: Growth of a graded $Si_{1-x}Ge_x$ layer, with $0 \leq x \leq 20\%$, followed by a strain-relaxed buffer layer of $Si_{0.2}Ge_{0.8}$ and capped with biaxially-tensile strained silicon.

top surface, with low surface roughness [1-3,7,17-19,19-21,58,67,68]. These layers are of sufficient quality for MOSFET fabrication, and biaxially-tensile strained silicon prepared in this way has shown enhanced low-field NMOS and PMOS mobilities [42].

While the SiGe relaxed substrate approach described above has been well developed, it suffers from several fundamental problems. The presence of a SiGe layer under the silicon channel modifies dopant diffusion and changes the thermal conductivity of the substrate. The thermal budget of the FET process must be limited in order to control germanium diffusion into the silicon channel, and the growth of a thick SiGe strain-relaxed buffer can be costly. But most importantly, the technique is limited to biaxial tensile strain, which has been shown (see Sec. 2.3.2) to have no enhancement effect on hole mobility at high vertical electric field. For these reasons, other strain-generation techniques should be investigated to simultaneously optimize both hole and electron mobility.

2.4.3. Process-Induced Strain

As transistor channel lengths, L, have been scaled down to less than 100 nm, it has become increasingly apparent that standard processing steps, such as source/drain silicide [69], metal gate deposition [27], and shallow trench isolation [70] can affect the channel strain and thus the FET performance. Such process-induced strain is often anisotropic and its magnitude strongly-dependent on transistor gate length and process parameters [5,27,69].

Recently, researchers have figured out how to exploit process-induced channel strain in order to create predominantly uniaxial tensile strain in n-MOSFETs and uniaxial compressive strain in p-MOSFETs, so that both electron and hole mobility are simultaneously enhanced, at least by a small amount [4,5,23,24,71-73]. Initially it was noted by NEC Corporation [71] that a silicon nitride (SiN_x) gate capping layer (see Fig 2.2), normally used an etch stop, induced mechanical stress in short channel (L_{eff}=90 nm) n- and p-MOSFETs. By changing the SiN_x deposition parameters, they were able to tune the nitride stress from compressive (-1.4 GPa) to tensile (+300 MPa). Compressive stress increased the transistor on-current, I_{ON}, for p-channel devices by 7% while decreasing I_{ON} for n-channel devices by 8%, compared to the tensile strain case. Hitachi took the process

one step further to eliminate the n-MOSFET performance degradation [72]. By implanting germanium into a compressively-strained nitride layer they were able to locally release the nitride stress in the area of NMOS devices. The PMOS devices, with a nitride cap but no germanium implant, showed an I_{ON} increase of 10% while the NMOS devices, with germanium implanted nitride, shown a 0 to 3 % increase in I_{ON}. They also noted that the results could be reversed ($I_{ON,pMOS}$ -1%, $I_{ON,nMOS}$ +12%) if a highly *tensile* SiN_x capping layer was used together with germanium implantation over the PMOS devices. This work showed, critically, the possibility of increasing (or at least, maintaining) both NMOS and PMOS mobilities on the same wafer by locally tuning the process-induced strain. However the germanium implant position is hard to control for sub-100 nm gate length devices.

Intel Corporation has taken the lead in process-induced strain in the last several years. As first reported in 2002 [24], and reported in full in 2003 [4], they have obtained simultaneous NMOS and PMOS mobility enhancement through a combination of different process-induced strain techniques. On n-MOSFETs, after source/drain salicidation a highly tensile SiN_x capping layer is deposited. For the short channel devices (L_{eff}=45 nm) used in Intel's 90-nm technology node logic process, the 75-nm SiN_x layer generates a uniaxial tensile stress of 200-300 MPa in the channel, and a compressive stress of 200-300 MPa in the out-of-plane <001> direction [73]. This uniaxial tensile stress results a 10% increase in the n-MOSFET saturation current, $I_{DS,sat}$ [5,23]. To enhance hole mobility, an entirely new technique, which builds on the heteroepitaxy process described in Sec. 2.4.1, was developed. After the sidewall spacer formation (see Fig. 2.2), the source/drain region of the substrate was recess etched. By selective epitaxy, SiGe was grown in the source/drain regions. Because the SiGe is heteroepitaxial to the silicon substrate, it has a large compressive stress, and it exerts this stress on the narrow channel region that lies between the SiGe source/drain. Using $Si_{1-x}Ge_x$ with x=17%, and by later neutralizing the effect of the SiN_x added for nFET mobility enhancement, 500-600 MPa of compressive stress was created in the silicon pFET channels (L_{eff}=50 nm), resulting in a hole mobility increase of 50% [5,73]. Most importantly, this mobility improvement was maintained at high electric field. This strained-silicon technology is currently being used in production for Intel's 90-nm technology node logic process.

By further optimizing the device design and fabrication processes and by decreasing L_{eff} to 35 nm, Intel has been able to improve μ_n by +40% and μ_p by +100% over unstrained silicon (with similar increases observed for $I_{DS,sat}$), for its 65-nm technology node logic process [6]. For a 65-nm low power process with longer gates (L_{eff} =55 nm), the pFET mobility enhancement is still +85% over unstrained silicon, slightly less than the 65-nm high-performance logic process but still a significant improvement over the 90-nm technology node [74]. IBM (White Plains, NY) has developed a similar process [25,26] for its 90-nm technology node that incorporates both the hybrid orientation approach discussed above in Sec. 2.1 and process-induced strain methods. Other researchers have succeeded in using dilute silicon carbon alloys in the source/drain regions to generate a uniaxial tensile strain in short-channel devices for electron mobility enhancement [75,76].

In summary, in recent years process-induced strain has been shown to have great promise for improving the performance of n- and p-MOSFETs simultaneously by using different processing techniques on n- and p-channel devices, resulting in opposite signs of channel strain. For nFETs, uniaxial tensile strain from highly tensile SiN_x capping layers has been shown to moderately increase the electron mobility, whereas for pFETs the hole mobility is increased by SiGe source/drains, which create a uniaxial compressive stress in the channel. Despite the significant progress in this area, process-induced strain has several notable limitations. First, the channel strain is a strong function of the process parameters and gate length, so that strain-specific process must be re-optimized for different technologies (as noted for Intel's 65-nm logic and low power processes [6,74]), and little strain-induced mobility enhancement can be expected for long channel (> 100 nm) transistors. Secondly, the process-induced strain is predominantly uniaxial in the plane of the silicon film, and can only be aligned parallel to charge carrier flow in the channel. For channels aligned to <110>, as seen above in Sec. 2.2, biaxial strain provides much stronger electron mobility enhancement than uniaxial strain. Therefore, the NMOS mobility enhancement caused by process-induced strain, at least using current approaches, will remain low, probably less than that observed for biaxial tensile strained silicon. A process that incorporates both biaxial and uniaxial strain on the same sample would thus be preferred.

2.4.2. Post-Process Mechanically-Induced Strain

Another method to generate strain in silicon MOSFETs is to add strain to the final FET substrate. In this approach, transistors are made using standard fabrication techniques. The substrate is thinned and mounted on a metal plate, which is bent cylindrically concave or convex to generate compressive or tensile silicon strain, as depicted in Fig. 2.18. Several authors have used similar techniques to generate uniaxial compressive and tensile strain in silicon, and investigate its effect on electron and hole mobility [77-83]. This technique allows the reversible application of either compressive or tensile strain in an arbitrary direction relative to the channel. However the amount of strain achieved is usually quite small (of the order of 0.1%), and only one type and direction of strain can be used for the entire sample (both nFETs and pFETs), which has already been seen to be disadvantageous. While useful for demonstrating strain-mobility relationships, it is unlikely that this method would be adopted for practical integrated circuits.

2.4.4. Strained Silicon-on-Insulator by Wafer Bonding

The final strain generation method to be discussed is that of wafer bonding. The basic process is shown in Fig. 2.19. A strained-silicon layer is grown on a donor wafer, and transferred by wafer bonding to a handle wafer. Variations on this process have been developed by several research groups [68,84-86], as well as the companies SOITEC [87,88] (S.O.I.TEC Silicon On Insulator Technologies, Bernin, France) and Amberwave [89,90] (AmberWave Systems Corporation, Salem, NH). These groups have used SiGe strain-relaxed layers on donor wafers to grow biaxially-tensile strain silicon as shown in Fig. 2.17. A separate handle wafer is coated with oxide, and the donor and handle wafers are bonded together. Either by mechanical grinding and etch back (chemical mechanical polishing, or CMP) or by use of a cleave plane, the majority of the donor wafer is removed. The remaining SiGe layer is typically removed by wet chemical etching or further CMP, leaving a biaxially-tensile strained-silicon layer on insulator. This technique represents an improvement over biaxially-tensile strained-silicon layers on SiGe relaxed buffer substrates. The removal of the SiGe layer under the strained silicon

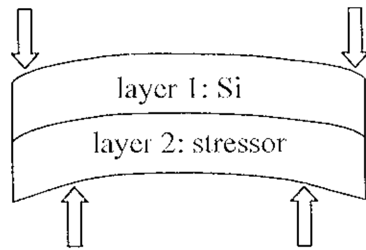


Figure 2.18: Schematic of a thin silicon substrate (layer 1) attached to a metal strip (layer 2), where the arrows indicate applied force. In this diagram, the top silicon surface is in tensile strain. If the force directions were reversed, the two layers would be concave and the silicon in compression.

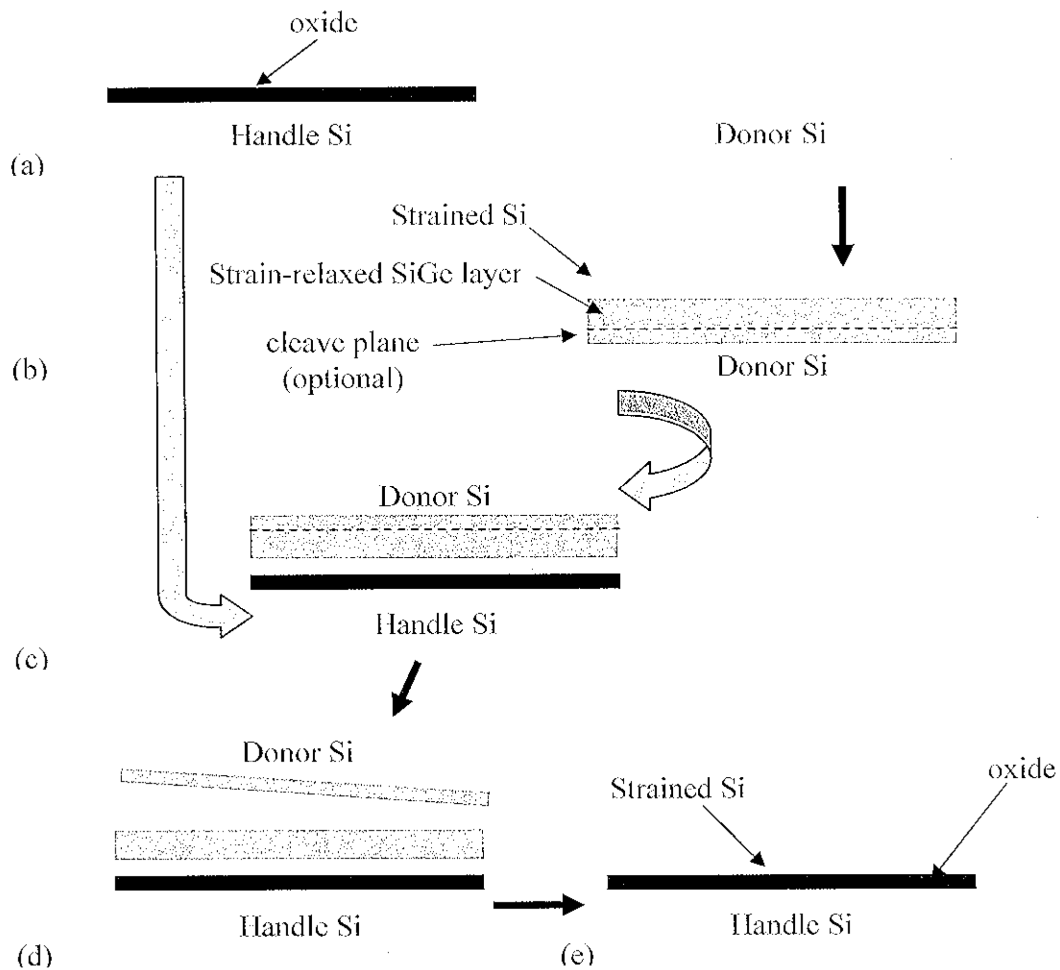


Figure 2.19: Basic wafer bonding process (a) coating a handle wafer with oxide; (b) growth of a strained-relaxed SiGe layer and strained-silicon layer on a handle wafer as shown in Fig. 2.17; (c) bonding of the donor and handle wafers; (d) removal of the majority of the donor silicon wafer; (e) removal of the SiGe layer, to leave a strained-silicon layer on oxide.

eliminates the propagation of misfit dislocation up from the SiGe buffer layer during subsequent thermal processing. The silicon-on-insulator (SOI) structure provides several advantages over bulk silicon substrates through the reduction of parasistic capacitances and easily-made shallow source/drain junctions. (Alternately, a relaxed SiGe layer can be transferred by wafer bonding and the strained silicon grown after layer transfer, to yield strained silicon on silicon-germanium-on-insulator (SGOI) [91-93], which does not provide the same benefits as SOI.) Ultimately, however, these technologies are limited by the original biaxially-tensile strained silicon source, which has been shown to provide no advantage for high-field PMOS mobilities.

Freescale Semiconductor (Austin, TX) has very recently used strained SOI in combination with process-induced strain control to obtain hybrid biaxial and uniaxial strains on the same wafer for pMOSFET mobility enhancement [94,95]. Initially the thin silicon layer is under full biaxial tensile strain, which is optimal for electron mobility enhancement. By reducing the distance between the gate and short trench isolation, strain along the channel length can relax. This leaves a predominantly uniaxial tensile strain in the plane of the film, perpendicular to the channel. Then, by using a compressive gate capping layer, as described above in Sec. 2.4.3, a compressive stress is applied parallel to the channel. The tensile strain perpendicular to the channel along with compressive strain parallel to the channel should maximum hole mobility (see Sec. 7.3.1), although the maximum $I_{DS,sat}$ improvement observed to-date in <110> p-channel MOSFETs for this hybridized-strain technique is low, about 15% [94,95]. The expected simultaneous performance improvement of nMOSFETs made in the original biaxial-strained silicon film has not yet been published. Freescale's work nicely demonstrates a method to use bonded strained-silicon-on-insulator for p-channel FET improvement, and elegantly uses an asymmetric (not uniaxial) in-plane strain. However this work also makes clear the difficulties of process-induced-strain techniques: for certain device and device-isolation geometries, $I_{DS,sat}$ was actually degraded by 10-15% (*e.g.*, in pFETs with narrow widths, the tensile strain along the width can relax, reducing μ.) Such strong dependence of transistor performance on device layout may be inevitable in aggressively scaled technologies, but it is certainly undesirable from a circuit-design perspective.

Alternate bonding techniques that generate uniaxial silicon strain have been developed by Silicon Genesis Corporation (San Jose, CA) [96,97]. The basic process, very similar to that outlined in Fig. 2.19, consists of four steps. First, on top of a silicon donor wafer are grown a etch stop layer, capped with the (unstrained) silicon device layer and (optionally) an amorphous oxide. A cleave plane is introduced below the etch stop layer by proton implant. (This process will be described more fully in Ch. 3.) Second, a handle wafer (possibly oxidized) is bonded at room temperature using plasma activation to the donor wafer. Third, the donor and handle wafers are separated by a mechanical cleaving process at room temperature. The donor wafer splits off at the cleave plane. Finally, the etch stop layer is removed and the silicon layer is smoothed, cleaned, and (sometimes) thickened.

In order to introduce uniaxial strain, two methods have been used. In the direct-transfer method, the donor and handle wafers are cylindrically bent, as shown in Fig. 2.18, during bonding and cleaving, with a <110> tangent direction. This process achieves compressive and tensile silicon strains of up to ±0.23% strain (±400 MPa). The amount of strain depends on the relative thickness of the donor and handle wafers, and on the bending radius. A second method uses an intermediate transfer substrate to obtain compressive and tensile stresses of up to 1.1 GPa. First, the donor wafer and its layers are transferred by the mechanical cleaving process onto an intermediate substrate coated with an adhesive. The etch stop layer is removed, and the silicon surface smoothed. The intermediate substrate is mechanically stretched and, in its stretched-state, is bonded to the handle wafer. After removing the adhesive and intermediate substrate, the uniaxially-strained-silicon layer is annealed and smoothed by CMP and thickened by epitaxial regrowth. Both processes result in uniaxially-strained silicon (compressive or tensile) that can be mounted either on a buried oxide, for an SOI structure, or directly onto a bulk Si substrate. The strain direction can be arbitrary (*i.e.*, <110> or <100>), but only a single type of strain is available across the entire wafer surface. While the "intermediate substrate" process results in impressively large uniaxial stresses, it is not yet clear if the process is applicable to today's 300-mm diameter silicon wafers, nor have the layers been shown to exhibit mobility enhancements (*i.e.*, the defect density and bonding quality is unknown).

2.5 Summary

In this chapter, it has been shown that strain can increase the electron and hole mobility in silicon, and thereby help advance the performance improvements required by Moore's law without further reductions in transistor gate length. The physical mechanisms of strain-induced electron and hole mobility enhancement, due to band splitting and band warping, leading to changes in the electron and hole effective conductivity masses and the relative charge carrier scattering rates, have been elucidated. Biaxial tensile strain, which has been widely used, increases both electron and hole mobilities at low vertical electric field, but provides no hole mobility enhancement at high fields. Conversely, uniaxial compressive stress has been shown to increase hole mobilities at both low and high fields. Clearly, for maximization of both electron and hole mobility, different types of strain are needed for nFETs and pFETs. Existing methods for generating biaxial and uniaxial strain have been surveyed. While several methods are well-developed and have been shown to improve hole or electron mobilities, none of the present approaches allow for biaxial and uniaxial strain to be present on the same wafer, as would be optimal, without imposing severe device geometry constraints.

In the next chapter, a novel process for generating uniaxially- and biaxially-stressed SiGe on a single wafer using a wafer bonding and layer transfer technique combined with lateral expansion of SiGe islands is presented. In subsequent chapters this method will be extended to obtain biaxially- and uniaxially-strained silicon on the same sample, and the resulting electron and hole mobility enhancements will be measured experimentally. Our process thus offers the unique possibility of simultaneously improving both hole and electron mobility.

Chapter 3

Generation of biaxially- and uniaxially-stressed SiGe on BPSG

Strain engineering of thin silicon and SiGe films is necessary for strain-induced mobility-enhancement in MOSFETs. In this chapter, biaxially- and uniaxially-stressed silicon-germanium (SiGe) on a compliant borophosphorosilicate glass (BPSG) insulator is fabricated. The strained-SiGe on BPSG structure is formed using a wafer bonding and layer transfer process. The SiGe film is patterned into islands and annealed at high temperature (> 700°C). Biaxial and uniaxial compressive stresses in SiGe can be obtained for square and rectangular islands, respectively, and under certain conditions, the SiGe film is completely relaxed. The stress is uniform over fairly large areas, and relaxed SiGe buffers are not used in the process. The magnitude of uniaxial stress generated by the process is dependent on the in-plane crystal-direction, and can be modeled accurately using the known mechanical properties of SiGe. In subsequent chapters, the strain relaxation process will be numerically modeled, and the process extended to bi-layers of Si/SiGe, in order to generate uniaxially tensile strained silicon.

3.1 Introduction to Strain Engineering Using Compliant Substrates

As described in Chapter 2, the band alignment and mobility of a semiconductor thin film are determined, in part, by the amount and type of film strain. For maximum mobility enhancement, the maximum possible strain should be used. According to Eqn. 2.28, for SiGe grown on bulk silicon, the $Si_{1-x}Ge_x$ strain increases linearly with germanium content, x. However as the germanium content increases, the critical thickness decreases, as shown in Fig. 2.16. For layers with $x > 0.5$, the critical thickness $h_{c,eq}$ is less than five nanometers. Similarly, the critical thickness for strained silicon on high germanium content SiGe layers will also be small. Layers above the critical thickness will have crystal defects which can degrade transistor performance. Therefore

in practice it is difficult to obtain by heteroepitaxy thick layers with large amounts of strain, due to the critical thickness constraint.

To circumvent this constraint, consider the following thought experiment. The SiGe/Si structure shown in Fig. 2.14 is inverted: a silicon epilayer is grown on top of a free-standing, unstrained, thin SiGe layer. As long as the SiGe layer is below its critical thickness, the silicon layer can be as thick as desired, and no misfit dislocations should occur in either layer. Moreover, the strain in the silicon will become close to zero when its thickness dominates the total Si/SiGe stack, as the thin SiGe will become compressively strained to accommodate it. This idea, of heteroepitaxy on a thin, compliant layer, was first proposed by Lo [98] in 1991. Using an energy equalization argument, Lo states that as the substrate layer is thinned to thickness h_{sub}, the effective critical thickness of the top epilayer, $h_{c,eff}$, increases according to

$$\frac{1}{h_{c,eff}} = \frac{1}{h_c} - \frac{1}{h_{sub}}, \quad (3.1)$$

where h_c is the critical thickness ($h_{c,eq}$ or $h_{c,MS}$) for an infinitely thick substrate. In Fig. 3.1, this equation is plotted with substrate thickness and effective critical thickness both normalized by h_c. By using a thin substrate layer, thick relaxed epilayers can be obtained without dislocations. For example, with $Si_{0.7}Ge_{0.3}$, the metastable critical thickness for an infinitely thick silicon substrate is 76 nm, whereas on a 100-nm Si membrane, the effective SiGe critical thickness is 320 nm. When the substrate thickness is less than h_c, then it should be possible to grow an infinitely thick, relaxed pseudomorphic layer on top without misfit dislocations. Lo [98] also showed that when $h_{sub} < 2h_c$, threading dislocations that are formed will tend to be pushed into the substrate. Such a substrate is said to be "compliant," because it allows a thick, relaxed epilayer to be grown without requiring defects.

As described so far, the preferred compliant substrate is a thin (less than one micrometer) single crystal membrane which is free-standing, *i.e.*, suspended such that it is not constrained by an underlying layer. There are two problems [99] with the practical realization of such thin film substrates: (1) the thin membrane requires mechanical support; and (2) the thin membrane must be able to expand or contract laterally to allow strain transfer from the epilayer into the substrate layer. One solution is the bonding of

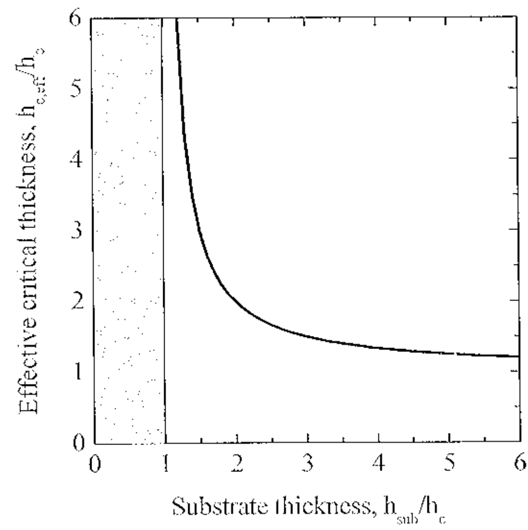


Figure 3.1: The normalized effective critical thickness, $h_{c,eff}/h_c$, vs normalized substrate thickness, h_{sub}/h_c, for epilayers grown on finite thickness substrates, from Eqn. 3.1. When the substrate thickness is less than the critical thickness, the effective critical thickness is infinite, as shown by the grey box. After Ref. [98].

the thin substrate layer to a "handle" wafer that is coated with an amorphous compliant layer [100,101] as shown in Fig. 3.2. The handle wafer mechanically supports the thin substrate layer. The amorphous compliant layer, in our case BPSG, can flow upon annealing to allow slippage of the substrate layer without introduction of dislocations. The next section reviews the requirements of the BPSG amorphous compliant layer used in our work, and describes the wafer bonding and layer transfer process used to obtain strained SiGe on BPSG. While compliant substrates were originally developed for thick, relaxed epilayers with low defect densities, the same processes can be used to generate biaxial, uniaxial and zero-stress epilayers all on the same substrate, as will be described later in this chapter.

3.2 Fabrication of Strained SiGe on Compliant BPSG

First, the BPSG handle wafer and donor wafer with SiGe/Si epitaxy must be prepared. The BPSG must meet certain requirements in order to act as an amorphous compliant layer as described above. Then, the two wafers are cleaned and bonded. Using the Smart-Cut™ separation technique[1] [102], the SiGe/Si layers are transferred from the donor wafer onto the BPSG handle wafer. Finally, the undesired donor wafer residue is removed by selective etch to obtain a thin, flat strained SiGe film on compliant BPSG. The process was initially developed by Karl Hobart at Naval Research Labs [103,104]. Details of the fabrication steps are given in Appendix A.

3.2.1 BPSG Requirements

As described in Sec 3.1, the use of an amorphous compliant layer and underlying handle wafer to support a bonded thin single-crystal film allows for strain engineering of the top film and other heteroepitaxial layers grown on it. The amorphous compliant layer must be compatible with the other processes, and must provide a smooth and clean surface for wafer bonding. An evident choice for silicon-based technologies is silicon dioxide. Initial experiments using SiGe epitaxy on bonded and etched-back silicon-on-

[1] Smart-Cut™ is a trademark of S.O.I.TEC Silicon on Insulator Technologies (Bernin, France).

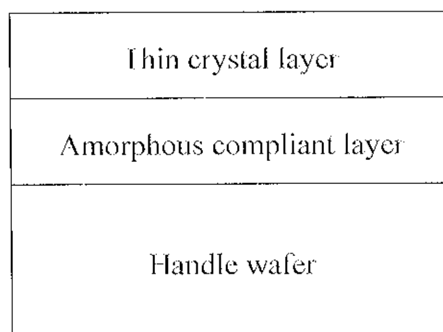


Figure 3.2: Schematic (not to scale) of a thin crystal layer bonded to a handle wafer that was coated with an amorphous compliant layer. The thin crystal layer, which can be bonded strained or unstrained, can be used as a template for growing a heteroepitaxial layer on top.

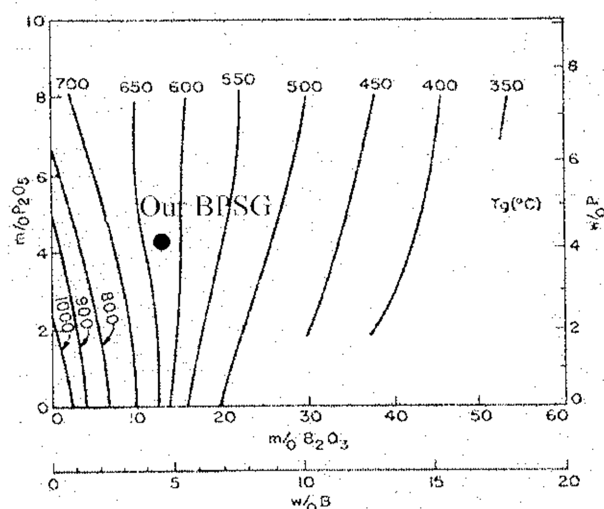


Figure 3.3: Glass transition temperature isotherms for BPSG, in degrees Celsius. The BPSG used this work has 4.4% (weight %) boron and 4.1% (weight %) phosphorus, as indicated by the black dot. The figure is from Ref. [105]. Reproduced with permission from *J. Electrochem. Soc.*, **132**, 409 (1985). Copyright 1985, The Electrochemical Society.

insulator (BESOI) [106] substrates showed that very high temperatures (> 950 °C) were needed to obtain strain transfer, from a top $Si_{1-x}Ge_x$ (x=14-15%) epilayer to the underlying silicon via viscous flow in thermally grown silicon dioxide [107,108]. This result is not unexpected as the glass transition temperature, T_g, is 1160 °C for pure SiO_2 [105]. The use of such high temperatures is undesirable, as they can cause the strained epitaxial film to develop dislocations before the compliant layer flows, as well as cause atomic diffusion between the various layers and along the film surface.

One way to lower the viscosity of silicon dioxide is to incorporate boron and/or phosphorus, that is, to make borophosphorosilicate glass (BPSG). BPSG is an amorphous matrix of SiO_2, P_2O_5 and B_2O_3 that can be deposited by chemical vapor deposition (CVD). It has been used widely in the integrated circuit industry as a passivation layer for back-end interconnect because it can reflow at moderate temperatures to smooth wafer topology and thereby improve metal step coverage [16]. The glass transition temperature, T_g, is defined as the point at which the viscosity, η, is equal to 10^{12} N·s·m^{-2}. For P_2O_5 and B_2O_3 the values of T_g are 580 and 450 °C, respectively, much lower than the T_g of thermal SiO_2 at 1160 °C. The lower T_g of boronic and phosphoric glass are thought to originate from their different molecular structures, and therefore different intra-molecular bonding, compared to SiO_2 [105,109]. In Fig. 3.3 the glass transition isotherms for BPSG are plotted [105]. As the B_2O_3 concentration increases, T_g decreases. Early experiments on relaxation of thin semiconductor layers on boron and oxygen implanted SiO_2 layers [110] and on borosilicate glass deposited by low pressure CVD (LPCVD) [101], showed significant reductions in the annealing temperatures required for glass reflow, as expected. As shown in Fig. 3.3, for B concentrations less about 5% w/o, increasing the P concentration further decreases T_g. Moreover, the addition of P decreases the T_g sensitivity to B concentration (visible in Fig. 3.3 as a wider spacing between constant-T_g curves), making BPSG process control simpler. However, both B and P concentrations need to be kept low (< ~7%) or the oxide will become hygroscopic. Such a layer will be unstable in room ambient humidity and can form boric or phosphoric acid [101,105].

In this work, the BPSG film has been deposited by CVD by Northrop Grumman (Linthicum, MD) in collaboration with Anthony Margarella, with 4.4% B and 4.1% P by weight. After deposition, the BPSG-coated wafers are annealed at 800 °C in a wet O_2

atmosphere for 1 hr, followed by a 30-min anneal at 900 °C in N_2, to densify the BPSG oxide, and are subjected to a final wet clean. These annealing steps may complicate the BPSG layer structure; see Sec. 4.2. The viscosity of the resulting film has been characterized [8,111] as 5.5×10^{10} 10^{12} N·s·m^{-2} at 750°C and 1.2×10^{10} 10^{12} N·s·m^{-2} at 800°C. These values are plotted in Fig. 3.4. Extrapolating exponentially, assuming [109] $\log\eta \sim 1/T$ (K^{-1}), a glass transition temperature of 667°C, at which $\eta=10^{12}$ N·s·m^{-2}, is predicted. The values of η from Refs. [111] and [8] as well as those extracted in this thesis, assume elastic constants of Si and SiGe near room temperature for experiments at 750-800°C. This leads to a slight overestimation (on the order of +10%) of the viscosity values [112-116]. Nonetheless, the extracted value for T_g of 667°C is close to the 625°C predicted in Ref. [105] and shown in Fig. 3.3.

Figure 3.5 shows the measured surface roughness of a 200-nm BPSG layer after post-deposition anneal and clean. The RMS surface roughness was measured by Atomic Force Microscopy (AFM) at the center of the 150-mm diameter wafer and again 10 mm from the wafer edge. Both locations show very low surface roughness of less than 0.3 nm for small AFM scan sizes (< 10x10 μm^2). (For larger scan sizes, it is more likely that surface particulate will be scanned and thus the areally averaged RMS surface roughness increases.) A typical AFM scan over 2x2 μm^2 is shown in Fig. 3.6. The BPSG is very smooth, equivalent to the surface as-received prime-grade silicon wafers. Such BPSG meets our requirements for a compliant layer: the deposition process is CMOS-compatible and results in a very smooth film with a low glass transition temperature.

3.2.2 Wafer Bonding and Layer Transfer Technique

The bonding and layer-transfer process has six steps: (1) preparation of the donor wafer and proton implant, (2) wafer cleaning, (3) initial bonding of the donor and handle wafers, (4) anneal to strengthen the bond, (5) layer transfer of donor epitaxial layers to handle wafer using Smart-Cut[TM], and (6) selective etch to remove remaining donor wafer layer and yield SiGe on BPSG structure. Figure 3.7 outlines the process flow, and Appendix A provides detailed recipes and notes.

A layer of 30-nm compressively strained $Si_{0.7}Ge_{0.3}$ is pseudomorphically grown on a sacrificial silicon substrate with (001) surface orientation [Fig. 3.7(b)]. The SiGe is

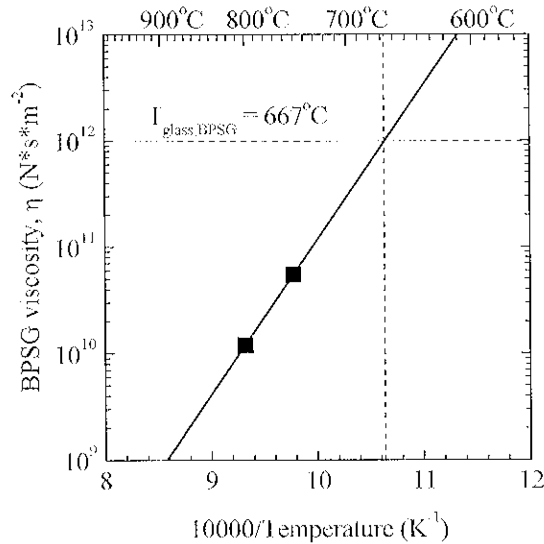


Figure 3.4: Measured BPSG viscosity values from Ref. [111], after [8]. By exponential extrapolation (the solid line), a glass transition temperature, T_g, of 667°C (indicated by the dotted lines) is expected for this BPSG.

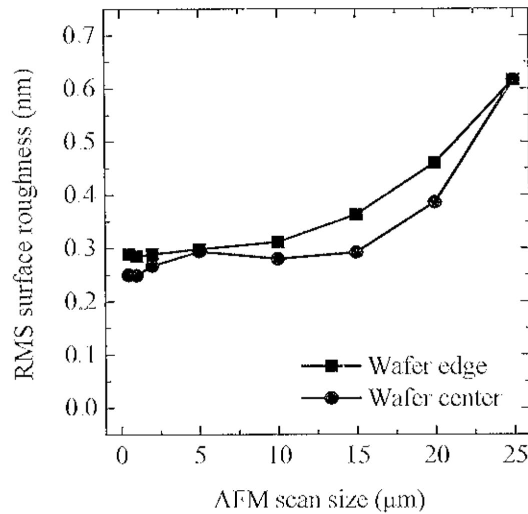


Figure 3.5: Surface roughness of 200-nm BPSG (batch #2, slot #4) after post-deposition anneal, before bonding. The RMS surface roughness was measured by AFM at the center of a 150-mm wafer (black squares), and 10 mm from the edge (red circles). The surface roughness increases with scan size due to the greater likelihood of scanning over dust or other surface particulates.

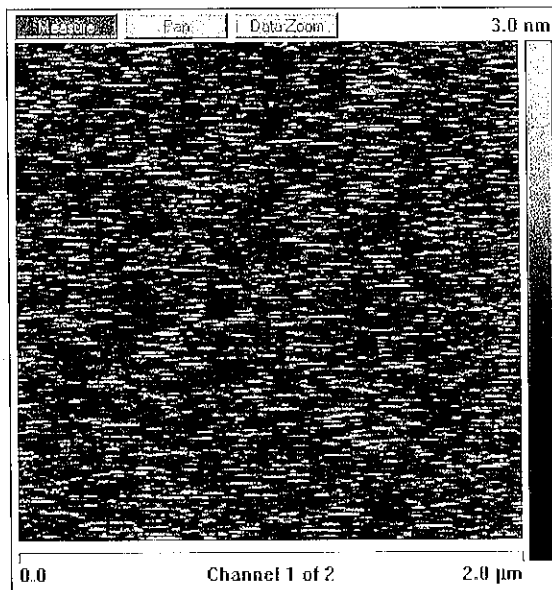


Figure 3.6: Typical AFM scan of BPSG surface roughness after image flattening. The wafer is the same as that characterized in Fig. 3.5. This scan was taken 10 mm from the edge of a 150-mm wafer, with a scan size of 2 µm x 2 µm and RMS surface roughness of 0.29 nm.

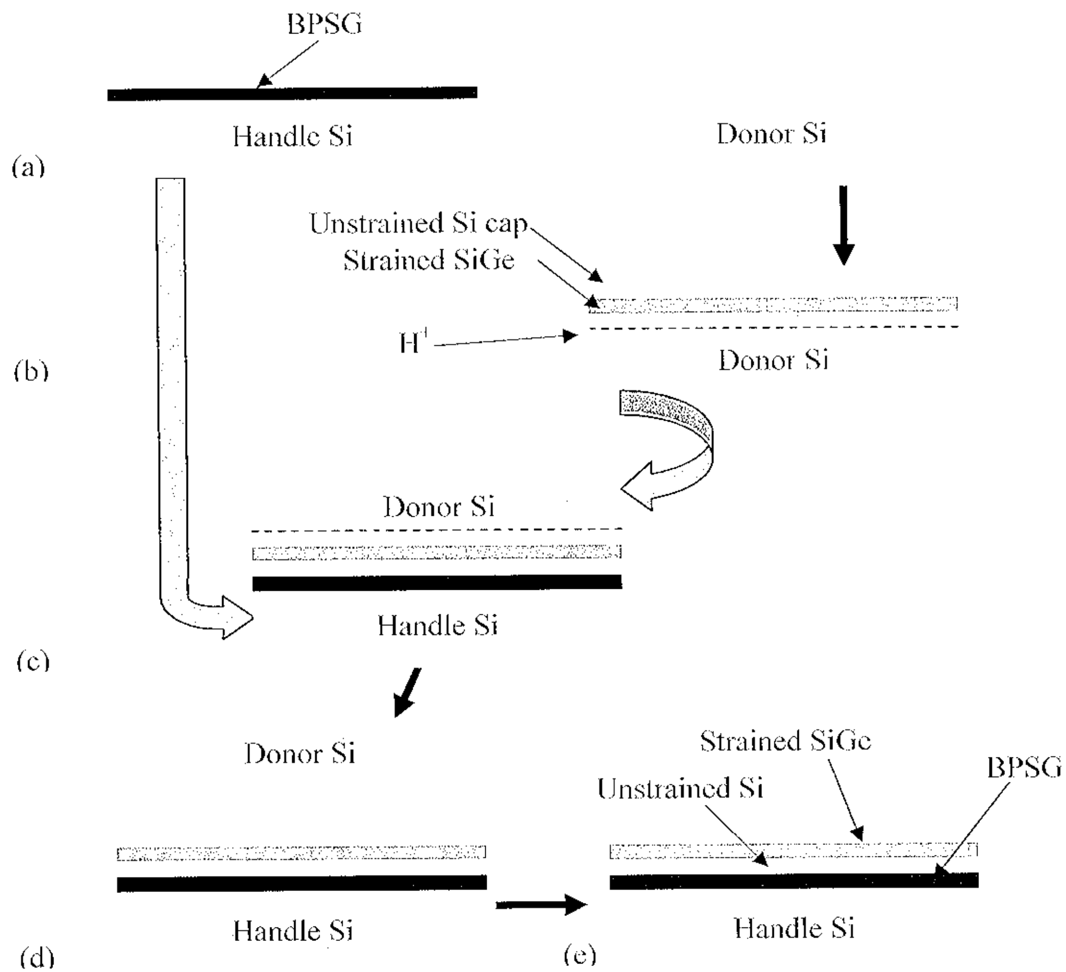


Figure 3.7: Layer transfer process for SiGe/Si bi-layer onto BPSG: (a) BPSG (borophosphorosilicate glass) is deposited on Si handle wafer; (b) Heteroepitaxy of SiGe/Si bilayer on silicon (001) donor wafer, followed by H^+ proton implant; (c) Cleaning and bonding, bond strengthening anneal; (d) Smart-Cut™ split and selective etch; (e) Final strained SiGe on unstrained Si on compliant BPSG structure.

fully strained with an initial biaxial strain, ε_0, of -1.2%. Since the layer is thin and strained, it can be grown with very low defect densities. The SiGe is capped with a thin (1-2 nm) unstrained silicon layer. The silicon cap layer is present so that the bonding surface is reproducibly hydrophilic. Lawrence Semiconductor Research Laboratory in Tempe, AZ performed the epitaxial CVD of SiGe/Si on 4" prime grade silicon wafers. Next, these donor wafers are ion implanted with H^+ of energy 150 to 180 keV and dose of 4.5×10^{16} cm^{-2}. The hydrogen implant is done by Implant Sciences Corp. (Wakefield, MA) and will be used to split off the donor epitaxial layers after wafer bonding, using the Smart-Cut™ process. Upon receipt, the donor wafer is cleaned with a 10-min 3:1 H_2SO_4:H_2O_2 30% etch at 125 °C. If desired, the BPSG wafer is cleaned for 16 min in SC-1 etchant (5:1:1 D.I. H_2O : 30% H_2O_2 : NH_4OH) at 40 °C and then annealed, to further densify and outgas the BPSG. For the samples used later in this chapter, the BPSG thickness is 1 μm.

Before wafer bonding, the wafers undergo a two-step cleaning process of a 2- to 5-min O_2 plasma followed by a 16-min SC-1 wet etch, rinse and dry. The two wafers are hydrophilically bonded [Fig. 3.7(c)] and an infrared image is taken to observe bond voids. Voids arise due to particles on the surface of the wafers that mechanically separate the two wafers and locally prevent bonding. Minimization of particles is thus a key requirement for successful wafer bonding [117]. An infrared image taken after initial bonding of one wafer pair is shown in Fig. 3.8(a). The areas of white-gray concentric arcs or rings indicate voids. To strengthen the bond, the wafers are annealed at 250-260 °C for between 4 and 14 hr. Note, the voids are not eliminated by this anneal [Fig. 3.8(b)]. Next, the bonded wafers are briefly heated to 400 °C or more to initiate the Smart-Cut™ layer transfer [102,118]. At these temperatures the hydrogen ions, trapped at the implantation depth, form H_2 gas which expands and causes thermal exfoliation of the top of the donor wafer. The donor wafer fractures parallel to the bonded interface with an audible crack [Fig. 3.7(d)]. A proton implant energy of 150-180 keV results in an implant depth of approximately 700-800 nm, which is approximately where the fracture occurs; the crack propagates well below the SiGe/Si epitaxial layers. The implant depth is confirmed by measuring with surface profilometry a step height of 715 nm from BPSG onto the as-split surface of the donor layers (wafer #021705A) with H+ at 150 keV. If the bond between

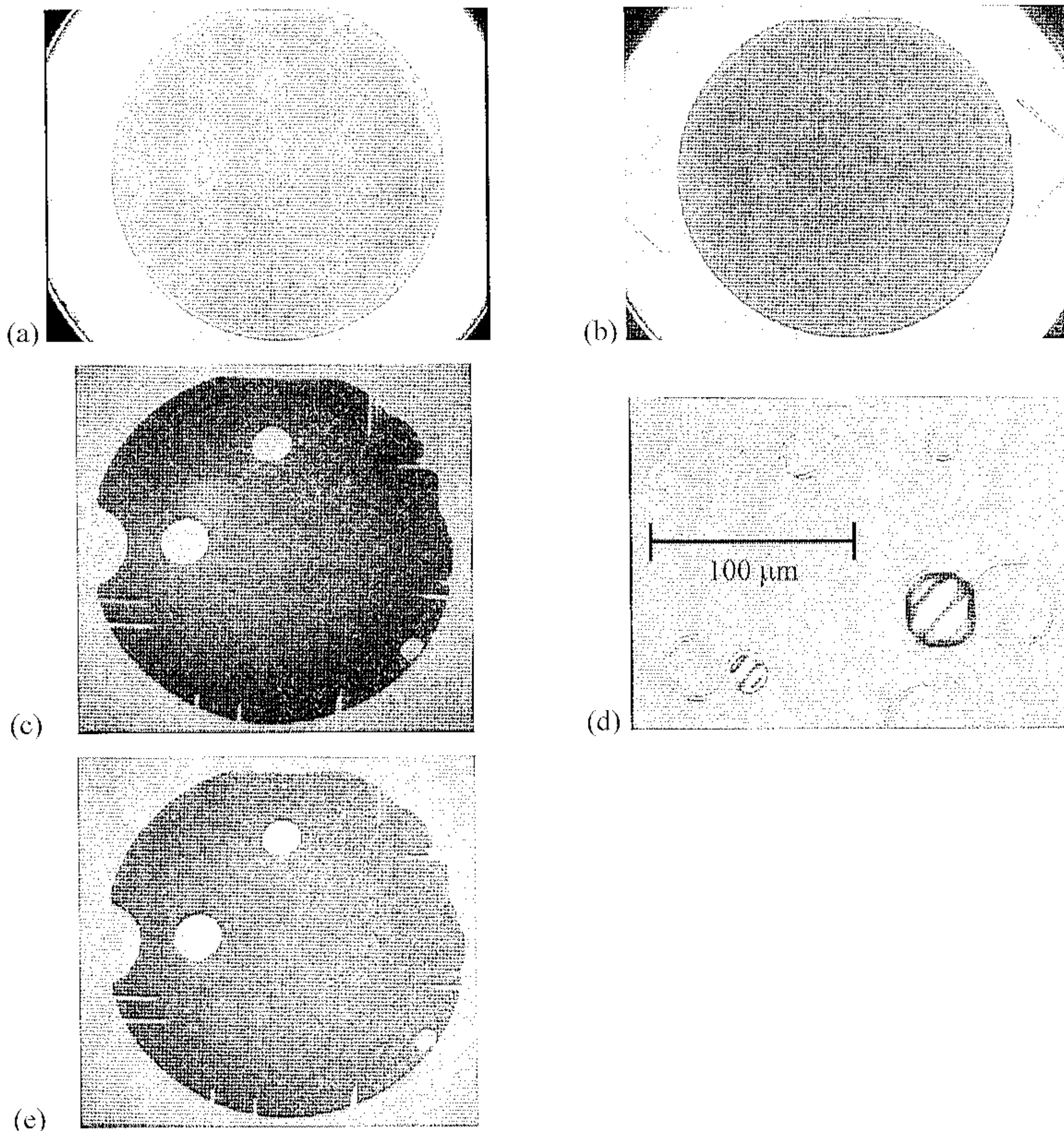


Figure 3.8: Infrared images of wafer #160219N (021005B): (a) after initial bond. The large white circle follows the outline of the 6" handle wafer while the center dark circle is the 4" donor wafer. The white spots are voids where the wafers failed to bond due to surface particulate. (b) after the 4-hr 250 °C bond strengthening anneal. The voids are in the same positions as in (a); (c) after the Smart-Cut^{TM} donor layer transfer, before the 650°C anneal. The layer did not transfer in the large void areas and in the numerous small voids, which are likely due to the SiN_x cap on the SiGe/Si donor wafer used in this case. (d) A picture of the small voids before the 650 °C anneal, with scale indicated. (e) after the 2-hr 650 °C anneal, before selective etch. On all pictures, shadows are artifacts from poor camera settings.

these layers and the BPSG on the handle wafer is strong enough, the layers will be successfully transferred to the handle wafer.

To measure typical bond strength, the crack-opening method is used. Two BPSG wafers were bonded to one another using a procedure similar to that described above, ending with a 4-hr bond strengthening anneal at 250 °C, *i.e.*, with no proton implant and no Smart-Cut^{TM} split. A razor blade of thickness, t_b, was used to slightly separate the two wafers at the edge. By measuring the size, t_g, of the resulting horseshoe-shaped opening shown in Fig. 3.9, the bond energy, γ, can be calculated [117]:

$$\gamma = \frac{3t_b^2 E_1 h_1^3 E_2 h_2^3}{16 t_g^4 \left(E_1 h_1^3 + E_2 h_2^3\right)}, \tag{3.2}$$

where E_i and h_i are the Young's modulus and thicknesses of the two bonded wafers. For two identical wafers, this equation reduces to

$$\gamma = \frac{3t_b^2 E h^3}{32 t_g^4} \tag{3.3}$$

For the wafer pair shown in Fig. 3.9 the average gap distance, t_g, was 1.60 cm. With $E_{(001)Si}$=148 GPa, t_b = 100 μm and h = 670 μm, γ = 0.637 J·m^{-2} or 637 ergs·cm^{-2}. This bond energy is not particularly high, but can be increased by modifying the pre-bond cleaning processes as described in Appendix A and Ref. [119,120]. By using a short O_2 plasma and short SC-1 etch just before bonding, the number of surface OH groups is greatly increased, aiding the formation of Si-O-Si bonds at the bond interface [119]. Implant and bond strengthening anneal conditions also significantly affect bond strength. As the proton implant dose increases, the anneal temperature required to achieve a strong bond decreases. However, to minimize crystal damage, the proton dose must be kept as low as possible. Transmission electron microscopy (TEM) of as-bonded Si and SiGe layers transferred by Smart-Cut^{TM} has shown no defects due to the hydrogen implantation for doses of $< 5\times10^{16}$ cm^{-2} [102,103]. In general, the bond strength should be as high as possible in order to ensure successful layer transfer and minimize the as-split surface roughness [120].

An infrared image of the wafer after layer transfer is shown in Fig. 3.8(c). The large void areas did not transfer. In this case numerous small voids have also formed

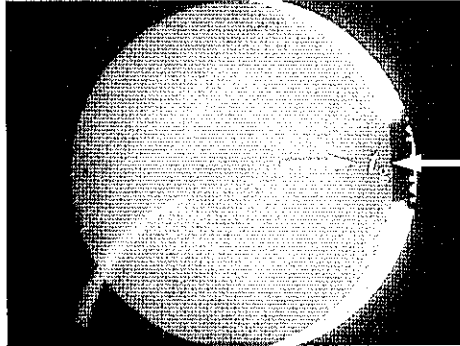


Figure 3.9: Infrared image of the "horseshoe" gap formed at the separation point between two 6" BPSG wafers that were bonded using the procedure of Appendix A. The wafers are separated by inserting a razor of known thickness. By measuring the horseshoe gap depth, t_g, indicated by the white arrows, the bond energy can be calculated according Eqn. 3.2 or 3.3. Two BPSG wafers were used because this crack-separation method requires the bonded wafers to be of identical diameter.

because the donor wafer was coated with LPCVD SiN_x. It is hypothesized that the SiN_x/BPSG bonding interface prevents outdiffusion of H_2 gas and thus trapped H_2 bubbles burst through the donor layer surface causing the small voids, shown in detail in Fig. 3.8(d). These small voids are usually not present when the SiN_x capping layer is eliminated, because in that case the H_2 gas (or its reactant H_2O) can diffuse along the Si/BPSG bonding interface and escape [119]. The RMS surface roughness of the as-split surface in areas where bonding is successful is measured to be 3.0 nm. A typical AFM scan is shown in Fig. 3.10.

After successful splitting, the bonded wafers were typically annealed at 550-650 °C for 1-2 hrs to further strengthen the bond. The infrared image after this high temperature anneal is shown in Fig 3.8(e); there are no visible changes to the bonded layer. Finally, the remaining the silicon from the original donor wafer is removed by a selective etch in 45% KOH: D.I. H_2O 1:5 after a brief dilute HF etch to make the silicon surface hydrophobic [Fig. 3.7(e)]. The KOH etch rate of silicon at 40 °C is approximately 100 nm/min. KOH is known to etch SiO_2 very slowly compared to silicon (Ref. [121] indicates a selectivity of ~180), and thus the BPSG layer remains intact. Moreover, the etch selectivity of silicon to $Si_{0.7}Ge_{0.3}$ is estimated to be ~20 [104]. Thus, the SiGe layer acts as an etch stop. The RMS surface roughness of the final 30-nm SiGe layer on 1-μm BPSG is measured by AFM to be 0.59 nm over a 5x5-μm^2 area and 0.93 nm over a 25x25–μm^2 area. A cross-section of this structure is shown in Fig. 3.11, where for simplicity the very thin (~1 nm) silicon layer between strained SiGe and BPSG is neglected.

3.3 Characterization of Film Strain and Crystal Orientation

As transferred, the SiGe film is still fully compressively strained. Accurate measurement of film strain is needed. In this section, two strain characterization methods will be described: (1) X-ray diffraction and (2) micro-Raman spectroscopy. Both are non-invasive, non-destructive techniques that require little sample preparation and yield repeatable data on average strain in the plane of the film. A variation on the latter method, polarized Raman spectroscopy, can be used to roughly determine the crystal

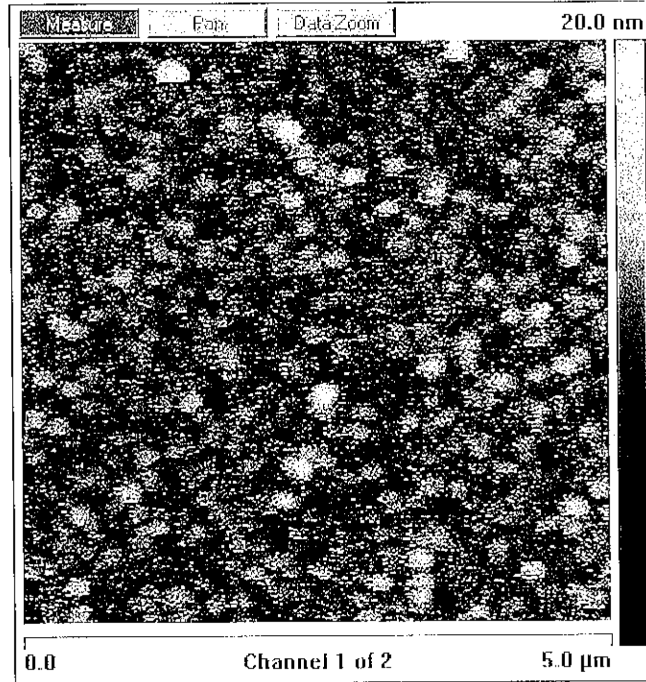


Figure 3.10: AFM scan of wafer #021705A after Smart-Cut™ layer transfer, before HF and KOH etches. The average RMS surface roughness for 5-µm x 5-µm scans at three different locations is 3.0 nm.

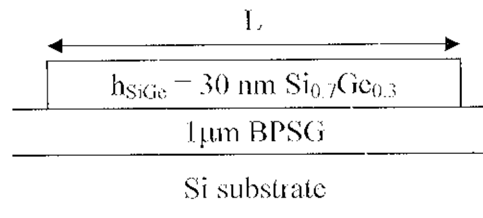


Figure 3.11: Schematic cross-section of the final strained 30-nm $Si_{0.7}Ge_{0.3}$ on 1-µm BPSG structure. The thin (1-2 nm) silicon layer under the SiGe has been omitted for simplicity. Here, the SiGe layer has been patterned into an island with dimension L.

orientation of the film. Electron back-scattering diffraction, or EBSD, can be used to determine crystal orientation to greater precision. These techniques will also be described.

3.3.1 X-ray Diffraction

X-ray diffraction (XRD) uses the periodicity of the crystal lattice to measure lattice spacing and thus strain. An incident x-ray beam elastically scatters off the Si or Ge atoms, which lie in crystal planes. Parallel crystal planes cause constructive interference of the reflected beam, which is detected as a strong increase in signal intensity. The Bragg condition [41] for constructive interference is

$$2d\,sin\theta = n\lambda\,, \tag{3.4}$$

where d is the spacing of adjacent crystal lattice planes, θ is the incident angle between the beam and crystal lattice plane, λ is the x-ray wavelength, and n is an integer, here taken to be one. The reflected x-ray intensity is measured as a function of angle, θ. A schematic of the experimental apparatus is shown in Fig. 3.12. The sample is rotated in the plane of the sample surface during measurement, as for powder measurements, in order to average out any z-axis misalignment between the bonded and substrate layers [122]. (Ideally the average of two measurements taken 180° apart would be used, instead of sample rotation. However for the apparatus used here (Rigaku MiniFlex XRD, Imaging and Analysis Center, Princeton University) precise discrete rotation is not possible and thus measurements were taken under continuous rotation.) For films with (001) surface orientation, the strongest peak to observe is from the (004) reflection, for which the perpendicular lattice constant, $a_\perp$, is equal to $4d$. Using $a_\perp$, the vertical strain in the film, $\varepsilon_\perp$, can be calculated from [122]

$$\varepsilon_\perp = \frac{a_\perp - a_r}{a_r} = -2\left(\frac{c_{12}}{c_{11}}\right)\left(\frac{a_{//,\,avg} - a_r}{a_r}\right) = -2\left(\frac{c_{12}}{c_{11}}\right)\varepsilon_{//,\,avg}\,. \tag{3.5}$$

The average strain in the plane of the film, $\varepsilon_{//,\,avg}$, defined in Eqn. 2.28 for the case of biaxial-strain, can then be calculated from the right-hand-side of Eqn. 3.5, where c_{11} and c_{12} are the elastic stiffness coefficients found in Table 3.1. Typical scans are shown in Fig. 3.13 for a silicon substrate and for a strained SiGe on BPSG on silicon sample. The

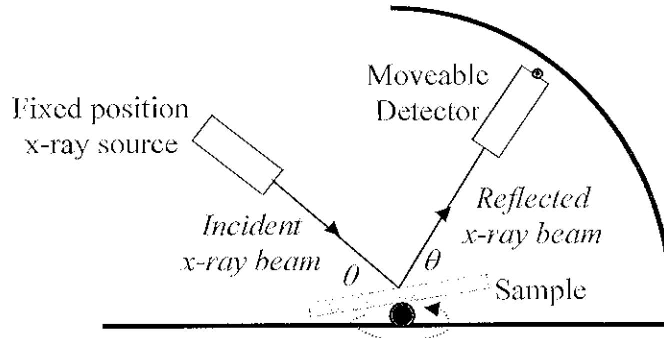


Figure 3.12: Schematic of x-ray diffraction apparatus. The x-ray beam (λ=1.540598 Å) is incident on the surface of the sample at angle θ, and is reflected by the same angle. The angle θ is scanned by tilting the sample and moving the detector. Throughout the measurement, the sample is rotated in the plane of the sample surface as indicated by the dotted arrow.

Material	p	q	ω_o (cm^{-1})	c_{11} (10^{10} N/m^2)	c_{12}	c_{44}	s_{11} (10^{-12} m^2/N)	s_{12}	s_{44}
Si	-1.85	-2.30	521.0	16.48	6.35	7.90	7.72	-2.15	12.66
$Si_{0.7}Ge_{0.3}$	-1.73	-2.20	501.5	15.39	5.89	7.53	8.24	-2.28	13.27

Table 3.1: For silicon and $Si_{0.7}Ge_{0.3}$, values at 300 K of: deformation potentials, p and q [123]; Raman frequencies for bulk materials, ω_o [124]; elastic stiffness constants, c_{11}, c_{12} and c_{44} [34], linearly interpolated between Si and Ge; and elastic compliance constants, s_{11}, s_{12} and s_{44}, calculated from the stiffness constants (see Ref. [125] and App. B.)

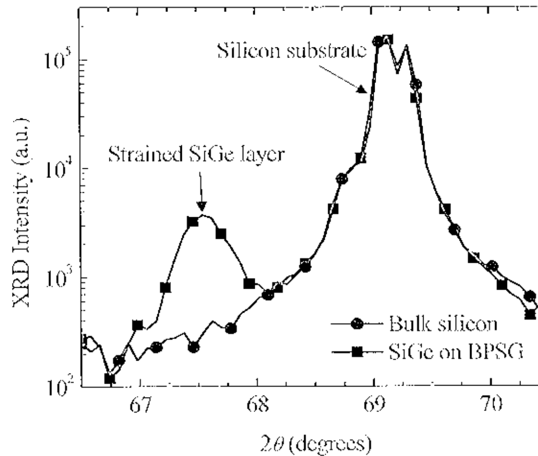


Figure 3.13: X-ray diffraction data for two samples: (a) a bulk silicon wafer and (b) 30-nm $Si_{1-x}Ge_x$ (assumed x=27%) on 95-nm BPSG. The double peaks around 2θ=69° are the first and second Cu Kα peaks from the silicon substrate, while the broad peak at 2θ=67-68° is from SiGe.

silicon peaks, which come from the bulk silicon substrate in both cases, indicate zero strain. The angle of the SiGe peak, 67.6°, (a shift of 1.5° from the silicon substrate peak) corresponds to biaxial compressive strain of -1.2% strain, assuming a Ge content of 27%. The measurement error is estimated to be ±0.04° or ±0.07% strain. Due to the large XRD beam size (0.1x1 cm^2), the areally averaged strain is measured. In order to measure strain in smaller geometries, micro-Raman spectroscopy is used, as described in the following section.

3.3.2 Micro-Raman Spectroscopy

In Raman spectroscopy, a beam of UV or visible light is inelastically scattered off the crystal. As the incident photon scatters, it creates or destroys a phonon (first-order Stokes or anti-Stokes Raman scattering, respectively). The change in wavenumber between the incident and reflected photons is measured. Because crystal momentum is conserved, peaks in the Raman spectra can be quantitatively correlated to a specific phonon, with peak position revealing (in our case) SiGe alloy content and film strain [124,126-131]. In micro-Raman spectroscopy, the incident light is a focused laser beam with a spot size of a few μm, enabling probing of small geometries. If the polarization of the incident and reflected beams is controlled, the peak magnitude can be used to determine the film's crystal orientation.

The Raman apparatus, in the lab of Prof. T. S. Duffy, Geosciences Dept., Princeton University, is drawn schematically in Fig. 3.14. It uses 488- or 514-nm wavelength light from an argon ion laser to generate anti-Stokes Raman scattering. Using the objective (labeled O1) and eyepiece (EP) to focus on the sample surface, the sample stage position is adjusted in 3D to align the laser beam position to specific locations on the sample. The incident laser beam has a spot size of ~3 μm. The liquid-nitrogen cooled CCD detector in a spectrometer collects the reflected photons. A computer acquires the spectra and, separately, a *matlab* program (detailed in Appendix C) is used to identify the wavenumber positions of the Raman peaks.

In Fig. 3.15 Raman spectra are shown for the same silicon substrate and strained SiGe on BPSG on silicon samples described above in Sec. 3.3.1. The spectral range plotted in Fig. 3.15 (490-530 cm^{-1}) corresponds to scattering with Si-Si phonons. In

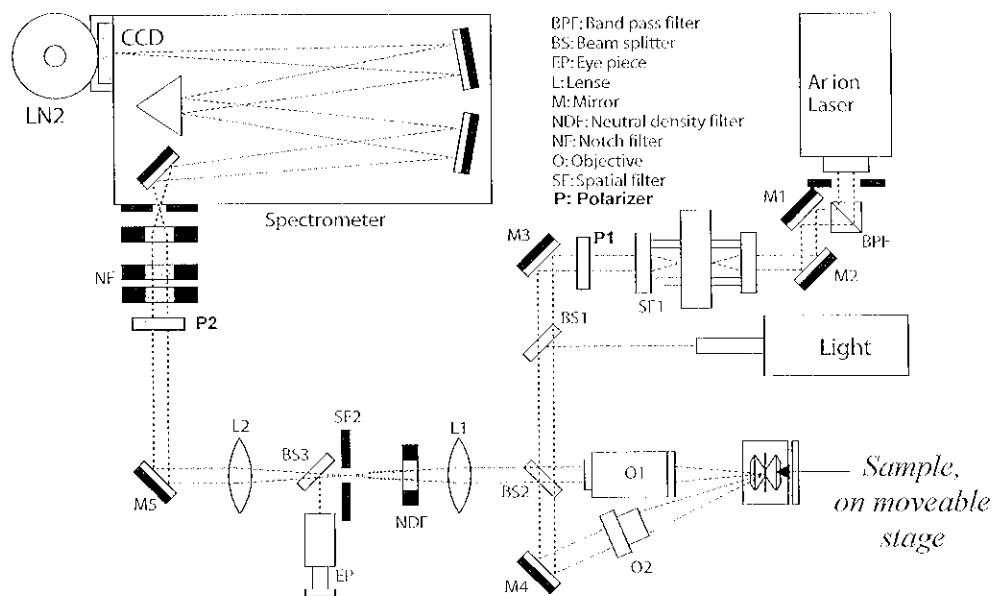


Figure 3.14: Schematic of micro-Raman spectroscopy apparatus, top-down view. The LN2 is the liquid nitrogen chamber, cooling the CCD detector; the remaining labels are described in the legend. The two polarizers, P1 and P2, are only used when determining crystal orientation by polarized Raman spectroscopy. Image courtesy of Prof. T. S. Duffy, Geosciences Dept., Princeton University.

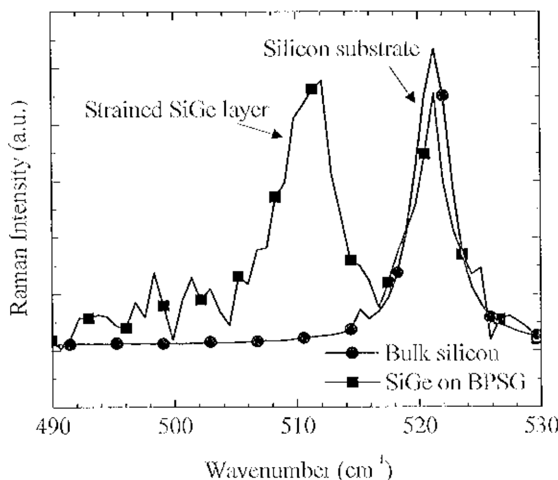


Figure 3.15: Micro-Raman spectroscopy from the same two samples described in Fig. 3.13: (a) a bulk silicon wafer and (b) 30-nm $Si_{1-x}Ge_x$ (assumed x=27%) on 95-nm BPSG. The peaks at 521 cm^{-1} are from the silicon substrate while the peak around 510 cm^{-1} is from the thin strained SiGe layer. The raw data of sample (a) has been vertically resized because it is of much greater amplitude than that of sample (b), due to the reflections in sample (b) at the BPSG interfaces which obscure the substrate signal [8]. The SiGe layer yields low Raman intensities because it is so thin.

$Si_{1-x}Ge_x$ alloys, peaks from Ge-Ge (~300 cm^{-1}) and Si-Ge phonons (~410 cm^{-1}) are also present, with amplitudes proportional [124] to x^2 and $2x(1-x)$, respectively, as shown in Fig. 3.16. For SiGe layers with low Ge content, these peaks have relatively low amplitudes compared to the Si-Si phonon peak which has an amplitude proportional to $(1-x)^2$. Thus, from this point forward only the Si-Si peaks will be considered.

The peaks from the bulk silicon substrates occur at the relaxed wavenumber, ω_o, of 521.0 cm^{-1}. The Si-Si phonon peak for relaxed (cubic) $Si_{0.7}Ge_{0.3}$ would be at 501.5 cm^{-1}; Ge content shifts the Si-Si peak of cubic crystals to lower wavenumbers. Strain likewise shifts peaks: tensile strain shifts peaks to lower wavenumbers, whereas compressive strain causes shifts to higher wavenumbers. From the measured Si-Si phonon peak location, $\omega_{measured}$, strain is calculated using [123]

$$\omega_{measured} = \omega_o + \omega_o\left(q - \frac{c_{12}}{c_{11}}p\right)\varepsilon_{//,avg}. \tag{3.6}$$

As in Sec. 3.3.1, $\varepsilon_{//,avg}$ is the average strain in the plane of the film, and values for p, q, c_{11} and c_{12} are given in Table 3.1. The wavenumber shift from unstrained to strained layers, $\Delta\omega$, is calculated from Eqn. 3.6 to be $\Delta\omega = \omega_{measured} - \omega_o = -771\ \varepsilon_{//,avg}$ for $Si_{0.7}Ge_{0.3}$, and for silicon $\Delta\omega = -826 \cdot \varepsilon_{//,avg}$.

In Fig. 3.15, the peak corresponding to strained SiGe occurs at 511.0 cm^{-1}. Thus the biaxial compressive strain in the $Si_{0.7}Ge_{0.3}$ layer is calculated to be -1.2%. The Raman measurement error is approximately ±0.5cm^{-1} or ±0.06% strain for the biaxial strain case. The strain values obtained by XRD and Raman on the same samples (Figs. 3.13 and 3.15) agree with one another within the measurement error, confirming the validity of the Raman measurements used throughout this work.

The Raman effect comes from the polarizability of valence electrons in the semiconductor crystal. Thus by controlling the polarization of the incident and scattered beams, the crystal orientation of the film can be determined [132-136]. In this work, the Raman backscattering geometry is used, where the incident and detected scattering beams are normal to crystal surface. Using this geometry on the (001) surface of cubic crystals, the intensity of the Raman peak varies [132,133] as

$$I \sim sin^2\left(\psi + \psi' + 2\psi_o\right), \tag{3.7}$$

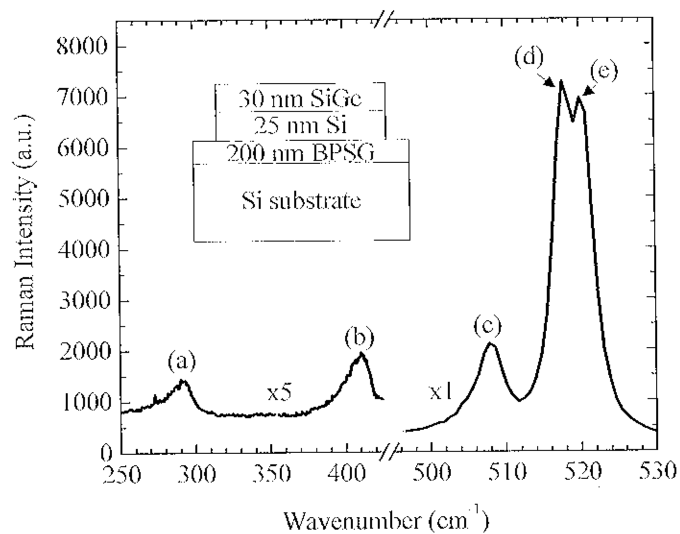


Figure 3.16: Full micro-Raman spectrum for a SiGe/Si bilayer on BPSG (schematic cross-section shown in the inset) showing five distinctive peaks. From the compressively strained SiGe layer, there are 3 peaks: (a) from Ge-Ge phonons; (b) from Si-Ge phonons; and (c) from Si-Si phonons. Peak (d) comes from Si-Si phonons in the epitaxial Si layer. This layer is in tension and thus exhibits a peak that is shifted to lower wavenumbers relative to the (unstrained) silicon substrate peak (e).

where φ and φ' are the angles between the laboratory x-axis (an arbitrary direction in the surface plane of the sample) and the polarization vector of the incident and scattered photons, respectively, and φ_o is the sample orientation to the lab x-axis. The polarization of the incident beam is fixed by the laser itself and is confirmed by placing a polarizer between the laser shutter and sample surface (P1 in Fig. 3.14). The polarization of the scattered beam that is measured is varied by placing a second polarizer just before the scattered beam enters the detector (P2 in Fig. 3.15).

The results are shown in Fig. 3.17, where Raman peak intensity is plotted vs the polarization angle of the scattered beam, φ'. The heavy solid line shows the relationship of Eqn. 3.7 while the data points connected by thinner lines (to guide the eye) correspond to measured intensities. Two samples were measured. The bulk silicon wafer, sample (a), shown by the inverted triangles, clearly follows the form of Eqn. 3.7 (with $\varphi + 2\varphi_o = 125^o$), as expected. The second sample, (b), consists of 30-nm $Si_{0.7}Ge_{0.3}$ on 25-nm Si on 200-nm BPSG on a silicon wafer. The strain in the top two layers has been altered by a relaxation anneal, as will be described later in this Chapter and in Ch. 5, so that the SiGe's compressive strain is reduced and the thin Si layer is in tension. The underlying bulk Si substrate remains unstrained. Thus, three different Si-Si phonon peaks are seen in the Raman spectrum for the SiGe epi, Si epi, and Si substrate, as shown in Fig. 3.16.

Comparing the Si substrate intensity curves for samples (a) and (b) in Fig. 3.17, it is clear that the crystal directions of the two sample substrates are aligned identically with respect to the sample mount. The bulk Si [sample (a)] crystal direction is known due to the wafer flat; the <110> direction is along the sample edge. In contrast, sample (b) was a small piece diced from a larger wafer, and thus the crystal direction was uncertain. By comparison of the two Si substrate curves, the crystal direction in the plane for the Si substrate of sample (b) is easily ascertained: in this case, the edges of the sample (b) substrate are also parallel to <110>. The measurement error of this method is estimated to be $\pm 5^o$.

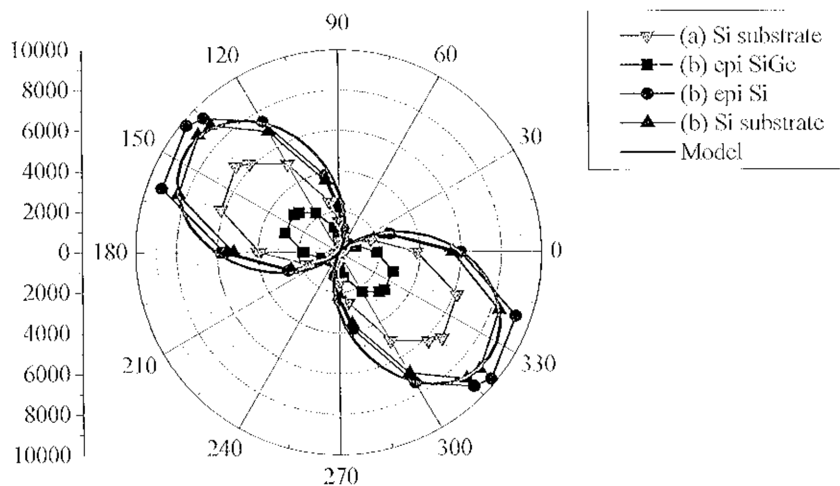


Figure 3.17: Raman peak intensities vs polarization angle of scattered beam, ψ', for two samples: (a) a bulk silicon wafer and (b) the SiGe/Si on BPSG structure shown in the Fig. 3.16 inset, compared to the model of Eqn. 3.7 with $\psi + 2\psi_o = 125^o$. Data were taken from 0 to 180° and then mirrored to complete the curve.

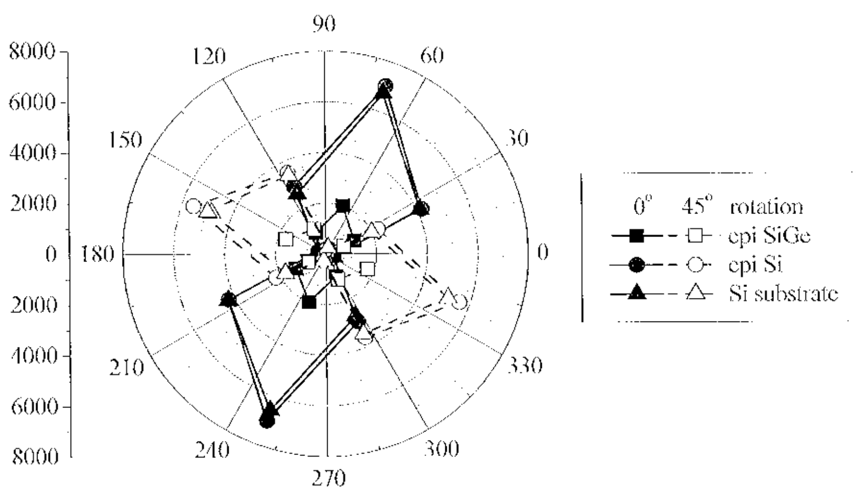


Figure 3.18: Raman peak intensities vs polarization angle of scattered beam for sample (b) of Fig. 3.17, with the sample as originally positioned (ψ_o=0°) and then rotated by 45° (to ψ_o–45°). The intensities of each layer rotate at twice the rotation angle of the sample (*i.e.*, by $2\psi_o$, or 90°), as expected from Eqn. 3.7.

More importantly, the crystal alignment of the top SiGe and Si epi films can also be determined. These films were attached to the BPSG by wafer bonding. The wafer bonding process does not intrinsically set any requirement for the alignment of the underlying substrate to the transferred epitaxial films; by rotating the donor and handle wafers so that the wafer flats are not parallel, the top layers could easily be misaligned to the substrate. However, the three sample (b) curves from Fig. 3.17 show clearly that the top epi layers and underlying substrate have the same angular dependence within ±5° and thus are properly aligned.

Finally, to further verify the parallel alignment of all three layers (epi SiGe, epi Si and Si substrate), more data was taken on sample (b). Measurements were made first with the sample as originally positioned, and secondly with the sample rotated by 45° in the sample mount. Figure 3.18 plots the resulting peak intensities vs polarization angle for these two cases. The peak intensities of all three curves rotate together. For sample rotation (φ_o) of 45°, the intensities rotate by 90° ($2\varphi_o$), as predicted from Eqn. 3.7. Thus one can be confident that all three layers are aligned in the same way, with (001) surfaces and <110> crystal directions parallel to the sample edges.

Polarized Raman spectroscopy can also used to measure multiple strain components (ε_{xx}, ε_{yy}, and ε_{zz}) in strained Si and SiGe layers [137,138]. There are three different **k**=0 phonons in the first Brillouin zone, and for anisotropic strain (where $\varepsilon_{xx} \neq \varepsilon_{yy} \neq \varepsilon_{zz}$) these three phonons give rise to three different Raman peaks, located at

$$\Delta\omega_1 = \frac{\omega_o}{2}\left[p\varepsilon_{xx} + q\left(\varepsilon_{yy} + \varepsilon_{zz}\right)\right] \quad (3.8a)$$

$$\Delta\omega_2 = \frac{\omega_o}{2}\left[p\varepsilon_{yy} + q\left(\varepsilon_{xx} + \varepsilon_{zz}\right)\right] \quad (3.8b)$$

$$\Delta\omega_3 = \frac{\omega_o}{2}\left[p\varepsilon_{zz} + q\left(\varepsilon_{xx} + \varepsilon_{yy}\right)\right]. \quad (3.8c)$$

By measuring the location of all three peaks, ε_{xx}, ε_{yy}, and ε_{zz} can be independently determined. However, as one might surmise by the discussion above, the peak intensities of these three peaks vary with crystal orientation and direction of the sample. For Raman in the backscattering geometry on cubic crystals with (001) surfaces, the intensities of the

$\Delta\omega_1$ and $\Delta\omega_2$ peaks vanish to zero. Thus, the only measurable peak is $\Delta\omega_3$ – Eqns. 3.8c and 3.6 are functionally identical via Eqn. 3.5 – which has an intensity given by Eqn. 3.7. For samples with other surface orientation, *e.g.*, (110), or in a non-backscattering Raman geometry [137,139], the degeneracy is removed and two or three peaks appear. Unfortunately the requirement of Raman micro-positioning precludes the use a non-backscattering geometry with our Raman apparatus, and all samples have (001) surfaces, so such measurements could not be made for this work.

3.3.3 Electron back-scattering diffraction

Another way to determine crystal orientation, with greater precision, is by electron back-scattering diffraction (EBSD) [140]. In a scanning electron microscope, an electron beam is directed onto the sample at a 15° incident angle with respect to the sample surface. The electrons are scattered by the atoms in the crystal lattice by Bragg reflection just as photons are scattered in x-ray diffraction, according to Eqn. 3.4. The scattered electrons hit a phosphor screen and create a diffraction pattern, which is imaged by a CCD detector.

A typical EBSD image is shown in Fig. 3.19(a). The white or light gray rays in the image are called Kikuchi bands. Each band represents a different diffracting crystal plane. The intersections of the bands correspond to specific crystal directions (that is, to intersections of crystal planes). Using the Hough transform to numerically analyze the image, the band intersections are indexed and the crystal orientation determined. In Fig. 3.19(b) the indexed image is shown for the sample of Fig. 3.11, which has edges aligned to <100> in the (001) surface plane. In practice, the sample surface is scanned over a 100×100-μm^2 area with step size of 1 μm to yield more than 10,000 points for statistical analysis. In Fig. 3.20, the results from a typical scan are shown for a sample with SiGe on Si on silicon nitride on BPSG (see schematic cross-section in Fig. 3.20). Each pixel in the left image of Fig. 3.20 represents an indexed EBSD diffraction pattern, as shown in Fig. 3.19(b). The color of the pixel displays the conformity of this diffraction pattern to the desired crystal direction, from 0 to 2° off axis. For this sample, the EBSD scan shows that the surface normal is <001> within 2.0°, and the sample edges are <110> within 2.1°,

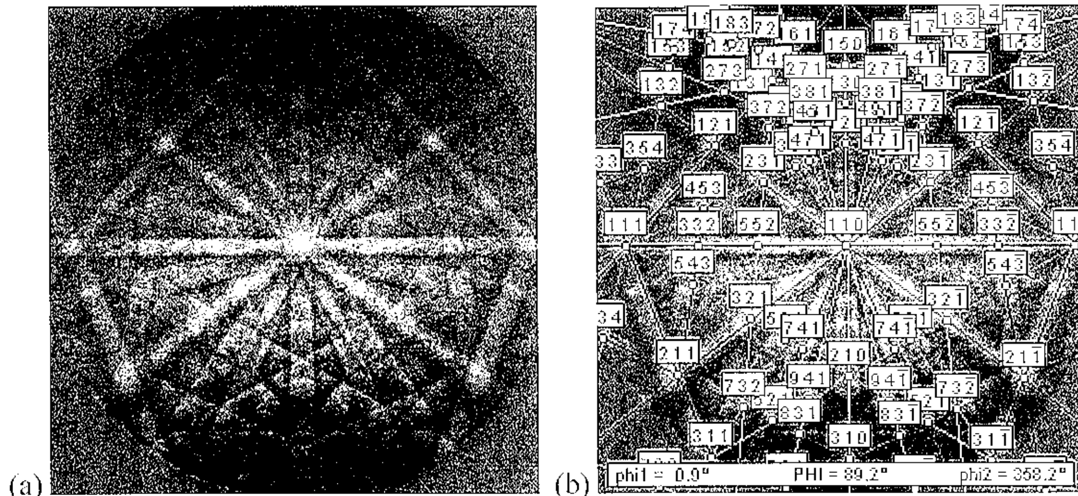


Figure 3.19: Electron backscatter diffraction (EBSD) patterns of the sample pictured in Fig. 3.11. Images (a) and (b) are identical, but (b) has been analyzed and labeled to show that the sample has a (001) surface and edges aligned to <100> crystal directions. Measurements and analysis are courtesy of Matt Nowell at EDAX/TSL (Draper, UT).

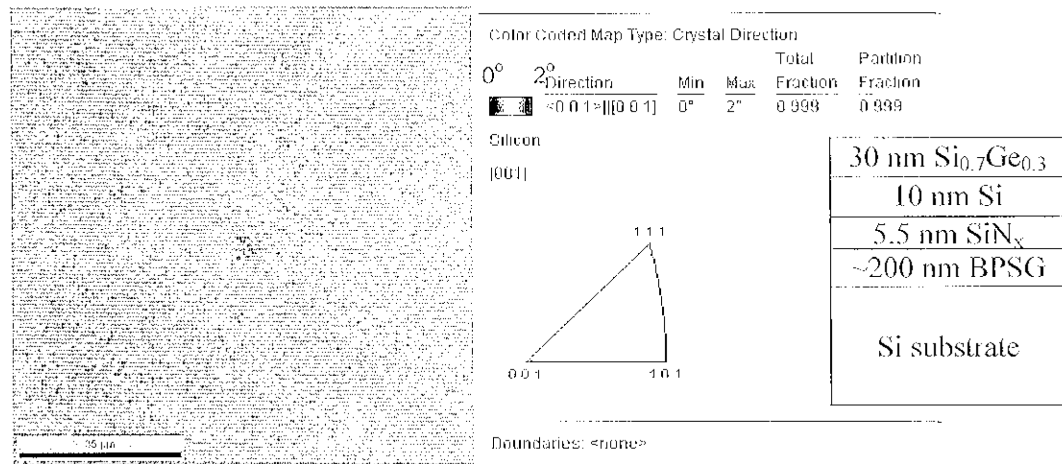


Figure 3.20: Scanned EBSD results across a 100x100-μm^2 area on a sample with a SiGe/Si/SiN_x/BPSG cross-section as shown on the right. The plot is color coded to show the alignment of the sample normal to the <001> crystal direction. In the top right corner are the statistical results: 99.9% of the > 10,000 data points indicate alignment within 2° of <001>. A similar set of data (not shown here) identifies the sample edges as aligned to <110> within 2.1° (again for 99.9% agreement). During the measurement, the center of the scan area is subjected to lengthy beam exposure and thus may be melted or thermally deformed, causing data non-uniformities [141]. The data was taken by Ron Witt at EBSD Analytical (East Lehi, UT).

both with 99.9% data agreement (for 100% agreement, the tolerance must be increased to 3.0°.)

Key advantages of EBSD over polarized Raman for characterization of crystal orientation are the statistical nature of the results due to automated data analysis, and the smaller amount of error (±2.0° vs ±5°). The EBSD technique also has a much lower penetration depth – the diffraction volume is limited to the top 10 to 50 nm of the sample [141] – so that the data comes solely from the thin SiGe and Si layers, and not from the silicon substrate.

3.4 Lateral Relaxation of SiGe Films on BPSG

This section describes the lateral relaxation of biaxially-strained SiGe films on BPSG to yield unstrained SiGe. The defect density of the unstrained SiGe films is measured.

3.4.1 Lateral Relaxation Used to Achieve Unstrained SiGe

After wafer bonding and layer transfer as in Fig. 3.7, the SiGe film on BPSG maintains its original biaxial compressive strain. From Eqn. 2.28, for $Si_{0.7}Ge_{0.3}$ the initial strain in the plane of the film, ε_o, is calculated to be -1.2%. The SiGe layer is patterned into islands using a 30-sec plasma reactive ion etch (RIE) with SF_6/O_2 – 60/20 at 100 mT and 100 W. This etch removes Si and SiGe, but etches BPSG very slowly. The resulting SiGe islands are square with island edge length L, as shown in cross-section in Fig. 3.11.

Upon a high-temperature anneal (> 700 °C), the BPSG becomes soft and can flow. This allows the SiGe to relax laterally to become less-compressively strained, as shown in Fig. 3.21. If annealed for a sufficiently long time, the SiGe will fully relax so the film island is strain and stress free, that is $\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{zz} = \sigma_{xx} = \sigma_{yy} = \sigma_{zz} = 0$, where ε_{ii} and σ_{ii} indicate component strain and stress, respectively [Fig. 3.22(a)]. In our process the SiGe film is not vertically constrained, so σ_{zz} always equals zero and ε_{zz} is determined by Hooke's law (see App. B for details.)

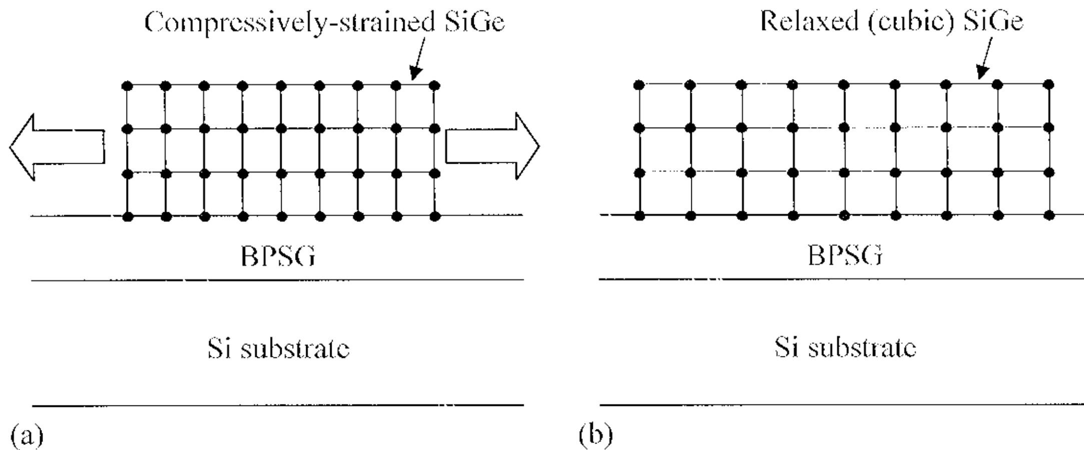


Figure 3.21: Cross-section of SiGe on BPSG, with the SiGe lattice represented schematically: (a) After layer transfer and island patterning. The SiGe film is biaxially-compressively strained. The white arrows show the force that drives lateral relaxation. (b) After a relaxation anneal. The flow of the BPSG allowed the SiGe island to relax by lateral expansion to its unstrained, cubic state.

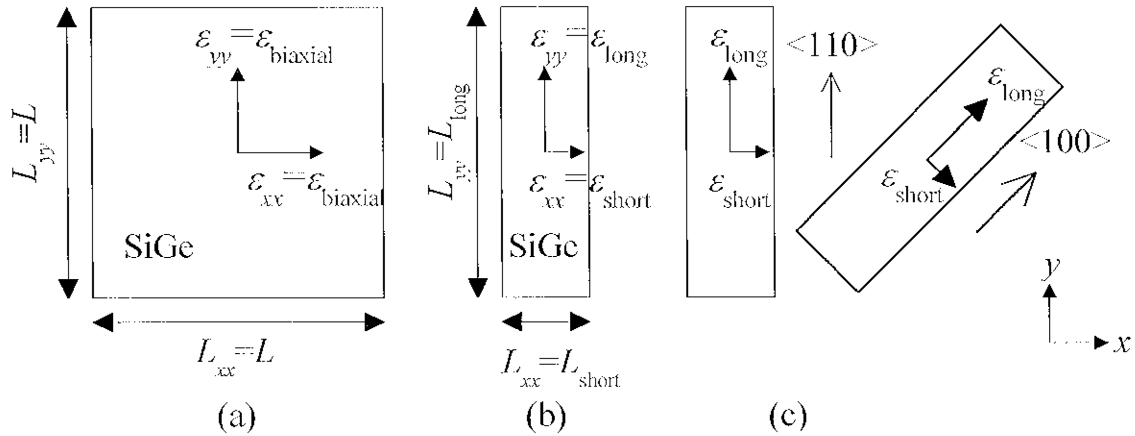


Figure 3.22: (a) Top view schematic of a square island, where $L=L_{xx}=L_{yy}$ and $\varepsilon_{biaxial}=\varepsilon_{xx}=\varepsilon_{yy}$; (b) Top view schematic of rectangular island and definition of strain directions in the short and long island directions, ε_{short} and ε_{long}, respectively; (c) Schematic of rectangular islands aligned to <100> and <110> crystal directions in the (001) surface plane of the film. The z-axis is defined normal to the film surface.

For square islands the strain at the center of the island before, during and after the relaxation anneal is always biaxial in the plane of the film, that is, ε_{xx} and ε_{yy} are equal [Fig. 3.22(a)]. Film relaxation begins at the island edge, and proceeds toward the island center. After an initial delay, the biaxial strain at the center of a square island decays exponentially with a time constant for lateral relaxation, τ_L, given by [142,143]

$$\tau_L = \frac{\eta L^2}{c'_{11} h_f h_g}, \tag{3.9}$$

where η is the viscosity of BPSG, L the square island edge length, h_f the SiGe film thickness, and h_g the BPSG glass thickness. The variable c'_{11} is the crystal-direction dependent SiGe film elastic stiffness coefficient. For the <100> crystal direction, $c'_{11} = c_{11}$, the value of which is given in Table 3.1. (For other crystal directions, see Appendix B.) The time-dependence of the relaxation process will be explored more fully in Chapter 4.

Because relaxation time depends on L^2, small islands will relax quickly, while the centers of large islands maintain their initial strain for a much longer time. To test this theory, square SiGe islands of various sizes were patterned on BPSG and the sample annealed in nitrogen for 5 min at 750 °C to initiate lateral relaxation. The SiGe strain after annealing was measured by micro-Raman spectroscopy. As shown in Fig. 3.23, SiGe islands of size 10 μm x 10 μm are fully relaxed, while the centers of large islands (> 60 μm x 60 μm) remain fully compressively strained with $\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_o = -1.2\%$. Islands with intermediate sizes are partially relaxed. AFM measurements of fully relaxed 10x10-μm^2 islands show a smooth surface with RMS surface roughness of 0.46 nm. Thus it is possible to achieve smooth, flat, and unstrained $Si_{0.7}Ge_{0.3}$ films of size ~100 μm^2 using lateral relaxation on BPSG.

Note that during the relaxation process the SiGe strain is biaxially-symmetric only along island diagonals and at their intersection, the island center. For example, at the island edge the film stress perpendicular to the edge is zero while the stress parallel to the edge is non-zero, leading to asymmetric biaxial strain, as shown in the images of ε_{xx} and ε_{yy} in Fig. 4.3. Of course, the strain is biaxially-symmetric across the entire island both

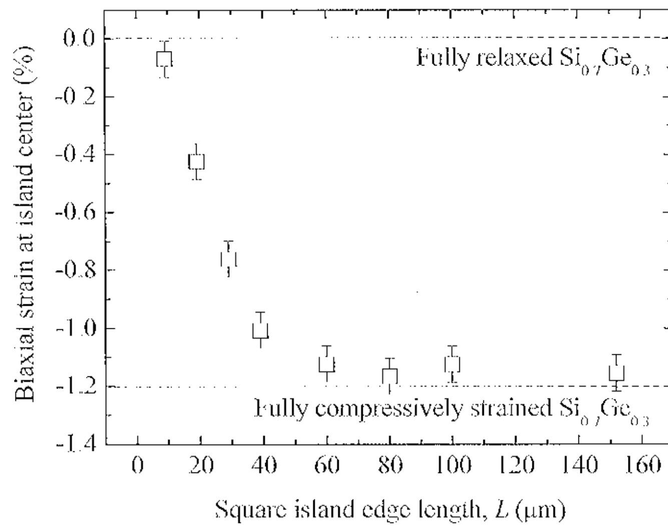


Figure 3.23: Biaxial strain in SiGe layer for squares of different sizes after a 5-min anneal at 750 °C. The sample has the cross-section shown in Fig. 3.11. The symbols are measured biaxial strain at island center, while the dashed lines indicate fully strained and fully relaxed $Si_{0.7}Ge_{0.3}$.

before annealing (full initial biaxial strain), and at the end of a long anneal (full relaxation, zero strain).

3.4.2 Defect Density of Relaxed SiGe films on BPSG

The wafer bonding, layer transfer and lateral relaxation process described above makes possible unstrained SiGe films without requiring, in theory, generation of crystal defects. This is a key advantage of this process over other relaxed SiGe methods, such as strain-relaxed buffers (described in Chapter 2), which rely on defect generation to relax the SiGe film. Crystal defects degrade the performance of optical and electronic devices by, for example, decreasing the minority carrier lifetime of Si and SiGe, decreasing the breakdown voltage of thermal oxides and decreasing charge carrier mobility. Thus, one would like to minimize the defect density of the films. In this section, the defect density of unstrained SiGe films on BPSG is measured, as was similarly done in Ref. [8].

Islands of 30-nm $Si_{0.7}Ge_{0.3}$ islands on 1-μm BPSG are patterned with a dry etch (see Fig. 3.11) and the sample is annealed at 800 °C for 1 hour in H_2/N_2 to relax the islands. The SiGe strain is measured by Raman spectroscopy to confirm that islands with L up to 30 μm are fully relaxed to the zero strain state. Ideally the defect density would be measured directly at this point, but traditional defect etches require thicker films. So, the SiGe film is epitaxially thickened before performing the defect etch.

On part of the sample, the SiGe islands and BPSG layers are removed by dilute HF etch to expose the underlying Si substrate. An 80-nm layer of $Si_{0.7}Ge_{0.3}$ is then regrown on the relaxed SiGe islands and on the Si substrate by rapid-thermal CVD (RTCVD) using dichlorosilane and germane gas precursors at 625 °C (sample #3446). The sample is annealed further at 800°C for 1 hour in H_2/N_2 to allow the SiGe grown on the silicon substrate to relax. Raman spectroscopy shows that the regrown SiGe is fully relaxed on the relaxed SiGe islands. On the Si substrate, the regrown SiGe is partially strained, as expected due to the lattice mismatch and subsequent anneal, which generated defects in the SiGe layer. A dilute Schimmel etch [144-147], HF (49%) : 0.3M CrO_3 4:5 of 60 sec was used to highlight the SiGe defects in both layers. Images taken using optical phase contrast (Nomarski) microscopy are shown in Fig. 3.24. On the left, in Fig. 3.24(a), a relaxed SiGe island is shown with 3 visible defects over a 40x40-μm^2 island,

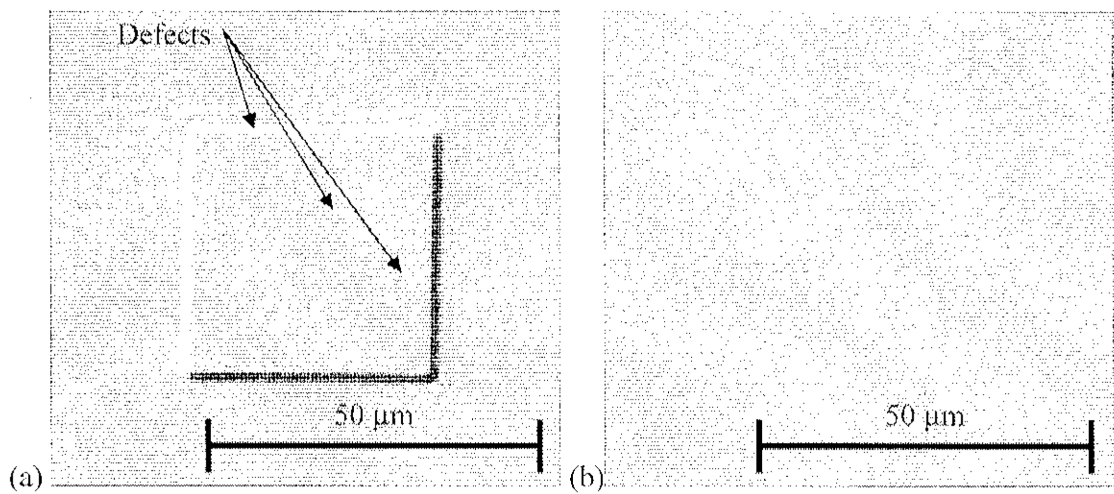


Figure 3.24: Optical phase-contrast microscope images of (a) 40x40-μm^2 island of unstrained 110-nm SiGe on BPSG with a defect density of ~1.9×10^5 cm^{-2}; (b) 80-nm SiGe grown on bulk Si with a defect density of ~2.5×10^6 cm^{-2}.

which corresponds to a defect density of 1.9×10^5 cm^{-2}. While this density is not particularly low, this image represents the highest defect density observed on SiGe islands. More typically, densities of a few 10^4 cm^{-2} were observed. In Fig. 3.24(b) a typical image of the partially relaxed SiGe film on the Si substrate is shown. The defect density is clearly much higher than on the fully relaxed SiGe islands; the defect density in this case is calculated to be 2.5×10^6 cm^{-2}. The commonly used process of thick SiGe buffer layers (~1 μm thick for $Si_{0.7}Ge_{0.3}$) to obtain relaxed SiGe substrates (as described in Sec. 2.4.1) has defect densities of 10^5 cm^{-2}, even after 15 years of work to reduce defect levels [67,68]. The typical relaxed SiGe layer defect densities for our technology are an order of magnitude lower than this, and further reductions may be possible by improving the cleanliness of our wafer bonding process.

As proposed, the unstrained SiGe layer resulting from the layer transfer and lateral relaxation process has a relatively low defect density and thus is suitable for use directly or as a heteroepitaxial substrate for strained silicon for optical or electronic device applications.

3.5 Generation of Uniaxial Stress in SiGe

3.5.1 Stripe Geometry for Uniaxial Stress

To obtain asymmetric strain in the plane of the film, a rectangular shape for the SiGe islands is chosen [148,149]. The rectangular island has edge lengths of L_{short} and L_{long}, as shown in Fig. 3.22(b). The strains in the two in-plane directions, ε_{xx} and ε_{yy}, are now notated ε_{short} and ε_{long}, with reference to the short and long dimensions of the island rectangle. Because the lateral relaxation time constant depends on the square of island edge length (Eqn. 3.9), during relaxation ε_{short} and ε_{long} will be different. Upon annealing, the SiGe layer will quickly relax in the short island dimension direction until the *stress* in the short direction, σ_{short}, equals zero. In contrast, in the long island dimension direction the island maintains its initial strain, $\varepsilon_{long} = \varepsilon_0$, for moderate anneals, since the relaxation process is very slow over the long island dimension. Due to the Poisson effect, the SiGe compression in the long direction will create a tensile strain in the short direction [149] as

sketched in Fig. 3.25. Thus the short direction of a rectangular island is not fully relaxed to zero strain as in the square geometry case (where $\sigma_{xx} = \varepsilon_{xx} = 0$) but is in tension (not compression), with

$$\varepsilon_{\text{short, single SiGe layer}} = -\varepsilon_{\text{long}}\nu \,, \tag{3.10}$$

where ν is Poisson's ratio of the SiGe film. Equation 3.10 corresponds to a state of zero stress in the short direction, $\sigma_{\text{short}}=0$. The strain in the long direction is still ε_o (compression) so that a uniform but highly anisotropic strain results over nearly the entire film. (The uniformity is examined in Chapter 4.) Note the difference in the square and rectangular boundary conditions for the single SiGe layer: A square island reaches a state of zero strain, where $\varepsilon_{xx} = \varepsilon_{yy} = \sigma_{xx} = \sigma_{yy} = 0$, while a rectangle first reaches a state of uniaxial stress, with $\sigma_{\text{short}} = 0$, $\varepsilon_{\text{long}} = \varepsilon_o$ and $\varepsilon_{\text{short}}$, $\sigma_{\text{long}} \neq 0$. The crystal-direction dependence of $\varepsilon_{\text{short}}$ is the focus of the next section.

3.5.2 Contribution of the Poisson Effect and its Crystal-Direction Dependence

Poisson's ratio is strongly dependent on crystal direction. For an arbitrary crystal-direction making an angle of θ with the <100> direction and laying in the (001) plane of the film, Poisson's ratio is given by (Refs. [125,150] and App. B)

$$\nu(\theta) = -\frac{2\cdot s_{12} + (s_{11} - s_{12} - s_{44}/2)\cdot\sin^2(2\theta)}{2\cdot s_{11} - (s_{11} - s_{12} - s_{44}/2)\cdot\sin^2(2\theta)}, \tag{3.11}$$

where s_{11}, s_{12} and s_{44} are the elastic compliance constants (given in Table 3.1), which are constant for a given material. The result is plotted in Fig. 3.26. For a $Si_{0.7}Ge_{0.3}$ layer in the two directions of interest, <100> and <110>, ν takes values of

$$\nu_{\text{SiGe,<100>}} = -\frac{s_{12}}{s_{11}} = 0.277 \quad \text{and} \quad \nu_{\text{SiGe,<110>}} = \frac{s_{11} + s_{12} - s_{44}/2}{s_{11} + s_{12} + s_{44}/2} = 0.054 \,. \tag{3.12}$$

The level of tensile strain in the direction parallel to the short island edge, $\varepsilon_{\text{short}}$, is directly dependent on ν via Eqn. 3.10. Thus the highest possible tensile strain should be achieved with islands aligned to <100>, because of the large value of ν in this direction. This angular dependence of strain will be verified in the next section.

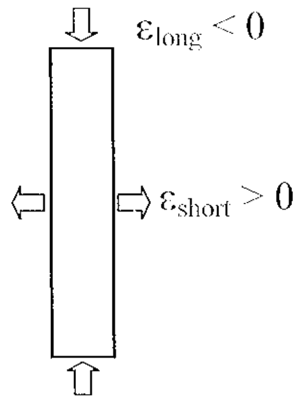


Figure 3.25: Schematic of the Poisson effect acting on a rectangular SiGe island: compression in the long direction results in extension in the short direction.

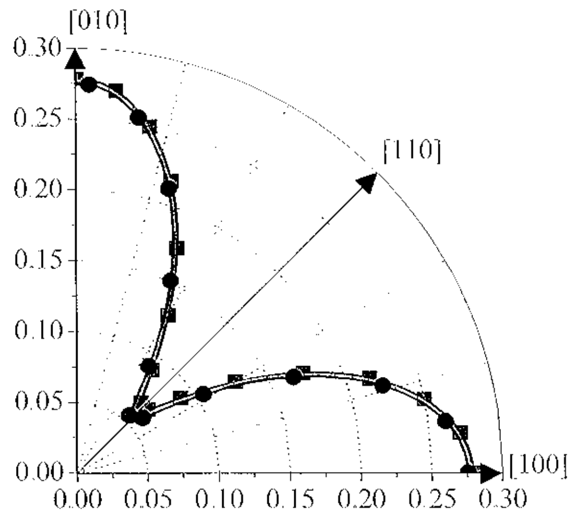


Figure 3.26: Poisson's ratio, ν, in silicon (red line with squares) and $Si_{0.7}Ge_{0.3}$ (black line with circles) vs crystal direction in the (001) plane, after Refs. [125] and [150].

3.5.3 Experimental Results for Uniaxial Stress

To verify this prediction, islands are patterned with a constant L_{long} of 150 μm and various values of L_{short}, from 5 to 150 μm. One set of islands is aligned with the edges of the islands parallel to the <100> in-plane crystalline direction while another set is aligned to <110>, as shown in Fig. 3.22(c). As in Fig. 3.11, the islands consist of a single layer of 30-nm $Si_{0.7}Ge_{0.3}$ on 1-μm BPSG. The islands are annealed for 5 min at 750 °C and strain at the island center is measured using micro-Raman spectroscopy.

By comparing Eqns. 3.6 and 3.8c using Eqn. 3.5, it is clear that Raman spectroscopy with normal incidence measures the average strain in the plane of the film,

$$\varepsilon_{//,avg} = \frac{\varepsilon_{xx} + \varepsilon_{yy}}{2} . \tag{3.13}$$

For the square island case of biaxial strain, $\varepsilon_{xx} = \varepsilon_{yy}$, and the biaxial strain level can be directly obtained. For rectangular islands, in order to calculate ε_{short} a value for ε_{long} must be assumed: here, one can safely assume that the strain in the long direction is pinned to its initial value, so $\varepsilon_{long} = \varepsilon_o$. This value for ε_{long} can be confirmed by measuring biaxial strain, $\varepsilon_{biaxial}$, for a large square island with edge length equal to the long rectangle dimension under the same annealing conditions, and assuming that the relaxation of the rectangle in the long direction and of the square are equal when $L_{square} = L_{long,\ rectangle}$. The simulations of Chapter 4 justify this assumption.

The measured strain in the short island direction is plotted as a function of short island edge length in Fig. 3.27. The assumed value of $\varepsilon_{long} = \varepsilon_o$ is also plotted. Narrow islands exhibit the largest changes in ε_{short} from its initial value of ε_o. For narrow islands with L_{short} less than 15 μm, the strain in the short direction is tensile as predicted. The narrowest islands, with L_{short} = 6 μm, have reached the maximum tensile strain state described by Eqn. 3.10, at which $\sigma_{short} = 0$. For wider islands with L_{short} greater than 70 μm, the strain in the short direction is still at its initial compressive value, ε_o; these islands are large enough that for this short anneal, their centers have not yet relaxed. For island sizes between these two extremes, the strain in the short direction is somewhere between ε_o and the maximum tensile ε_{short}, indicating that the relaxation process is in progress.

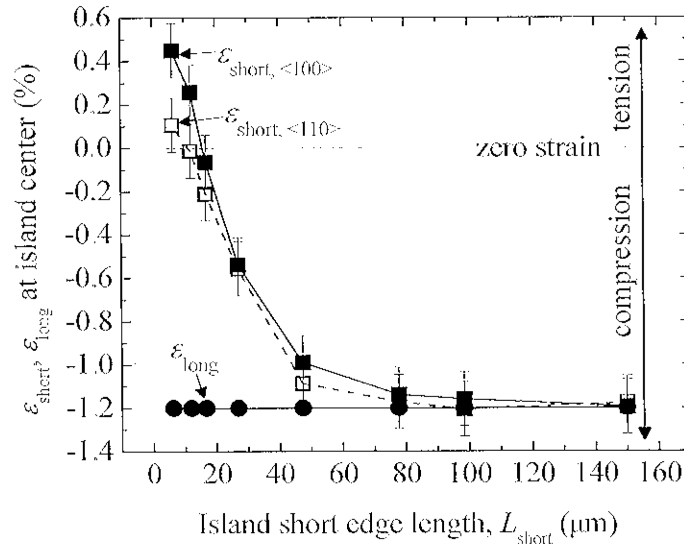


Figure 3.27: Measured anisotropic SiGe strain at the center of rectangular islands with L_{long} = 150 μm vs short island edge length after a 5-min anneal at 750 °C. The layer structure is shown in Fig. 3.11. Symbols indicate ε_{short} for islands aligned along <110> (open squares) and along <100> (solid squares); the lines are drawn from point to point. In both cases, ε_{long} (solid circles) is fixed at ε_0=-1.2% by the large L_{long} that prevents relaxation. The measurement error for uniaxial strain is ±0.12%.

In Fig. 3.23, we saw that 10-μm x 10-μm square SiGe islands were fully relaxed after a 5-min anneal at 750°C. Therefore we expect that for the same anneal the narrowest rectangles, with L_{short} = 6 μm, will readily reach a zero stress (tensile strain) state in the short direction, as observed in Fig. 3.27. Making the rectangles narrower would not be expected to further increase ε_{short}; the maximum possible value given by Eqn. 3.10 has already been achieved by the 6-μm x 150-μm islands.

Figure 3.27 shows that for narrow rectangular islands, the Poisson effect gives rise to a significant level of tension in the short direction, while the long direction remains compressively strained; the in-plane strain is highly anisotropic. This is contrast to the square island case, in which the equilibrium state is zero in-plane strain, as seen for the smallest island in Fig. 3.23. Where squares of narrow dimension rapidly become fully relaxed to $\varepsilon_{biaxial} = 0$, narrow rectangles quickly become tensilely strained in the short direction, with $\sigma_{short} = 0$ and $\varepsilon_{short} = -\varepsilon_o \nu$, while maintaining compressive strain, $\varepsilon_{long} = \varepsilon_o$, in the long direction. Moreover, for narrow islands where the short direction has reached the zero-stress state, the magnitude of tension in the short direction is strongly dependent on crystal direction. For the narrowest islands shown in Fig. 3.27, tensile strains of ε_{short} =+0.45% and +0.11% for rectangles aligned to <100> and <110>, respectively, are measured. These measured values compare favorably to the predicted strains for these two cases, from Eqns. 3.10 and 3.12, of ε_{short} =+0.33% and +0.06%. The measurements qualitatively confirm the prediction that tension is maximal in the <100> direction, and for the first time show the uniaxial strain dependence on rectangle alignment with respect to the crystal direction.

Using similar samples, the average areal strain over an array of identically-sized rectangular islands has been measured by x-ray diffraction of the (004) reflection. The x-ray diffraction rocking curves and micro-Raman spectra for these samples are shown in Fig. 3.28. As with Raman measurements, in x-ray diffraction the average in-plane strain is measured, and a value for ε_{long} must be assumed to calculate ε_{short}. Three samples are used – one unpatterned and unannealed, a second patterned into islands but not annealed, and the third patterned and annealed to reach the σ_{short}=0 state. The measurements show that patterning does not change island strain, but annealing patterned islands has a strong

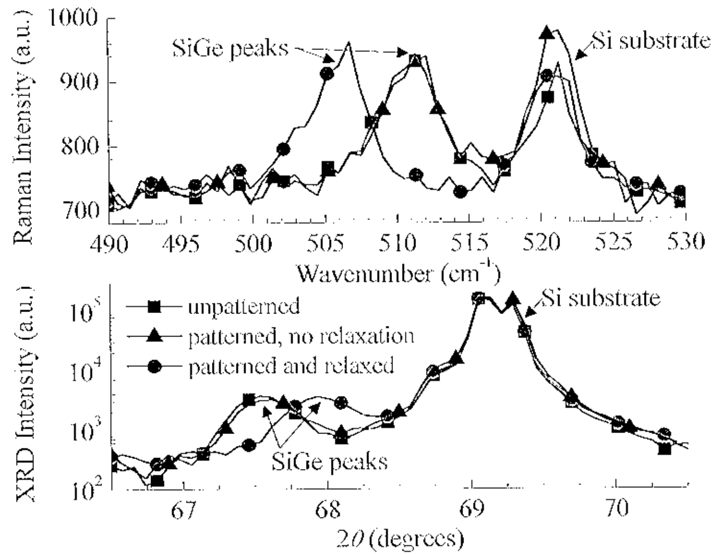


Figure 3.28: A comparison of micro-Raman spectroscopy and x-ray diffraction measurements of uniaxial strain. The samples have 30-nm $Si_{1-x}Ge_x$ (assumed x=27%) on 95-nm BPSG. One sample remains unpatterned (The unpatterned sample data were shown previously in Figs. 3.13 and 3.15.) On the other two samples, the SiGe layer is patterned into an array of rectangles 10 μm x 8700 μm with 2-μm gaps between the rectangles. One of the patterned samples is annealed 3.5 h at 900°C to create uniaxial Si strain. (The viscosity of this BPSG is much higher than normal, thus the high temperature and lengthy anneal.) Only the annealed sample shows a change of in-plane strain.

Sample	Strain in the short island direction, ε_{short}, measured by:	
	micro-Raman spectroscopy	X-ray diffraction
Unpatterned	-1.16%	-1.18%
Patterned, with no anneal	-1.19%	-1.11%
Patterned and annealed	+0.06%	+0.10%

Table 3.2: Strain in the short island direction, ε_{short}, under various conditions, extracted from the raw data shown in Fig. 3.28. Unpatterned and patterned samples with no anneal have the original compressive strain, ε_0, of approximately −1.2%. Patterned samples that have been annealed show a small amount of tensile strain, +0.1%, in the short island direction, as expected when σ_{short}=0 for islands aligned to <110>. A full description of the samples is given in the Fig. 3.28 caption. For each sample, the Raman and x-ray results agree within the measurement error of ±0.06-0.07%, although some disagreement is to be expected as Raman measures strain at the island center whereas x-ray diffraction measures areally averaged strain.

effect as island relaxation takes place. The extracted strain values measured by Raman and x-ray diffraction, tabulated in Table 3.2, agree with one another within the measurement error, again confirming the validity of the Raman measurements used throughout this work.

Finally, Hooke's law can be used to convert between stress and strain in a thin film:

$$\sigma_{biaxial} = \frac{E}{1-\nu}\varepsilon_{biaxial}, \tag{3.14a}$$

$$\sigma_{short} = \frac{E}{1-\nu^2}\left(\varepsilon_{short} + \nu\varepsilon_{long}\right), \text{ and} \tag{3.14b}$$

$$\sigma_{long} = \frac{E}{1-\nu^2}\left(\nu\varepsilon_{short} + \varepsilon_{long}\right). \tag{3.14c}$$

Like Poisson's ratio, ν, the Young's modulus, E, depends on crystal direction, as shown in Fig. 3.29 and described in detail in Refs. [125] and [150] and App. B. For $Si_{0.7}Ge_{0.3}$, $E_{<100>}$=121 GPa and $E_{<110>}$=158 GPa. Using Eqn. 3.14 and the measured biaxial and uniaxial strains shown in Fig. 3.23 and 3.27, the biaxial and uniaxial stress generated in SiGe has been plotted in Figs. 3.30 and 3.31.

The initial biaxial compressive stress of the SiGe film is -2.0 GPa, as calculated from Eqn. 3.14a using ε_0=-1.2%. Figure 3.30 clearly shows that the square islands go from a state of full biaxial stress to a fully-relaxed, zero-stress state. For rectangular islands, in Fig. 3.31 it is observed that the short island direction quickly reaches a zero stress state, $\sigma_{short} = 0$, as expected, putting the island under uniaxial stress. Using Eqns. 3.10 and 3.14, the expected uniaxial stress in the long direction can be calculated: $\sigma_{long} = E\,\varepsilon_{long}$. Because of the crystal-direction dependence of E, the maximum uniaxial SiGe stress value will also be direction-dependent: $\sigma_{long,<100>}$=-1.5 GPa while $\sigma_{long,<110>}$=-1.9 GPa, as shown in Fig. 3.31.

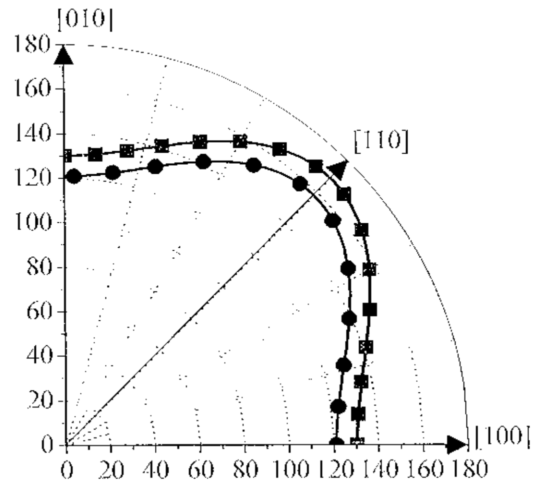


Figure 3.29: Variation of Young's modulus, E, in the (001) crystal plane for silicon (red line with squares) and $Si_{0.7}Ge_{0.3}$ (black line with circles), after Refs. [125] and [150].

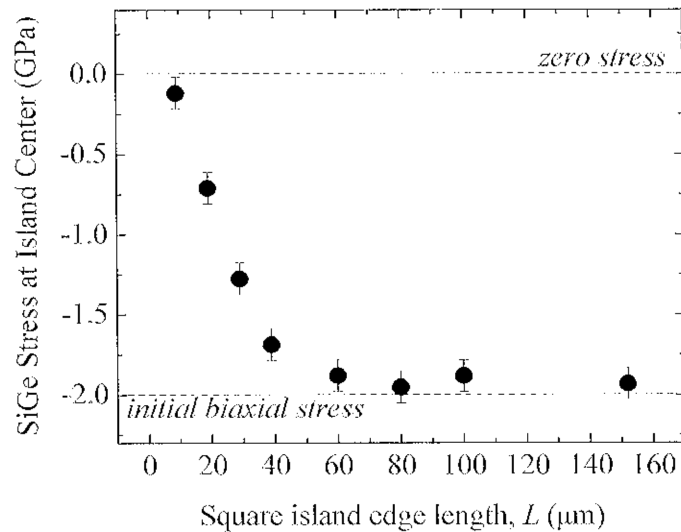


Figure 3.30: Biaxial stress in square islands, calculated from the measured strain values shown in Fig. 3.23 using Eqn. 3.14a. The smallest islands have relaxed to almost zero stress.

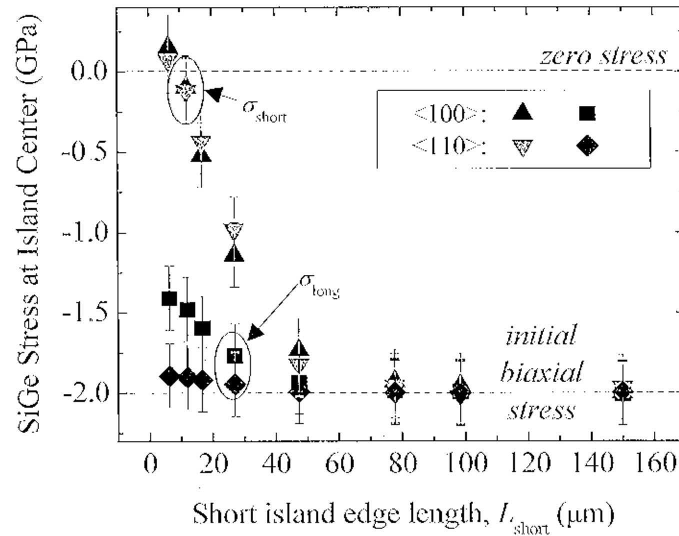


Figure 3.31: Stress in the short, σ_{short}, and long, σ_{long}, directions of rectangular islands aligned to <100> and <110>, as indicated. The stress values are calculated from the measured strain values shown in Fig. 3.27 using Eqns. 3.14b and 3.14c. Note that both sets of islands relax to a zero-stress state in the short direction, but exhibit different uniaxial stress (σ_{long}) values, due to the variation of Young's modulus, E, with crystal-direction, as described in Sec. 3.5.3.

3.6 Summary

Compliant substrates can be used to engineer the strain in semiconductor films. In this chapter, a wafer bonding and layer transfer process has been described to obtain strained SiGe on compliant BPSG. Through island patterning and lateral relaxation, the SiGe film stress can be varied from full biaxial stress, to uniaxial stress, to fully relaxed (zero stress, zero strain). Raman spectroscopy or x-ray diffraction are used to measure film strain, while electron backscatter diffraction or polarized Raman spectroscopy are used to determine film crystal orientation. In the case of zero strain, the measured defect density of the relaxed SiGe films is ~ 10^5 cm^{-2}. In the uniaxial stress case the SiGe film is under highly anisotropic strain, with compressive strain in the long island direction and tensile strain in the short island direction. The level of tensile strain in the short direction is determined by the crystal direction of the island edges, with maximum tension for islands aligned to <100>. The island geometry plays a key role in determining the film strain. In the next chapter, numerical modeling of the relaxation process is used to investigate the relationship between island geometry and film stress or strain, in order to maximize the strain asymmetry achieved, to assess strain uniformity, and to determine process windows for asymmetric strain generation.

Chapter 4

Dynamic Model of Lateral Expansion and Process Window for Maximum Asymmetry of Strain

The previous chapter describes a wafer-bonding and layer-transfer process used to generate biaxial, uniaxial and zero stress in thin SiGe films on BPSG insulator through lateral expansion of SiGe islands. The island geometry determines the type and amount of SiGe stress: square islands result in biaxial stress while rectangular islands result in uniaxial stress. Upon further annealing, all SiGe islands will eventually expand to a fully-relaxed, zero-strain, zero-stress state. That is, the uniaxial stress state is temporary; it persists until the stress in the long island direction relaxes. Thus careful modeling is necessary to predict and control the generation of uniaxial stress. This dynamic lateral expansion process has been modeled previously for square islands [142] and compared to experimental data in Ref. [111]. In this chapter, the numerical model is extended to rectangular island geometries, and the results are used to determine the requirements on island shape and annealing process windows to obtain maximum strain asymmetry and to assess strain uniformity across the islands.

4.1 Introduction to the Lateral Expansion Model

As described and measured in Ch. 3, a SiGe film that is initially strained under biaxial compression can relax via lateral expansion to achieve biaxial, uniaxial and zero-stress states. The island geometry and dimensions were shown to strongly affect the type and level of strain achieved. Small square islands quickly relax to zero strain and zero stress, while larger square islands slowly expand in both in-plane directions to a reduced level of biaxial strain (Figs. 3.23 and 3.30). Narrow, long rectangular islands exhibit different strain in the two island directions: the short island dimension quickly relaxes to

a state of zero stress (and tensile strain), while the long island dimension maintains its initial compressive strain, resulting in a state of uniaxial stress (Figs. 3.25, 3.27 and 3.31).

The lateral expansion that drives these different strain states begins at the edges of the island and proceeds toward the center of the island. Shown in Fig. 4.1 is measured SiGe strain, ε_{short}, across a 50-μm x 150-μm island after a 15-min anneal at 750°C. The lateral relaxation process has just begun: the edges of the island have partially relaxed while the middle of the island is still in the initial, fully-strained state. It is the lateral nature of this process that makes relaxation dependent on island size. In Ref. [142], Huang developed a model to describe the lateral expansion of thin compressively strained films and applied this model to square island geometries. This section describes this model.

Huang's model [142] assumes a single elastic film, in our case, $Si_{0.7}Ge_{0.3}$, which initially is under biaxial compression ε_o. A schematic of the structure is shown in Fig. 4.2, where the SiGe layer has thickness h_f and the BPSG layer has thickness h_g and viscosity η. The SiGe island has in-plane stresses of σ'_{xx}, σ'_{yy} and σ'_{xy}, and shear tractions of τ'_x and τ'_y at the SiGe/BPSG interface. When the island is in equilibrium [142,151]

$$\frac{\partial \sigma'_{xx}}{\partial x} + \frac{\partial \sigma'_{xy}}{\partial y} = \frac{\tau'_x}{h_f}, \text{ and} \tag{4.1a}$$

$$\frac{\partial \sigma'_{xy}}{\partial x} + \frac{\partial \sigma'_{yy}}{\partial y} = \frac{\tau'_y}{h_f}. \tag{4.1b}$$

For a cubic crystal, the stress-strain relations from Hooke's law (see App. B.) are:

$$\sigma'_{xx} = c'_{11}\varepsilon'_{xx} + c'_{12}\varepsilon'_{yy} + c'_{13}\varepsilon'_{zz}, \tag{4.2a}$$

$$\sigma'_{yy} = c'_{12}\varepsilon'_{xx} + c'_{11}\varepsilon'_{yy} + c'_{13}\varepsilon'_{zz}, \tag{4.2b}$$

$$\sigma'_{zz} = c'_{13}\varepsilon'_{xx} + c'_{13}\varepsilon'_{yy} + c'_{33}\varepsilon'_{zz}, \text{ and} \tag{4.2c}$$

$$\sigma'_{xy} = 2c'_{44}\varepsilon'_{xy}$$

(Here the c'_{16} and c'_{26} terms of Eqn. B.20 have been neglected since they are equal to zero in the two directions of interest, <100> and <110>.) The local strain is calculated as the initial strain plus the local film displacement:

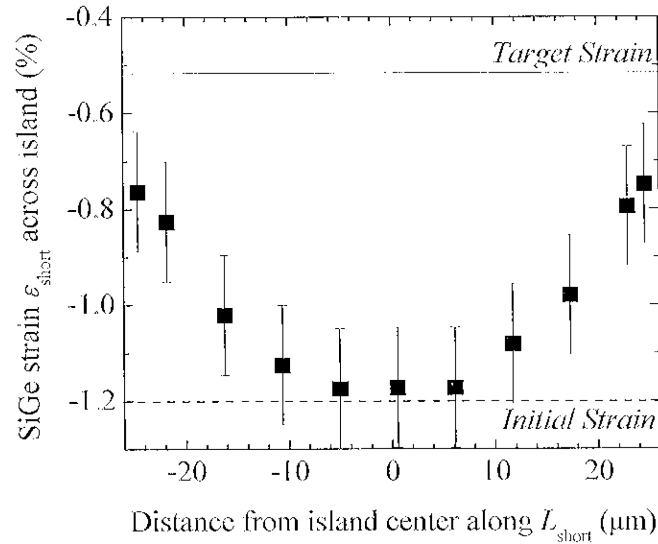


Figure 4.1: Measured SiGe strain in the short island direction, ε_{short}, across a 50-μm x 150-μm island aligned along <110> after a 15-min anneal at 750°C. The relaxation begins at the edges of the island and proceeds toward the center, which for this short anneal remains in the initial, fully strained state. The island structure consists of 30-nm $Si_{0.7}Ge_{0.3}$ on 25-nm silicon on ~200-nm BPSG on a silicon handle wafer. The target strain in SiGe is predicted assuming no relaxation in the long dimension and zero net force in the short direction for the Si/SiGe bi-layer. The bi-layer structure limits the target strain to approximately -0.5%, instead of the zero-stress state observed for single SiGe layer islands (see Ch. 5 for further discussion).

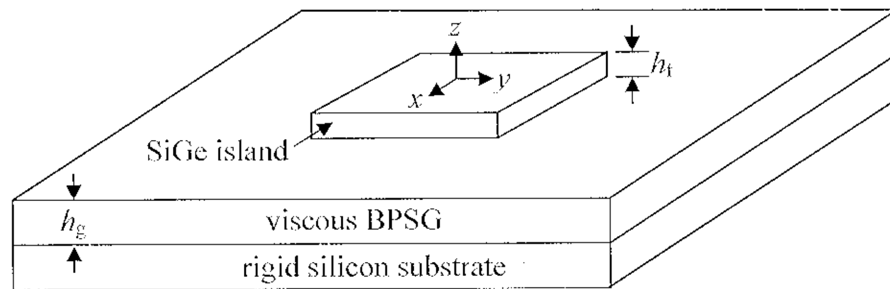


Figure 4.2: Schematic of a SiGe island on viscous BPSG on a rigid silicon substrate, after Ref. [142]. The thicknesses of the SiGe film and BPSG layer are indicated by h_f and h_g, respectively.

$$\varepsilon_{xx}(x,y,t)=\varepsilon_{\text{short}}(x,y,t)=\frac{\partial u(x,y,t)}{\partial x}+\varepsilon_0 \tag{4.3a}$$

$$\varepsilon_{yy}(x,y,t)=\varepsilon_{\text{long}}(x,y,t)=\frac{\partial v(x,y,t)}{\partial y}+\varepsilon_0 \tag{4.3b}$$

$$\varepsilon_{xy}(x,y,t)=\frac{1}{2}\left(\frac{\partial u(x,y,t)}{\partial y}+\frac{\partial v(x,y,t)}{\partial x}\right) \tag{4.3c}$$

The underlying layer, in our case BPSG glass, is assumed to be linearly viscous, with the flow velocity linear from the bottom to the top surface:

$$\tau'_x=\frac{\eta}{h_g}\frac{\partial u}{\partial t}, \text{ and} \tag{4.4a}$$

$$\tau'_y=\frac{\eta}{h_g}\frac{\partial v}{\partial t}, \tag{4.4b}$$

where u and v are the displacements of the SiGe film in the short direction, x, and in the long direction, y, respectively. One can assume that the SiGe film and BPSG expand together at their interface, that the SiGe film, which has an (001) surface, stays flat during lateral expansion, and that the stress in the direction out of the film plane is zero, $\sigma'_{zz} = 0$.

Solving Eqns. 4.1-4.4 under these assumptions, the displacements u and v of the SiGe film as a function of time, t, are defined by two diffusion equations:

$$\frac{\partial u}{\partial t}=\frac{c'_{11}h_f h_g}{\eta}\left[\left(1-\frac{c'^2_{13}}{c'_{11}c'_{33}}\right)\frac{\partial^2 u}{\partial x^2}+\left(\frac{c'_{44}}{c'_{11}}\right)\frac{\partial^2 u}{\partial y^2}+\left(\frac{c'_{12}+c'_{44}}{c'_{11}}-\frac{c'^2_{13}}{c'_{11}c'_{33}}\right)\frac{\partial^2 v}{\partial x\partial y}\right] \tag{4.5a}$$

$$\frac{\partial v}{\partial t}=\frac{c'_{11}h_f h_g}{\eta}\left[\left(1-\frac{c'^2_{13}}{c'_{11}c'_{33}}\right)\frac{\partial^2 v}{\partial y^2}+\left(\frac{c'_{44}}{c'_{11}}\right)\frac{\partial^2 v}{\partial x^2}+\left(\frac{c'_{12}+c'_{44}}{c'_{11}}-\frac{c'^2_{13}}{c'_{11}c'_{33}}\right)\frac{\partial^2 u}{\partial x\partial y}\right], \tag{4.5b}$$

where c' are the crystal direction-dependent elastic stiffness coefficients (see discussion below and in App. B). These equations come from the theory, outlined above, of Ref. [142], where minor modifications have been made to generalize the equations for a film island aligned to an arbitrary crystal direction. Clearly these equations are directly linked to the definition of the lateral relaxation time constant in Eqn. 3.9; the coefficient on the right side of the equations, $c'_{11}h_f h_g/\eta$, is equal to $\tau_L^{-1}L^2$. For the case of solely uniaxial stress ($\sigma_{xx}=\sigma_{xy}=\sigma_{yz}=\sigma_{zz}=0$; $\sigma_{yy}\neq 0$), Eqn. 4.5b reduces to

$$\frac{\partial v}{\partial t} = \frac{E' h_f h_g}{\eta} \frac{\partial^2 v}{\partial y^2} \tag{4.6}$$

where E' is the direction-dependent Young's modulus described by Eqns. B.13-B.15 and Fig. 3.29. Therefore under uniaxial stress, the exponential time constant can be defined as

$$\tau_l = \frac{\eta L^2}{E' h_f h_g}, \tag{4.7}$$

as discussed elsewhere [8,152]. In this thesis, we will maintain the τ_L definition given in Eqn. 3.9, which is more generally applicable, where c'_{11} replaces E' in Eqn. 4.7. The physical meaning of τ_l will be discussed in Sec. 4.2.

Boundary conditions require that at the island edges, the in-plane stresses are zero. Equations 4.5a and 4.5b can be solved numerically using the finite difference method for the in-plane displacements, $u(x,y,t)$ and $v(x,y,t)$, and the strains calculated directly from Eqns. 4.3. If needed, the film stress can be calculated from Eqns. 3.14. The *matlab* script, which I adapted from Huang's original to implement the numerical solution of the partial differential equations (Eqn. 4.5) for rectangular islands, can be found in App. C. The simulations were done assuming island edges parallel to the <100> crystal direction, although results for other crystal directions can be obtained by simple linear scaling: the elastic stiffness coefficients, c'_{ij}, in Eqns. 3.9 and 4.5 should be replaced by the direction-dependent equations given in App. B. For SiGe islands aligned to the <110> direction, this results in a ~15% reduction in τ_L, and a reduction in the maximum asymmetry of strain, that is, a lower tensile strain in the short island direction for the same compressive strain in the long direction, as described by Eqns. 3.10-3.12.

4.2 Lateral Expansion of Square Islands

Using this model, cross-sections of normalized strain across the island mid-section, $\varepsilon_{biaxial}/\varepsilon_o$, are plotted in Fig. 4.3a for various normalized anneal times, t/τ_l. As described above, the simulations are for square islands of a single $Si_{0.7}Ge_{0.3}$ layer that is initially fully compressively strained, with island edges aligned to the <100> crystal direction. Strain and time values are normalized so that the results may be applied to

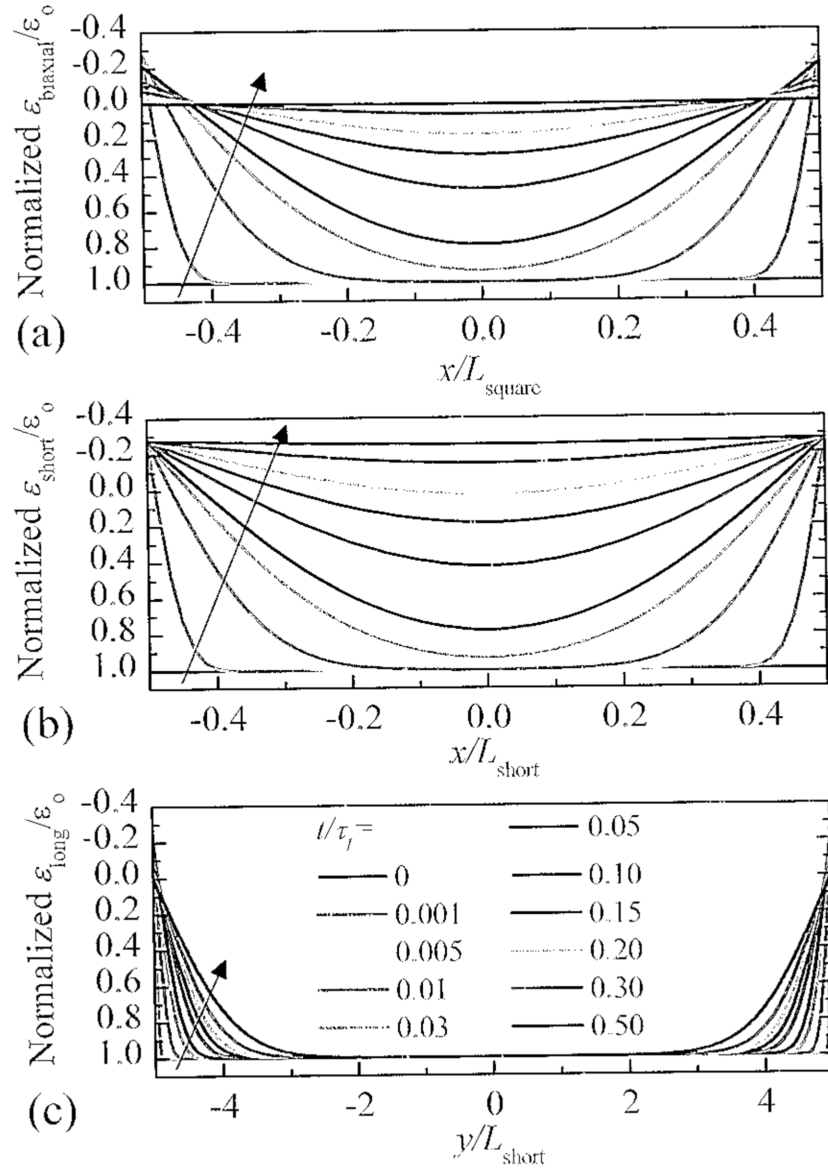


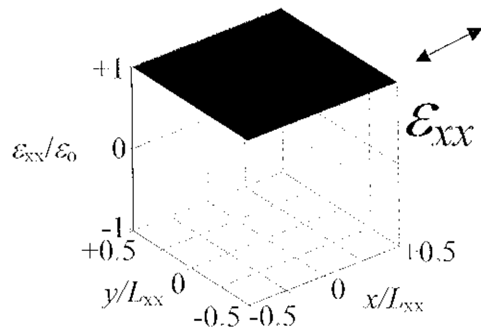
Figure 4.3: Numerical simulations of the strain across an island cross section for various anneal times. The arrows indicate the direction of the evolution of the curves in time. The plots show cross sections of: (a) a square island (after Ref. [142]); (b) the short direction; and (c) the long direction of a rectangular island with L_{long}=10 L_{short}. Note the different horizontal axis scales in (b) and (c). The islands simulated are a single $Si_{0.7}Ge_{0.3}$ layer initially fully compressively strained, aligned to <100>. The strain, plotted on the vertical axis, is normalized so the initial compressive strain equals 1.0. Each line represents a different anneal time t, normalized by τ_l, the lateral relaxation time of Eqn. 3.9 where L=L_{square}=L_{short}. In the directions of the arrows, t/τ_l increases from 0 to 0.5.

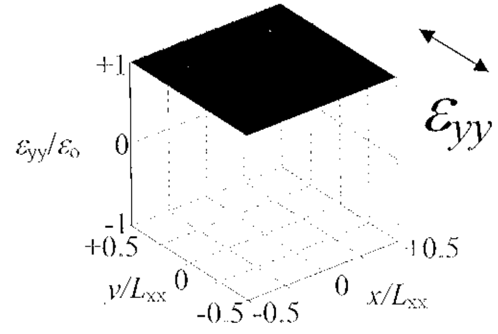
arbitrary structures. The vertical axis indicates normalized strain, with 1.0 being fully compressive ($\varepsilon=\varepsilon_o$), zero being unstrained ($\varepsilon=0$), and negative numbers indicating tension. Anneal time, t, increases in the direction of the arrow according to the values specified in the legend of Fig. 4.3c. In Fig. 4.3a, the square island edges quickly relax. Gradually, the center also relaxes. The final simulated state, at $t/\tau_L=0.5$, is zero strain across the island, precisely $\varepsilon_{biaxial}/\varepsilon_o = 0.0040 \pm 0.0035$ (average ± standard deviation). A similar plot for square islands was shown in Ref. [142]; it is shown here to contrast with the results to be presented for rectangular islands.

The dynamic relaxation of square islands is plotted in three-dimensions in Fig. 4.4. The strain state is indicated by the colorbar: red is full compressive strain while light blue is zero strain. The curves of Fig. 4.3a represent cross-sectional slices across the middle of the island (*i.e.*, ε_{xx} at $y=0$) of these 3-D plots. The images on the left and right represent strain in the two in-plane directions, ε_{xx} and ε_{yy}, respectively (see Fig. 3.22 for definitions of ε_{xx} and ε_{yy}). The lateral relaxation process is vividly displayed: the edges of the island relax first, then as the anneal progresses the relaxation reaches the center, until the entire island is relaxed to zero strain. Clearly, for square islands the relaxation in both in-plane directions is identical, $\varepsilon_{xx}=\varepsilon_{yy}=\varepsilon_{biaxial}$.

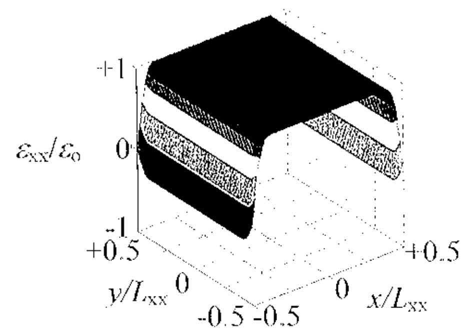
The normalized strain at the island center, $\varepsilon_{biaxial}/\varepsilon_o$ at $x=y=0$, is plotted vs normalized time, t/τ_l, in Fig. 4.5 (pink line and diamond symbols). As seen in Figs. 4.3 and 4.4, the film strain at the center of a square island relaxes to zero by $t/\tau_L=0.5$. This zero-strain equilibrium state is maintained indefinitely upon further annealing as shown by the flat line of $\varepsilon_{biaxial}/\varepsilon_o$ for $t/\tau_l>0.5$ in Fig. 4.5. For a typical sample with 30-nm $Si_{0.7}Ge_{0.3}$ on 1-μm BPSG annealed at 750°C ($\eta=5.5\times10^{10}$ N·s·m^{-2}), the value of $\tau_l \cdot L^{-2}$ from Eqn. 3.9 is ~11 sec·μm^{-2}. So, the full relaxation of a 20-μm x 20-μm island after a 40-minute anneal at 750°C is expected.

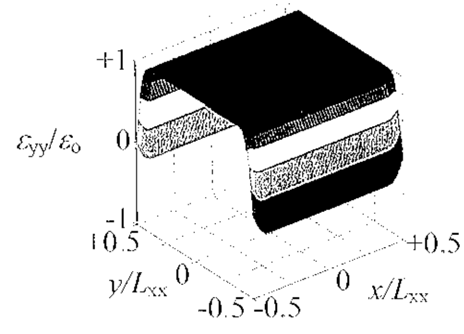
The experimental strain data shown previously in Fig. 3.23 are compared to the model in Fig. 4.6. Two different values of viscosity, $\eta=2.5\cdot10^{10}$ and $5.5\cdot10^{10}$ N·s·m^{-2}, are used in plotting the model results, with much better fitting obtained with $\eta=2.5\cdot10^{10}$ N·s·m^{-2}. Previously [111], a similar fitting procedure of experimental data to the 2-D lateral relaxation model yielded a BPSG viscosity value at 750°C of $\eta=5.5\cdot10^{10}$ N·s·m^{-2}.



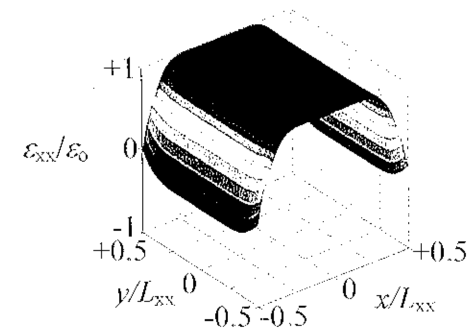


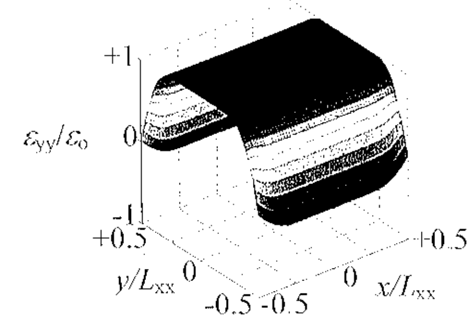
(a) $t/\tau_L = 0$



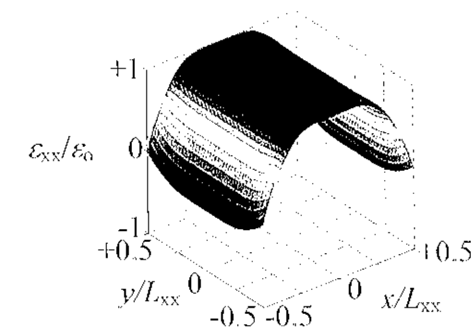


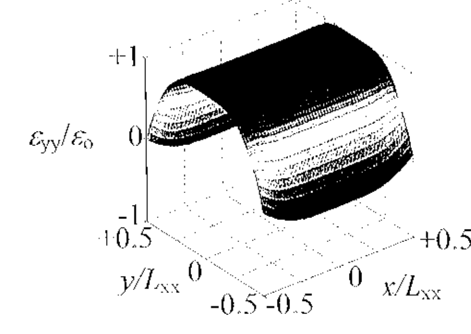
(b) $t/\tau_L = 0.001$



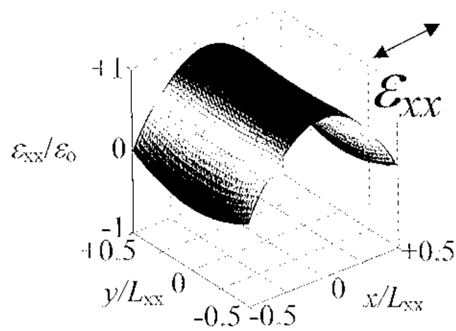


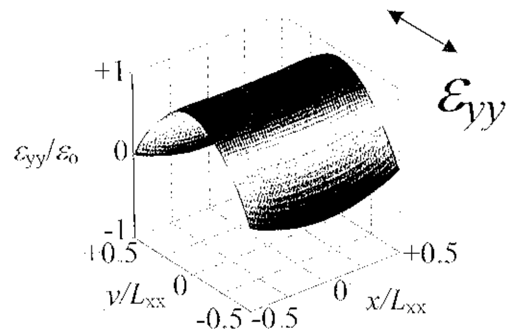
(c) $t/\tau_L = 0.005$



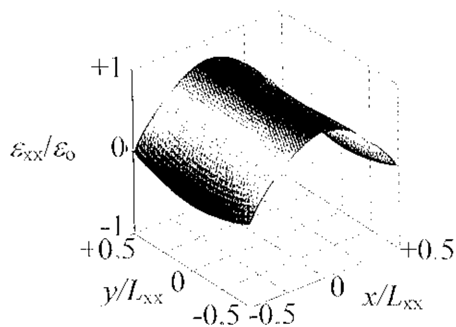


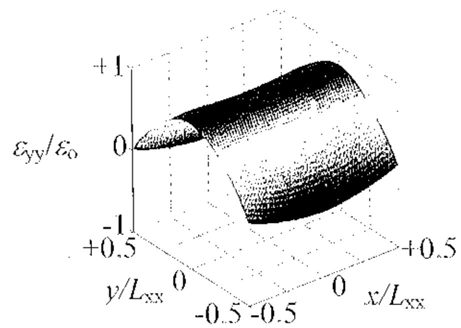
(d) $t/\tau_L = 0.01$



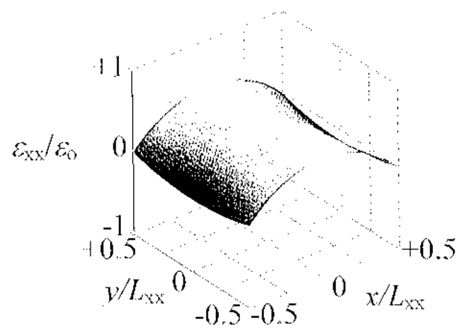


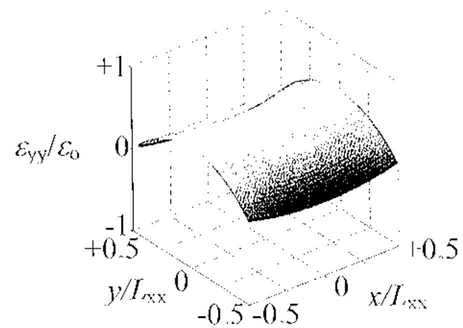
(e) $t/\tau_L = 0.03$



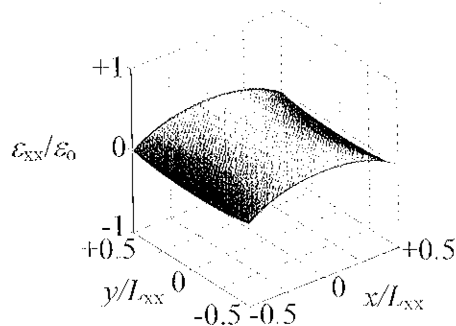


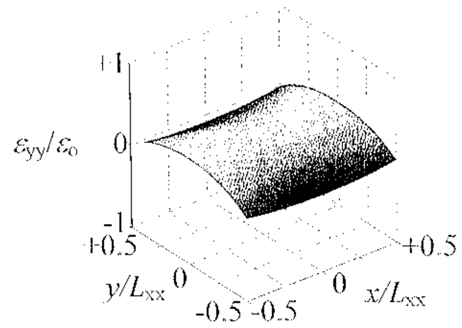
(f) $t/\tau_L = 0.05$



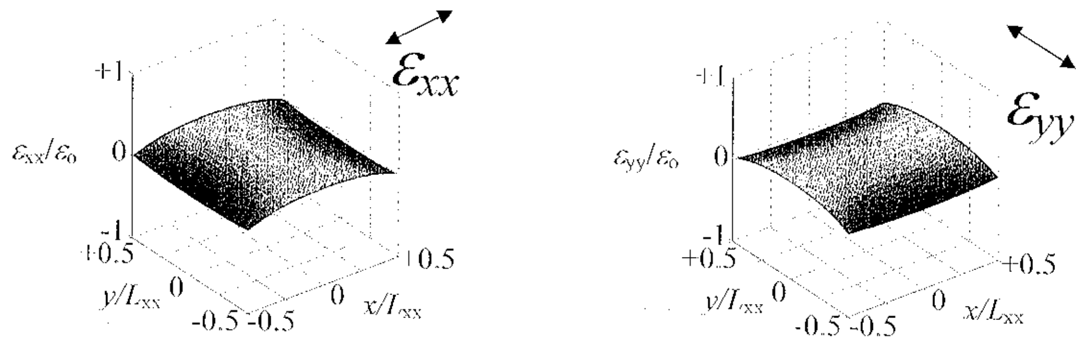


(g) $t/\tau_L = 0.10$

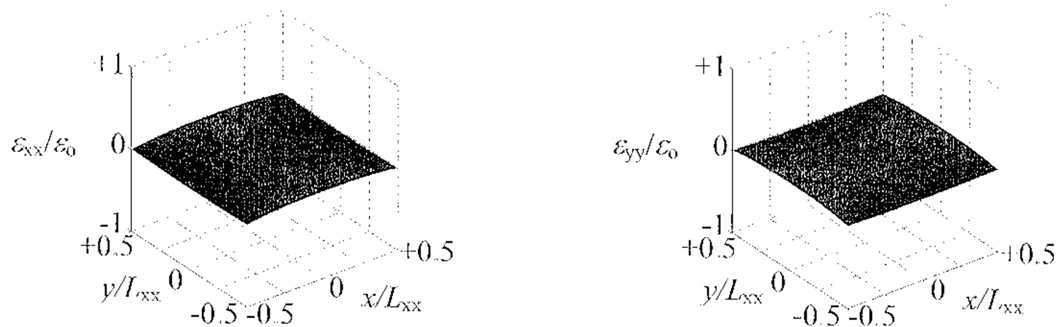




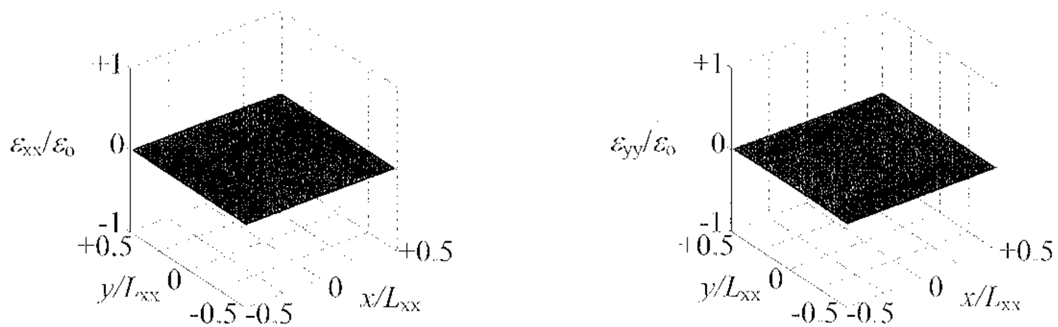
(h) $t/\tau_L = 0.15$



(i) $t/\tau_L = 0.20$



(j) $t/\tau_L = 0.30$



(k) $t/\tau_L = 0.50$

Figure 4.4: Numerical model of strain in a square island vs relaxation anneal time. The island consists of a single $Si_{0.7}Ge_{0.3}$ layer initially fully compressively strained. Each pair of images represents, on the left, strain in the x-direction and, on the right, strain in the y-direction at a specific time. The z-axis indicates normalized strain, where red indicates full compression and blue indicates tension (see color bar). Time is normalized by τ_L, the lateral relaxation constant of Eqn. 3.9. At each stage, the relaxation is exactly symmetric in the x and y directions. At the final simulation time the island has zero strain.