n-type silicon-germanium based terahertz quantum cascade lasers

Alexander Valavanis

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
School of Electronic and Electrical Engineering

July 2009

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement

For my family

"Do not believe anything merely on the authority of your teachers and elders. Do not believe in traditions because they have been handed down for many generations. But after observation and analysis, when you find that anything agrees with reason and is conducive to the good and benefit of one and all, then accept it and live up to it."

Gautama Buddha
Kalama Sutta, c. 500 BC.

Acknowledgements

I would like to express my gratitude to the many people who have encouraged and assisted me throughout the course of this work. Firstly, I thank my supervisors Rob Kelsall and Zoran Ikonić for introducing me to this fascinating field of research and for their support through many hours of enlightening meetings and discussions. Thanks are due also to the Engineering and Physical Sciences Research Council (EPSRC) for funding my studies.

I am indebted to the staff and students of the Institute of Microwaves and Photonics. In particular, to Paul Harrison for his advice during my application to study, to my collaborators Leon Lever and Craig Evans for their valuable contributions to our joint publications and to Margaret Quirke and Louise Redmonds for countless hours of administrative support. I also thank my colleagues past and present for many useful discussions: Anil Asokan, Marco Califano, Ivana Savić, Nenad Vukmirović and Goran Isić. Special words of thanks are due to Rob, Zoran and Leon for their time spent proof reading this thesis.

These acknowledgements would be incomplete without mentioning my friends and family. I'm forever grateful to you for coping with me through my monologues about charge transport simulations and my late nights at the office. Your support has been the keystone of my work.

Alex Valavanis, July 2009

Abstract

Terahertz quantum cascade lasers (THz QCLs) have many potential applications, including detection of skin tumours, and of illicit drugs and explosives. To date, all THz QCLs use III–V compound semiconductors, but silicon (Si)-based devices could offer significant benefits. The high thermal conductivity of Si may allow higher operating temperatures, removing the need for large and costly cryogenic coolers, and the non-polar nature of Si may allow a wider range of emission frequencies. The mature Si processing technology may reduce fabrication costs and ultimately allow integration of THz QCLs with mainstream semiconductor electronics.

This work investigates the suitability of a range of Si-based material configurations for THz QCL design. An effective mass/envelope function model of the electronic bandstructure is developed, taking into account the effects of strain and crystal orientation. Scattering models for Coulombic interactions, structural imperfections and interactions with phonons (lattice vibrations) are developed and used to predict the electron distribution, current density and gain in a range of device designs. The effect of nonabrupt interface geometries is investigated and the effect of intervalley mixing upon the emission spectrum is considered. It is shown that germanium/germanium–silicon heterostructures offer much better prospects for THz QCL development than silicon/silicon–germanium systems and can yield sufficient optical gain to overcome the threshold for copper–copper waveguides.

List of publications

The following journal and conference papers were published by the author during the course of the present work. The asterisk (*) denotes first-authorship. Citations (excluding self-citations) of these publications are listed under each item where applicable.

Journal papers

- * A. Valavanis, Z. Ikonić, and R. W. Kelsall, “Growth variation effects in SiGe-based quantum cascade lasers,” *J. Opt. A*, vol. 11, no. 5, p. 054012, 2009
- * A. Valavanis, Z. Ikonić, L. Lever, C. Evans, and R. Kelsall, “Substrate orientation and alloy composition effects in n-type SiGe quantum cascade structures,” in *5th IEEE Int. Conf. on group IV Photon.*, pp. 251–252, 2008
- * A. Valavanis, L. Lever, C. A. Evans, Z. Ikonić, and R. W. Kelsall, “Theory and design of quantum cascade lasers in (111) n-type Si/SiGe,” *Phys. Rev. B*, vol. 78, no. 3, p. 035420, 2008
    - G. W. Carter prize awarded for the best written paper accepted for publication in 2008/9 by a postgraduate student in the School of Electronic and Electrical Engineering at the University of Leeds.
- L. Lever, A. Valavanis, Z. Ikonić, and R. W. Kelsall, “Simulated [111] Si-SiGe terahertz quantum cascade laser,” *Appl. Phys. Lett.*, vol. 92, no. 2, p. 021124,

2008

  - Cited by refs. [5] and [6].

- * A. Valavanis, Z. Ikonić, and R. W. Kelsall, “Intersubband carrier scattering in *n*- and *p*-Si/SiGe quantum wells with diffuse interfaces,” *Phys. Rev. B*, vol. 77, p. 075312, Feb. 2008
  - Cited by ref. [8].
- * A. Valavanis, Z. Ikonić, and R. W. Kelsall, “Intervalley splitting and intersubband transitions in *n*-type Si/SiGe quantum wells: Pseudopotential vs. effective mass calculation,” *Phys. Rev. B*, vol. 75, no. 20, p. 205332, 2007
  - Cited by refs. [5], [10], [11] and [12].

Conference papers

- * A. Valavanis, L. Lever, C. Evans, Z. Ikonić, and R. Kelsall, “Low electric field silicon-based THz quantum cascade laser employing *L*-valley intersubband transitions,” in *UK Semicond.*, Inst. of phys., 2009
- * A. Valavanis, Z. Ikonić, L. Lever, C. Evans, and R. Kelsall, “Substrate orientation and alloy composition effects in *n*-type SiGe quantum cascade structures,” in *The 5th Annual Conf. on Group IV Photon.*, IEEE LEOS, 2008
- * A. Valavanis, Z. Ikonić, and R. Kelsall, “Intersubband transitions in *n*-type group IV quantum cascade lasers,” in *Int. quantum cascade lasers school and workshop*, (Monte Verita, Switzerland), 2008
- * A. Valavanis, L. Lever, C. Evans, Z. Ikonić, and R. Kelsall, “Theory and design of (111) oriented Si/SiGe quantum cascade lasers,” in *QEP-18*, Inst. of phys., 2008

- * A. Valavanis, L. Lever, Z. Ikonić, and R. Kelsall, “Growth variation effects in Si/SiGe quantum cascade lasers,” in *QEP-18*, Inst. of phys., 2008
- * A. Valavanis, L. Lever, C. Evans, Z. Ikonić, and R. Kelsall, “Crystal orientation and waveguide geometry effects in n-type Si/SiGe quantum cascade lasers,” in *UK Semicond.*, Inst. of phys., 2008
- * A. Valavanis, L. Lever, Z. Ikonić, and R. Kelsall, “n-type Si/SiGe quantum cascade structures,” in *The 9th Int. Conf. on Intersubband Transitions in Quantum Wells*, 2007
- * A. Valavanis, Z. Ikonić, and R. Kelsall, “Intervalley mixing in two-dimensional n-type Si/SiGe heterostructures,” in *The 9th Int. Conf. on Intersubband Transitions in Quantum Wells*, 2007
- * A. Valavanis, Z. Ikonić, and R. Kelsall, “n-type Si/SiGe quantum cascade lasers,” in *Mini-Symposium on Si-based optoelectronics*, The Rank Prize Funds, 2006

List of abbreviations

BTC	Bound-to-continuum
CB/VB	Conduction/valence band
CVD	Chemical vapour deposition
CMOS	Complementary metal–oxide–semiconductor
CSL	Chirped superlattice
DVEMA/EMA	(Double valley) effective mass approximation
EPM	Empirical pseudopotential model
FCC	Face-centred cubic
FTIR	Fourier transform infrared (spectroscopy)
FWHM	Full-width at half-maximum
LED	Light-emitting diode
LH/HH/SO	Light-hole, heavy-hole, spin-orbit splitting (valence band)
LLL/ULL	Lower/upper laser level
LO/LA	Longitudinal optic/acoustic (phonon)
MBE	Molecular beam epitaxy
ML	Monolayer
NEGF	Nonequilibrium Green's function
TEM	Transmission electron microscopy
TO/TA	Transverse optic/acoustic (phonon)
QCL	Quantum cascade laser
QW	Quantum well
RP	Resonant phonon

List of symbols

Fundamental constants

$c = 3.00\times10^{8}\,\mathrm{m/s}$	Speed of light in vacuum
$e = 1.60\times10^{-19}\,\mathrm{C}$	Elementary charge
$\varepsilon_0 = 8.85\times10^{-12}\,\mathrm{F/m}$	Vacuum permittivity
$\hbar = 1.05\times10^{-34}\,\mathrm{J\,s}$	Reduced Planck constant
$k_B = 1.38\times10^{-23}\,\mathrm{J/K}$	Boltzmann constant
$m_e = 9.11\times10^{-31}\,\mathrm{kg}$	Rest mass of free electron

Geometric notation

R, R'	Interfacial and crystallographic coordinate systems
$\boldsymbol{R}$	Displacement (3-dimensional)
$\boldsymbol{r}$, r	Displacement, distance (in-plane)
U	Transformation matrix, $\mathsf{U} : R \to R'$
z	Position normal to interfaces

Thermal properties

T	Lattice temperature
T_e	Electron temperature

Mechanical notation

E	Energy
E_k, $\hat{T}$	Kinetic energy
$\boldsymbol{k}$, k, g	Wavevector
m	Rest mass
t	Time
θ	Direction of motion
u	Bloch function
ψ, ϕ	Wavefunction
ω	Angular frequency

Elasticity

C, C_{ij}	Elastic constant tensor (Voigt notation)
c, c_{ijfg}	Elastic constant tensor (full notation)
ε, ε_{ij}	Strain
σ, σ_{ij}	Stress
ν	Poisson ratio

Design specification

d	Doping profile
F	Average electric field
L_w, L_p	Width of quantum well/period of QCL
N_{2D}	Total donor density
S	Length of QW in monolayers
x	Ge fraction

Material properties

a, b_{bow}	Lattice constant, bowing parameter
a_v	Valence band hydrostatic deformation potential
Δ_{so}	Spin-orbit splitting
E_c, E_v	Energy of conduction/valence band edge
E_g	Bandgap energy
ε, ε_r	Permittivity, dielectric constant
$\boldsymbol{G}$	Reciprocal lattice vector
k_0	Wavevector of conduction band minimum
m_c^*, m_v^*	Effective mass at conduction/valence band edge
m_l, m_t	Effective mass along longitudinal/transverse axis of valley
m_q, m_d	Quantisation/density-of-states effective mass
n_{val}	Number of equivalent valleys
ρ^{2D}	Two-dimensional density of states
ρ	Mass density
v_s	Speed of sound
Ξ_d, Ξ_u	Conduction band deformation potentials (dilation, shear)

Interdiffusion

D	Diffusion constant
ΔE_{pk}	Peak shift in subband separation
L_d	Interdiffusion length
L_{pk}	L_d giving peak shift in subband separation
L_w	L_d giving 50% increase in scattering rates

Electronic properties

D	Electric displacement field
D_{ij}	Dipole matrix element
ΔE	Valley splitting
E_F	Quasi-Fermi energy
f_{FD}, P	Fermi-Dirac occupation probability
$\hat{H}$, H	Hamiltonian operator
n	Total electron density
Q	Total charge density
ρ	Charge density
ρ_e	Electron density
$\hat{U}$	Valley splitting potential
V_v	Energy splitting between conduction band valleys
V_b	Conduction band offset
V_{crystal}	Potential due to crystal
V_F	Potential due to internal/external electric fields
V_{range}	Usable range of energies in quantum well

Scattering parameters

A	Cross-sectional area
Δ, Λ	Interface roughness height, correlation length
ΔV_{ad}	Alloy disorder potential
F_{ij}, B_{if}	Interface roughness scattering matrix element
$\boldsymbol{q} = (q, \theta_q)$	Scattering vector (magnitude, angle)
τ	Lifetime
$\hat{V}$, V_{ij}	Perturbation potential, matrix element
W_{ij}, $W_{ij \to fg}$	Scattering rate
Ω	Volume containing single lattice site

Coulombic scattering

$A_{ij\to fg}$	Electron-electron scattering matrix element
I_{if}	Coulombic scattering matrix element
J_{if}	Ionised impurity scattering matrix element
k_F	Fermi wavevector
Π_{ii}	Polarisation factor
q_{TF}	Thomas-Fermi screening vector

Phonon scattering

Δ_{ac}	Acoustic phonon deformation potential
D_0, D_1	Optical phonon deformation potential
F_{if}, G_{if}, H_{if}	Phonon scattering matrix elements
f_{MB}, N_q	Maxwell-Boltzmann occupation factor
n_{dest}	Number of destination valleys
ω_q	Phonon angular frequency

Gain and current density

α_m, α_w	Mirror/waveguide loss
f_{ij}	Oscillator strength
γ_{ij}	Half-width at half maximum (linewidth)
g	Gain coefficient
G	Gain
G_{th}	Threshold gain
Γ	Modal overlap factor
J	Current density
λ	Wavelength
L_{ij}	Lineshape (Lorentzian)
μ_g	Gain figure of merit
n	Refractive index
R	Facet reflectivity
σ	Optical cross-section

Contents

Chapter 1

Introduction

This work discusses the theory and design of n-type silicon-based quantum cascade lasers (QCLs) emitting terahertz (THz) frequency radiation. Numerous texts have been written, describing all aspects of QCLs from the most fundamental underlying quantum theory to their applications in commercial products. This introduction provides an overview of the most relevant material, including the basic concepts behind QCLs and the possible applications for THz radiation sources. This chapter also explains the motivation behind the use of silicon-based materials. As a prerequisite, it is assumed that the reader has some familiarity with semiconductor electronics.

1.1 Interband semiconductor lasers

Light interacts strongly with electrons if the incident photon energy $\hbar\omega$ is close to the separation between a pair of electron energy levels $E_{21} = E_2 - E_1$ as shown in fig. 1.1. Either of the following processes can occur depending on the initial state of the electron. If the electron is initially in the upper energy level [fig. 1.1(a)] then *stimulated emission* occurs: the electron drops to the lower energy level and a new photon is emitted with the same frequency, phase and direction of propagation as the incident photon. Alternatively, if the electron is initially in the lower energy state

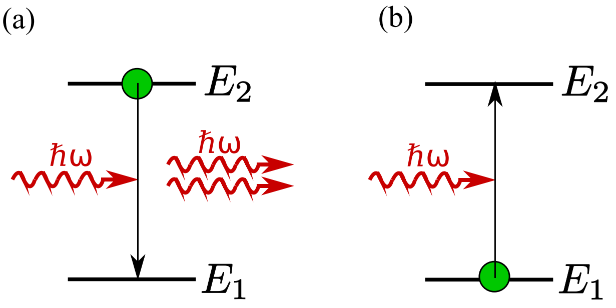


Figure 1.1: (a) Stimulated emission occurs when a photon interacts with an excited electron. The electron drops to a lower energy level and a second photon is emitted in phase coherence with the incident photon. (b) An electron can be excited to a higher energy level by absorbing a photon of the appropriate energy.

[fig. 1.1(b)], it may absorb the incident photon and become excited to the upper energy level. A system in thermal equilibrium has a larger number of electrons in lower energy states than in higher energy states and absorption of photons occurs more readily than emission. Light is, therefore, absorbed as it propagates through the system.

A laser, however, is a non-equilibrium system in which electrons are *pumped* (using an external light source or electrical current source) into higher energy states to create a *population inversion*, where more electrons exist in the upper laser level than in the lower laser level. As a result, stimulated emission occurs more readily than absorption and light amplification (or optical gain) occurs as photons pass through the device.

In conventional laser diodes [fig. 1.2(a)], the optical transition occurs across the energy bandgap which separates the conduction and valence band states. This is known as an *interband* transition and the emission energy of the device is approximately equal to the bandgap of the semiconductor.

Figure 1.2(b) shows a *quantum well* (QW) device structure, which offers significant advantages over conventional bulk interband lasers. In QW lasers, a very thin layer of semiconductor is “sandwiched” between two layers of a different semicon-

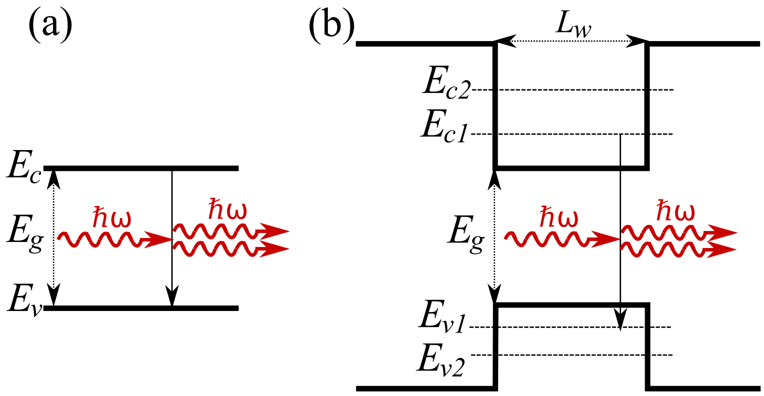


Figure 1.2: (a) In a conventional interband laser, optical transitions occur between the conduction and valence band edge and the emission energy is approximately equal to the bandgap. (b) In a quantum well laser, the optical transition occurs between conduction and valence subbands and the emission energy is greater than the bandgap of the well material.

ductor which trap electrons inside a potential well in the central layer. The QW is so thin that electrons are subject to quantum confinement effects and may only occupy a set of discrete energy *subbands*. If the confining potential is deep, the permitted conduction band energies are approximately given by

$$E_{c,n} \approx E_c + \frac{\hbar^2\pi^2 n^2}{2m_c^* L_w}, \tag{1.1}$$

and the valence band energies by

$$E_{v,n} \approx E_v - \frac{\hbar^2\pi^2 n^2}{2m_v^* L_w}, \tag{1.2}$$

where m_c^* and m_v^* are the effective masses of electrons in the conduction and valence band respectively, L_w is the width of the central layer of the QW, E_c and E_v are the conduction and valence band edges. $E_{c,n}$ and $E_{v,n}$ are the quantised sets of conduction and valence subband minima. The emission energy between the first conduction subband and the first valence subband is now given by

$$\hbar\omega \approx E_g + \frac{\hbar^2\pi^2}{2L_w}\left(\frac{1}{m_c^*} + \frac{1}{m_v^*}\right). \tag{1.3}$$

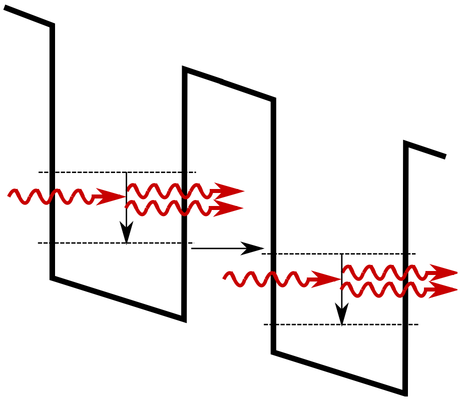


Figure 1.3: Schematic representation of optical amplification in a biased superlattice. The thick black line represents the conduction band edge in a periodic heterostructure. The electron undergoes an optical transition in each quantum well and tunnels through to the next well.

Thus, the emission energy may be increased by reducing the thickness of the QW.

The bandgap is of fundamental importance in determining the emission frequency of interband devices. Although shorter wavelengths can be achieved by reducing the well width of QW lasers, the low frequency limit is always determined by the bandgap. This presents a problem, as the most commonly used optoelectronic semiconductors (GaAs and InP based materials) cannot emit below near-infrared frequencies. Mid-infrared frequencies are accessible using more exotic lead salt compounds[22], but these materials may be expensive, difficult to process and poorly understood[23].

1.2 Quantum cascade lasers

The low energy limit may be overcome by designing structures which exploit transitions between pairs of conduction subbands rather than transitions across the bandgap. These *intersubband transitions* allow much lower emission frequencies than those achievable in interband lasers. In 1971, Kazarinov and Suris[24] proposed that a population inversion, and hence light amplification, could be achieved

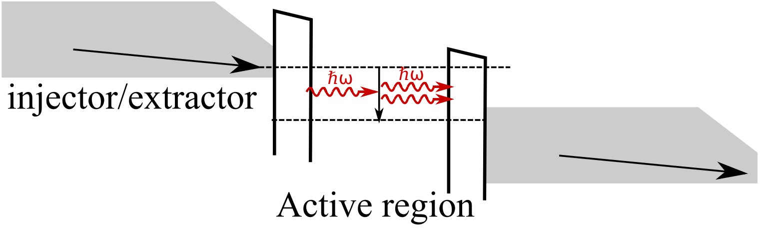


Figure 1.4: Generalised schematic of a quantum cascade laser. All energy levels in the figure lie within the conduction band.

from a long chain of coupled QWs (a superlattice) placed in an electric field as illustrated in fig. 1.3. The principle behind such a device is that electrons transition from a state in one QW to a state in the next, cascading “downstream” with the electric field and emitting a photon at each stage. However, lasing was not achieved experimentally due to formation of non-uniform electric fields and the inability to establish a stable current through the device. Optical emission was also hindered by the emission of LO-phonons (lattice vibrations)[25].

Despite these early setbacks, optical amplification from a biased superlattice has become the operating principle for quantum cascade lasers (QCLs), illustrated in fig. 1.4. QCLs build upon the biased superlattice concept by using several coupled QWs of varying width in each period of the structure, rather than a single quantum well. Each period is split into an *active region* and an *injector/extractor region*. The optical transition occurs between a pair of energy levels in the active region. The purpose of the injector/extractor region is to efficiently extract electrons from the lower laser level in one period and inject electrons into the upper laser level of the next period downstream.

In 1994, a group at Bell Labs demonstrated the first QCL[26], using a GaInAs/AlInAs heterostructure on an InP substrate. The injector/extractor region was designed to efficiently draw electrons from the lower laser level by exploiting rapid LO-phonon scattering—the same effect which proved detrimental in biased superlattices. Within a year of their invention, continuous wave emission[27] had been demonstrated. Soon afterwards, in 1996, Faist et al reported the first room temperature operation of a mid-infrared QCL[28].

An alternative “bound–to–continuum” QCL design was first demonstrated in 1997, which uses a superlattice as its injector/extractor region[29]. Here, the rapid scattering within a superlattice is used to establish a population inversion. This type of device was found to be well suited to long emission wavelengths and has the ability to handle larger operating currents. In 1998, the first GaAs/AlGaAs based QCL was developed[30], showing that QCLs may be constructed in more than one material system. Indeed, the longest wavelength mid-infrared emission from a QCL ($\lambda = 15\,\mu$m) was obtained from a GaAs/AlGaAs device[31]. However, InP-based devices have several advantages over GaAs-based devices at mid-infrared frequencies. The larger conduction band offset reduces leakage currents in short wavelength devices, allowing emission at wavelengths as short as $3\,\mu$m[32]. The higher thermal conductivity of InP allows heat to be removed effectively from the active region and mid-infrared high-intensity (204 mW), continuous-wave optical output is achievable at room temperature[33]. By contrast, the best room-temperature GaAs-based mid-infrared QCLs only yield around 50 mW pulsed-mode emission[34]. The shortest emission wavelengths have been achieved by using InAs/AlSb heterostructures upon InP substrates. The large conduction band offset (~ 2.1 eV) has allowed emission wavelengths as short as $2.75\,\mu$m to be realised[35].

1.3 Terahertz radiation

The terahertz frequency range usually refers to the part of the electromagnetic spectrum with frequencies between 300 GHz and 10 THz. Terahertz radiation interacts strongly with a number of physical systems including small molecules such as water, and electronic transitions in semiconductors, making it useful in a variety of applications[36]. Terahertz features in the spectra of thermal emissions from stellar dust clouds have proved useful in the search for extrasolar planets[37]. Terahertz spectroscopy is also developing rapidly with applications such as electron mobility measurements in semiconductors and dielectrics[38] and carrier lifetime measurements in heterostructures[39]. Terahertz imaging has been used to investigate early detection of skin tumours[40] and systems have been developed to test the manufacturing quality of pill coatings and to detect concealed weapons[41].

In order to realise the full potential of these applications, a compact, coherent, continuous-wave radiation source is required. However, a so-called "terahertz gap" exists in the technology. Electronic sources, such as Gunn oscillators and Schottky diode multipliers work well at radio or microwave frequencies but their power output rolls off at higher frequencies due to the nonzero carrier transit times and the capacitance of the devices[42, 43]. On the other hand, conventional semiconductor lasers are restricted to high frequency emission by the material bandgap, as described in the preceding sections.

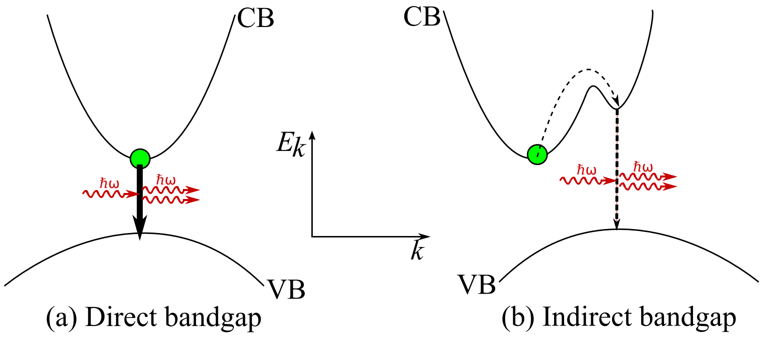
Early methods of bridging the terahertz gap included broadband sources such as photoconductive antennas[44]. Subpicosecond laser pulses are used to excite electron–hole pairs within the photoconductor, which are excited by an applied electric field. The resulting transient currents give rise to broadband radiation around the terahertz region. An alternative nonlinear optics approach, using optical rectifiers (also driven by ultrafast laser pulses) gives relatively low output powers but very high bandwidths[45]. High-powered, tunable narrowband sources are extremely important for spectroscopy[46]. A common approach to narrowband emission involves frequency upconversion of microwave sources using chains of Schottky diode multipliers[47]. High-powered coherent narrowband emission can be achieved with methanol or hydrogen cyanide lasers, however these are large devices, requiring kilowatt power supplies[46]. Even higher emission powers are achievable using free-electron lasers with energy-recovering linear accelerators. Electrons are accelerated to a high velocity, before passing through a strong, spatially undulating magnetic field. Light is emitted by the oscillating charge and confined by mirrors, resulting in optical amplification and laser action[48].

Compact, coherent terahertz emitters first became viable in 2002 with the first demonstration of a THz QCL[49]. Although the performance of THz QCLs is not yet at the same level as mid-infrared devices, significant advances have been made in recent years. Devices are now capable of operating in ambient temperatures of up to 186 K[50]. Output powers up to 248 mW (pulsed mode) and 138 mW (continuous wave) are achievable[51], and devices have been demonstrated with emission frequencies ranging from 1.2 THz[52] (or 950 GHz in a strong magnetic field[53]) up to 4.9 THz[54]. Unlike mid-infrared devices, almost all THz QCLs have been based upon GaAs, although an InP-based device has been demonstrated which operated up to 45 K[55].

Despite the remarkable success with GaAs and InP based terahertz QCLs, there are several serious limitations. Firstly, the maximum operating temperature remains so low that large and expensive cryogenic cooling systems are required. This severely limits the practicality of THz QCLs for any application requiring portability. Secondly, light with a frequency close to the natural vibration frequencies (optical phonon frequencies) of the crystal lattice is absorbed strongly by III–V materials due to the polar nature of the chemical bonds. This leads to a forbidden *Reststrahlen band* of optical energies, which lies between the mid-infrared and terahertz regions.

1.4 Silicon lasers

Silicon is the dominant material in the semiconductor industry. It has convenient properties, such as the ability to form an electrically insulating SiO_2 layer and the enormous investment in Si technology has resulted in very large scale integration and much lower device costs compared with other semiconductors[56]. Unfortunately, Si is a poor material for constructing interband lasers due to its *indirect bandgap* (fig. 1.5). This means that the electrons in Si must gain sufficient energy to populate the higher energy Γ valley which lies directly above the valence band maximum before light is emitted, which makes interband lasing almost impossible to achieve.



(a) Direct bandgap

(b) Indirect bandgap

Figure 1.5: Schematic illustration of direct and indirect bandgap materials. The vertical axis represents the energy of the electron, while the horizontal axis represents the wavevector. (a) In direct bandgap materials, the conduction band minimum and the valence band maximum lie at the same wavevector and low energy electrons can undergo optical transitions directly across the bandgap. (b) In indirect bandgap materials, the conduction band minimum and valence band maximum lie at different wavevectors. Electrons must gain enough energy to populate the much higher energy Γ valley, which lies directly above the valence band maximum before undergoing a radiative transition.

Despite the challenges, Si-based lasers are extremely desirable[57]. A Si laser could potentially be integrated with mainstream CMOS electronics to obtain a *photonic integrated circuit*[58, 59] allowing optical signals to be generated and controlled on a single chip. Si has a higher thermal conductivity than compound semiconductors ($1.3\,\mathrm{W\,cm^{-1}\,K^{-1}}$ in Si *c.f.* $0.55\,\mathrm{W\,cm^{-1}\,K^{-1}}$ in GaAs and $0.68\,\mathrm{W\,cm^{-1}\,K^{-1}}$ in InP[60]), which could allow heat to be dissipated more effectively from devices, allowing higher temperature operation. Additionally, Si is a non-polar material, meaning that there is no forbidden Reststrahlen band and potentially opening up a wider range of emission energies. The absence of polar LO-phonon interactions in silicon may also improve the population inversion achievable at high temperatures.

Several notable approaches have been taken towards a silicon laser. Phonon assisted optical emission has been used to overcome the indirect bandgap limitiation. In this process, the change in electron wavevector as it crosses the bandgap is balanced by the emission of one or two phonons. Such interactions are normally weak, but structures have been demonstrated which enhance light–matter interactions around the appropriate wavelength. Free-carrier absorption of photons is minimised in such structures by using a very low doping density, and undesirable nonradiative recombination processes are avoided by using exceptionally high quality materials. Efficient Si-based LEDs have been developed using this principle[61].

An alternative approach involves the use of erbium doping in Si diodes. This introduces extra energy levels in the depletion region, which allow a direct recombination with a wavelength of $1.54\,\mu$m [62], making them suitable for fibre communications technology. LEDs[63] and optical amplifiers[64] have been developed based on this principle.

Most other approaches rely on confining electrons in one or more dimensions in order to change the energy band structure and thus relax the momentum conservation issues. Porous Si crystals have been used [65] to create quantum wires and dots. Si nanocrystals may also be embedded in a layer of SiO_2[66], and light amplification has been demonstrated in such systems[67]. An alternative approach uses small dislocations to confine carriers[68, 69].

A breakthrough in Si optoelectronics came in 2004 with the demonstration of the first Si laser by Intel[70], which uses stimulated Raman scattering. Continuous wave operation followed in 2005[71]. Recent research (also by Intel) has used a ring resonator to increase the output power to 30 mW[72]. However, this approach requires optical pumping from an external laser and to date, no electrically driven Si-based lasers have been demonstrated. Additionally, all the devices described above emit light around near-infrared wavelengths, and therefore leave the useful terahertz region (described in the previous section) unexplored.

As the QCL does not employ interband transitions, the indirect bandgap of silicon is not an issue for optical emissions. Indeed, research into Si-based QCLs has already yielded very promising results. All experimental work to date has involved intersubband transitions within the valence band of SiGe/Si heterostructures. Mid-infrared electroluminescence from a p-type SiGe/Si quantum cascade structure was first observed in 2000[73], and terahertz emitting devices were first demonstrated in 2002[74]. Investigations into the electronic behaviour of these structures have been conducted,[75, 76], giving a greater understanding of the material systems. However, the complexity of the valence band structure makes the design of *p*-type SiGe/Si QCLs a complicated process and lasing has never been observed.

Initially, research into *n*-type QCL structures was dismissed as the large effective mass of the Si conduction band requires the layers of semiconductor in the QCL to be extremely thin and would be detrimental to the optical gain of the device. However, the conduction band edge in silicon has an almost parabolic dispersion relation, meaning that transition energies are almost independent of electron wavevector. This could potentially give a lower spectral linewidth than *p*-type structures, and hence greater peak gain. In fact, in recent years, attention has switched noticeably toward *n*-type structures. A number of theoretical investigations have proposed methods to overcome the issue of the large conduction band effective mass. An approach developed by the author of the present work makes use of (111) oriented Si/SiGe structures[3, 4], in which the conduction band effective mass is comparable to the effective hole mass in *p*-SiGe/Si structures. Recently, structures using germanium quantum wells with SiGe barriers upon Si substrates have gained attention. The effective mass is roughly half that of holes in *p*-SiGe/Si structures and QCL designs have been proposed for electrons either close to the Ge conduction band edge[77–79] or in the higher energy Γ valley[80]. These theoretical investigations, along with the advances in crystal growth technology (described in chapter 6), lay the foundations for the development of the first silicon-based quantum cascade laser.

1.5 Thesis structure

The primary aim of this work is to investigate the suitability of SiGe-based heterostructures for THz QCLs. A secondary aim is to use the results of this investigation to develop promising preliminary device designs. Chapter 2 describes the energy bandstructure of SiGe thin films, taking into account the effects of strain and crystal orientation and investigates whether sufficiently deep quantum wells are achievable for THz QCL designs. Chapter 3 then describes the effective masses of electrons in these alloy films and shows how energy subbands (and wavefunctions) can be found in a heterostructure, taking the effect of a nonuniform space-charge distribution into account. The effect of *intervalley mixing* (due to the complicated structure of the conduction band in Si or Ge) upon the energy bandstructure is considered in chapter 4.

In chapter 5, Fermi's Golden Rule is used to model each of the significant scattering mechanisms in SiGe-based heterostructures and scattering rates are determined for simple quantum well structures using a variety of material configurations. Chapter 6 then describes practical issues concerning the growth of heterostructures, including strain-balancing to ensure mechanical stability, and the effect of diffuse interfaces upon the bandstructure and scattering rates. Chapter 7 shows how the scattering rate equations may be used to find the energy distribution of electrons and describes how gain and current density may be calculated for THz QCLs. In chapter 8, a range of THz QCL designs are presented and simulated. The performance of the devices is analysed, and related to the properties of the different material configurations. Chapter 9 summarises the results of previous chapters and presents proposals for further work.

Chapter 2

Silicon–germanium band structure

Quantum wells (QW)s are very narrow regions of low potential surrounded by regions of higher potential. QWs are the building blocks of quantum cascade lasers and in this chapter it is shown that sufficiently deep QWs can be constructed from silicon–germanium (SiGe) alloys.

Electrons in crystalline solids may only occupy states within certain bands of energy. The lowest *conduction band* is the focus of the present investigation. The conduction band energy varies with respect to wavevector, and most of the electron population lies in a small number of “valleys” at the bottom of the band. The dependence of the valley energies upon the amount of Ge in a SiGe alloy is investigated in this chapter.

QWs may be constructed by sandwiching two different SiGe alloys together. The difference in atomic spacing between Si and Ge crystals causes strain in the QW and this shifts the energy of the valleys. Strain also leads to certain sets of valleys being split in energy.

The *usable* range of energies within QWs is limited by two factors: the depth of the QW itself, and the energy separation between sets of conduction band valleys.

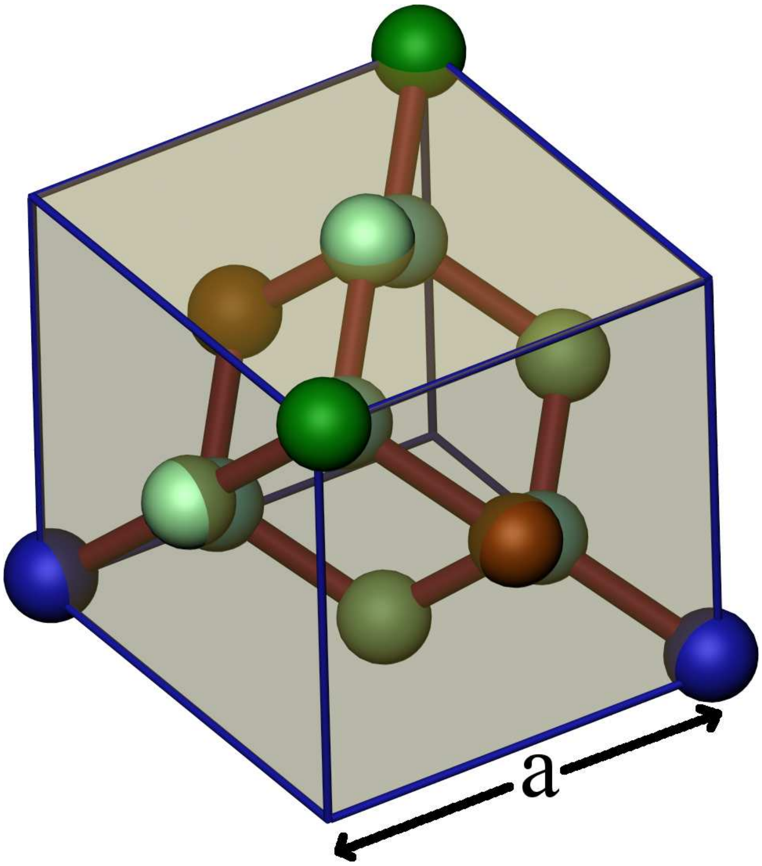


Figure 2.1: Diamond lattice structure. The lattice constant is marked as a in the figure[81].

It is shown that the best energy range is obtained by using either (111) oriented Si or (001) Ge as the well material.

2.1 General properties

Si and Ge atoms form a diamond lattice structure (fig. 2.1). This is a face-centred cubic (FCC) structure with a diatomic basis and a single lattice constant a describes its dimensions. Additional atoms are located at displacements of $\left(\frac{a}{4}, \frac{a}{4}, \frac{a}{4}\right)$ from each FCC atom, meaning that each unit lattice cell contains four atomic monolayers. The model solid approximation[82] treats crystals of SiGe alloys as if they were

Constant	Si	Ge	Unit
a	0.5431[1]	0.5633[1]	nm
Δ_{so}	44[2]	296[3]	meV
C_{11}	165.773[4]	128.528[4]	GPa
C_{12}	63.924[4]	48.260[4]	GPa
C_{44}	79.619[4]	66.799[4]	GPa
$(\Xi_d + \frac{1}{3}\Xi_u - a_v)^\Delta$	1.72[5]	1.31[5]	eV
$(\Xi_d + \frac{1}{3}\Xi_u - a_v)^L$	-3.12[5]	-2.78[5]	eV
Ξ_u^Δ	9.16[5]	9.42[5]	eV
Ξ_u^L	16.14[5]	15.13[5]	eV

Table 2.1: Material parameters for Si and Ge.

made uniformly of virtual atoms with characteristics which can be interpolated between those of Si and Ge, and the diamond lattice structure is therefore preserved. Important material parameters for Si and Ge are given in table 2.1.

A lattice mismatch of around 4% exists between Si and Ge and the lattice constants of relaxed $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ alloys are given by

$$a(x) = a_{\mathrm{Si}}(1-x) + a_{\mathrm{Ge}}x - b_{\mathrm{bow}}x(1-x). \tag{2.1}$$

A bowing constant of $b_{\mathrm{bow}} = 0.2733\,\mathrm{pm}$[83] is included to account for nonlinearity[88].

---

[1] Reference [83].

[2] Reference [84].

[3] Reference [85].

[4] Reference [86].

[5] Reference [87].

2.2 Bloch wave model

The electric potential in a crystal is periodic in real space and Bloch's theorem states that the wavefunction for electrons is of the form

$$\psi_{\boldsymbol{k}'}(\boldsymbol{R}) = u_{\boldsymbol{k}'}(\boldsymbol{R})\mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{R}}, \tag{2.2}$$

where $\boldsymbol{k}'$ is the Bloch wavevector. In other words, it is a plane wave modulated by a *Bloch function*, $u_{\boldsymbol{k}'}(\boldsymbol{R})$ which shares the same periodicity as the crystal lattice[89].

If the structure has a lattice vector $\boldsymbol{a}$, then a reciprocal lattice vector $\boldsymbol{G}_n$ may be defined such that $\boldsymbol{G}_n \cdot \boldsymbol{a} = 2\pi n$. Any Bloch wavevector may then be written as $\boldsymbol{k}' = \boldsymbol{k} + \boldsymbol{G}_n$, where $k < \frac{\pi}{a}$ and n is an integer. The first Brillouin zone may be defined as a primitive cell of the reciprocal lattice, and all unique solutions to the Schrödinger equation may be mapped to bands of energies within this region.

The Bloch model is not strictly applicable to systems in an external electric field as this breaks the periodicity of the crystal potential. In QCLs however, the electric field introduces only a small shift in potential across a crystalline monolayer and it is reasonable to assume that the Bloch model approximates the system well.

2.2.1 Brillouin zone for diamond structure crystals

The Brillouin zone for a diamond lattice is a truncated octahedron. A number of important symmetry points and directions in the Brillouin zone are illustrated in fig. 2.2. The most important symmetry points are labelled as $\Gamma = (0,0,0)$, $X = \frac{2\pi}{a}\{1,0,0\}$ and $L = \frac{\pi}{a}\{1,1,1\}$[90]. The X points are connected to Γ along the Δ directions, while the L points are connected to Γ along the Λ directions.

2.3 Bulk Si and Ge bandstructure

The energy of electrons within a bulk semiconductor crystal may be calculated as a function of wavevector using atomistic methods such as the pseudopotential

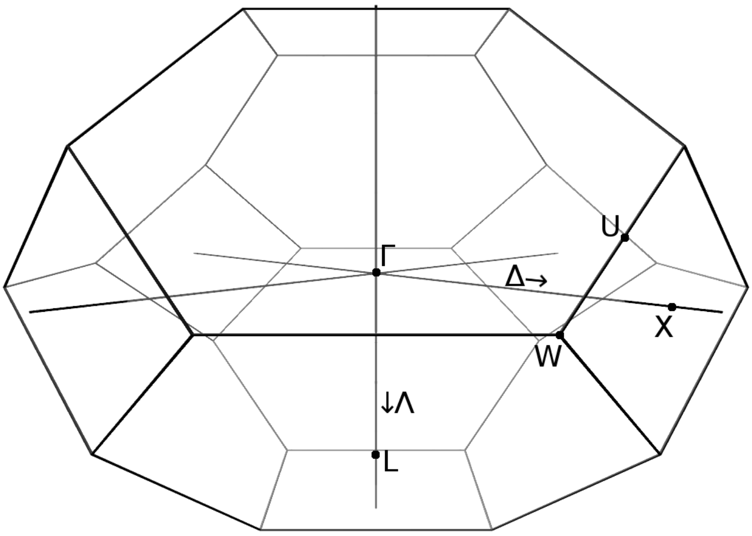


Figure 2.2: Brillouin zone with symmetry points and directions shown.

model[92]. The energy of electrons is plotted along various directions in reciprocal space in fig. 2.3.

The energy bands below the bandgap are referred to as valence bands. They are derived from bonding configurations and are completely filled with electrons at a temperature of 0 K. The bands above the bandgap, referred to as conduction bands, are derived from anti-bonding configurations and are completely empty at 0 K.

In this section, the general structure of the valence and conduction bands is discussed, and a “model solid” (or “virtual crystal”) approximation is used to give a simple yet accurate description of the lowest conduction band energies.

2.3.1 The valence band

Under normal conditions (*e.g.* in doped semiconductors at finite temperatures) the valence bands are highly occupied by electrons. Charge transport is therefore better modelled by considering the relatively small number of unoccupied states or “holes”. The valence band is split into light-hole (LH) and heavy-hole (HH) bands, which are degenerate at Γ. A third, lower energy (SO) band is separated from the LH and

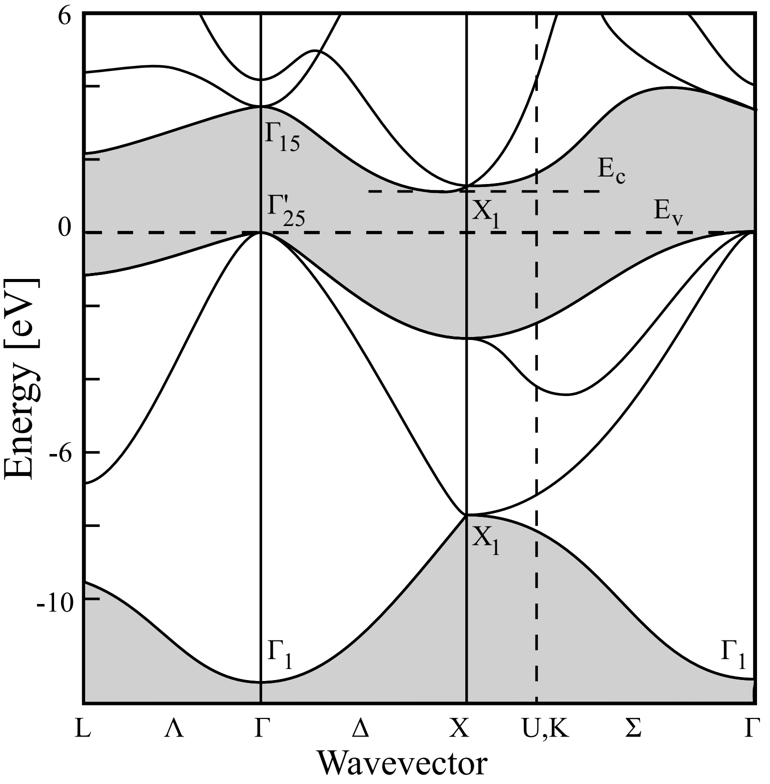


Figure 2.3: Electronic band structure for bulk Si. Energies are shown as a function of wavevector, along each of the significant directions in reciprocal space. Adapted from ref. [91] after local pseudopotential data from ref. [92].

HH bands at Γ by an energy ΔV_{SO} due to spin-orbit splitting. It follows that the valence band maximum is given by

$$E_v^{\mathrm{bulk}} = \overline{E_v} + \frac{1}{3}\Delta_{\mathrm{SO}}, \tag{2.3}$$

where $\overline{E_v}$ is the average of the LH, HH and SO maxima.

Van de Walle showed that the difference between $\overline{E_v}$ in a pair of SiGe alloys is almost independent of strain and crystal orientation, whereas the conduction band minima are highly sensitive to these effects[87]. This quantity is therefore a useful reference for investigating the locations of conduction band minima.

Rieger and Vogl determined an interpolation scheme for the average valence band offset between a film of $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ and a substrate material[93],

$$\Delta\overline{E_v} = (0.47 - 0.06x_s)(x - x_s), \tag{2.4}$$

where x_s is the Ge fraction in the substrate.

2.3.2 The conduction band

SiGe alloys are indirect band gap materials as the lowest conduction band minima are located at different wavevectors from the valence band maximum. The range of energies around a conduction band minimum is referred to as a *valley* and the location of the lowest valleys depends upon the composition of the $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ alloy. For Ge compositions below 85%, the lowest valleys exist in each of the six Δ-directions, about 85% of the way toward the X-points[93]. For Ge compositions above 85%, the lowest valleys are at the eight L-points. Each L-valley lies at the boundary between neighbouring Brillouin zones, so only half of each valley can be attributed to each zone.

The conduction band edge in an unstrained alloy is located by adding the indirect bandgap of the material E_g to the valence band edge,

$$E_c^{\mathrm{bulk}} = E_v^{\mathrm{bulk}} + E_g. \tag{2.5}$$

To a good approximation, the bandgaps for the Δ and L valleys in an unstrained alloy are given (in eV) by[94]

$$\begin{aligned} E_g^{\Delta} &= 1.155 - 0.43x + 0.0206x^2 \\ E_g^{L} &= 2.010 - 1.270x. \end{aligned} \tag{2.6}$$

The model solid approximation has thus been used to calculate the energies of the conduction band minima in an unstrained alloy relative to the average valence band in a substrate material.

2.4 Strain effects

As shown in table 2.1, the lattice constant of Ge is around 4.2% greater than that of Si. When a thin film of $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ is grown upon a thick $\mathrm{Si}_{1-x_s}\mathrm{Ge}_{x_s}$ substrate, the film deforms to match the lattice constant of the substrate along the interface. If $x < x_s$, the film exhibits in-plane tensile strain, whereas if $x > x_s$ the strain is compressive.

This section details the calculation of strain in (001) and (111) orientated films. It is shown that strain has a large effect upon the electronic behaviour of the heterostructure.

2.4.1 Strain tensors

The elements of a strain tensor ε give the expansion (or contraction) of the layer in a given direction relative to its original size in another direction. The indices of the tensor elements refer to orthogonal directions in the *interface* coordinate system $R = (x, y, z)$, where the z-direction is perpendicular to the interface.

As crystal growth is assumed to be isotropic over the xy plane, the in-plane strain is given by[95]

$$\varepsilon_{11} = \varepsilon_{22} = \varepsilon_{\parallel} = \frac{a_s - a}{a}, \tag{2.7}$$

where a is the lattice constant of the unstrained layer and a_s is that of the substrate. It is also assumed that stress on the film does not cause shearing or "kinks" either in-plane or vertically, so all off-diagonal strain tensor elements are zero[96].

Hooke's law relates stresses σ to strains via a rank four elastic constant tensor c, where

$$\sigma_{ij} = \sum_{f,g=1}^{3} c_{ijfg}\varepsilon_{fg}. \tag{2.8}$$

External stress exists only across the plane of the interface, so $\sigma_{33} = 0$ can be substituted into equation 2.8 to obtain the strain perpendicular to the interface

$$\varepsilon_{33} = -\frac{c_{3311} + c_{3322}}{c_{3333}}\varepsilon_{\|}. \tag{2.9}$$

The elastic constants in R are therefore required in this expression.

This introduces a problem, as the experimental values for lattice constants (table 2.1) were obtained in a *crystallographic* coordinate system $R' = (x', y', z')$, in which the principal crystallographic axes are used. The shorthand Voigt notation is used here, in which the first and second pairs of subscripts in c'_{ijfg} are rewritten as single subscripts in a rank two tensor C such that $\{11, 22, 33, 23, 31, 12\} \rightarrow \{1, 2, 3, 4, 5, 6\}$[97].

It can be shown[97] that symmetry in cubic crystals greatly simplifies the C matrix, such that only three independent elements exist

$$\mathsf{C} = \begin{pmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{pmatrix}. \tag{2.10}$$

In this work, SiGe heterostructures are considered in both the (001) and (111) orientations. For the (111) orientation, the crystal lattice is misaligned with the interface and hence $R \neq R'$. A transformation matrix $\mathsf{U} : R \rightarrow R'$ is therefore required[96, 98] to transform between the coordinate systems, such that

$$\mathsf{U} = \begin{pmatrix} \cos\alpha\cos\beta & -\sin\alpha & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta & \cos\alpha & \sin\alpha\sin\beta \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}, \tag{2.11}$$

where the Euler angles (α, β) are defined as rotations clockwise around the z-axis and the new y-axis respectively. For the (111) orientation, the Euler angles are $(\alpha, \beta) = (45°, 54.74°)$. Substituting these values into the transformation matrix gives

$$\mathsf{U}^{(111)} = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ -\sqrt{\frac{2}{3}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}. \tag{2.12}$$

For completeness, it should be noted that $\mathsf{U}^{(001)} = \mathsf{I}_3$, (the third order identity matrix). Due to the trivial nature of this transformation however, it is omitted in the following derivation.

The elastic constants in the interface coordinate system may now be obtained using the relation[96]

$$c_{\alpha\beta\gamma\delta} = \sum_{ijkl} U_{i\alpha} U_{j\beta} U_{k\gamma} U_{l\delta} c'_{ijkl}. \tag{2.13}$$

Substituting these values into equation 2.9 yields the strain perpendicular to the interface

$$\begin{aligned} \varepsilon_{33}^{(001)} &= -\frac{2C_{12}}{C_{11}} \varepsilon_{\|} \\ \varepsilon_{33}^{(111)} &= -\frac{2C_{11} + 4C_{12} - 4C_{44}}{C_{11} + 2C_{12} + 4C_{44}} \varepsilon_{\|}. \end{aligned} \tag{2.14}$$

As with the elastic constants, experimental data for the electronic behaviour of strained SiGe are known only in the crystallographic coordinate system. For (111) oriented systems, it is therefore necessary to transform the strain tensor using the relation[96]

$$\varepsilon'_{\alpha\beta} = \sum_{ij} U_{\alpha i} U_{\beta j} \varepsilon_{ij}. \tag{2.15}$$

The (001) strain tensor is unchanged, as both coordinate systems are identical. However, the (111) strain tensor is transformed to

$$\varepsilon^{(111)'} = \begin{pmatrix} 4C_{44} & C_\alpha & C_\alpha \\ C_\alpha & 4C_{44} & C_\alpha \\ C_\alpha & C_\alpha & 4C_{44} \end{pmatrix} \frac{\varepsilon_\parallel}{C_\beta}, \tag{2.16}$$

where $C_\alpha = -(C_{11} + 2C_{12})$ and $C_\beta = C_{11} + 2C_{12} + 4C_{44}$.

2.4.2 Bandstructure effects

Having determined the strain tensors for thin films, it is now possible to return to the model solid approximation, and calculate the effect on bandstructure. The shift in energy for a given conduction band minimum k is related to strain by a deformation potential tensor Ξ[96] such that

$$\Delta E_c^k = \sum_{ij} \Xi_{ij}^k \varepsilon_{ij}. \tag{2.17}$$

In the Δ and L valleys of cubic crystals, only two deformation potential tensor elements are independent: Ξ_d, which relates to pure dilations and Ξ_u, which relates to pure shears. It can be shown[96] that the total shift in band energy for a valley aligned with the unit vector $\boldsymbol{a}$ is given by

$$\Delta E_c^k = \Xi_d^k \operatorname{Tr} \varepsilon + \Xi_u^k \boldsymbol{a}^{\mathrm{T}} \varepsilon \boldsymbol{a}. \tag{2.18}$$

This expression can be separated into two terms, one of which is common to all valleys and represents the effect of hydrostatic strain[96]

$$\Delta E_c^{k,\mathrm{Hyd}} = \left(\Xi_d^k + \frac{1}{3} \Xi_u^k \right) \operatorname{Tr} \varepsilon, \tag{2.19}$$

where the deformation potentials may be grouped into a single "hydrostatic deformation potential" $a_c^k = \Xi_d^k + \frac{1}{3}\Xi_u^k$[99]. It is important to note that this deformation potential relates the strained conduction band potential to its unstrained value rather than to the well-known valence band potential discussed previously. By using this expression directly, the relative potentials of the L and Δ valleys are therefore left unknown.

A better method therefore requires knowledge of the change in bandgap. This is given by the difference between the shift in conduction and valence band potentials $\Delta E_g^{k,\mathrm{Hyd}} = \Delta E_c^{k,\mathrm{Hyd}} - \Delta E_v^{\mathrm{Hyd}}$. Using a similar approach to that for the conduction band, the valence band shift is given by

$$\Delta E_v^{\mathrm{Hyd}} = a_v \,\mathrm{Tr}\, \varepsilon, \tag{2.20}$$

and hence the deformation potential for the bandgap is $\Xi_d^k + \frac{1}{3}\Xi_u^k - a_v$. Theoretical and experimental values for this deformation potential are given in table 2.1.

The second strain dependent term is specific to each valley and represents the uniaxial strain effects. Due to symmetry, uniaxial strain effects are absent in the Δ valleys in (111) oriented crystals and in L valleys in (001) oriented crystals. They are significant, however, in Δ valleys in the (001) orientation and result in the energy shifts

$$\begin{aligned} \Delta E_c^{\Delta_4,\mathrm{Uni}} &= \frac{1}{3}\Xi_u^{\Delta}(\varepsilon_{11} - \varepsilon_{33}) \\ \Delta E_c^{\Delta_2,\mathrm{Uni}} &= \frac{2}{3}\Xi_u^{\Delta}(\varepsilon_{33} - \varepsilon_{11}), \end{aligned} \tag{2.21}$$

for the valleys with their major axes in the xy-plane and the z-direction respectively. For L valleys in (111) oriented crystals, the energy shifts are given by

$$\begin{aligned} \Delta E_c^{L_1,\mathrm{Uni}} &= 2\Xi_u^{L}\varepsilon_{12} \\ \Delta E_c^{L_3,\mathrm{Uni}} &= -\frac{2}{3}\Xi_u^{L}\varepsilon_{12}, \end{aligned} \tag{2.22}$$

for the valley oriented in the (111) direction and the other three L valleys respectively. The result of the uniaxial strain is therefore that the degeneracy of the valleys is lifted (unless symmetry rules forbid this).

Having found the conduction band minima in a bulk alloy, and the strain dependent energy shift, it is now possible to express the energy of the conduction band minima in a strained film as

$$E_c = E_c^{\mathrm{bulk}} + \Delta E_g^{\mathrm{Hyd}} + \Delta E_c^{\mathrm{Uni}}. \tag{2.23}$$

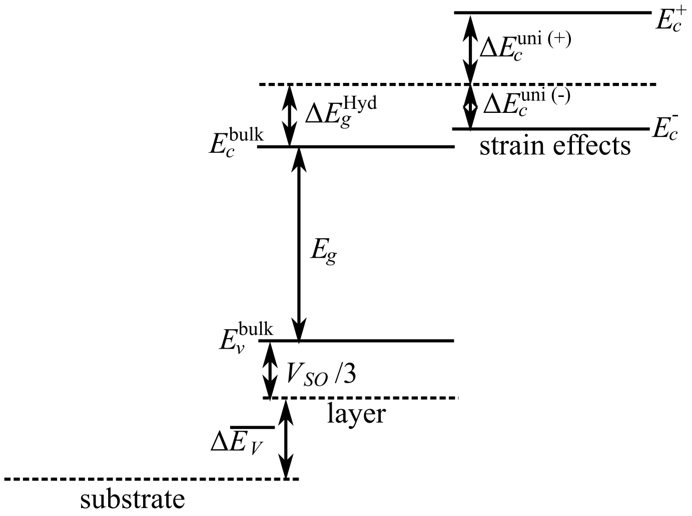


Figure 2.4: Schematic for calculation of conduction band offset in SiGe heterostructures. Adapted from image first published in ref. [3].

Fig. 2.4 summarises this calculation. First, the average valence band energy is determined relative to that of the substrate as a reference. The highest valence band is located one-third of the spin-orbit splitting above this energy in an unstrained alloy. The bandgap is added to this to find the conduction band minima in the bulk material. The hydrostatic strain effect on the bandgap is added to shift all conduction band valleys identically. Finally, uniaxial strain effects are added (if applicable) to lift the degeneracy between valleys.

2.5 Usable energy range

The previous section showed how the conduction band offset between a pair of alloys may be determined and our attention may now turn to quantum wells constructed from a "sandwich" of two different alloys. In this section, Si/Ge/Si and Ge/Si/Ge QWs in the (001) and (111) orientations are investigated as possible components for QCL design.

For this investigation, a "usable energy range" for QCL design is defined as

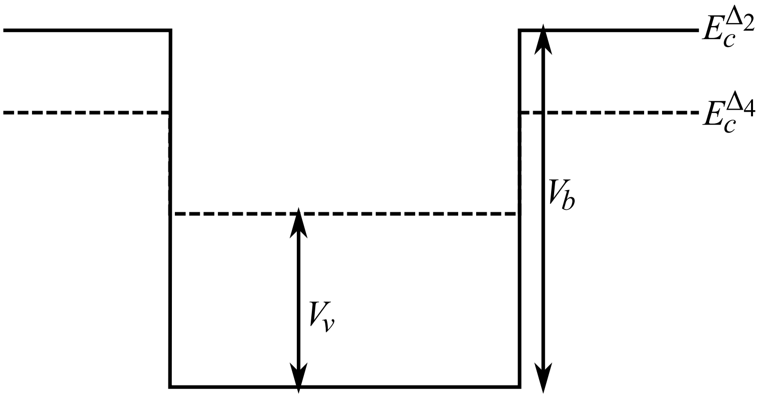


Figure 2.5: Schematic representation of the usable range of energies in a (001) oriented Si/SiGe QW (not to scale). The depth of the well is denoted V_b, while the separation between conduction band valleys is denoted V_v.

the lowest range of energies containing bound states from a single set of conduction band valleys. This definition of "usable" contains two criteria which are illustrated in fig. 2.5 and may be explained as follows.

Firstly, bound states exist only within the range of energies V_b inside quantum wells. This is given by the conduction band offset (for a given valley set) between the barrier and the well materials. The lowest valley set contains most of the conduction band electron density and is therefore considered the best candidate for QCL design in the present work.

Secondly, optical emission is an intravalley process. To reduce competition from nonradiative intervalley scattering processes (see chapter 5), the lowest conduction band valley set must be separated by a large energy V_v from the second lowest set.

The usable range of energies is therefore determined by the more restrictive of the two constraints

$$V_{\text{range}} = \min(V_b, V_v). \tag{2.24}$$

However, the emission of photons and injection and extraction of electrons must all take place within this range of energies in QCLs. Experiments have shown that this limits the maximum photon energy $\hbar\omega_{\text{max}}$ to around half the confining potential[100],

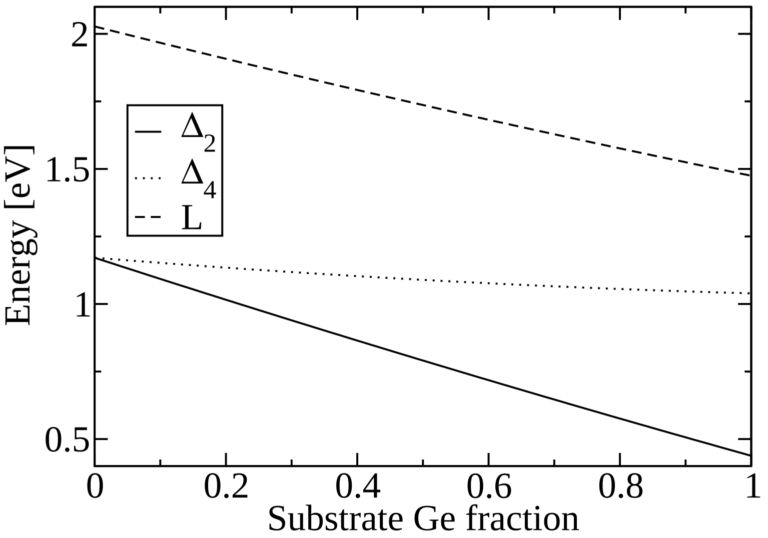


Figure 2.6: Energies of conduction band valleys in a (001) Si film as a function of substrate Ge fraction.

giving $\hbar\omega_{\mathrm{max}} \approx \min(V_b, V_v)/2$.

2.5.1 (001) orientation

The conduction band minima for (001) oriented Si and Ge films upon substrates of varying composition are shown in fig. 2.6 and fig. 2.7 respectively.

In Ge/Si/Ge QWs, electrons are confined within the central Si film. Fig. 2.6 shows that the Δ_2 valleys are the lowest energy set in Si films. The confining potential is given by the difference in Δ_2 potential in the Ge and Si films

$$V_b = E_{\mathrm{Ge}}^{\Delta_2} - E_{\mathrm{Si}}^{\Delta_2} \approx 0.66 - 0.07x_s. \tag{2.25}$$

The Δ_4 valleys are second lowest in energy in the Si film, and their separation from the Δ_2 valleys is given by

$$V_v = E_{\mathrm{Si}}^{\Delta_4} - E_{\mathrm{Si}}^{\Delta_2} \approx 0.6x_s. \tag{2.26}$$

The separation between the valleys is the dominant constraint for (001) oriented

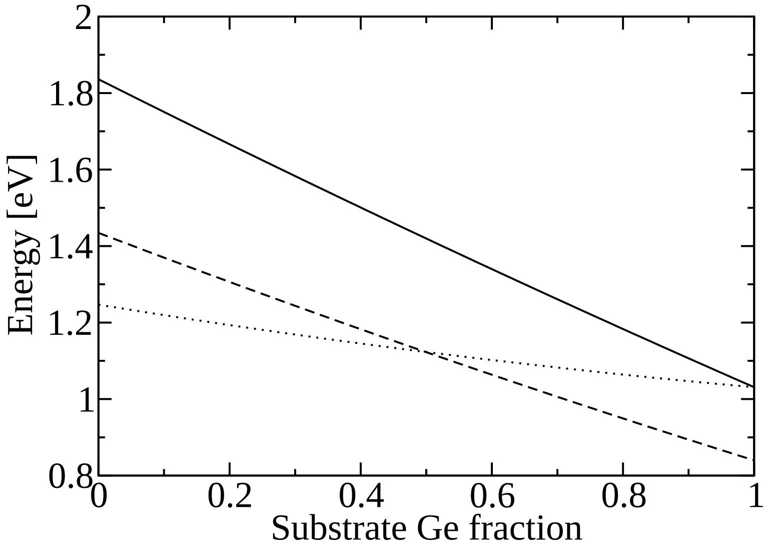


Figure 2.7: Energies of conduction band valleys in a (001) Ge film as a function of substrate Ge fraction. Legend is identical to that in fig. 2.6

Ge/Si/Ge QWs when the substrate Ge fraction is lower than 50%. In chapter 6, this is shown to be a realistic range for mechanically stable QCLs with Si well regions.

In the Ge films within (001) oriented Si/Ge/Si QWs, the L valleys are lowest in energy if the substrate Ge fraction is greater than 50%. Again, this is a realistic range for mechanically stable QCLs with Ge well regions. The confining potential is given by the difference in L valley potential between the Si and Ge layers

$$V_b = E_{\mathrm{Si}}^{L} - E_{\mathrm{Ge}}^{L} \approx 0.593 + 0.042 x_s. \tag{2.27}$$

The Δ_4 valleys are second lowest in energy in the Ge film, and their separation from the L valleys is

$$V_v = E_{\mathrm{Ge}}^{\Delta_4} - E_{\mathrm{Ge}}^{L} \approx 0.379 x_s - 0.189. \tag{2.28}$$

For realistic substrate alloys ($x_s > 50\%$) the valley separation provides the dominant constraint.

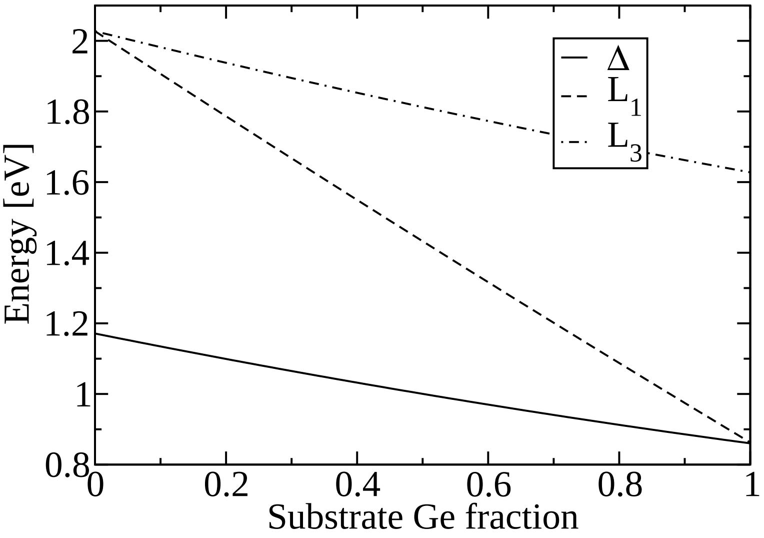


Figure 2.8: Energies of conduction band valleys in a (111) Si film as a function of substrate Ge fraction.

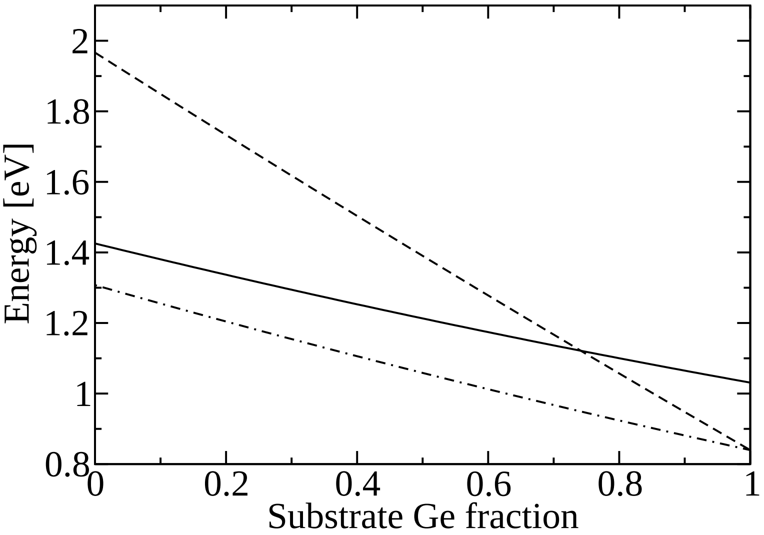


Figure 2.9: Energies of conduction band valleys in a (111) Ge film as a function of substrate Ge fraction

2.5.2 (111) orientation

The conduction band minima for (111) oriented Si and Ge films are shown in fig. 2.8 and fig. 2.9 respectively.

In (111) oriented Ge/Si/Ge QWs, the Δ valleys remain degenerate and are lowest in energy in the Si film. The confining potential is given by

$$V_b = E_{\mathrm{Ge}}^{\Delta} - E_{\mathrm{Si}}^{\Delta} \approx 0.25 - 0.17x_s. \tag{2.29}$$

The L_1 valley is second lowest in energy in the Si film and its separation from the Δ valleys is

$$V_v = E_{\mathrm{Si}}^{L_1} - E_{\mathrm{Si}}^{\Delta} \approx 0.85 - 0.85x_s. \tag{2.30}$$

The confining potential provides the dominant constraint in this case.

Finally, in (111) oriented Si/Ge/Si wells, the L_3 valleys are lowest in energy and the band offset is given by

$$V_b = E_{\mathrm{Si}}^{L_3} - E_{\mathrm{Ge}}^{L_3} \approx 0.72 + 0.068x_s. \tag{2.31}$$

The L_1 valley is second lowest in energy in the Ge film (for realistic substrate compositions of $x_s > 75\%$) and its separation from the L_3 valleys is

$$V_v = E_{\mathrm{Ge}}^{L_1} - E_{\mathrm{Si}}^{L_3} \approx 0.66 - 0.66x_s. \tag{2.32}$$

The dominant constraint is therefore due to the splitting between L valleys.

2.6 Conclusion

The energies of the conduction band valleys have been found in Si and Ge films, accounting for strain and crystal orientation effects. The bottom of the conduction band lies in the Δ valleys in Si films, and in the L valleys in Ge. To avoid intervalley scattering, the usable energies in a QW are limited to those between the lowest two sets of conduction band valleys.

Configuration	Usable energy range [eV]
(001) Ge/Si/Ge	$0.6x_s$
(001) Si/Ge/Si	$0.379x_s - 0.189$
(111) Ge/Si/Ge	$0.25 - 0.17x_s$
(111) Si/Ge/Si	$0.66 - 0.66x_s$

Table 2.2: Maximum usable energy range within Si and Ge QWs in the (001) and (111) orientations.

The range of usable energies for QCL design in each of the QW configurations considered in this chapter is summarised in table 2.2. In each case, the usable energy range depends upon the substrate Ge fraction, due to the effect of strain upon the bandstructure.

In structures with a Si well region, a low substrate Ge fraction ($xs \sim 10\%$) is required for mechanical stability (see chapter 6). For (001) oriented structures, a 10% Ge substrate gives an energy range of only 60 meV and a maximum emission energy of $\hbar\omega_{\mathrm{max}}$=30 meV. The usable energy range is much larger in the (111) orientation and $\hbar\omega_{\mathrm{max}}$=117 meV.

In structures with a Ge well region, a high substrate Ge fraction ($xs \sim 90\%$) is required. For (001) oriented structures, a 90% Ge fraction gives an energy range of 152 meV and $\hbar\omega_{\mathrm{max}}$=75 meV. The energy range is much smaller in the (111) orientation, where $\hbar\omega_{\mathrm{max}}$=33 meV.

It is important to note that these results give the *maximum* achievable band offset, which correspond to Si wells with Ge barriers or *vice versa.* The growth of such structures is likely to prove challenging as discussed in chapter 6, and in practice a SiGe/Si/SiGe or GeSi/Ge/GeSi configuration is preferable. The conduction band offset in such structures will be reduced correspondingly. In summary, the most promising results were predicted for (111) Si wells, with $\hbar\omega_{\mathrm{max}}$=117 meV (29 THz, 10 μm) or for (100) Ge wells, with $\hbar\omega_{\mathrm{max}}$=75 meV (19 THz, 16.5 μm). This indicates that light emission throughout the mid-infrared and terahertz frequency range should be possible with SiGe based QCLs.

Chapter 3

Heterostructures

The spatially varying potential in a QCL leads to quantum confinement of electrons, meaning that they may only occupy certain permitted states. The electrons are (nearly) free to move parallel to the interfaces in QCLs and it is shown in this chapter that the permitted states are grouped into sets of energy subbands.

The total energy of the system is given by the *Hamiltonian*, which contains the potential energy of the crystal, the kinetic energy of the electron and the effect of electric fields. The internal electric fields due to uneven charge distribution are included via a self–consistent solution of the Poisson and Schrödinger equations.

3.1 Time-independent Schrödinger equation

The time-independent Schrödinger equation may be written:

$$\hat{H}\psi_n(\boldsymbol{R}) = E_{n,\boldsymbol{k}}\psi_n(\boldsymbol{R}) \tag{3.1}$$

where $\psi_n(\boldsymbol{R})$ is the wavefunction in three dimensions and $E_{n,\boldsymbol{k}}$ is the total energy of the n^{th} quantised state at a given wavevector. $\hat{H}$ is the *Hamiltonian operator*, which gives the total energy of the system and takes the form,

$$\hat{H} = \hat{T} + V_{\text{crystal}} + V_F. \tag{3.2}$$

$\hat{T}$ is the kinetic energy of an electron, V_{crystal} is the potential due to the crystal lattice and V_F describes the effect of external and internal electric fields. The terms in the Hamiltonian are discussed in more detail in the following sections.

The Hamiltonian above describes all the *major* effects upon the electron. Subtler effects, such as interactions with charged particles, phonons and structural imperfections have a much weaker effect on the electron than the other terms[101]. They are therefore omitted from solutions of Schrödinger's equation and are treated as scattering mechanisms in chapter 5 instead.

3.1.1 Effective mass approximation

The kinetic energy of a nonrelativistic particle in free space is given by

$$E_k(\boldsymbol{k}) = \frac{\hbar^2 k^2}{2m}, \tag{3.3}$$

where $\boldsymbol{k}$ is its wavevector and m is its rest mass. This gives a parabolic dispersion relation between wavevector and energy.

The dispersion in the Si or Ge conduction band is quite complex, with multiple minima around the Δ and L valleys, whose relative energies are strongly affected by strain. However, atomistic simulations have shown that the dispersion is approximately parabolic near to each conduction band minimum[97]. An *effective mass* may be introduced into eqn. 3.3 to describe the dispersion in a particular direction

$$m_i^* = \hbar^2 \left[\frac{\partial^2 E_k(k_i)}{\partial k_i^2}\right]^{-1}, \tag{3.4}$$

where k_i is the component of the wavevector in the direction $\hat{\boldsymbol{r}}_{\boldsymbol{i}}$ relative to the valley minimum.

Due to its even symmetry, the polynomial expansion of the kinetic energy contains only even terms, yielding

$$E_k(k_i) \approx a_2 k_i^2 + a_4 k_i^4 + a_6 k_i^6 \ldots, \tag{3.5}$$

where a_m are constant coefficients. For small wavevectors, only the first term is significant and the dispersion relation is analogous to that of a free electron. The

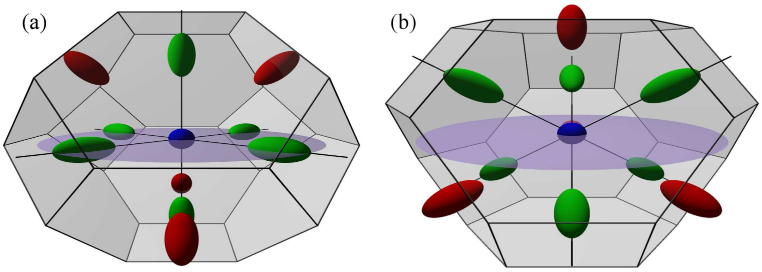


Figure 3.1: Constant energy surfaces (not to scale) for conduction band valleys within Brillouin zone. Δ valleys are shown in green, L in red and Γ in blue. The surfaces are plotted on axes within the crystallographic coordinate system. The purple disc is aligned with the (001) plane in figure (a) and with the (111) plane in figure (b).

energy is

$$E_k(k_i) \approx \frac{\hbar^2 k_i^2}{2m_i}. \tag{3.6}$$

Thus, a different effective mass describes the dispersion in each direction.

In describing the three-dimensional dispersion relations for Si and Ge conduction band valleys, it is helpful to define a valley coordinate system $P = (x^p, y^p, z^p)$, in which x^p is aligned with the major axis of the valley[102]. Equipotential surfaces of the Δ and L valleys are prolate spheroids[103] described as

$$E_k = \frac{\hbar^2}{2}\left[\frac{(k_x^p)^2}{m_l} + \frac{(k_y^p)^2 + (k_z^p)^2}{m_t}\right], \tag{3.7}$$

where m_l and m_t are the longitudinal and transverse effective masses respectively and wavevectors are expressed relative to the valley minima. The equipotential surfaces for each valley are shown in fig. 3.1.

The effective masses in Si and Ge have been found using experimental techniques such as cyclotron resonance[103] and theoretical methods such as empirical pseudopotential modelling[93]. Their values are stated in table 3.1.

Using a similar method to that described in section 2.4.1, Rahman *et al* [102] showed that the effective mass may be translated to the interface coordinate system, giving m_x and m_y for dispersion parallel to the interfaces and m_z perpendicular to

Material	Valley	m_l	m_t
Si	Δ	0.91[1]	0.19[2]
	L	1.7[3]	0.12[3]
Ge	Δ	0.95[3]	0.2[3]
	L	1.64[4]	0.08[2]

Table 3.1: Effective masses for conduction band valleys in Si and Ge, expressed as a multiple of the rest mass of a free electron.

Material system	Valley	m_q	m_d
(001) Si	Δ_4	m_t=0.19	$\sqrt{m_l m_t}$=0.42
	Δ_2	m_l=0.916	$m_t = 0.19$
(111) Si	Δ	$\frac{3m_l m_t}{2m_l+m_t}$=0.26	$\sqrt{\frac{m_t(2m_l+m_t)}{3}}$=0.36
(001) Ge	L	$\frac{3m_l m_t}{2m_l+m_t}$=0.12	$\sqrt{\frac{m_t(2m_l+m_t)}{3}}$=0.30
(111) Ge	L_1	m_l=1.64	m_t=0.08
	L_3	$\frac{9m_l m_t}{8m_l+m_t}$=0.089	$\frac{\sqrt{m_t(8m_l+m_t)}}{3}$=0.34

Table 3.2: Quantisation and two-dimensional density-of-states effective masses of conduction band valleys in (001) and (111) Si and Ge, using expressions derived from Ref [102]. Masses are expressed relative to the rest mass of a free electron.

the interfaces.

The next section describes how the solution to the Schrödinger equation may be split into two problems: the quantisation of electrons in the z direction, and the nearly-free dispersion in the x, y plane. It is therefore apparent that different effective masses are required for the two problems: the quantisation effective mass

---

[1] Reference. [104]

[2] Reference. [103]

[3] Reference. [93]

[4] Reference. [105]

$m_q = m_z$ for the former and the density-of-states effective mass m_d for the latter. The density-of-states effective mass is a directional average of the in-plane masses, $m_d = \sqrt{m_x m_y}$[102, 106, 107]. The values for these parameters are summarised in table 3.2.

3.1.2 Two-dimensional approximation

The potential in a heterostructure varies significantly in the direction perpendicular to the interfaces, due to the the bias potential and the conduction band offset between materials. Ideally, the interfaces in a QCL are perfectly flat, giving a constant potential over the x–y plane.[1] The electrons are therefore effectively free in two dimensions.

As such, the wavefunction may be decoupled into a plane-wave component parallel to the interfaces and a bound component $\psi(z)$ perpendicular to the interfaces. This yields the expression

$$\psi(\boldsymbol{R}) = \frac{1}{\sqrt{A}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \psi(z) u_{\boldsymbol{k}}(\boldsymbol{R}), \tag{3.8}$$

where $\boldsymbol{k}$ is the in-plane wavevector, $\boldsymbol{r}$ is the in-plane position, $u_{\boldsymbol{k}}(\boldsymbol{R})$ is a Bloch function and A is the cross-sectional area of the system.

The Bloch function varies on an atomic scale whereas the bound component varies much more slowly, on the scale of the heterolayers. The Bloch function may be omitted from the wavefunction, leaving a slowly varying *envelope function*,

$$\psi(\boldsymbol{R}) = \frac{1}{\sqrt{A}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \psi(z). \tag{3.9}$$

It can be shown that this gives an accurate mesoscopic description of structures with slowly modulated conduction band potentials[108]. In QCLs, there are abrupt changes in conduction band potential at each heterojunction, which reduces the accuracy of the approximation. However, good agreement exists with atomistic

---

[1] In reality, small fluctuations in the interface location exist and these are shown to cause scattering of electrons in chapter 5.

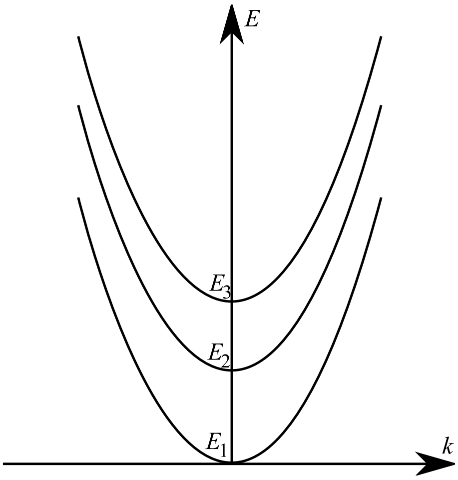


Figure 3.2: Schematic representation of energy of states in a two-dimensional system as a function of wavevector parallel to interfaces. The relative spacing between subband minima depends upon the geometry of the system, and is not shown to scale.

simulations of subband energies in superlattices with layer thicknesses comparable to those in QCLs[97]. Additionally, the effective mass/envelope function approximation is much more computationally efficient than atomistic methods, making it more suitable as a design tool for QCLs.

Substituting eqn. 3.9 into eqn. 3.1 allows the plane-wave components to be eliminated, leaving

$$\hat{H}\psi(z) = E_n\psi(z), \tag{3.10}$$

where the total energy of a state is given by

$$E_{n,\boldsymbol{k}} = E_n + \frac{\hbar^2}{2m_d}\left(k_x^2 + k_y^2\right). \tag{3.11}$$

The total energy of a state (fig. 3.2) is therefore the sum of a discrete component E_n, due to quantum confinement in the growth direction, and a continuum of energies due to momentum parallel to the interfaces, $\frac{\hbar^2 k^2}{2m_d}$. The conduction band in such a system is therefore discretised into a set of subbands and electrons may occupy any subband whose minimum lies below the total electron energy. The one-dimensional kinetic energy operator in the Hamiltonian for the one-dimensional Schrödinger equation (eqn. 3.10) may be expressed as

$$\hat{T} = -\frac{\hbar^2}{2m_q}\frac{\mathrm{d}^2}{\mathrm{d}z^2} \tag{3.12}$$

3.1.3 Boundary conditions

The present work is concerned with finding bound states in QCLs, rather than arbitrary conduction band potentials. The following discussion therefore focuses on the boundary conditions required for solving this specific set of problems.

Strictly, there are no truly bound states in a QCL when an electric field is applied, as electrons at any energy can ultimately tunnel out of the structure. However, only a finite number of discrete *quasi-bound* states have high probability densities within the QCL. In the range of realistic electric fields for a QCL ($F < 100\,\mathrm{kV/cm}$), the rate at which quasi-bound electrons tunnel out of the device is assumed to be very low, and in this work, they are treated as being fully bound.

To solve Schrödinger's equation, it is necessary to find a suitable pair of boundary conditions. However, this is a non-trivial task due to the complicated geometry of the system. Internal and external electric fields break the periodicity of the potential, making periodic boundary conditions unrealistic. "Hard-wall" conditions at the edges of a period are also unrealistic as a significant proportion of the wavefunction may extend beyond a single period.

A superior, although more computationally expensive solution requires three periods of the structure, with box boundary conditions at the edges.

$$\psi_i(0) = \psi_i(3L_p) = 0, \tag{3.13}$$

where L_p is the length of a period of the structure. This is still not adequate, as the solutions should be invariant through a translation to a neighbouring period. The three period solution does not provide this translational invariance as the box boundary conditions affect an electron in the outer periods more than one in the central period.

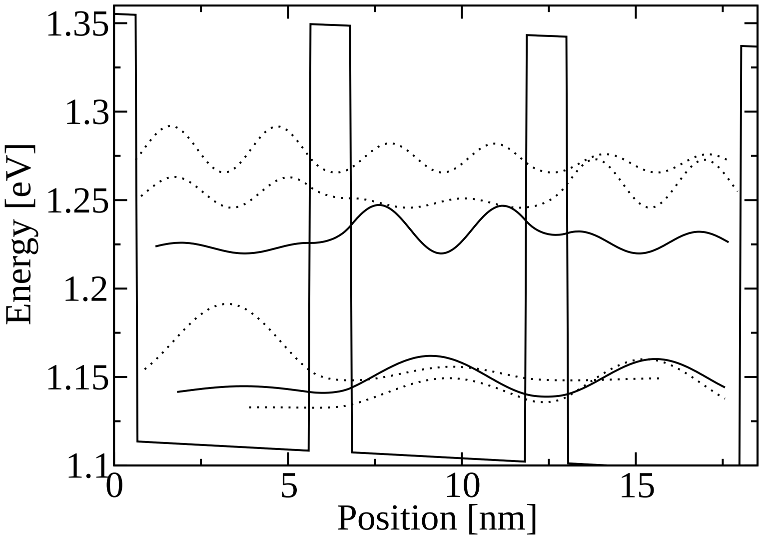


Figure 3.3: Δ-valley subbands determined for a (111) oriented structure with 5 nm Si wells and 1.2 nm Ge barriers on a 15% Ge substrate, using box boundary conditions over three periods. An external electric field of 10 kV/cm was applied. The central states (solid lines) appear slightly different from those in the outer periods (dotted lines).

This is illustrated in Fig. 3.3, which shows Δ valley states in a sample (111) oriented structure. The states in the outer periods appear slightly different from the central states due to the boundary conditions.

The central solutions are assumed to be most accurate as the hard boundary conditions are "padded" by the outer periods of the structure, allowing the wavefunction amplitude to decay in a more realistic manner. Translational invariance may be achieved via a three step process:

1. solve Schrödinger's equation for three periods of the structure;
2. reject any solutions which are not localised in the central period;
3. translate the remaining (central) solutions to the outer periods using the relations $\psi_{i+n}(z) = \psi_i(z - L_p)$ and $E_{i+n} = E_i - L_pF$.

Several methods of identifying central states were tested in this work. Firstly, the *median point* z_m of a wavefunction can be determined as

$$\int_0^{z_m} |\psi_1(z)|^2 \, \mathrm{d}z = \frac{1}{2}, \tag{3.14}$$

which represents the point at which an electron has identical probability of being found on either side. If this point lies within the central period, the state is assumed to be centrally localised.

Secondly, the *modal period* p_m is defined as the period which contains the largest part of the electron probability density, *i.e.*

$$p_m = \max_{i=0}^{2} \left[\int_{iL_p}^{(i+1)L_p} |\psi_1(z)|^2 \, \mathrm{d}z \right], \tag{3.15}$$

where z_i is the centre point of period p.

Finally, an *overlap method* may be used. This method begins by calculating the subbands in a single period of the QCL with box boundary conditions. It is then assumed that each state ϕ_i in the single period corresponds to a similar state ψ_m in the centre of a three period structure. The overlap σ_{im} between these states is therefore larger than that between state ϕ_i and any other state from the three period structure ψ_j such that

$$\sigma_{im} = \max_j \left[\int_0^{L_p} \phi_i(z)\psi_j(z - z_L)\,\mathrm{d}z \right]. \tag{3.16}$$

While none of the methods is infallible, the overlap method has proved most successful for weakly biased symmetric structures, and the average coordinate methods are more suited to heavily biased asymmetric structures.

3.1.4 Finite difference solution

Equation 3.10 may be rewritten as

$$-\frac{\hbar^2}{2m_q}\frac{\mathrm{d}^2\psi(z)}{\mathrm{d}z^2} + V(z)\psi(z) = E\psi(z), \tag{3.17}$$

where $V(z)$ is the total potential at each point. To solve the problem numerically, a set of N regularly spaced sampling points, $z_0, z_1, \ldots, z_{N-1}$ may be defined for the system. The spacing between each point is $\delta z = z_{i+1} - z_i$ and the value of the wavefunction at each point may be expressed in the shorthand notation $\psi_i = \psi(z_i)$.

The finite difference approximation for a second derivative is

$$\psi_i'' \approx \frac{\psi_{i+1} - 2\psi_i + \psi_{i-1}}{\delta z^2}, \tag{3.18}$$

and Schrödinger's equation may be rewritten in the discretised form

$$-\frac{\hbar^2}{2m_q}\left[\frac{\psi_{i+1} - 2\psi_i + \psi_{i-1}}{\delta z^2}\right] + V_i\psi_i = E\psi_i. \tag{3.19}$$

This can be expressed as a first-order difference equation of the form

$$a\psi_{i-1} + b_i\psi_i + a\psi_{i+1} = E\psi_i, \tag{3.20}$$

where the coefficients are defined as

$$\begin{aligned} a &= -\frac{\hbar^2}{2m_q\delta z^2} \\ b_i &= \frac{\hbar^2}{m_q\delta z^2} + V_i. \end{aligned} \tag{3.21}$$

Box boundary conditions may be imposed by stating that $\psi_{-1} = \psi_N = 0$, where z_{-1} and z_N are the points just beyond the start and end of the QCL respectively. This gives boundary conditions of

$$\begin{aligned} b_0\psi_0 + a\psi_1 &= E\psi_0 \\ a\psi_{N-2} + b_{N-1}\psi_{N-1} &= E\psi_{N-1}. \end{aligned} \tag{3.22}$$

The difference equation can be written for each sampling point in the system and packed into a matrix,

$$\begin{array}{ccccc} \begin{pmatrix} b_0 & a & 0 & \cdots & 0 \\ a & b_1 & a & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & a & b_{N-2} & a \\ 0 & \cdots & 0 & a & b_{N-1} \end{pmatrix} & \begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_{N-2} \\ \psi_{N-1} \end{pmatrix} & = & E & \begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_{N-2} \\ \psi_{N-1} \end{pmatrix} \\ \mathsf{H} & \boldsymbol{\psi} & = & E & \boldsymbol{\psi} \end{array} \tag{3.23}$$

The *Hamiltonian matrix operator*, H in equation 3.23 is symmetric and tridiagonal, which allows efficient linear algebra routines to be used[2]. The eigenvalues, E_n of H give the energies of the subband minima and the corresponding eigenvectors $\psi_n(z)$ give the components of the wavefunction in the growth direction for each subband.

Since the Hamiltonian matrix is Hermitian, all of its eigenvalues (subband energies) are real and the eigenvectors (wavefunctions) form an orthonormal set such that

$$\psi_i^{\mathrm{T}} \cdot \psi_j = \delta_{ij}, \tag{3.24}$$

where δ_{ij} is the Kronecker delta function. This is equivalent to the numerical solution of the overlap integral of a pair of normalised wavefunctions, which is given by

$$\int \psi_i(z)\psi_j(z)\mathrm{d}z = \delta_{ij}. \tag{3.25}$$

---

[2] The LAPACK3 library for FORTRAN 77 and the GNU Scientific Library for C/C++ were used in this work

3.2 Electric field effects

When driven by a constant voltage source, a potential V_F exists across a period of the QCL in addition to the conduction band potential. As well as experiencing a net change in potential across the period, electrons are attracted toward highly doped regions and repelled from regions of high electron density. This causes internal variations in the electric field and Poisson's equation must be used to find the true form of V_F. Poisson's equation in one dimension is given by

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[\varepsilon(z)\frac{\mathrm{d}V_F(z)}{\mathrm{d}z}\right] = -\rho(z), \tag{3.26}$$

where ρ is the charge density. If the potential at the start of a period is chosen as a reference, the boundary conditions may be specified as

$$\begin{aligned} V_F(0) &= 0 \\ V_F(L_p) &= -FL_p, \end{aligned} \tag{3.27}$$

where F is interpreted as an *average* electric field.

Substituting $D(z) = -\varepsilon(z)\frac{\mathrm{d}}{\mathrm{d}z}V_F(z)$ and integrating both sides of equation 3.26 gives

$$D(z) = Q(z) + D(0), \tag{3.28}$$

where the total charge density from the start of the structure up to z is given by

$$Q(z) = \int_0^z \rho(z')\mathrm{d}z'. \tag{3.29}$$

It follows that

$$V_F(z) = -\int_0^z \frac{\mathrm{d}z'}{\varepsilon(z')}\left[Q(z') + D(0)\right], \tag{3.30}$$

where the boundary condition $V_F(0) = 0$ has been applied.

Substituting in the second boundary condition gives

$$D(0) = -\frac{FL_p + \int_0^{L_p}\frac{\mathrm{d}z' Q(z')}{\varepsilon(z')}}{\int_0^{L_p}\frac{\mathrm{d}z'}{\varepsilon(z')}}, \tag{3.31}$$

which can then be used in equation 3.30 to find the potential profile across the device.

3.2.1 Charge distribution

If the bound states in the system and their populations are known, the spatial charge distribution may be found.[3] In a system under thermal equilibrium, the probability of a state being filled is given by Fermi-Dirac statistics

$$f_{FD}(E,T) = \frac{1}{\exp\left[\frac{E-E_F(T)}{k_BT}\right]+1}, \tag{3.32}$$

where $E_F(T)$ is the *Fermi energy* for an electron with temperature T and k_B is Boltzmann's constant. At the Fermi energy there is a 50% chance of a state being occupied and the state occupation decreases as energy increases.

In a QCL there is a potential difference across the structure and electrons flow between its contacts. The Fermi energy therefore varies along the device. It is shown in chapter 5 that intrasubband scattering rates are much greater than intersubband scattering rates. It may therefore be assumed that electrons quickly settle to a quasi-equilibrium within each subband. A modified form of the Fermi-Dirac distribution now describes the occupation probability of states within a single subband i

$$f_{FD,i}(E_k,T_{e,i}) = \frac{1}{\exp\left[\frac{E_k-E_{F,i}(T_{e,i})}{k_BT_{e,i}}\right]+1}. \tag{3.33}$$

The Fermi energy has been replaced here by a set of quasi-Fermi energies, which sets the total population for each subband. An effective electron temperature T_e describes the spread of electrons within each subband. Although this electron temperature may be different in each subband, good agreement with experimental data is achieved for III-V terahertz and mid-infrared QCLs if it is treated as a constant[109, 110]. Finally, for simplicity all energy levels are expressed relative to the subband minimum in this work.

The density of states in a two-dimensional system is given by[97]

$$\rho^{2D} = \frac{m_d}{\pi\hbar^2}, \tag{3.34}$$

---

[3] A method for finding the populations is described in chapters 5 and 7.

and the total number of electrons in subband i is therefore

$$n_i = \frac{m_d}{\pi\hbar^2}\int_0^\infty \frac{\mathrm{d}E_k}{\exp\left[\frac{E_k - E_{F,i}(T_e)}{k_B T_e}\right] + 1}. \tag{3.35}$$

Solution of this integral gives

$$n_i = \frac{m_d k_B T_e}{\pi\hbar^2}\left\{\frac{E_{F,i}(T_e)}{k_B T_e} + \ln\left[1 + \mathrm{e}^{-\frac{E_{F,i}(T_e)}{k_B T_e}}\right]\right\}. \tag{3.36}$$

For a given subband population and electron temperature, the quasi-Fermi energy may be determined by numerical solution of this expression.

The density of electrons in a given subband at a particular spatial point is

$$\rho_{e,i}(z) = n_i|\psi_i(z)|^2. \tag{3.37}$$

Assuming that the entire electron population is contained in a single valley set, the total density of electrons at a given point is found by summing this expression over all subbands:

$$\rho_e(z) = n_{\mathrm{val}}\sum_i \rho_{e,i}(z), \tag{3.38}$$

where n_{val} is the number of equivalent valleys.

For a modulation doped system with a doping profile $d(z)$, the total number of donor atoms in a period of the system is

$$N_{2D} = \int_0^{L_p} d(z)\mathrm{d}z. \tag{3.39}$$

Assuming that all donor atoms are ionised, charge neutrality requires that the number of donor ions equals the total number of electrons in all subbands

$$\int_0^{L_p} \rho_e(z)\mathrm{d}z = N_{2D}. \tag{3.40}$$

At a given point, the charge density $\rho(z)$ is given by

$$\rho(z) = e\left[d(z) - \rho_e(z)\right], \tag{3.41}$$

where e is the fundamental charge constant.

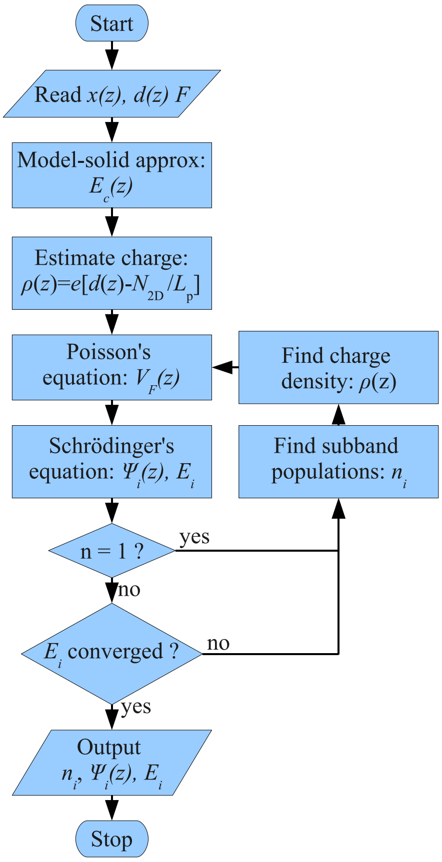


Figure 3.4: Flowchart for solution of the coupled Poisson and Schrödinger equations. "$n = 1$" denotes the first iteration of the self-consistent Poisson–Schrödinger calculation.

3.3 Self-consistent solution

It has been shown that Poisson's equation may be used to find the effect of a spatial charge distribution upon the potential profile in a QCL. Also, Schrödinger's equation can be used to find the permitted states for a given potential profile, and the charge distribution can then be determined using a rate equations approach (chapter 7).

The Poisson and Schrödinger equations are therefore coupled, and neither can be solved using information that is available *ab initio*. The nonlinear nature of the problem also prevents a simple solution through linear algebra. Instead, a set of initial conditions must be provided and an iterative process used to obtain a self-consistent solution to the two equations.

As a first approximation, it is assumed here that electrons are evenly distributed through the device such that $\rho_e(z) \approx N_{2D}/L_p$, giving a charge distribution of

$$\rho(z) = e[d(z) - N_{2D}/L_p]. \tag{3.42}$$

An estimate of the field effect upon the Hamiltonian may be found by substituting this into Poisson's equation. The resulting potential may then be used in Schrödinger's equation to determine the electronic bandstructure.

In turn, the scattering rates may be found (chapter 5) and rate equations may be solved (chapter 7) to find a better estimate of the charge distribution across the period. The process may be repeated until the subband populations converge. This approach is in good agreement with previous self-consistent models[111, 112] and is illustrated in fig. 3.4.

3.4 Doping profiles

A non-uniform charge distribution across a device has been shown to yield a spatially varying internal electric field. It follows that the doping profile within a QCL strongly affects its electronic behaviour. QCLs have been demonstrated with sheet doping densities as high as $N_{2D} = 1.0 \times 10^{12}\,\mathrm{cm}^{-2}$[113], and with a variety of differ-

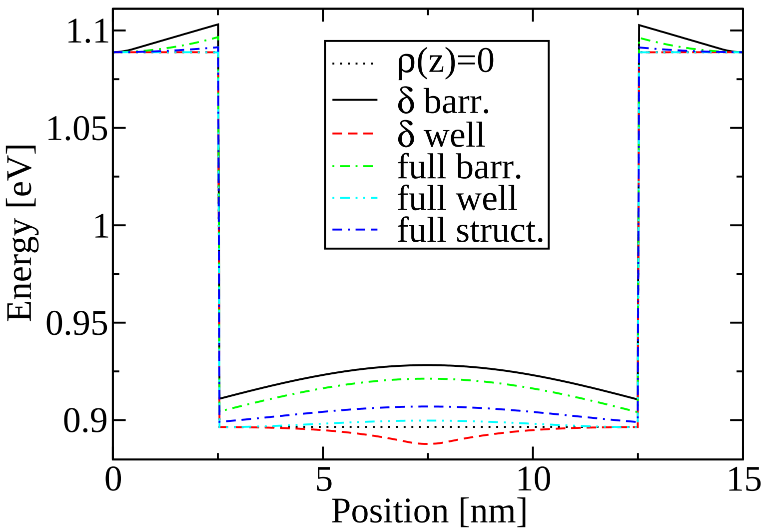


Figure 3.5: Total confining potential in a (001) oriented 10 nm Ge quantum well with 5 nm $Ge_{0.7}Si_{0.3}$ barriers. The dotted line corresponds to an uncharged system. The other lines represent a sheet doping density of $1\times10^{12}\,\mathrm{cm}^{-2}$ per period. Results are shown for δ doping (*i.e.* all dopants in a thin layer) in the middle of the well or barrier, and for dopants spread evenly over the entire well region, the entire barriers or the entire structure.

ent doping profiles. Useful insight into doping effects may be gained by investigating the internal electric fields in a quantum well.

The system modelled for this investigation was a 10 nm layer of Ge sandwiched between a pair of 2.5 nm, 70% Ge barriers in the (001) orientation. A substrate Ge fraction of 89% was found to provide mechanical stability. At low temperatures in an equilibrium system, it is assumed here that all carriers lie in the lowest L-valley subband. Fig. 3.5 shows the total confining potential for a sheet doping density of $1\times10^{12}\,\mathrm{cm}^{-2}$ per period using a variety of doping profiles.

In the absence of space-charge effects, the energy spacing between the lowest pair of subbands was found to be 60.3 meV. Fig. 3.6 shows however that this value varies significantly with doping. In the figures, δ-doping refers to the confinement of all donor ions to a very thin layer. This approach has been used both within the

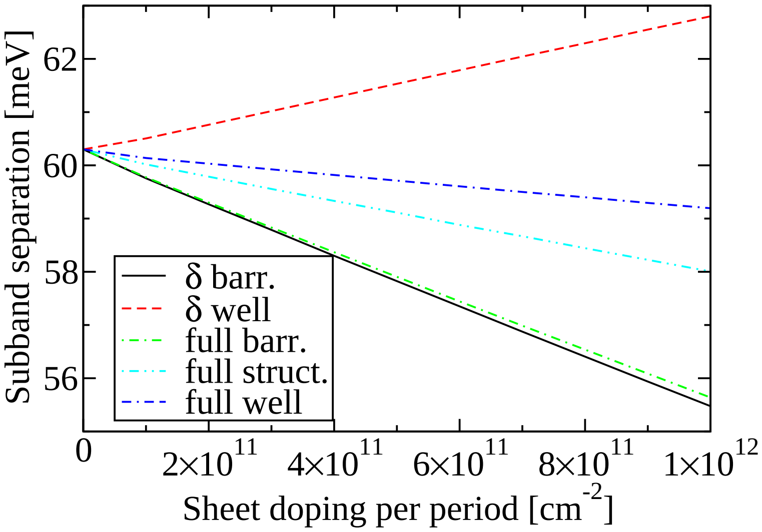


Figure 3.6: Energy separation between lowest pair of subbands as a function of sheet doping density for the quantum well in fig. 3.5. Each line corresponds to a different doping profile, as described in the caption of fig. 3.5.

barriers[114] and wells[115] of a QCL. Fig. 3.5 shows that attraction of electrons toward a 1 nm thick, highly doped region (*i.e.* δ doping) in the centre of the well causes a significant localised dip in the confining potential. When δ doping is applied to the centre of the barrier region, this causes a significant decrease (up to 4.8 meV) in subband separation, while in the well there is a smaller increase of up to 2.5 meV. The weaker effect in the well is due to the peak of the electron probability density coinciding with the δ-doped region.

QCLs have also been demonstrated with dopants spread evenly through a single layer[49, 116] or through several layers [27, 30]. Fig. 3.5 shows that in all cases this reduces the effect of doping upon the subband separation. The effect of spreading the dopants across the full width of the barrier is shown to be little different from that of δ-doping within the barrier. This occurs because the bulk of the effect upon subband energies is due to the large build-up of negative charge in the centre of the well. Conversely, fig. 3.6 shows that doping the well region leads to a greatly reduced effect upon subband energies. This is due to the areas of positive and negative charge coinciding.

3.5 Conclusion

In this chapter, an effective mass/envelope function description of a SiGe-based heterostructure was developed. It was shown that the effective masses are anisotropic and vary significantly between each of the valleys. The states within quantum confined systems were shown to be approximately two-dimensional in nature, and to lie within a discrete set of continuous subbands.

Electric fields due to an external voltage source and the spatial variation of charge density was shown to significantly alter the confining potential, and a self-consistent solution of the Poisson and Schrödinger equations was developed. Finally, it has been shown that the choice of doping profiles within the structure is important. In particular, δ doping leads to noticeable local extrema in the confining potential and has a strong effect on the separation between pairs of subbands, while spreading dopants over a larger area (particularly in the wells) reduces this effect.

The conclusion regarding doping profiles should be taken with some caution. Although it is tempting to reduce the effect on subband separation by doping the well regions, it is important to note that ionised impurity scattering (discussed in chapter 5) will be maximised for states localised in these wells. For this reason, QCLs tend to avoid doping in the active region, where nonradiative scattering effects are undesirable.

Chapter 4

Valley splitting

In the preceding chapters, the electron energy dispersion relation for wavevectors perpendicular to the growth plane was assumed to be a single parabola. This allows an effective mass to be defined, and a simple set of solutions to Schrödinger's equation followed. It has been shown, however, that multiple valleys exist in the SiGe conduction band and in (001) oriented Si, this leads to there being *two* parabolic valleys in the dispersion relation at $k_{\|} = 0$.

The single valley EMA therefore cannot provide a completely adequate description of the quantised states in (001) Si QWs. The single valley solutions can, however, be used as a basis for determining the double valley result. In this chapter, it is shown that each single-valley solution is split into a doublet when two valleys exist. This phenomenon is known as valley splitting.

In this chapter, two approaches are used to investigate the effect. A double-valley effective mass approximation (DVEMA) allows a quick and efficient solution for symmetric systems, while an empirical pseudopotential method (EPM) provides a more general atomistic description of systems. A reasonably good agreement between the two methods is found, with the former being much faster computationally. Although the DVEMA is defined for symmetric potentials, it is shown to remain approximately correct for slightly asymmetric structures such as a double quantum well, allowing qualitative conclusions to be drawn about valley splitting in QCLs.

The splitting is shown to be an oscillatory function of well width, almost independent of in-plane wave vector, and an increasing function of the magnitude of interface gradient. Intersubband optical transitions are investigated under both approximations and it is shown that in most cases valley splitting causes linewidth broadening, although under extreme conditions, transition line doublets may result.

4.1 Review of previous investigations

Valley splitting has been observed experimentally in Shubnikov-de Haas oscillation measurements in high magnetic fields[117–124] with energy splitting up to a few meV. Boykin *et al* presented a tight-binding model of the ground state splitting in a biased square quantum well with both hard-wall and cyclic boundary conditions [125–127]. The ground state splitting in an unbiased square well was found to be approximately

$$\Delta E_1 \approx \frac{16\pi^2 u}{(S+2)^3} \sin\left(\frac{\phi_{\min}}{2}\right) \left|\sin\left[(S+2)\frac{\phi_{\min}}{2}\right]\right|, \tag{4.1}$$

where $\phi_{\min} = k_0 a$, and k_0 denotes the position of the valley minimum in the Brillouin zone, a is the lattice constant, S is the number of atomic monolayers in the quantum well and u is a fitting constant. The model shows that the ground state splitting oscillates with well width; the frequency being dependent on the location of the valley minima. Similar results have been obtained for the two lowest subbands in an unbiased well by Chiang [128] in an anti-bonding orbital model, and by Nestoklon [129] in a slightly different tight-binding model.

Valley splitting in an electric field has been considered by modelling a triangular QW. Although an effective mass model by Sham [130] proposed that the splitting is simply proportional to the applied field, Boykin *et al* [131] and Grosso *et al* [132] show that the splitting is a non-linear function of both the well width and the electric field.

4.2 Double valley effective mass approximation

Modifications to the EMA have allowed valley splitting to be described in an infinite square well [133–135] and a finite square well with impurity states [136]. Ting and Chang's Double Valley Effective Mass Approximation (DVEMA) [137] provides an elegant self-contained description of valley splitting in symmetric systems. A *valley splitting potential* $\hat{U}(z)$ is added to the Hamiltonian, giving

$$\hat{H}(z) = -\frac{\hbar^2}{2m_q}\frac{\mathrm{d}^2}{\mathrm{d}z^2} + \hat{V}(z) \pm \hat{U}(z). \tag{4.2}$$

There are two equivalent Δ_2-valleys at $k_{\parallel} = 0$, centred around the wavevectors $k_{\perp} \pm k_0$ in (001) Si. The wavefunction must therefore be composed of basis functions of the form

$$|k\rangle = a^{+}|k + k_0\rangle + a^{-}|k - k_0\rangle. \tag{4.3}$$

where $|a^{+}| = |a^{-}| = 1/\sqrt{2}$[138]. In a symmetric potential, the basis functions must be either even or odd symmetric combinations of the single valley states, giving[137]

$$|k_{\pm}\rangle = \frac{1}{\sqrt{2}}\left(|k + k_0\rangle \pm |k - k_0\rangle\right). \tag{4.4}$$

The complete wavefunction is obtained by summing over all basis states:

$$|\psi_{\pm}\rangle = \sum_k \phi_{\pm}(k)\,|k_{\pm}\rangle. \tag{4.5}$$

The matrix elements of the envelope potential operator in the Hamiltonian are written in the basis defined by Eq. (4.4) as

$$V_{nm} = \langle k_n|\hat{V}|k_m\rangle = \frac{1}{2}\left(\langle k_n + k_0| \pm \langle k_n - k_0|\right)\hat{V}\left(|k_m + k_0\rangle \pm |k_m - k_0\rangle\right). \tag{4.6}$$

Rearranging this expression yields

$$\begin{aligned} V_{nm} &= \frac{1}{2}\left(\langle k_n + k_0|\,V\,|k_m + k_0\rangle + \langle k_n - k_0|\,V\,|k_m - k_0\rangle\right) \\ &\pm \frac{1}{2}\left(\langle k_n + k_0|\,V\,|k_m - k_0\rangle + \langle k_n - k_0|\,V\,|k_m + k_0\rangle\right). \end{aligned} \tag{4.7}$$

Table 4.1: Pseudopotential parameters[139].

Parameter	a_1	a_2	a_3	a_4	a_5	a_6
Si	212.1372	2.2278	0.6060	-1.9720	5.0	0.3
Ge	108.9024	2.3592	0.7400	-0.3800	5.0	0.3

Using the discrete to continuous approximation, $V_{nm} = \tilde{V}(k_m - k_n)$, the intervalley envelope term can be written as

$$\tilde{V}_{\pm}(k) = \tilde{V}\left(k\right) \pm \frac{1}{2}\left[\tilde{V}\left(k-2k_0\right) + \tilde{V}\left(k+2k_0\right)\right], \tag{4.8}$$

where $\tilde{V}(k)$ is the Fourier transform of the conduction band-edge envelope potential. The real-space form of the intervalley envelope function is found by taking the inverse Fourier transform of this result and the splitting potential is therefore extracted as

$$\begin{aligned} \hat{U}(z) &= \frac{1}{2}\mathcal{F}^{-1}\left\{\tilde{V}\left(k-2k_0\right) + \tilde{V}\left(k+2k_0\right)\right\} \\ &= V(z)\cos(2k_0 z), \end{aligned} \tag{4.9}$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform.

4.2.1 Empirical pseudopotential calculation

An empirical pseudopotential method (EPM) is used here to calculate electronic states in Si/SiGe based QWs, and provide a comparison with the DVEMA. As the EPM (in common with tight-binding) is an atomistic method, valley splitting is automatically included, and no special modifications are required. The supercell implementation of the EPM was used, with a continuous atomic formfunction, $V(g)$. The "modified Falicov" formfunction described by Friedel *et al* [139] was selected:

$$V(g) = \frac{a_1\left(g^2 - a_2\right)}{1 + \mathrm{e}^{\left[a_3\left(g^2 - a_4\right)\right]}} \cdot \frac{1}{2}\left[\tanh\left(\frac{a_5 - g^2}{a_6}\right) + 1\right]. \tag{4.10}$$

Fischetti and Laux[140] and Ikonić[141] have shown that this formfunction gives reasonable agreement with experimental data for both bulk Si and Ge band structure and for band discontinuities at the interface. A cut-off energy of 4.5 Ry was used, which gives an acceptable number of plane waves for accurate and rapid computation with all the structures considered. The parameters for Si and Ge are given in Table 4.1, and the virtual crystal approximation was used for the alloy.

The EPM can be used for structures with either abrupt interfaces or graded compositions. In the latter case, the interface grading is piecewise constant (*i.e.* with a minimum length equal to that of a single diatomic unit cell). Individual layers are given the required alloy compositions. In contrast, the DVEMA uses a continuous potential profile, though features smaller than the width of a unit cell have little practical meaning.

It is important to note that effective mass based calculations (like the DVEMA) can never fully reproduce the results of microscopic EPM modelling. This is because the DVEMA only handles four bulk states explicitly, while the EPM implicitly includes many evanescent states, stemming from remote bulk bands. Furthermore, the location of indirect valleys may vary between the bulk materials used in the well and barrier—a situation which is difficult to handle with effective mass methods. Given that remote bands are usually less important than the bands from which the quantised states are derived, one can expect *reasonable* accuracy from the DVEMA. This may be validated by comparison against the EPM calculation.

4.3 Numerical results and discussion

DVEMA and EPM calculations were performed for a range of Si/SiGe QWs. The Ge fraction in the substrate was fixed at 20%. In the EPM calculations the total length of the structure (*i.e.* the supercell period) which includes the well and barrier layers, was set to a fixed value of 35 diatomic unit cells (henceforth denoted as 35 ML). This maintained a constant number of plane-waves in the pseudopotential basis set, and avoided fluctuations in the results caused by variable size of basis. This is important since valley splitting is relatively small on the energy scale covered by EPM [104].

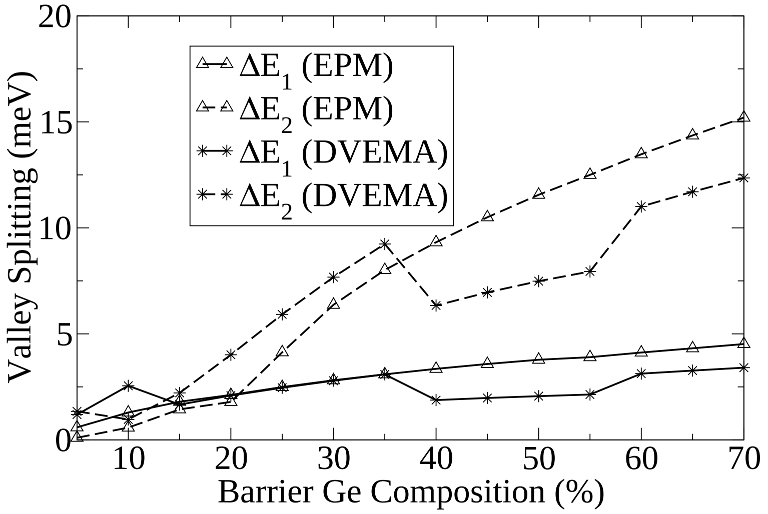


Figure 4.1: Valley splitting, ΔE in the lowest two subbands of an 8 ML quantum well as a function of barrier composition.

4.3.1 Finite square well

The first set of calculations was for a simple square QW with abrupt interfaces. Fig. 4.1 shows the influence of barrier composition (potential height) on the splitting of the lowest two subbands, obtained by both the DVEMA and EPM calculations, for a fixed, 8 ML wide Si QW. The confining potential increases almost linearly with the Ge content in the barriers [93]. The results show that valley splitting increases with confining potential. Both models are broadly in agreement, with the most obvious discrepancy being the discontinuities in the DVEMA plot.

The effect of well width upon valley splitting was investigated next, for a structure with a fixed barrier composition of 50% Ge. The well width was varied between 1 ML and 25 ML in a supercell of total length 35 ML. This leaves a minimum 10 ML barrier region, which ensures that neighbouring QWs are decoupled under periodic boundary conditions. It also represents a realistic range of well widths for QCLs. Figure 4.2 shows the EPM and DVEMA results. As predicted by equation 4.1, the valley splitting is a decaying oscillatory function of well width, originating from interference of the basis components reflecting at the QW interfaces. The DVEMA

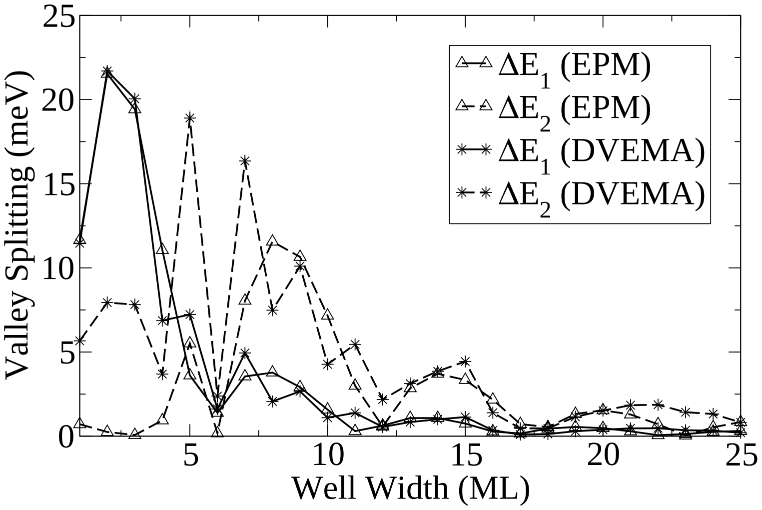


Figure 4.2: Valley splitting in a Si/Si$_{0.5}$Ge$_{0.5}$ QW as a function of well width

shows good agreement with the EPM results for the envelope of the splitting, while the oscillatory component is approximately correct. For very small well widths however, the results for higher subbands deviate from the theory, as the upper subband is unbound.

Boykin's[125] tight-binding model considers wide QWs, with a low-energy ground state and effectively infinite barriers. The decaying oscillatory form is, however, the same as that obtained from DVEMA calculations. Equation 4.1, extracted from the tight-binding model, predicts a period of around 6 ML as observed in fig. 4.2. Setting $u \approx 3$ yields a match in the amplitude between the three models. This figure is somewhat higher than the value given in the reference above, although the DVEMA results apply to loosely bound states in a finite square well as opposed to being derived from bulk dispersion characteristics.

4.3.2 Influence of in-plane wavevector

As the in-plane wavevector increases, states draw an increasing amount of their character from higher conduction bands, giving rise to nonparabolicity in the dispersion relation. The nonparabolicity results in $k_{\|}$ dependent valley splitting. The EPM

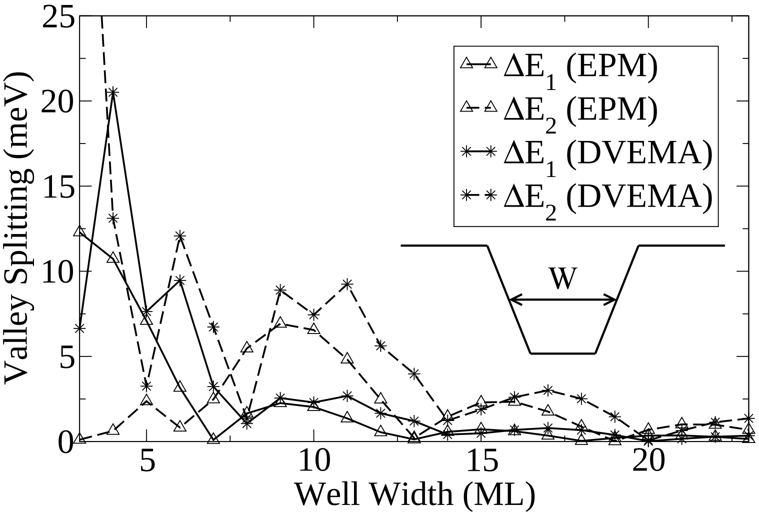


Figure 4.3: Valley splitting in lowest two subbands as a function of well width in a QW with 3-step linear graded interfaces. The inset shows the general structure of a linear graded QW.

calculation automatically accounts for this, but the assumption has been made in the EMA that the conduction band edge is parabolic.

In the EPM calculation for a 10 ML QW, the splitting in the first and second subbands increased approximately linearly by 17% and 8% respectively, when $k_{\|}$ changed from zero to 10% of the Brillouin zone edge (i.e. in the range with non-negligible electron occupancy at any reasonable temperature). This implies that $k_{\|}$-dependent corrections in the DVEMA are not mandatory.

4.3.3 Graded barrier potential

In real Si/SiGe QWs, surface segregation effects are well documented [142]. This refers to the "preference" of Ge atoms to exist on the surface of the material rather than in the bulk during molecular beam epitaxial (MBE) growth, leading to a decrease in the magnitude of the Ge composition gradient at the nominal interfaces. It is therefore unrealistic to model a Si/SiGe QW as having abrupt interfaces. The effect of graded interfaces on subband splitting is therefore considered. The linear-

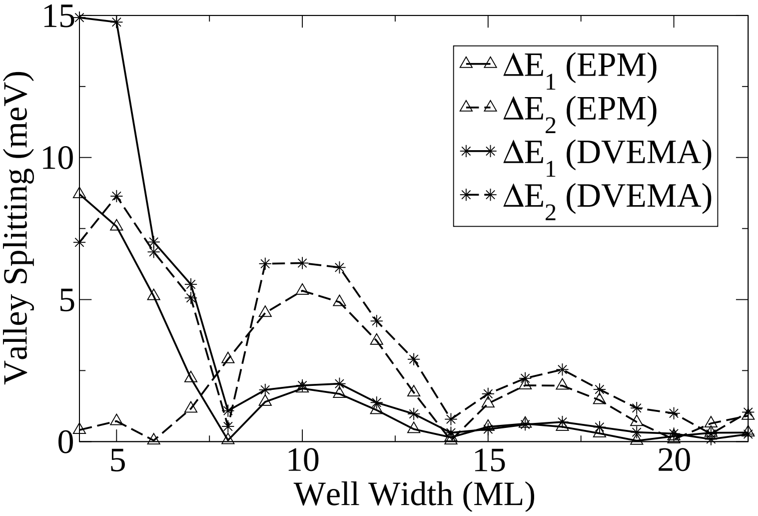


Figure 4.4: Valley splitting in lowest two subbands as a function of well width in a quantum well with 4-step linear graded interfaces.

graded structure shown in the inset of Fig. 4.3 is modelled first.

Within the EPM calculation, the linear-graded interfaces on either side of the QW are modelled as 3-step piecewise-linear, *i.e.* the interfaces spread across three ML, with Ge content of 17%, 33% and 50% sequentially. The results are shown in figure 4.3. The well width is defined as the full-width at half-maximum (FWHM) of the envelope potential. The results of the DVEMA and EPM are in good agreement for larger well widths—those which allow for more than a single bound state. The plots show that the oscillatory component of the valley splitting is unchanged, although the envelope decreases in magnitude. This is because graded interfaces have reduced large-wave-vector Fourier components in the envelope potential, which mix the two $\Delta_\perp$ valleys and hence the splitting is generally smaller.

As the width of the graded interfaces increases, the splitting is further reduced, as shown in Fig. 4.4 for a 4-step graded interface with Ge content of 13%, 25%, 38% and 50%, sequentially. Again, there is a very good agreement between the two models.

A linear graded interface is a somewhat idealised model as experimental evidence

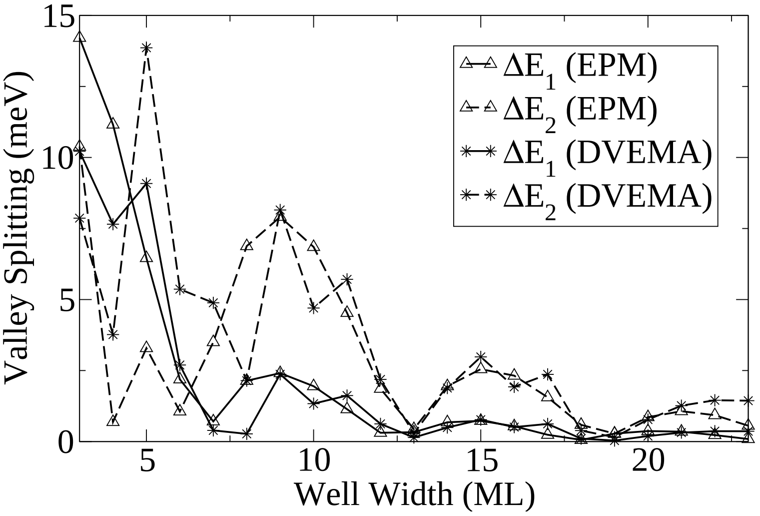


Figure 4.5: Valley splitting of the lowest two subbands as a function of well width in a quantum well with 3-step non-linear graded interfaces.

shows that the interface profile is closer to an error function. A 3-step grading with Ge concentrations of 13%, 38% and 50% is therefore used as an approximation to a typical interface composition. The results of EPM and DVEMA calculation are shown in Fig. 4.5.

The magnitude of the splitting is somewhat larger than for the case of linear grading, apparently because the potential gradient at the interface is now larger over a wide range of energies, thus corresponding to a steeper linear-grading at the energies of the first and second subband minima. The DVEMA results are again in close agreement with the EPM results.

4.3.4 Double quantum well

Next, the valley splitting in a double QW structure is considered. This relatively simple structure may provide sufficient design freedom for an optically pumped intersubband laser. It also allows the DVEMA to be tested for asymmetric structures where the simple cosine modulated splitting potential is no longer strictly applicable. The simulated structure, shown in the inset of Fig. 4.6, has a fixed 1 ML well

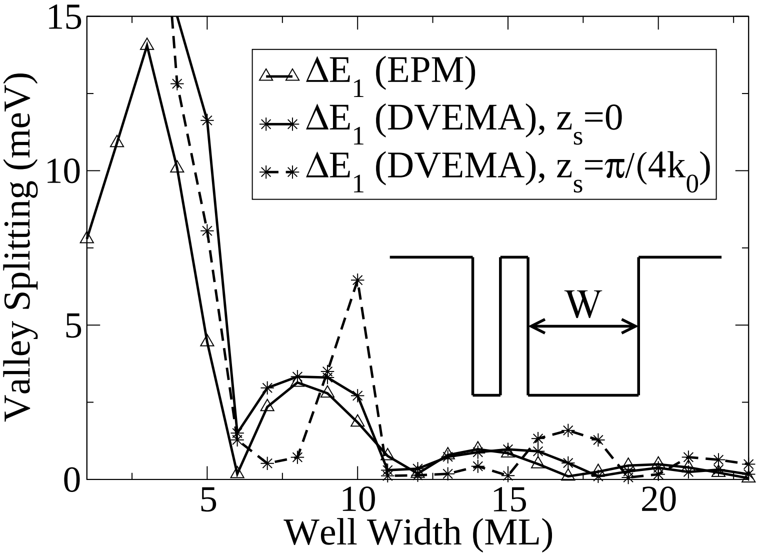


Figure 4.6: Valley splitting in lowest subband as a function of second well width in the double QW structure (inset). The results are shown for EPM and DVEMA. In the case of the DVEMA, results are shown for two different origins for the symmetric approximation.

separated from the second well by a 1 ML, 50% Ge barrier. All other parameters are unchanged.

The results for the EPM and DVEMA calculations are shown in Fig. 4.6. In this case, the structure is assumed to deviate only slightly from the square well, and therefore the periodic structure may be considered approximately symmetric about the $z = 0$ position (*i.e.* the left-hand side of the structure shown in the inset of Fig. 4.6). The axis of symmetry, z_s is therefore set at this point. As the structure only contains a relatively small perturbation from a symmetric quantum well, the DVEMA and EPM results are still in good agreement. The splitting energy is again lower than the simple square well case, since the left hand side (with a thin well and a thin barrier) can be viewed as a "soft", nonabrupt interface. Fig. 4.6 also shows the DVEMA results when the axis of symmetry is shifted to $z_s = \frac{\pi}{4k_0}$ such that the splitting potential becomes sine-modulated as opposed to cosine-modulated. This represents worst-case selection of the axis of symmetry, if $z_s = 0$ is assumed to be the best. The oscillatory component of the valley splitting now appears out of phase with the EPM results, although the envelope of the oscillations is approximately correct. The symmetric approximation is therefore dependent on the origin of the coordinate system. However, a good estimate of the magnitude of the valley splitting is possible, even with a poor choice of origin.

4.3.5 Electric field effects

In the previous section, it was shown that qualitatively correct information may be drawn for slightly asymmetric structures. In a QCL, however, the applied electric field makes it impossible to define an axis of pseudo-symmetry. The EPM may still be used however, to examine how the valley splitting varies under these conditions.

Fig 4.7 shows that in general the splitting increases with electric field. At high fields, the confinement is weakened on one side of the QW. Splitting becomes a linear function of the electric field and is only weakly dependent on well width, in agreement with Sham[130]. At low electric fields, the well width has a significant

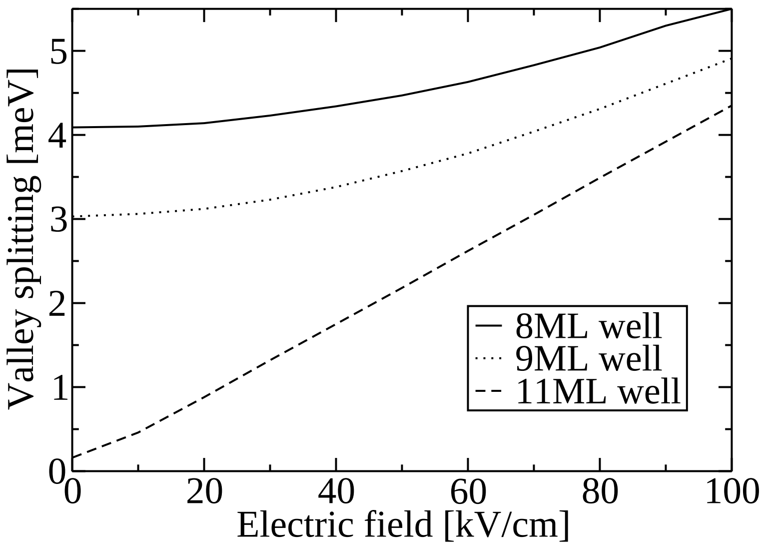


Figure 4.7: Valley splitting in the lowest subband of 8, 9 and 11 ML wide QWs as a function of electric field.

effect as predicted by Friesen.[138]

4.3.6 Intersubband optical transitions

Dipole matrix elements, $D_{21} = \langle 2|z|1\rangle$ were calculated for intersubband transitions in the square well (section 4.3.1). The results from the EPM and DVEMA simulations are shown in figure 4.8 along with the separation of the transition energies. The difference between the dipole matrix elements is small and approaches zero as the transition energies converge. This implies a similar magnitude of spectral contribution from each pair of valley-split states. The two methods are in close agreement for lower well widths, with the DVEMA predicting larger matrix elements at higher widths. In most cases, when considering valley splitting of states, the permitted optical transitions are from the upper “excited state” to the upper “ground state” and between the two lower states. However, when close to the splitting minima (at well widths of 17 ML and 23 ML), the converse situation sometimes applies with the EPM (fig. 4.9). The DVEMA always finds transitions to be of upper→upper

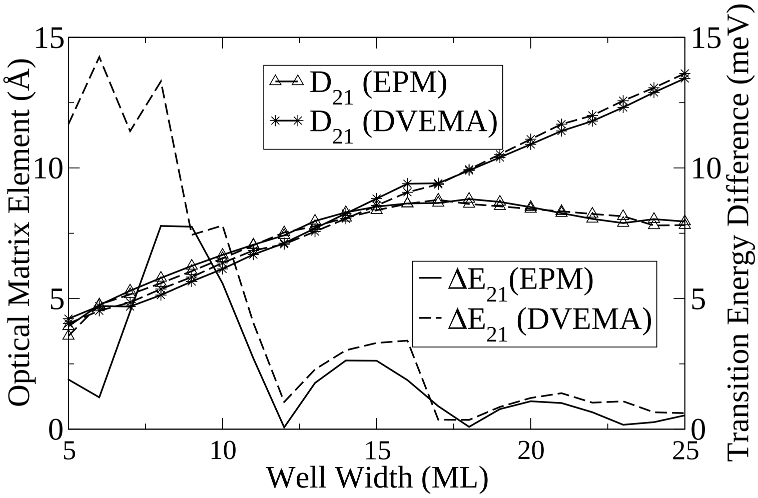


Figure 4.8: Dipole matrix elements as a function of well width for the permitted transitions between the first and second subbands in the finite square QW considered in section 4.3.1. The difference between the two permitted transition energies is also displayed.

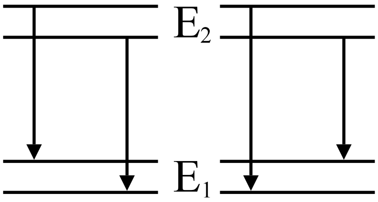


Figure 4.9: Permitted optical transitions in a square QW are usually between the two upper or lower valley-split states (left). When close to splitting minima however, this situation is sometimes reversed in EPM simulation (right).

and lower→lower character. Transitions exhibit linewidth broadening by interface roughness and carrier scattering, typically of the order 5–10 meV. In the majority of cases, valley splitting is relatively small and will only cause an increase in linewidth broadening by the amount shown in figure 4.8. However, when the valley splitting is large (for example at around 8 ML well width), a transition line doublet may become apparent.

4.4 Conclusion

The DVEMA method presented by Ting and Chang [137] has been extended to model intervalley-mixing in any symmetric structure. DVEMA and EPM methods have been used to calculate Δ-valley subband splitting in a range of symmetric and asymmetric Si/SiGe heterostructures, with both abrupt and graded interfaces.

The results of the two methods are in good agreement with each other and with published tight-binding results for symmetric structures, with the DVEMA demanding less than 0.5% of the computational run-time of EPM. Valley splitting of up to around 10 meV was predicted for abrupt-interface square QWs in the range of well widths of interest for Si intersubband devices [73, 143–147]. The valley splitting was shown to be a decaying, oscillatory function of well width at small electric fields. In large fields, the valley splitting was shown to be linearly related to the field and only weakly dependent on well width.

Unfortunately, the EPM is too cumbersome a tool for QCL simulation due to the great length of the structure, and the DVEMA cannot accurately model QCLs due to their asymmetry. Qualitative predictions of the effect of valley splitting upon QCLs are possible however.

It has been shown that valley splitting is likely to cause spectral linewidth broadening, although transition line doublets may form for small well widths, with both valley-split states contributing equally to the spectrum. The effect of surface segregation was modelled by considering both linear and non-linear composition grading at the interfaces. This was found to reduce the valley splitting, as it is dependent upon the potential gradient at the interfaces.

Chapter 5

Scattering mechanisms

In order to calculate the population of each state in a QCL, it is necessary to consider the transfer of electrons between states. Electrons may scatter by interacting with phonons, other electrons, or structural irregularities in the QCL. This chapter describes a general method for calculating scattering rates and applies it to each of the specific mechanisms.

5.1 Coherent and incoherent transport

Charge transport processes may be described as either coherent or incoherent. Incoherent processes transport electrons, one at a time, by independent scattering events. The time between scattering events is much longer than the time taken for the events themselves, *i.e.* the events may be considered effectively instantaneous.

Coherent effects such as quantum tunnelling through a barrier occur when a population of electrons is in phase coherence. In such processes, wave packets evolve smoothly over time between their initial and final states. The wave packet oscillates back and forth across the barrier at the *Rabi frequency* and the oscillations are damped by *dephasing* processes[148], leading to a steady state solution after a *dephasing time*.

Fermi's golden rule (described in the next section) allows a computationally efficient calculation of incoherent scattering rates but more detailed quantum theoretical methods are required to describe coherent effects. Nonequilibrium Green's function (NEGF)[149] approaches have shown promising results but the computational burden is too large for NEGF to be incorporated into a QCL design tool. As a result, a compromise must be found between the computational runtime and the comprehensiveness of the transport model.

Density matrix calculations offer a slightly less computationally expensive approach to coherent transport phenomena, but complete density matrix models are still far too cumbersome for use as a QCL design tool on standard computer hardware. Density matrix calculations, therefore, typically only consider charge transport across the injection barrier of a QCL and model the remainder of the device using either a thermal charge distribution[53] or a heavily simplified incoherent transport model[148].

In the present work, coherent transport is neglected entirely, allowing much quicker simulation times. This is justified by noting that reasonably good agreement has been achieved between experimental results and Boltzmann or rate equation based models of THz QCLs in III–V systems[49, 150]. The advantage of this approach is that incoherent transport can be considered more completely than in any of the partially-coherent calculations, accounting for all significant scattering mechanisms and charge carrier heating.

Naturally, there are some limitations to this fully incoherent model. In particular, the injection barrier thickness only affects the current density weakly and large current spikes occur, corresponding to the appearance of spatially-extended states. These effects are avoided in coherent injection models[148] and are discussed in more detail in chapter 7. An investigation into coherent transport in Si-based QCLs is proposed as further work in 9.

5.2 Fermi's golden rule

Each scattering event in QCLs may be explained in terms of an interaction between an electron and an electric potential. These potentials are much smaller than the conduction band offset and may arise from random processes. It is therefore reasonable to treat these potentials as small perturbations to the Schrödinger solution described in previous chapters[90]. The complete Hamiltonian for a system thus contains a large, time-invariant component $\hat{H}_0$ describing the unperturbed system and a small scattering potential $\hat{V}(t)$. The time-dependent Schrödinger equation is then

$$\left[\hat{H}_0 + \hat{V}(t)\right] |\psi(t)\rangle = \mathrm{i}\hbar \frac{\partial}{\partial t} |\psi(t)\rangle \,. \tag{5.1}$$

If the eigenstates of the unperturbed system are denoted $|j\rangle$, then the corresponding eigenvalues E_j represent the possible outcomes of a measurement of electron energy. The full time-dependent envelope functional form of the wavefunction is given by

$$|j(t)\rangle = |j\rangle \exp\left(-\frac{\mathrm{i}E_j t}{\hbar}\right) = \frac{1}{\sqrt{A}} \psi_j(z) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \mathrm{e}^{-\frac{\mathrm{i}E_j t}{\hbar}}. \tag{5.2}$$

When subjected to perturbations, the state "blurs" over time and is a time-varying, weighted sum of the unperturbed states

$$|\psi(t)\rangle = \sum_j a_j(t) |j\rangle \mathrm{e}^{-\frac{\mathrm{i}E_j t}{\hbar}}, \tag{5.3}$$

where $a_j(t)$ are time-dependent weightings. Eqn. 5.1 becomes:

$$\sum_j a_j(t) \hat{H}_0 |j(t)\rangle + \sum_j a_j(t) \hat{V}(t) |j(t)\rangle = \mathrm{i}\hbar \sum_j \left[a_j(t) \frac{\partial}{\partial t} |j(t)\rangle + \frac{\mathrm{d}a_j(t)}{\mathrm{d}t} |j(t)\rangle \right] \tag{5.4}$$

The first terms on either side of the expression form the time-dependent Schrödinger equation for the unperturbed system and therefore can be cancelled out. For a transition to a final state $|f\rangle$, the transition energy is given by $E_{fj} = E_f - E_j$ and the Schrödinger equation may be rewritten as

$$\sum_j a_j(t) \langle f | \hat{V}(t) | j \rangle \mathrm{e}^{\frac{\mathrm{i}E_{fj} t}{\hbar}} = \mathrm{i}\hbar \sum_j \frac{\mathrm{d}a_j(t)}{\mathrm{d}t} \langle f | j \rangle. \tag{5.5}$$

By noting the orthonormality of states, and defining $\langle f|\hat{V}(t)|j\rangle = V_{jf}(t)$, this yields

$$\frac{\mathrm{d}a_f(t)}{\mathrm{d}t} = \frac{1}{\mathrm{i}\hbar}\sum_j a_j(t)V_{jf}(t)\mathrm{e}^{\frac{\mathrm{i}E_{fj}t}{\hbar}}. \tag{5.6}$$

If the scattering rates are low, then electrons are likely to remain in their initial state $|i\rangle$ between two measurements. The approximation may therefore be made that $a_j(t) \approx \delta_{ij}$ and the (much smaller) probability of final state occupation is

$$|a_f(t)|^2 = \frac{1}{\hbar^2}\left|\int_0^t V_{if}(t')\mathrm{e}^{\frac{\mathrm{i}E_{fi}t'}{\hbar}}\,\mathrm{d}t'\right|^2 \tag{5.7}$$

5.2.1 Static scattering potentials

For time-independent scattering potentials, *i.e.* $V_{if}(t) = V_{if}$, Eqn. 5.7 simplifies to

$$|a_f(t)|^2 = \frac{|V_{if}|^2 t^2}{\hbar^2}\,\mathrm{sinc}^2\left(\frac{E_{fi}t}{2\hbar}\right). \tag{5.8}$$

After initial transients have died away, it can be shown[90] that this reduces to

$$\lim_{t\to\infty}|a_f(t)|^2 = \frac{2\pi|V_{if}|^2 t}{\hbar}\delta(E_{fi}). \tag{5.9}$$

The final state is therefore only occupied if the scattering is elastic. Differentiating this probability yields the scattering rate, W_{if} between an initial and a final state

$$W_{if} = \frac{2\pi}{\hbar}\left|V_{if}\right|^2\delta(E_{fi}). \tag{5.10}$$

This is known as Fermi's golden rule for static scattering potentials.

In two-dimensional systems, it is important to note that the energy eigenvalues lie within subbands and actually take the form

$$E_j(k) = E_j + \frac{\hbar^2 k^2}{2m_d}, \tag{5.11}$$

where E_j is the minimum of subband j. It is therefore useful to consider an expanded form of Fermi's golden rule,

$$W_{if}(k_i,k_f) = \frac{2\pi}{\hbar}\left|V_{if}\right|^2\delta\left[E_{fi} + \frac{\hbar^2}{2}\left(\frac{k_f^2}{m_{d,f}} - \frac{k_i^2}{m_{d,i}}\right)\right]\Theta(k_f^2). \tag{5.12}$$

The Θ here represents the Heaviside step function, which ensures that the final state lies within the final subband.

Assuming an identical density-of-states effective mass in the initial and final subbands allows the δ function to be rewritten in terms of in-plane wave vector,

$$W_{if}(k_i, k_f) = \frac{2\pi m_d}{\hbar^3 k_\alpha} |V_{if}|^2 \, \delta(k_f \pm k_\alpha)\Theta(k_f^2), \tag{5.13}$$

where $k_\alpha = \sqrt{k_i^2 - 2m_d E_{fi}/\hbar^2}$ is the final wavevector required for energy conservation.

The total scattering rate from a particular state in the initial subband to all states in the final subband may be obtained by summing over all final wavevectors,

$$W_{if}(k_i) = \frac{2\pi m_d}{\hbar^3 k_\alpha} \sum_{k_f} |V_{if}|^2 \, \delta(k_f \pm k_\alpha)\Theta(k_f^2) \tag{5.14}$$

Due to the continuity of the final states, the sum may be transformed to an integral, giving

$$\begin{aligned} W_{if}(k_i) &= \frac{A m_d}{2\pi\hbar^3 k_\alpha} \iint k_f \mathrm{d}k_f \mathrm{d}\theta \delta(k_f - k_\alpha)\Theta(k_f^2) \left|V_{if}\right|^2, \\ &= \frac{A m_d}{2\pi\hbar^3}\Theta(k_\alpha^2) \int \mathrm{d}\theta |V_{if}(k_\alpha)|^2, \end{aligned} \tag{5.15}$$

where A is the area of the 2D system.

A final improvement to this model may be made by noting that scattering cannot occur if the final state is occupied, due to Pauli exclusion. Scattering rates are therefore overestimated slightly, and can be corrected by multiplying the "raw" rate by the probability of the final state being vacant. If the intrasubband scattering rates are much faster than intersubband rates, then electrons in the final subband can be assumed to have settled to a thermal quasi-equilibrium and Fermi-Dirac statistics apply. Hence, the reduced scattering rate is given by

$$W_{if}(k_i) = \frac{A m_d}{2\pi\hbar^3}[1 - P_f(k_\alpha)]\Theta(k_\alpha^2) \int \mathrm{d}\theta |V_{if}(k_\alpha)|^2, \tag{5.16}$$

where $P_f(k_\alpha) = f_{\mathrm{FD}}(k_\alpha, E_{\mathrm{F},f}, T_e)$ is the final state occupation probability.

5.2.2 Time-varying scattering potentials

Sinusoidal perturbations include electron–phonon interactions and the scattering potential takes the form

$$V_{if}(t) = V_{if}\left(e^{j\omega_0 t} + e^{-j\omega_0 t}\right), \tag{5.17}$$

where ω_0 is the angular frequency of the oscillating potential. The solution of eqn. 5.7 becomes

$$\begin{aligned} |a_f(t)|^2 &= \frac{|V_{if}|^2 t^2}{\hbar^2}\left[\operatorname{sinc}^2\frac{(E_{fi}-\hbar\omega_0)t}{2\hbar} + \operatorname{sinc}^2\frac{(E_{fi}+\hbar\omega_0)t}{2\hbar}\right. \\ &+ \left. 2\cos\omega_0 t \operatorname{sinc}\frac{(E_{fi}-\hbar\omega_0)t}{2\hbar}\operatorname{sinc}\frac{(E_{fi}+\hbar\omega_0)t}{2\hbar}\right]. \end{aligned} \tag{5.18}$$

Taking the steady-state limit ($t \to \infty$), the steady-state scattering rate is given by

$$W_{if} = \frac{2\pi}{\hbar}|V_{if}|^2\left[\delta(E_{fi}-\hbar\omega_0) + \delta(E_{fi}+\hbar\omega_0)\right]. \tag{5.19}$$

Scattering is therefore non-zero only for two final energy states. In each case, the electron has either gained or lost $\hbar\omega_0$ in energy, which corresponds to the absorption or emission of a particle (*e.g.* a phonon) respectively.

Rewriting this in terms of wave vector and summing over the final subband gives a total scattering rate of

$$W_{if}(k_i) = \frac{2\pi}{\hbar}|V_{if}|^2\left[\Theta(k_{\alpha+}^2) + \Theta(k_{\alpha-}^2)\right], \tag{5.20}$$

where the final in-plane wave vectors for absorption (+ sign) and emission (- sign) interactions are

$$k_{\alpha\pm}^2 = \frac{2m(-E_{fi} \pm \hbar\omega_0)}{\hbar^2} + k_i^2. \tag{5.21}$$

5.2.3 Average scattering rate

Although the scattering rate has been derived for a given initial wavevector, it is also useful to find its average value. Assuming Fermi-Dirac statistics as before, this average is[97]

$$\overline{W_{if}} = \frac{\int_{k^-}^{k^+} k_i \mathrm{d}k_i P_i(k_i) W_{if}(k_i)}{\pi n_i} \tag{5.22}$$

Strictly, the integral is performed over the range $k_i = (0, \infty)$. However, an analytical solution is usually impossible to achieve and an upper limit for numerical integration is required. The Fermi–Dirac distribution is given by

$$f_{\mathrm{FD}}(E_k, T_e) = \frac{1}{\exp\left(\frac{E_k - E_F}{k_B T_e}\right) + 1}. \tag{5.23}$$

If the Fermi energy is located well below the subband minimum, (which is the case for subbands with low populations at temperatures above a few Kelvin), this simplifies to the Maxwell–Boltzman distribution

$$f_{\mathrm{MB}}(E_k, T_e) = \exp\left(\frac{-E_k}{k_B T_e}\right). \tag{5.24}$$

It can now be shown that

$$\int_0^{4.6 k_B T_e} \mathrm{d}E f_{\mathrm{MB}}(E_k, T_e) = 0.99 \int_0^{\infty} \mathrm{d}E f_{\mathrm{MB}}(E_k, T_e), \tag{5.25}$$

meaning that around 99% of the subband population is contained in an energy range of $4.6\, k_B T_e$ above the subband minimum, making $E_{k,\max} = 4.6 k_B T_e$ a suitable cut-off energy for calculations. The corresponding wave vector is $k^+ = \sqrt{9.2 m_d k_B T_e}/\hbar$. At low temperatures, the quasi-Fermi energy may lie above the subband minimum and in this case, the cut-off energy may be set as $E_{k,\max} = 4.6 k_B T_e + E_F$.

The lower limit may be increased above zero by noting that $W_{if}(k_i)$ is nonzero only when $k_\alpha^2 > 0$, giving the result

$$k^- = \begin{cases} 0, & E_{fi} \leq 0 \\ \sqrt{2 m_d E_{fi}}/\hbar, & E_{fi} > 0. \end{cases} \tag{5.26}$$

5.3 Coulombic scattering

The overall effect of the charge distribution in a QCL upon its bandstructure was included in the Hamiltonian in previous chapters. This represents the cumulative effect of all the charged particles in an average system. On a microscopic scale however, an electron interacts strongly with nearby charged particles via Coulombic forces. In the following sections, expressions are derived for the scattering rates due to donor ions and other electrons. Computing these expressions can be somewhat demanding and methods for optimising the numerical solution are presented.

5.3.1 Ionised impurity scattering

The Coulombic electrostatic potential for an electron in the vicinity of a single positive ion is[151]

$$\hat{V}(\boldsymbol{r}-\boldsymbol{r_0}, z-z_0) = -\frac{e^2}{4\pi\epsilon\sqrt{|\boldsymbol{r}-\boldsymbol{r_0}|^2+(z-z_0)^2}}, \tag{5.27}$$

where $(z, \boldsymbol{r})$ and $(z_0, \boldsymbol{r_0})$ are the electron and ion positions respectively. The time-independent envelope wavefunctions were defined in chapter 3 as

$$|j\rangle = \frac{1}{\sqrt{A}}\mathrm{e}^{\mathrm{i}\boldsymbol{k_j}\cdot\boldsymbol{r}}\psi_j(z) \tag{5.28}$$

The scattering matrix element is given by

$$V_{if} = \langle f|\hat{V}|i\rangle = \frac{1}{A}\int \psi_f^*(z)\psi_i(z)\,\mathrm{d}z \iint \mathrm{e}^{\mathrm{i}(\boldsymbol{k_i}-\boldsymbol{k_f})\cdot\boldsymbol{r}}\,\mathrm{d}^2\boldsymbol{r}. \tag{5.29}$$

Making the substitutions $\boldsymbol{u} = \boldsymbol{r} - \boldsymbol{r_0}$ and $\boldsymbol{q} = \boldsymbol{k_f} - \boldsymbol{k_i}$, the scattering potential matrix element for a single impurity may be rewritten as

$$V_{if}(\boldsymbol{q}, \boldsymbol{r_0}, z_0) = \frac{1}{A}\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r_0}}\int \psi_f^*(z)\psi_i(z)\,\mathrm{d}z \iint \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{u}}\hat{V}(u, z, z_0)\,\mathrm{d}^2\boldsymbol{u}. \tag{5.30}$$

The double integral is a two-dimensional Fourier transform, with an analytical solution. This can be solved to give the square of the single-impurity scattering matrix element[152],

$$|V_{if}(q, z_0)|^2 = \left(\frac{e^2}{2\epsilon A q}\right)^2 |I_{if}(q, z_0)|^2, \tag{5.31}$$

where $I_{if}(q, z_0) = \int \psi_f^*(z)\mathrm{e}^{-q|z-z_0|}\psi_i(z)\,\mathrm{d}z$.

A QCL has multiple impurities spread through its volume, with concentration varying over z_0. The total matrix element is found by multiplying the single impurity result by the number of impurities in a thin slice through the system and integrating. The number of impurities in a slice of width $\mathrm{d}z_0$ is $d(z_0)A\mathrm{d}z_0$, where $d(z_0)$ is the dopant concentration, giving a total matrix element of[153]

$$|V_{if}(q)|^2 = \frac{e^4}{4\epsilon^2 A q^2} \int d(z_0)|I_{if}(q, z_0)|^2 \, \mathrm{d}z_0. \tag{5.32}$$

Substituting into Eqn. 5.16 gives

$$W_{if}(k_i) \;=\; \frac{m_d e^4}{4\pi\hbar^3\epsilon^2}[1 - P_f(k_\alpha)]\Theta(k_\alpha^2) \int_0^\pi \mathrm{d}\theta \frac{J_{if}(q_\alpha)}{q_\alpha^2}, \tag{5.33}$$

where $q_\alpha^2 = k_i^2 + k_\alpha^2 - 2k_i k_\alpha \cos\theta$ and $J_{if}(q_\alpha) = \int d(z_0)|I_{if}(q_\alpha, z_0)|^2 \, \mathrm{d}z_0$. The limits of the integration over θ have been halved due to the even, periodic nature of q_α.

An improvement to this model may be introduced by including the effect of screening by a two-dimensional electron gas. Local electrons are strongly attracted towards a positive donor ion, making its charge appear weaker to remote electrons. Davies shows that this may be included by incorporating the Thomas-Fermi screening wavevector, q_{TF} in Eqn. 5.33 to obtain[90]

$$W_{if}(k_i) = \frac{m_d e^4}{4\pi\hbar^3\epsilon^2}[1 - P_f(k_\alpha)]\Theta(k_\alpha^2) \int_0^\pi \frac{J_{if}(q_\alpha)\mathrm{d}\theta}{(q_\alpha + q_{\mathrm{TF}})^2}, \tag{5.34}$$

where

$$q_{\mathrm{TF}} = \frac{m_d e^2}{2\pi\epsilon\hbar^2}. \tag{5.35}$$

Ionised impurity scattering presents a computational challenge as solving the three nested integrals using the trapezium rule requires a vast number of iterations. The scattering vector q_α depends on the scattering angle θ and the initial wavevector k_i and recalculating J_{if} on-the-fly for each (θ, k_i) combination would be extremely inefficient. Instead, a table of $J_{if}(q_\alpha)$ against q_α may be calculated before commencing integration. The required values are then obtained by interpolation.

Further optimisation is possible by noting that each value of J_{if} requires an integration over a function of I_{if}. A small decrease in the computation time for I_{if} therefore leads to a much greater decrease in the computation time for J_{if}. An elegant solution involves splitting the integrand into a pair of functions with semi-infinite domains i.e.

$$e^{-q_\alpha|z-z_0|} = \begin{cases} e^{q_\alpha(z_0-z)}, & z \geq z_0 \\ e^{q_\alpha(z-z_0)}, & z < z_0. \end{cases} \tag{5.36}$$

This reduces the computation time significantly by removing the need to evaluate the modulus function.

It follows that the matrix element may be rewritten in the form

$$I_{if}(q_\alpha, z_0) = \frac{1}{e^{q_\alpha z_0}} C_{if}^-(q_\alpha, z_0) + e^{q_\alpha z_0} C_{if}^+(q_\alpha, z_0), \tag{5.37}$$

where

$$\begin{aligned} C_{if}^-(q_\alpha, z_0) &= \int_{-\infty}^{z_0} \psi_f^*(z) e^{q_\alpha z} \psi_i(z)\, dz \\ C_{if}^+(q_\alpha, z_0) &= \int_{z_0}^{\infty} \psi_f^*(z) \frac{1}{e^{q_\alpha z}} \psi_i(z)\, dz. \end{aligned} \tag{5.38}$$

$C_{if}^{\pm}(q_\alpha, z_0)$ may be tabulated against z_0 before calculating I_{if}. Pre-calculating the exponential terms and the product of the wavefunctions increases the computational efficiency even further.

The ionised impurity scattering rate in a (001) Ge quantum well with Si barriers is shown as a function of well width in fig. 5.1. Dopants were spread evenly through the well region. Scattering rates were calculated from the second to first subband as a function of subband separation, assuming an electron temperature of 100 K. As the well width increases, the subband separations decrease and the scattering vectors required for energy conservation are therefore quite small. The denominator in the integral over θ shrinks and the scattering rate increases slightly. An approximately linear scaling is seen with respect to doping density.

Figure 5.2 shows that ionised impurity scattering depends heavily upon the doping profile. The highest rates occur when dopants are spread evenly through the well region (as in fig. 5.1) due to the high overlap between the wavefunctions in the doped region. Moving the dopants to the barrier reduces the scattering rates by

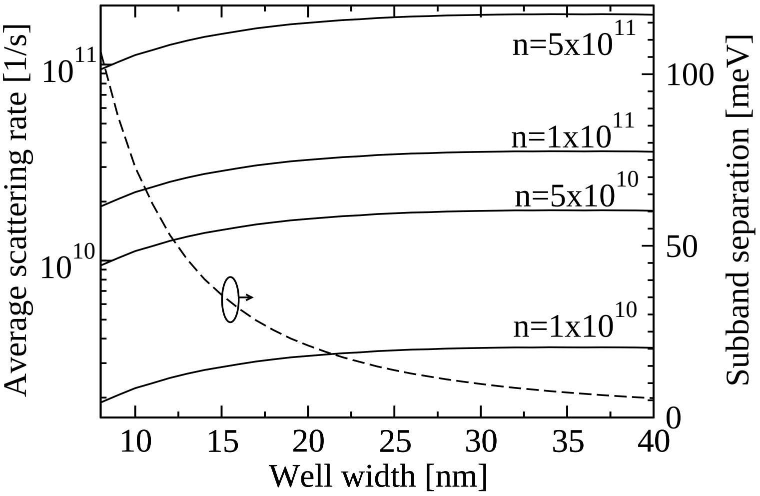


Figure 5.1: Average ionised impurity scattering rate from the second to first subband as a function of well width for a (001) Ge QW with 10 nm thick Si barriers. Dopants are spread evenly through the well region at concentrations in the range 1×10^{10}–5×10^{11} cm^{-2}. An electron temperature of 100 K was assumed. The subband populations were assumed to be equal prior to scattering and the separation between the subbands is plotted on the same axes.

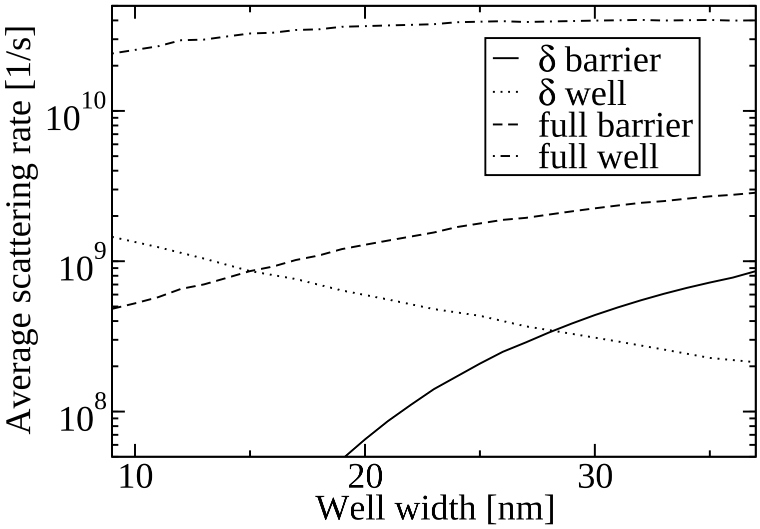


Figure 5.2: Average ionised impurity scattering rate from the second to first subband as a function of well width for a (001) Ge QW with 10 nm thick Si barriers for various doping profiles. All parameters were set identical to those in fig. 5.1, except for the doping concentration, which was fixed at $1\times10^{11}\,\mathrm{cm}^{-2}$. Results are shown for dopants spread throughout the full well or barrier region and for dopants confined to a 1 nm thick region in the well or barrier (δ doping).

a factor of around 50 and δ doping at the edges of the barrier region reduces the scattering rate even further. It is interesting to note that δ doping in the centre of the well also gives very low scattering rates. This is due to the dopants lying around a node (*i.e.* zero probability density) in the upper wavefunction. In an asymmetric system such as a QCL, it is likely that this effect would be less significant.

5.3.2 Electron–electron scattering

As with the case of ionised impurity scattering, the interaction between a pair of electrons is governed by the Coulomb potential,

$$\hat{V}(\boldsymbol{r}-\boldsymbol{r'},z-z') = \frac{e^2}{4\pi\varepsilon\sqrt{|\boldsymbol{r}-\boldsymbol{r'}|^2+(z-z')^2}}, \tag{5.39}$$

where $(z,\boldsymbol{r})$ and $(z',\boldsymbol{r'})$ are the locations of the first and second electrons respectively. The state of the two-electron system is given by

$$\begin{aligned} |ij\rangle &= \frac{1}{A}\psi_i(z)\psi_j(z')\mathrm{e}^{\mathrm{i}(\boldsymbol{k_i}\cdot\boldsymbol{r}+\boldsymbol{k_j}\cdot\boldsymbol{r'})} \\ |fg\rangle &= \frac{1}{A}\psi_f(z)\psi_g(z')\mathrm{e}^{\mathrm{i}(\boldsymbol{k_f}\cdot\boldsymbol{r}+\boldsymbol{k_g}\cdot\boldsymbol{r'})}, \end{aligned} \tag{5.40}$$

where the subscript pairs i,j and f,g denote the inital and final states respectively. Momentum conservation was ignored in the calculation of ionised impurity scattering, due to the enormous difference between the electron mass and that of the ion bonded into the crystal lattice. In the case of electron–electron scattering however, it must be included. As such,

$$\boldsymbol{k_i}+\boldsymbol{k_j} = \boldsymbol{k_f}+\boldsymbol{k_g} \tag{5.41}$$

and the scattering vector is given by $\boldsymbol{q}=\boldsymbol{k_f}-\boldsymbol{k_i}=\boldsymbol{k_j}-\boldsymbol{k_g}$.

The scattering matrix element is now written as

$$\begin{aligned} V_{ij\to fg}(q) &= \frac{e^2}{4\pi\varepsilon A^2}\int \mathrm{d}z\psi_f^*(z)\psi_i(z)\int \mathrm{d}z'\psi_g^*(z')\psi_j(z') \\ &\times \iint \mathrm{d}^2\boldsymbol{r'}\iint \mathrm{d}^2\boldsymbol{u}\frac{\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{u}}}{\sqrt{u^2+(z-z')^2}}\delta(\boldsymbol{k_i}+\boldsymbol{k_j}-\boldsymbol{k_f}-\boldsymbol{k_g}), \end{aligned} \tag{5.42}$$

where the substitution $\boldsymbol{u} = \boldsymbol{r} - \boldsymbol{r'}$ has been made. Solving the Fourier transform and the integral over $\boldsymbol{r'}$ gives

$$V_{ij\to fg}(q) = \frac{e^2}{2\varepsilon Aq} A_{ij\to fg}(q) \times \delta(\boldsymbol{k_i} + \boldsymbol{k_j} - \boldsymbol{k_f} - \boldsymbol{k_g}), \tag{5.43}$$

where

$$A_{ij\to fg}(q) = \int \mathrm{d}z \psi_f^*(z)\psi_i(z) \int \mathrm{d}z' \psi_g^*(z')\psi_j(z')\mathrm{e}^{-q|z-z'|}. \tag{5.44}$$

Following the method of Smet[154], which is very similar in form to that for ionised impurities (albeit considerably lengthier), the scattering rate is given by[1]

$$\begin{aligned} W_{ij\to fg}(\boldsymbol{k_i}) &= \frac{e^4 m_d}{(4\pi\hbar)^3\varepsilon^2} \int_0^\infty k_j \mathrm{d}k_j P_j\left(k_j\right) \\ &\times \; \Theta\left(k_{ij}^2\right) \int \mathrm{d}\alpha \int \mathrm{d}\theta \frac{|A_{ij\to fg}(q_\alpha)|^2}{q_\alpha^2} \Theta\left(q_\alpha^2\right), \end{aligned} \tag{5.45}$$

where α is the angle between the initial wavevectors and θ is the angle between a pair of relative wavevectors $\boldsymbol{k_{ij}} = \boldsymbol{k_j} - \boldsymbol{k_i}$ and $\boldsymbol{k_{fg}} = \boldsymbol{k_g} - \boldsymbol{k_f}$. q_α is the magnitude of the scattering vector which yields energy and momentum conservation and is given by

$$(2q_\alpha)^2 = 2k_{ij}^2 + \Delta k_0^2 - 2k_{ij}\sqrt{k_{ij}^2 + \Delta k_0^2}\cos\theta, \tag{5.46}$$

where

$$\Delta k_0^2 = \frac{4m_d}{\hbar^2}(E_i + E_j - E_f - E_g) \tag{5.47}$$

$$k_{ij}^2 = k_i^2 + k_j^2 - 2k_i k_j \cos\alpha. \tag{5.48}$$

The scattering rate equation above assumes that the final states are always vacant, which is a very good approximation for low electron temperatures, and reduces the computational demand significantly. The Thomas-Fermi screening approach described for ionised impurity scattering has been shown to overestimate the screening significantly for electron–electron scattering and a better approach considers screening due to electrons within the initial subband[154]. The relative permittivity is

---

[1] Smet's expression has been modified slightly to make the requirement for real wavevectors explicit

replaced by

$$\varepsilon_{\mathrm{sc}}(q,T_e) = 1 + \frac{2\pi e^2}{4\pi\epsilon q}\Pi_{ii}(q,T_e)A_{ij\to fg}(q). \tag{5.49}$$

At low-temperatures, the polarisation factor is given by[155]

$$\Pi_{ii}(q,0) = \frac{m_d}{\pi\hbar^2}\left[1 - \Theta(q-2k_F)\sqrt{1-\left(\frac{2k_F}{q}\right)^2}\right], \tag{5.50}$$

where the Fermi wavevector is $k_F = \sqrt{2\pi n_i}$.

Substituting the screening permittivity into eqn. 5.45 gives

$$\begin{aligned} W_{ij\to fg}(\boldsymbol{k_i}) &= \frac{e^4 m_d}{(4\pi\hbar)^3\varepsilon_0^2}\int k_j \mathrm{d}k_j \Theta\left(k_j^2\right)P_j\left(k_j\right) \\ &\times \Theta\left(k_{ij}^2\right)\int \mathrm{d}\alpha \int \mathrm{d}\theta F_{ijfg}(q_\alpha), \end{aligned} \tag{5.51}$$

where

$$F_{ij\to fg}(q_\alpha) = \left[\frac{|A_{ij\to fg}(q_\alpha)|}{q_\alpha \varepsilon_{\mathrm{sc}}(q_\alpha)}\right]^2 \Theta\left(q_\alpha^2\right). \tag{5.52}$$

The computational implementation of electron–electron scattering is more demanding even than that of ionised impurity scattering as there are now five nested integrals. A set of optimisation methods developed in the course of the present work dramatically improves the situation however. Firstly, $F_{ij\to fg}(q_\alpha)$ depends only on the magnitude of the scattering vector and has no direct dependence on scattering angles or initial wavevectors. As such, it may be tabulated before performing the integrals rather than recalculating them on-the-fly. In addition to this, $q\varepsilon_{\mathrm{sc}}$ depends only on the initial subband index and therefore need only be recalculated once for each initial subband.

The calculation time of $A_{ij\to fg}$ may be improved substantially by rewriting it as

$$A_{ij\to fg}(q) = \int \mathrm{d}z \psi_{if}(z) I_{jg}(q,z), \tag{5.53}$$

where the double-subscripted wavefunctions denote a product and the I_{jg} matrix element is identical to that used for ionised impurity scattering. The optimised calculation of I_{jg} described in the previous section gives a factor of over 100 improvement in the calculation speed. Nevertheless, if electron–electron scattering

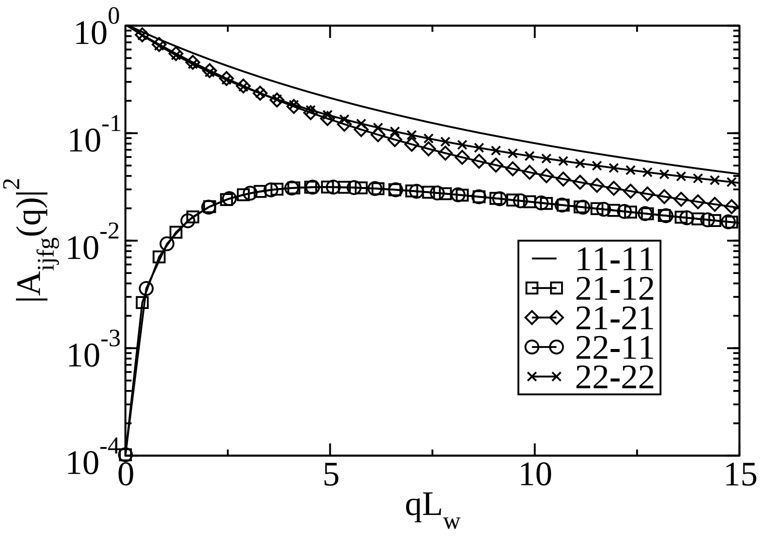


Figure 5.3: Squared electron-electron scattering matrix elements as a function of scattering vector × well width for each of the possible interactions in a 10 nm wide infinite QW.

calculations are included in a QCL simulation, they account for around 90% of the total simulation runtime.

For a system with N subbands, there are N^4 possible scattering interactions. However, not all of these have a significant effect on charge transport. The interactions may be categorised into three sets. Firstly, there is a set of intrasubband scattering mechanisms, which may take the form $W_{ii\to ii}$ or $W_{ij\to ij}$. Here, both electrons stay in their original subbands (which may be different from each other). These interactions are always permitted regardless of the scattering vector, due to the continuity of final states within the subband and are therefore much more rapid than other interactions. Fig. 5.3 shows that the scattering matrix element is much larger for intrasubband interactions than for other interactions. This adds weight to the earlier assumption that intrasubband scattering rapidly causes carriers to settle to quasi-thermalised distributions between intersubband scattering events.

The next set of interactions are purely intersubband, where both electrons change

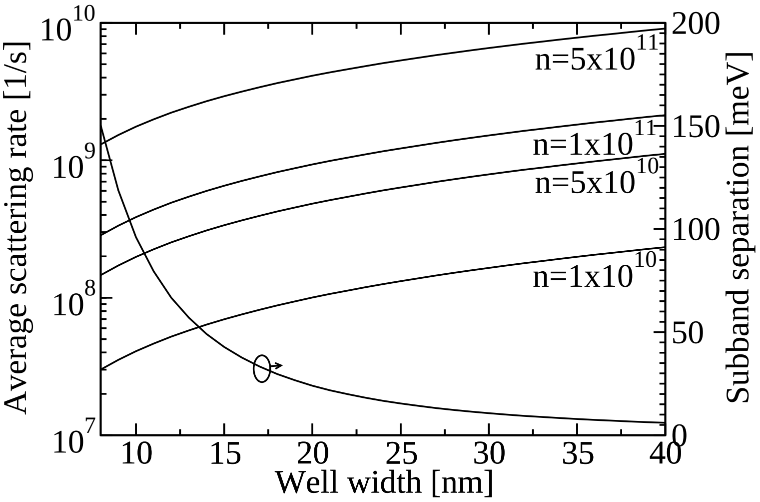


Figure 5.4: Average, electron–electron scattering rate from the second to first subband in a (001) Ge QW identical to that described in the previous section. The effective one-electron scattering rate is shown for total carrier densities in the range 1×10^{10} – $5 \times 10^{11}\,\mathrm{cm}^{-2}$.

subband. These may take the form $W_{ij\to fg}$ or $W_{ij\to ji}$. The former leads to a change in the subband populations, whereas the latter does not. When the subband separation is large (100 meV in fig. 5.3), a large scattering vector is required unless the electron has very large kinetic energy within the initial subband. The scattering matrix element is therefore very low at small scattering vectors. Finally, Auger transitions exist in which only one electron changes subband. These take the form $W_{ii\to ij}$ or $W_{ij\to jj}$. Smet[154] shows that these processes are forbidden in symmetric systems and are likely to be small otherwise.

In order to calculate the effect of these two-electron interactions on subband populations, an effective one-electron scattering rate may be defined by summing over all the second electron subbands

$$W_{if} = W_{ii\to ff} + \sum_{j,g} W_{ij\to fg}. \tag{5.54}$$

The additional $W_{ii\to ff}$ term accounts for the fact that two electrons are transferred between subbands by such processes. Fig. 5.4 shows that the effective one-electron scattering rate decreases as a function of well width. As the Auger processes are negligible, the rate is dominated by the $W_{22\rightarrow 11}$ and $W_{21\rightarrow 12}$ transitions. As the well width decreases, the subband separation and hence the scattering vectors increase and the q_α term in the denominator of $F_{ij\rightarrow fg}$ leads to a decrease in scattering rate.

Both ionised impurity and electron–electron scattering rates scale approximately linearly with respect to dopant concentration. There is a fundamental difference, however, in that ionised impurity scattering depends directly upon the spatial distribution of dopants whereas electron–electron scattering does not. The variations of ionised impurity (fig. 5.1) and electron–electron (fig. 5.4) scattering rates with respect to well width have similar functional forms when dopants are spread evenly throughout the well, although the ionised impurity scattering rates are around 100 times larger. It can be deduced from fig. 5.2, however, that electron–electron scattering may exceed ionised impurity scattering for δ-doping profiles.

5.4 Scattering from structural imperfections

The calculations in chapter 2 assumed that the conduction band potential was constant throughout each layer of a QCL and that the interfaces were perfectly formed. In reality, the conduction band potential is derived from overlapping atomic potentials and the position of interfaces may vary significantly across the growth plane. These deviations from the ideal system both act as perturbations to the Schrödinger solution and the resulting scattering rates are derived in the following sections.

5.4.1 Alloy disorder scattering

Si and Ge atoms are distributed randomly through an alloy and the crystal potential varies stochastically on an atomic scale, Bloch waves do not form and it is impractical to solve Schrödinger's equation. This difficulty was avoided in the model solid approximation by replacing the Si and Ge atoms with a uniform array of virtual SiGe atoms. Although the virtual crystal model matches the average potential, a discrepancy exists on the atomic scale. These microscopic fluctuations act as a perturbation to the system, and induce scattering.

If potentials due to individual Si and Ge atoms are $V_{\mathrm{Si}}(\boldsymbol{R})$ and $V_{\mathrm{Ge}}(\boldsymbol{R})$ respectively then the total potential at a point in the crystal is given by summing the contributions due to the individual atoms. The complete set of lattice sites in the system, R_A is given by $R_A = R_{\mathrm{Si}} \cup R_{\mathrm{Ge}}$, where R_{Si} and R_{Ge} are sites occupied by Si and Ge atoms respectively. This yields a total potential at point $\boldsymbol{R}$ of

$$V(\boldsymbol{R}) = \sum_{\boldsymbol{R_i} \in R_{\mathrm{Ge}}} V_{\mathrm{Ge}}(\boldsymbol{R} - \boldsymbol{R_i}) + \sum_{\boldsymbol{R_i} \in R_{\mathrm{Si}}} V_{\mathrm{Si}}(\boldsymbol{R} - \boldsymbol{R_i}). \tag{5.55}$$

The potential of a virtual atom is taken to be

$$V_{\mathrm{va}}(\boldsymbol{R}) = x(z) V_{\mathrm{Ge}}(\boldsymbol{R}) + [1 - x(z)] V_{\mathrm{Si}}(\boldsymbol{R}), \tag{5.56}$$

where $x(z)$ is the probability of a lattice site at a given position being occupied by a Ge atom. The virtual crystal potential is therefore

$$V_{\mathrm{vc}}(\boldsymbol{R}) = \sum_{\boldsymbol{R_i} \in R_A} \{x(z) V_{\mathrm{Ge}}(\boldsymbol{R}) + [1 - x(z)] V_{\mathrm{Si}}(\boldsymbol{R})\}. \tag{5.57}$$

The difference between the actual, and virtual crystal potentials gives the perturbation potential,

$$V_p(\boldsymbol{R}) = \sum_{\boldsymbol{R_i} \in R_{\mathrm{Ge}}} [1 - x(z)] \, \Delta V(\boldsymbol{R} - \boldsymbol{R_i}) - \sum_{\boldsymbol{R_i} \in R_{\mathrm{Si}}} x(z) \Delta V(\boldsymbol{R} - \boldsymbol{R_i}), \tag{5.58}$$

where $\Delta V(\boldsymbol{R}) = V_{\mathrm{Ge}}(\boldsymbol{R}) - V_{\mathrm{Si}}(\boldsymbol{R})$ is the difference between the Si and Ge atomic potentials.

By assuming that an atom's contribution to the total potential is only distinguishable over a small radius, it can be shown that

$$\Delta V(\boldsymbol{R}) \approx \Omega \Delta V_{\mathrm{ad}} \delta(\boldsymbol{R}), \tag{5.59}$$

where $\Omega = a^3/8$ is the volume containing a single lattice site and ΔV_{ad} is a constant alloy scattering potential[156]. The matrix element for scattering is therefore

$$\begin{aligned} V_{if} &= \frac{a^3 \Delta V_{\mathrm{ad}}}{8A} \iiint \mathrm{d}^3 \boldsymbol{R} \psi_f^*(z) \psi_i(z) \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} \\ &\times \left\{ \sum_{\boldsymbol{R_i} \in R_{\mathrm{Ge}}} [1 - x(z)]\, \delta(\boldsymbol{R} - \boldsymbol{R_i}) - \sum_{\boldsymbol{R_i} \in R_{\mathrm{Si}}} x(z) \delta(\boldsymbol{R} - \boldsymbol{R_i}) \right\} \\ &= \frac{a^3 \Delta V_{\mathrm{ad}}}{8A} \sum_{\boldsymbol{R_i} \in R_{\mathrm{Ge}}} [1 - x(z_i)]\, \psi_f^*(z_i) \psi_i(z_i) \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r_i}} \\ &- \frac{a^3 \Delta V_{\mathrm{ad}}}{8A} \sum_{\boldsymbol{R_i} \in R_{\mathrm{Si}}} x(z_i) \psi_f^*(z_i) \psi_i(z_i) \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r_i}}. \end{aligned} \tag{5.60}$$

The mean square matrix element is then given by

$$\langle |V_{if}|^2 \rangle = \frac{a^3 \Delta V_{\mathrm{ad}}^2}{8A} \int \mathrm{d}z |\psi_f(z)|^2 x(z) [1 - x(z)] |\psi_i(z)|^2. \tag{5.61}$$

Substituting this into Fermi's golden rule gives

$$W_{if}(k_i) = \frac{m_d a^3}{8\hbar^3} [1 - P_f(k_\alpha)] \Theta(k_\alpha^2) \Delta V_{\mathrm{ad}}^2 \int \mathrm{d}z |\psi_f(z)|^2 x(z) [1 - x(z)] |\psi_i(z)|^2. \tag{5.62}$$

As a first estimate, the alloy disorder potential ΔV_{ad} may be taken as the Si–Ge band offset but the calculated rates may differ somewhat from those observed experimentally. There is, however, no consensus in the literature on a more appropriate value.

Apart from the requirement for real and unoccupied final states, alloy disorder scattering is constant with respect to wavevector. For electrons scattering to lower subbands, all final wavevectors are real and the scattering rate is therefore approximately constant over the entire range of initial wavevectors. Instead, the rate is determined primarily by the overlap of the wavefunctions within regions of high alloy disorder. The term $x(z)[1 - x(x)]$ in the matrix element is maximised when $x = 0.5$ (*i.e.* an equal mixture of Si and Ge) and drops to zero in pure Si and Ge.

Figure 5.5 shows the L-valley scattering rates in a QW with pure Si barriers and wells composed of a material with variable Ge fraction x_w. In such a structure,

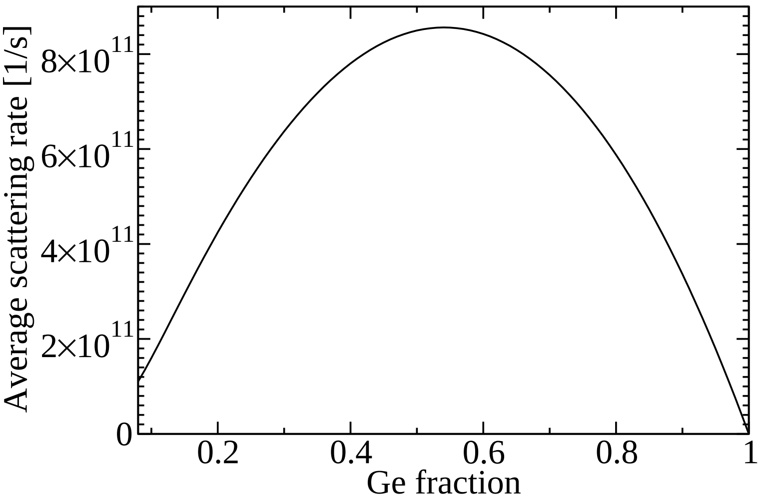


Figure 5.5: Average alloy disorder scattering rate from the second to first L-valley subband in a 10 nm wide QW of variable composition with pure Si barriers

alloy scattering only occurs in the well region. The figure shows that the scattering rate drops to zero when $x_w = 0$ or 1 as expected but the maximum occurs at $x_w = 0.55$ rather than 0.5. This shift may be understood by considering the probability density of electrons within the alloy region. As x_w increases, the quantum well becomes deeper and electrons are better confined within the well. In other words, the $|\psi_f(z)|^2|\psi_i(z)|^2$ term in the scattering matrix element becomes larger within the alloy region, which shifts the peak scattering rate to a larger value of x_w.

These results show that nonradiative alloy disorder scattering in a QCL, is minimised when either pure Si or Ge is used as the well material, rather than an alloy. In principle, it could be eliminated entirely by using Si or Ge as the barrier material too. However, as discussed in chapter 6, growth of alternating Si and Ge epilayers is challenging, especially in the (111) orientation. Consequently, a Si/SiGe or Ge/GeSi configuration is more realistic for QCLs.

Figure 5.6 shows the alloy disorder scattering rate when pure Ge wells are surrounded by $\mathrm{Ge}_{x_b}\mathrm{Si}_{1-x_b}$ barriers with variable x_b. Here, the maximum scattering rate

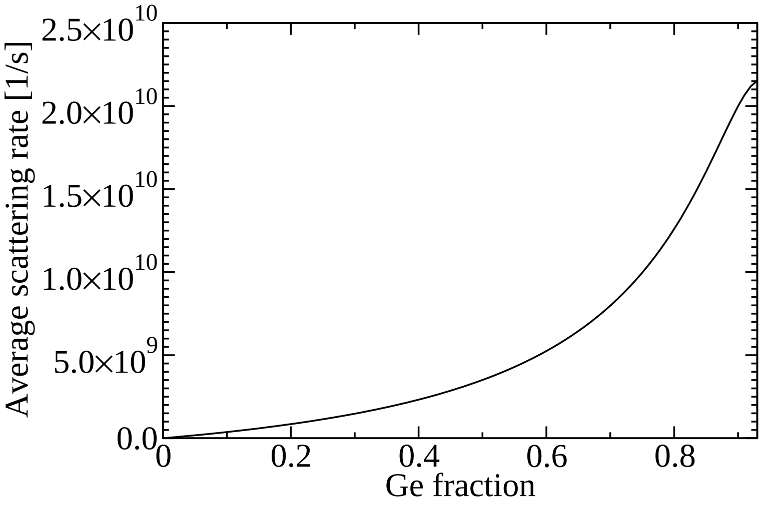


Figure 5.6: Average alloy disorder scattering rate from the second to first L-valley subband in a 10 nm wide Ge QW with variable barrier composition

occurs when $x_b > 0.93$[2]. This large shift away from $x_b = 0.5$ may again be explained by considering the effect on the conduction band offset. As the barrier Ge fraction increases, the quantum well becomes shallower and electrons are able to leak much further into the alloy region in the barriers.

5.4.2 Interface roughness scattering

Real heterojunctions are not perfectly flat and random fluctuations in the interface location cause small shifts in the conduction band potential as shown in fig. 5.7. This leads to the Schrödinger solution varying slightly depending upon the in-plane location of the electron[157, 158].

The fluctuations are usually assumed to have a Gaussian Fourier transform $\Delta_z(\boldsymbol{r})$ with height Δ and correlation length Λ[76, 155, 159] such that

$$\langle \Delta_z(\boldsymbol{r}) \Delta_z(\boldsymbol{r'}) \rangle = \Delta^2 \exp\left(-\frac{|\boldsymbol{r} - \boldsymbol{r'}|}{\Lambda^2} \right). \tag{5.63}$$

---

[2] The graph is plotted over the range where at least two subbands exist within the well. For $x_b > 0.93$, only a single subband exists.

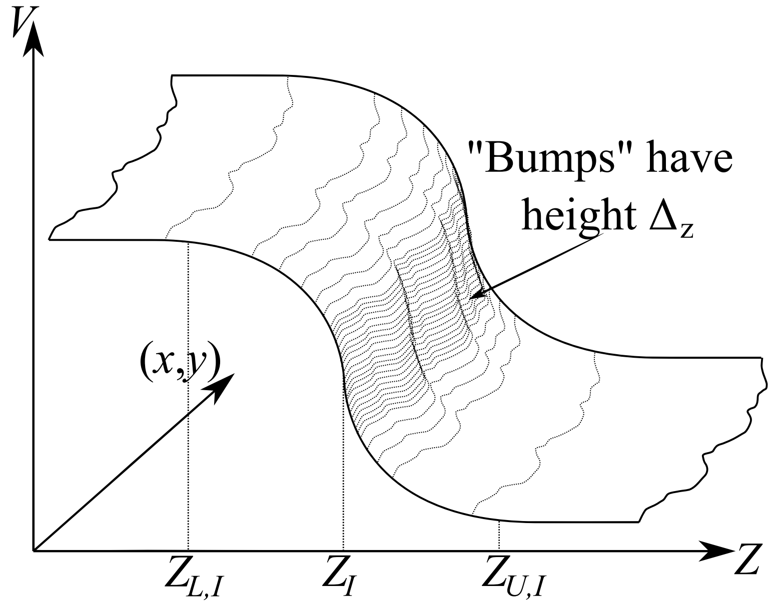


Figure 5.7: Schematic representation of interface roughness at a diffuse heterojunction. The actual conduction band potential deviates from the nominal value depending upon in-plane location, due to random variations in the interface location.

Most previous models have been limited to perfectly abrupt interface geometries[160], and good agreement with experimental results has been achieved[39]. A more general approach has been developed as part of the present work, which allows interface roughness scattering rates to be calculated for *any* interface geometry.

The perturbing potential around an interface I is given by

$$\begin{aligned} \hat{V}_I(\boldsymbol{R}) &= \mathrm{rect}\left(\frac{z - z_I}{z_{U,I} - z_{L,I}}\right) V[\boldsymbol{z} - \Delta_z(\boldsymbol{r})] - V(\boldsymbol{R}) \\ &\approx -\Delta_z(\boldsymbol{r}) \frac{\mathrm{d}V(z)}{\mathrm{d}z} \mathrm{rect}\left(\frac{z - z_I}{z_{U,I} - z_{L,I}}\right), \end{aligned} \tag{5.64}$$

where the interface is defined as a shift in conduction band over the range $(z_{L,I}, z_{U,I})$, centred about a nominal location z_I.

The scattering matrix element is therefore

$$V_{if,I} = -\frac{F_{if,I}}{A} \iint \mathrm{d}^2\boldsymbol{r} \mathrm{e}^{i\boldsymbol{q}\cdot\boldsymbol{r}} \Delta_z(\boldsymbol{r}), \tag{5.65}$$

where

$$F_{if,I} = \int_{z_{L,I}}^{z_{U,I}} \mathrm{d}z \psi_f^*(z) \frac{\mathrm{d}V(z)}{\mathrm{d}z} \psi_i(z). \tag{5.66}$$

The mean-square of the scattering matrix element is now given by

$$\begin{aligned} \left\langle |V_{if}(z_I)|^2 \right\rangle &= \left| \frac{F_{if}(z_I)}{A} \right|^2 \iint \mathrm{d}^2\boldsymbol{r} \iint \mathrm{d}^2\boldsymbol{r'} \Delta_z(\boldsymbol{r}) \Delta_z(\boldsymbol{r'}) \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{u}} \\ &= \frac{|F_{if,I}\Delta\Lambda|^2 \pi}{A} \mathrm{e}^{-q^2\Lambda^2/4} \end{aligned} \tag{5.67}$$

where $\boldsymbol{u} = \boldsymbol{r} - \boldsymbol{r'}$.

Substituting this into Fermi's golden rule gives the scattering rate due to the perturbation at a single interface,

$$W_{if,I}(k_i) = \frac{|F_{if,I}\Delta\Lambda|^2 m_d}{\hbar^3} \Theta(k_\alpha^2)[1 - P_f(k_\alpha)] \int_0^\pi \mathrm{d}\theta \mathrm{e}^{-q_\alpha^2\Lambda^2/4}. \tag{5.68}$$

The integral may be solved quasi-analytically using result 3.339 of Gradshteyn and Ryzhik[161],

$$\int_0^\pi \mathrm{e}^{a\cos\theta} \mathrm{d}\theta = \pi I_0(a) \tag{5.69}$$

where $I_0(a)$ is the regular modified cylindrical Bessel function of zeroth order. The scattering rate at a given interface is now

$$W_{if,I}(k_i) = \frac{\pi m_d (\Delta\Lambda)^2}{\hbar^3} \beta(k_i) \left| F_{if,I} \right|^2, \tag{5.70}$$

where

$$\beta(k_i) = \mathrm{e}^{-(k_i^2+k_\alpha^2)\Lambda^2/4} I_0\left(\frac{k_i k_\alpha \Lambda^2}{2}\right) \Theta(k_\alpha^2)[1 - P_f(k_\alpha)] \tag{5.71}$$

The total scattering rate, assuming that the roughness of separate interfaces is completely uncorrelated, is obtained by the summation

$$W_{if}(k_i) = \frac{\pi m_d (\Delta\Lambda)^2}{\hbar^3} \beta(k_i) \sum_I \left| F_{if,I} \right|^2, \tag{5.72}$$

The general result derived above may be checked against previous models which assume that interfaces are perfectly abrupt[7]. Here, the potential at an interface is a Heaviside step function $V(z) = V_0 \Theta(z - z_I)$, where V_0 is the step in conduction band potential at the interface. The perturbation is now given by

$$\hat{V}_I(\boldsymbol{R}) = V_0 \,\mathrm{rect}\left\{ \frac{1}{\Delta_z(\boldsymbol{r})} \left[z - \left(z_I + \frac{\Delta_z(\boldsymbol{r})}{2} \right) \right] \right\}. \tag{5.73}$$

For a small perturbation, $\Delta_z(\boldsymbol{r}) \to 0$, and the perturbing potential becomes

$$\hat{V}_I(\boldsymbol{R}) = V_0 \Delta_z(\boldsymbol{r}) \delta(z - z_I), \tag{5.74}$$

i.e. the perturbation acts only at the interface. The scattering matrix element is therefore

$$V_{if,I} = \frac{V_0}{A} \psi_f^*(z_I) \psi_i(z_I) \iint \mathrm{d}^2\boldsymbol{r} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} \Delta_z(\boldsymbol{r}), \tag{5.75}$$

which is the expression seen in previous works[155]. This can be used to specify an approximate version of the matrix element in eqn. 5.66

$$F_{if,I} \approx V_0 \psi_f^*(z_I) \psi_i(z_I), \tag{5.76}$$

which is valid in the limit of abrupt interfaces.

Figure 5.8 shows that the scattering rate increases with roughness height and is maximised at correlation lengths around $\Lambda = 30$ Å. The parameters have been fitted

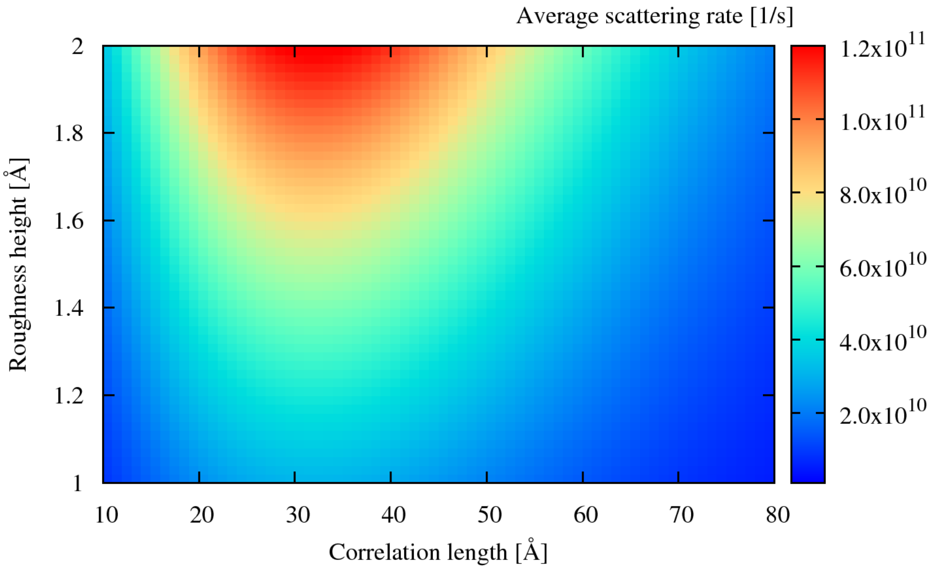


Figure 5.8: Interface roughness scattering rates from second to first subband in a 10 nm wide (001) Ge QW with 10 nm wide $Si_{0.8}Ge_{0.2}$ barriers as a function of the roughness height and correlation length.

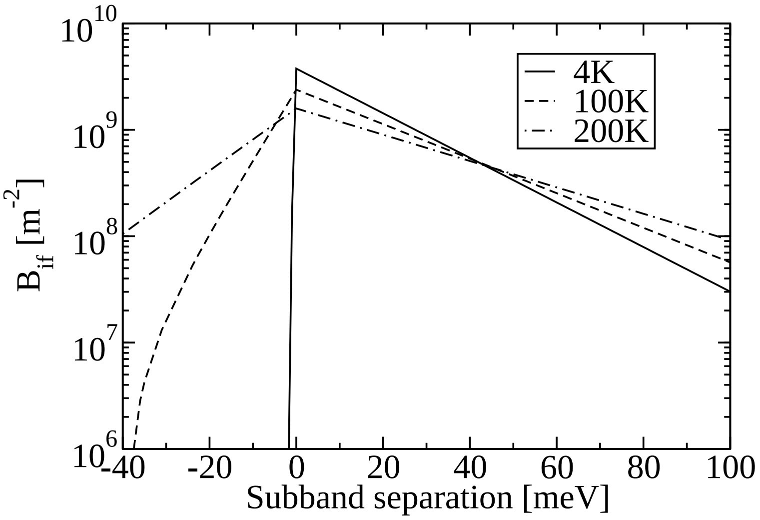


Figure 5.9: Interface roughness scattering matrix element B_{if} as a function of E_{if} for an arbitrary (001) oriented Ge/GeSi heterostructure with Λ=50 Å. Positive subband separations indicate that the initial subband minimum exceeds that of the final subband.

to experimental data in several publications, with Δ typically in the range 1 to 5 Å and Λ between 50 and 300 Å[158, 162, 163].

In a recent study, Califano showed that a choice of Δ=1.2–1.5 Å and Λ=40–60 Å matches experimental data for asymmetric *p*-type Si/SiGe heterostructures[39]. The combination $\Delta, \Lambda = (1.4, 50)$ Å (yielding the median scattering rate from Califano's parameter space) was chosen as an estimate for the *n*-type systems considered in the present work.

To analyse the scattering behaviour it is helpful to rewrite the average rate (eqn. 5.22) in the form

$$\overline{W_{if}} = \frac{m_d(\Delta\Lambda)^2}{n_i\hbar^3} B_{if} \sum_I |F_{if,I}|^2 , \tag{5.77}$$

where

$$B_{if} = \int_{k^-}^{k^+} k_i \mathrm{d}k_i P_i(k_i)\beta(k_i). \tag{5.78}$$

The B_{if} function (fig. 5.9) depends upon the transition energy and is independent of the heterostructure geometry. Its peak value occurs when both initial and final

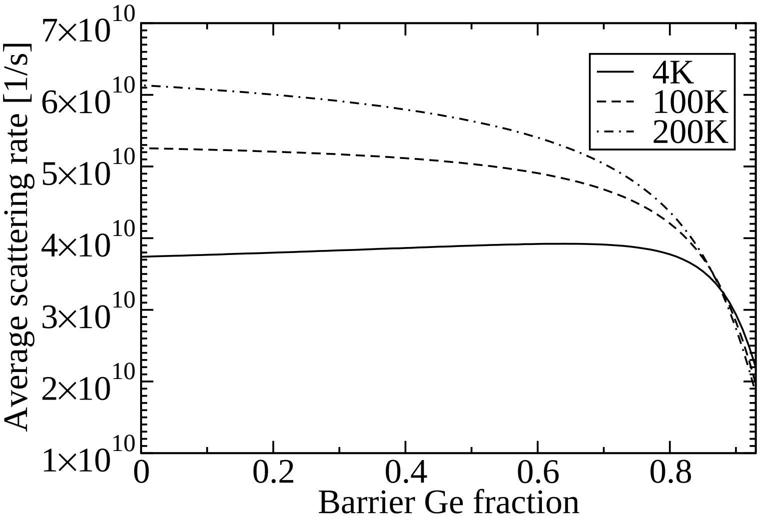


Figure 5.10: Average interface roughness scattering rate in a 10 nm wide Ge QW with Λ=50 Å as a function of barrier Ge fraction at electron temperatures of 4,100 and 200 K.

subbands are at the same energy and the scattering rates decrease approximately exponentially as the subband separation grows. The function drops much more sharply when the initial subband lies below the final subband. This is because elastic scattering is forbidden for initial states near the subband minimum, which contain most of the carrier population.

The complete effect of structure geometry upon the average scattering rate can be broken down into two factors. Firstly, the B_{if} factor is maximised when subband separations are small, as discussed above. Secondly, the $F_{if,I}$ terms are maximised when wavefunctions overlap strongly with interfaces. The wavefunction overlap is greatest when barriers are thin, when the conduction band offset is small or when the subband minimum lies close in energy to the barrier potential. Fig. 5.10 shows that there is little change in the scattering rate as the barrier Ge increases from 0 to around 70% as the decrease in conduction band offset is balanced by an increase in wavefunction overlap. For $x_b \gtrsim 70\%$, the scattering rate decreases as the conduction band offset drops toward zero.

5.5 Electron–phonon scattering

Vibrations in the crystal lattice may be represented by a quasi-particle called the *phonon.* These vibrations cause periodic variations in strain, which in turn perturbs the conduction band potential and induces scattering.

5.5.1 Intravalley acoustic phonon scattering

For quantum wells made of pure Si or Ge, it is possible to model the crystal lattice as a mass-spring system. For longitudinally oscillating atoms, the displacement of atom j is related to that of neighbouring atoms by[90]

$$m\frac{\mathrm{d}^2u_j}{\mathrm{d}t^2} = K[(u_{j-1} - u_j) - (u_j - u_{j+1})], \tag{5.79}$$

where m is the mass of each atom, and K is the elastic constant of the bonds between the atoms. This is referred to as a longitudinal acoustic (LA) phonon.

Bloch's theorem states that $u_{j+1} = u_j\mathrm{e}^{\mathrm{i}qa}$, where q is the wavevector of the phonon and a is the bond length. Assuming that solutions take the plane-wave form $u_j = U_0 \exp[\mathrm{i}(qaj - \omega_q t)]$, the dispersion relation becomes

$$\omega_q = 2\sqrt{\frac{K}{m}}\left|\sin\frac{qa}{2}\right|. \tag{5.80}$$

For small wavevectors, this dispersion relation is approximately linear *i.e.* $\omega_q = v_s q$, where $v_s = a\sqrt{K/m}$ is the speed of sound in the material. For wavevectors close to the Brillouin zone edge however, the phonon frequency becomes fixed at $\omega_q = \omega_0 = 2\sqrt{K/m}$.

As the atoms are undergoing simple harmonic motion, their average kinetic and potential energies are both $mU_0^2\omega_q^2/4$. For a system of volume Ω with mass-density ρ, the number of atoms is given by $N = \Omega\rho/m$ and the total energy is given by

$\Omega\rho U_0^2\omega_q^2/2$. Equating this to the energy of the quantised system gives the amplitude

$$U_0 = \sqrt{\frac{2\hbar}{\Omega\rho\omega_q}}. \tag{5.81}$$

For long wavelengths, there is almost a continuous distribution of atoms and

$$u(z) \approx U_0 \mathrm{e}^{\mathrm{i}(qz-\omega_q t)}. \tag{5.82}$$

The strain may be written as a derivative

$$\varepsilon(z) = \frac{\partial u}{\partial z} = \mathrm{i}U_0 q \mathrm{e}^{\mathrm{i}(qz-\omega_q t)}. \tag{5.83}$$

In chapter 2, it was shown that strain leads to a shift in the conduction band potential. In a zeroeth order model, the potential is assumed to vary proportionally to the strain, such that

$$\hat{V}(z) = D_{\mathrm{ac}}\varepsilon(z) = \mathrm{i}D_{\mathrm{ac}}U_0 q \mathrm{e}^{\mathrm{i}(qz-\omega_q t)}, \tag{5.84}$$

where D_{ac} is a deformation potential. These oscillations in the conduction band potential act as a perturbation to the Hamiltonian and hence lead to scattering.

In addition to the longitudinal acoustic (LA) phonon described so far, there are transverse acoustic (TA) phonons which represent vibrations perpendicular to the direction of wave propagation. At small wavevectors, these acoustic phonons are approximately identical, and a single deformation potential accurately describes their effect on the Hamiltonian.

Price[164] shows that in general, a two-dimensional electron system interacting with a three-dimensional phonon has a scattering rate of the form

$$W_{if} = \int C^2(q)|G_{if}(q_z)|^2 \mathrm{d}q_z, \tag{5.85}$$

where $C(q)$ is a coupling constant and $G_{if}(q_z)$ is an overlap integral given by

$$G_{if}(q_z) = \int \mathrm{d}z \psi_f^*(z)\psi_i(z)\mathrm{e}^{\mathrm{i}q_z z}. \tag{5.86}$$

For wavevector independent interactions, $C(q)$ may be taken outside the integral, giving

$$W_{if} = C^2 \int |G_{if}(q_z)|^2 \mathrm{d}q_z. \tag{5.87}$$

To a good approximation, the substitution $\int |G_{if}(q_z)|^2 \mathrm{d}q_z \approx 2\pi F_{if}$ can be made[165], where

$$F_{if} = \int \mathrm{d}z \psi_f^2(z) \psi_i^2(z). \tag{5.88}$$

The resulting scattering rate takes the form $W_{if} = 2\pi C^2 F_{if}$. Although the coupling constant C is uniquely defined for each type of electron–phonon interaction, it is proportional to the phonon population in all cases. The number of phonons is given by the Bose-Einstein distribution

$$N_q(\omega_q) = \left[\exp\left(\frac{\hbar\omega_q}{k_B T}\right) - 1\right]^{-1}, \tag{5.89}$$

where T is the temperature of the crystal lattice.

For acoustic phonon interactions, the phonon wavevector is small and the equipartition approximation, $\hbar\omega_q \ll k_B T$, can be applied. The Bose-Einstein distribution simplifies to

$$N_q(\omega_q) = \frac{k_B T}{\hbar\omega_q}. \tag{5.90}$$

It was noted above that the acoustic phonon frequency is approximately $\omega_q = v_s q$ at small wavevectors, which gives the final expression

$$N_q(\omega_q) = \frac{k_B T}{\hbar v_s q}. \tag{5.91}$$

The combined absorption and emission rate for low-wavevector acoustic phonons is[165]

$$W_{if}(k_i) = \Theta(k_\alpha^2) \frac{m_d D_{ac}^2 k_B T}{\hbar^3 \rho v_s^2} F_{if} \left[1 - P_f(k_\alpha)\right]. \tag{5.92}$$

Deformation potentials of 1.2 eV for Si and 2.5 eV for Ge have been found to give good agreement with experimental results[166]. The other parameters are given for Si (Ge) as ρ=2.33 (5.32) gcm^{-3} and v_s=5870 (3810) ms^{-1}[166].

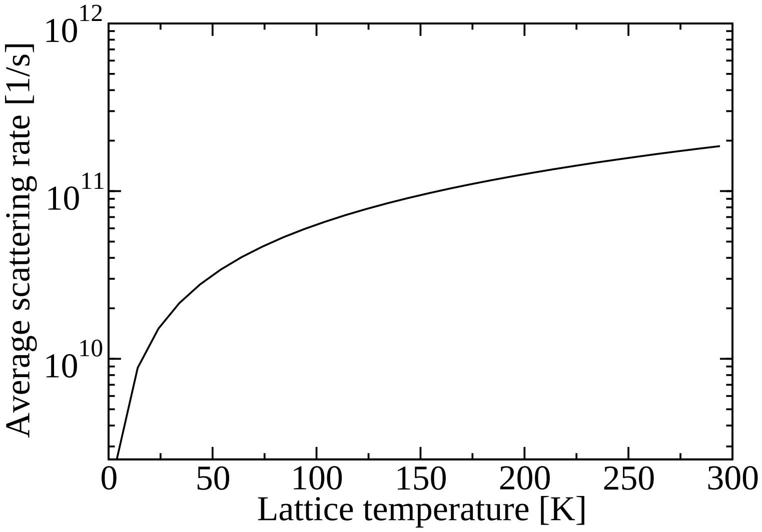


Figure 5.11: Average acoustic phonon scattering rate as a function of lattice temperature in a Ge QW identical to that in the previous section.

Fig. 5.11 shows the expected linear increase in scattering rate with respect to lattice temperature. The rate is independent of subband separation for downward transitions as the final wavevector will always be real.

5.5.2 Optical and intervalley phonon scattering

The acoustic phonon model considered so far is only valid for acoustic phonons with low wavevectors. For optical phonons and any phonons with wavevectors near the Brillouin zone edge, the linear approximation $\omega_q = qv_s$ breaks down. Instead, the gradient of the phonon dispersion curve is small and it is better to approximate the phonon frequency as a constant value ω_0. Due to the symmetry of the system, optical phonon processes are forbidden in Δ valleys but are significant in L valleys[167].

The same general form of the scattering rate expression exists as for acoustic phonons, but the coupling coefficient is derived from an optical phonon deformation potential $\boldsymbol{D_0}$. This deformation potential differs from D_{ac} in that it describes the shift in conduction band potential as a function of atomic displacement rather than strain, such that $V = \boldsymbol{D_0} \cdot \boldsymbol{u}$. The optical phonon energy may be quite large in comparison with k_BT and the equipartition approximation no longer applies. The phonon population is therefore included explicitly in the scattering rate expression.

The scattering rate now becomes[165]

$$W_{if}^{\mp}(k_i) = \Theta(k_{\alpha^\pm}^2)\frac{n_{\text{dest}} m_d D_0^2}{2\hbar^2 \rho \omega_q}\left[N_q(T) + \frac{1}{2} \mp \frac{1}{2}\right] F_{if}\left[1 - P_f(k_{\alpha^\pm})\right], \tag{5.93}$$

where n_{dest} is the number of destination valleys for the process. The upper sign of the ± and ∓ symbols refer to the absorption of a phonon while the lower sign refers to an emission.

Intervalley scattering may occur between pairs of subbands in two different valley sets. For the low-lying Δ and L valleys considered in this work, the $\Delta \leftrightarrow \Delta$, $\Delta \leftrightarrow L$ and $L \leftrightarrow L$ interactions may all affect the subband populations. When uniaxial strain induced degeneracy splitting is taken into account, this leads to an enormous number of possible interactions. In the present work, the problem is simplified considerably by only including intervalley scattering within a degenerate valley set. This approximation is justified by noting that separations of at least 60 meV exist between non-degenerate valleys in SiGe based QCLs (chapter 2), which significantly reduces the respective scattering rates.

Intervalley scattering in Ge

The distance between any pair of L valleys in reciprocal space is identical and therefore only one intervalley phonon wavevector need be considered. The number of destination valleys for $L \to L$ scattering is $n_{\text{dest}} = 3$. To the author's best knowledge no attempt has been made to fit the contributions of individual phonon branches to experimental data. Instead, a phenomenological approach has been taken, in which the combined contributions of all phonon branches have been treated as a single zeroeth order interaction with energy $\hbar\omega_q = 27.56$ meV and a deformation potential of $D_0 = 3.0 \times 10^8$ eVcm^{-1}[168].

Figure 5.12 shows the $L \to L$ scattering rates in a Ge QW as a function of well

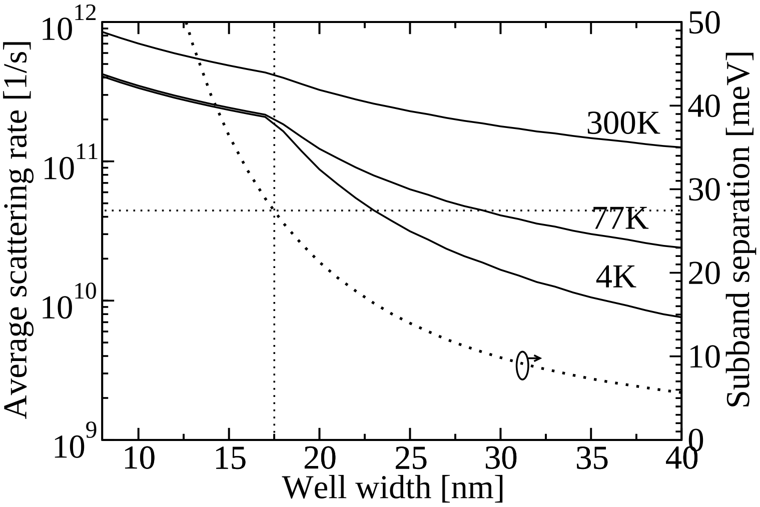


Figure 5.12: Average intervalley scattering rates from second to first subband in a Ge QW with Si barriers as a function of subband separation. Results are shown for lattice temperatures of 4, 77 and 300 K, assuming that the electron temperature is given by $T_e = T+4\,\mathrm{K}$. The fine dotted line shows that a “kink” in the scattering rate curve occurs when the subband separation equals the intervalley phonon energy of $\hbar\omega_q = 27.56\,\mathrm{meV}$.

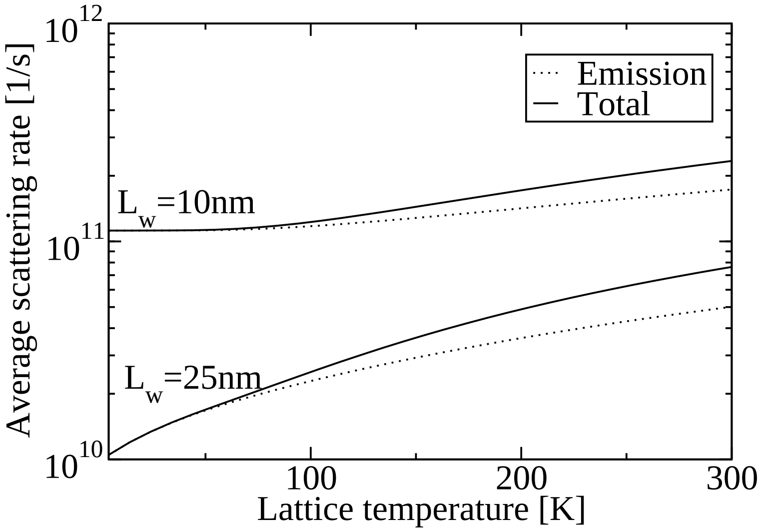


Figure 5.13: Average intervalley scattering rate from second to first subband in a Ge QW with Si barriers as a function of lattice temperature. Results are shown for well widths of 10 nm ($E_{21} > \hbar\omega_q$) and 25 nm ($E_{21} < \hbar\omega_q$).

width. A "kink" in the curve occurs when the subband separation equals the phonon energy (when $L_w = 17.5\,\text{nm}$). This is because phonon emission is only permitted when the final state is real. For subband separations below the phonon energy, electrons with low initial wavevector may not take part in phonon emission and the total scattering rate drops rapidly.

Figure 5.13 shows that the phonon scattering rates increase with temperature. At low temperature, $\lim_{T\to 0}[N_q(T)] \approx 0$, meaning that absorption rates are extremely low and only phonon emission is permitted. As the lattice temperature increases, the phonon absorption becomes much more significant.

Intravalley scattering in Ge

Due to symmetry, intravalley optical phonon scattering is permitted in Ge L valleys, but not in the Si Δ valleys[167]. For intravalley processes, n_{dest}=1. A deformation potential of $D_0 = 3.5 \times 10^8\,\text{eVcm}^{-1}$ and a phonon energy of $\hbar\omega_q = 37.04\,\text{meV}$ has

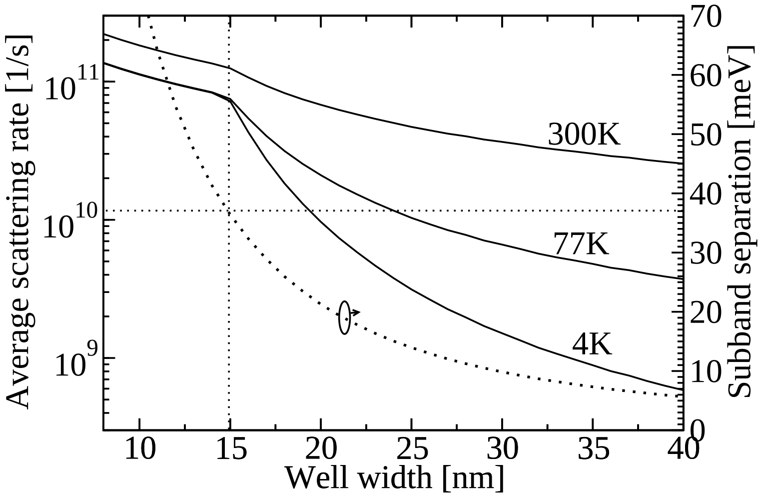


Figure 5.14: Average intravalley optical phonon scattering rate from second to first subband in a Ge QW as a function of well width.

been shown to give a good match to experimental data[166]. Figure 5.14 shows that the scattering rates take the same form as for intervalley processes, but with a smaller magnitude. The kink in the curve is also shifted due to the change in phonon energy.

Intervalley scattering in Si

The difference in wavevector between a pair of Si Δ valleys may be denoted g or f depending on their relative locations in reciprocal space, as illustrated in fig. 5.15. g-transitions transfer electrons to the opposite valley along the same crystallographic axis whereas f-transitions transfer electrons to the nearest valley along a $\{110\}$ direction. For g-transitions $n_{\text{dest}} = 1$ while for f-transitions $n_{\text{dest}} = 4$.

Symmetry selection rules state that only high-energy phonons are permitted within this zero order model, and that all other interactions are forbidden. The permitted intervalley phonon interactions are g-LO ($\hbar\omega_q = 63.2\,\text{meV}$), f-LA ($\hbar\omega_q = 46.3\,\text{meV}$) and f-TO ($\hbar\omega_q = 59.1\,\text{meV}$) and a single deformation potential of $D_0 = 3.4 \times 10^8\,\text{eVcm}^{-1}$ has been shown to give good agreement with experimental data

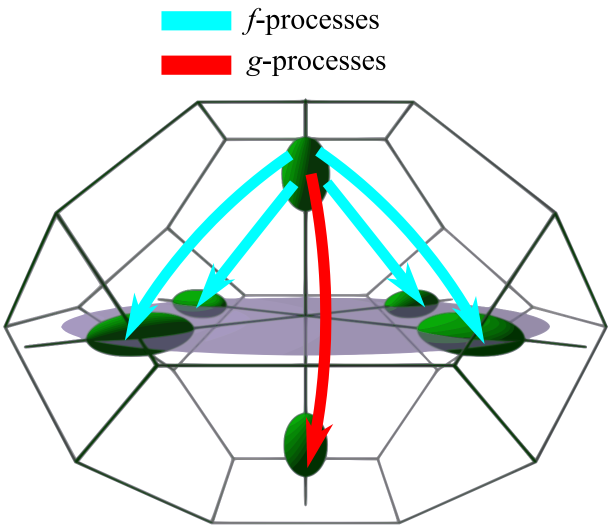


Figure 5.15: Schematic illustration of f- and g-type intervalley scattering processes between Δ valleys.

for all three processes[169].

Ferry[170] noted that the “forbidden” low-energy interactions can be included by expanding the deformation potential model to include first order terms in phonon wavevector. The first order optical deformation potential D_1 relates the shift in conduction band potential to the first derivative of atomic displacement (*i.e.* strain) and is therefore analogous to the D_{ac} deformation potential. The first order scattering rate expression is given by[171]

$$\begin{aligned} W_{if}^{\mp}(k_i) &= \Theta(k_{\alpha\pm}^2)\frac{n_{\mathrm{dest}}m_{d,f}D_1^2}{\hbar^2\rho\omega_q}\left[N_q(T)+\frac{1}{2}\mp\frac{1}{2}\right] \\ &\times \left[\frac{e}{\hbar^2}(E_{k,i}m_{d,i}+E_{k,f}m_{d,f})F_{if}-\frac{H_{if}}{2}\right][1-P_f(k_{\alpha\pm})], \end{aligned} \tag{5.94}$$

where

$$H_{if} = \int \mathrm{d}z\psi_{if}(z)\frac{\mathrm{d}^2\psi_{if}(z)}{\mathrm{d}z^2} \tag{5.95}$$

The first-order phonon interactions in Si are g-TA ($\hbar\omega_q = 11.4\,\mathrm{meV}$), g-LA ($\hbar\omega_q = 18.8\,\mathrm{meV}$) and f-TA ($\hbar\omega_q = 21.9\,\mathrm{meV}$), all with a deformation potential of $D_1 = 3.0\,\mathrm{eV}$[169].

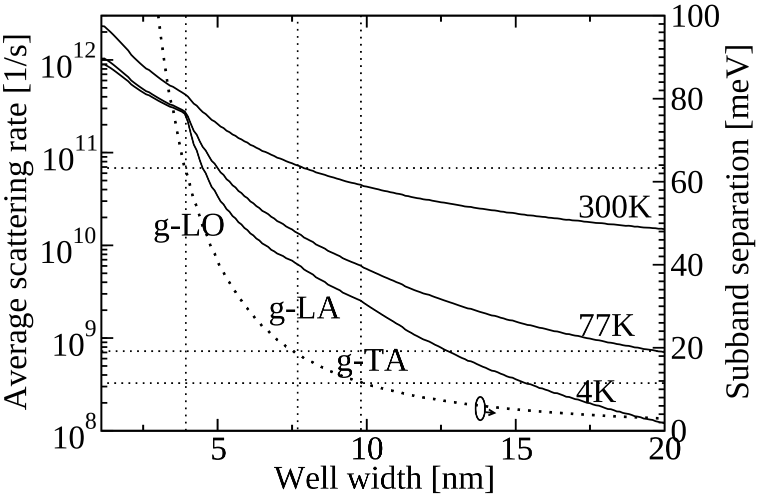


Figure 5.16: Average $\Delta_2 \rightarrow \Delta_2$ scattering rates from second to first subband in (001) Si. Dotted lines show that "kinks" in the scattering rate coincide with a subband separation equal to each of the g-phonon energies.

In (001) oriented Si, the Δ_2 valleys are lowest in energy and it is assumed that only $\Delta_2 \rightarrow \Delta_2$ intervalley interactions are significant in this work. As such, only g-transitions are considered. Figure 5.16 shows the total scattering rate as a function of well width in a Ge/Si/Ge well. A kink in the plot occurs at $L_w = 3.9\,\text{nm}$, which corresponds to $E_{21} = 63.2\,\text{meV}$ (the g-LO phonon energy). The first order g-TA and g-LA processes are much slower than g-LO, and only give rise to very small features in the scattering plot.

In the (111) orientation, all six Δ valleys are degenerate and therefore both g and f processes are important. Figure 5.17 shows the total intervalley scattering rate in a (111) oriented well as a function of the well width. Again, features can be seen in the plot corresponding to each of the zeroeth-order processes, while the first-order processes have a much weaker effect.

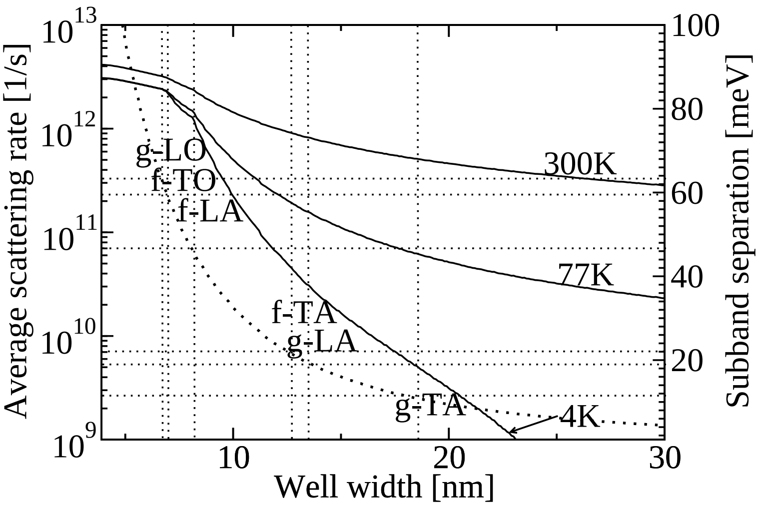


Figure 5.17: Average $\Delta \rightarrow \Delta$ scattering rates from second to first subband in a (111) oriented Ge/Si/Ge well. Dotted lines show that "kinks" in the scattering rate coincide with a subband separation equal to each of the g-phonon energies.

5.6 Total scattering rates

It is possible to calculate the total intersubband scattering rate by summing over all possible processes, giving

$$\frac{1}{\tau_{if}} = \sum_{p} \overline{W_{if}^{p}}, \tag{5.96}$$

where p denotes a particular scattering process. In a QCL, the wavefunctions overlap multiple QWs and numerous parameters are required to determine the scattering rates between two subbands. It is therefore impractical to present a full analysis in the present work. Instead, some basic properties of the total scattering rate are determined for simple QW structures in this chapter. The observable effects of scattering in a QCL (*i.e.* gain and current density) are then discussed in chapters 7 and 8.

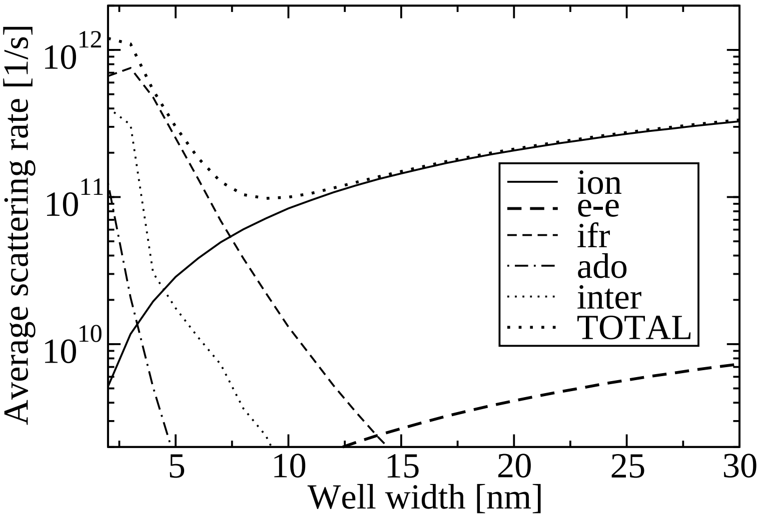


Figure 5.18: Scattering rates as a function of well width in a (001) oriented Si-based QW with 50% Ge barriers. The lattice temperature was fixed at $T = 4\,\mathrm{K}$ and the electron temperature at $T_e = 24\,\mathrm{K}$. Dopants were spread evenly through the well region at a concentration of $n_{2\mathrm{D}} = 1 \times 10^{11}\,\mathrm{cm}^{-3}$. The abbreviations in the legend are as follows: "ion"=ionised impurity, "ado"=alloy disorder, "e-e"=electron–electron, "ifr"=interface roughness, "inter"=intervalley phonon.

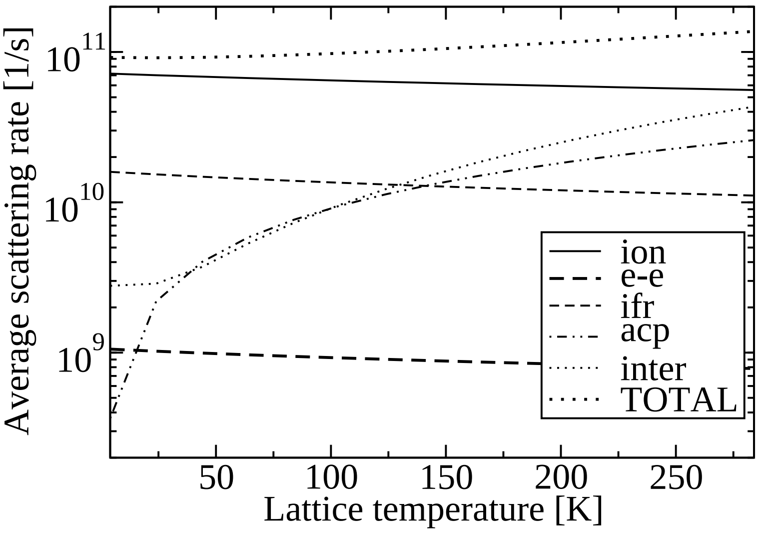


Figure 5.19: Scattering rates as a function of lattice temperature in a 9.2 nm wide (001) oriented Si QW with 50% Ge barriers at $T_e = T + 20\,\mathrm{K}$. Dopants are spread evenly through the well region at a concentration of $n_{2\mathrm{D}} = 10^{11}\,\mathrm{cm}^{-2}$. The abbreviation “acp” refers to acoustic phonon scattering.

5.6.1 (001) Si/SiGe quantum wells

Scattering was simulated in a Si well with 10 nm wide $\mathrm{Si}_{0.5}\mathrm{Ge}_{0.5}$ barriers and dopants spread throughout the well region at a concentration of $n_{2\mathrm{D}} = 10^{11}\,\mathrm{cm}^{-2}$. The width of the well was varied and the resulting scattering rates are plotted in fig. 5.18.

Ionised impurity scattering dominates for wide wells due to the large region of overlap between the wavefunctions and the dopants. The converse is true for interface roughness scattering, which becomes the dominant mechanism for well widths below 7.5 nm. Intervalley phonon scattering is also significant in narrow wells as the large energy separation permits all initial states to take part in phonon emission interactions. A local minimum in the total scattering rate occurs at well widths around 9.2 nm. This corresponds to a subband separation of $E_{21} = 13.5\,\mathrm{meV}$ or 3.3 THz.

Figure 5.19 shows the effect of lattice temperature upon the total scattering rate

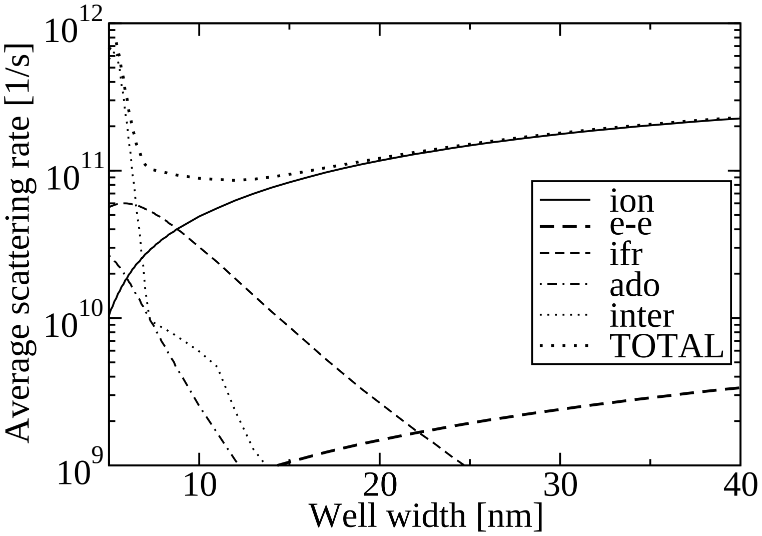


Figure 5.20: Scattering rate as a function of well width in a (111) oriented Si-based QW with 50% Ge barriers and dopants spread evenly through the well region at a concentration of $n_{2\mathrm{D}} = 1 \times 10^{11}\,\mathrm{cm}^{-3}$. A lattice temperature of 4 K and an electron temperature of 24 K were used. The abbreviation "ado" refers to alloy disorder scattering. All other line styles are as defined in fig. 5.18

for a 9.2 nm well. It can be seen that phonon scattering interactions increase with lattice temperature and represent a significant proportion of the total scattering rate for $T \gtrsim 100\,\mathrm{K}$. The other scattering rates decrease slightly due to final state blocking.

5.6.2 (111) Si/SiGe quantum wells

Figure 5.20 shows the scattering rates as a function of well width in a (111) $\mathrm{Si_{0.5}Ge_{0.5}/Si/Si_{0.5}Ge_{0.5}}$ QW with 10 nm thick barriers at a lattice temperature of 4 K. Rapid f-LA phonon interactions give rise to large scattering rates in wells narrower than 7 nm. The reduction in barrier potential reduces the interface roughness scattering rate compared with the (001) oriented system. The minimum scattering rate lies between well widths of 7 and 15 nm, which corresponds to a subband

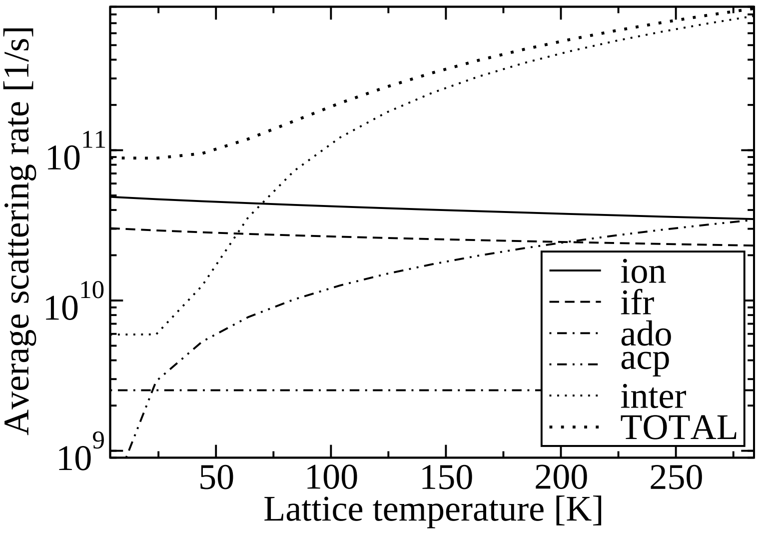


Figure 5.21: Scattering rates as a function of lattice temperature in a 10 nm wide (111) oriented Si QW with 50% Ge barriers. Dopants were spread evenly through the well region at a concentration of $n_{2D} = 10^{11}\,\mathrm{cm}^{-2}$.

separation of 13–40 meV or an emission energy of 3.1–9.7 THz. Ionised impurity scattering causes a rise in the total scattering rate in wider wells.

Figure 5.21 shows the scattering rates in a 17 nm (111) oriented $\mathrm{Si_{0.5}Ge_{0.5}/Si/Si_{0.5}Ge_{0.5}}$ well as a function of lattice temperature, assuming $T_e = T + 20\,\mathrm{K}$. Intervalley phonon scattering overtakes ionised impurity scattering as the dominant process for $T > 70\,\mathrm{K}$.

5.6.3 (001) Ge/GeSi quantum wells

Figure 5.22 shows the scattering rates in a $\mathrm{Si_{0.5}Ge_{0.5}/Ge/Si_{0.5}Ge_{0.5}}$ QW as a function of well width at T=4 K and T_e=24 K. Optical and intervalley phonon scattering dominates at well widths below 17 nm, while ionised impurity scattering is fastest at higher well widths. A local minimum in scattering rate occurs at well widths around 20 nm, which corresponds to a subband separation of 19.3 meV or 4.7 THz.

Figure 5.23 shows the scattering rates in a 20 nm well. It can be seen that

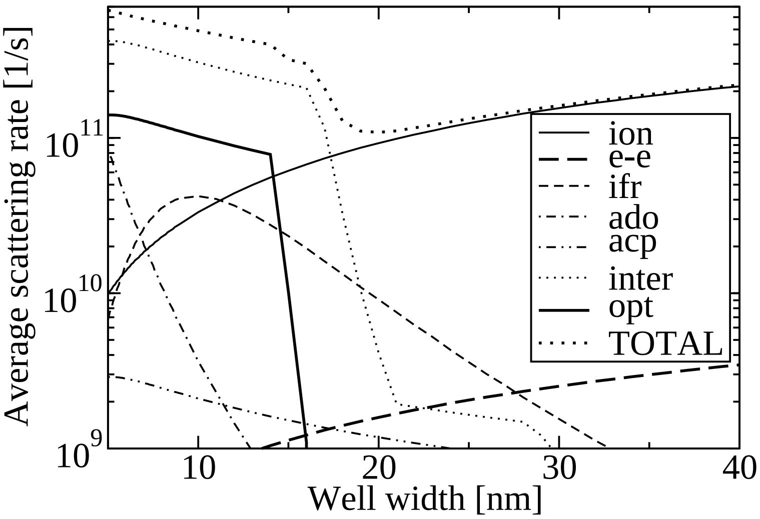


Figure 5.22: Scattering rates as a function of well width in a (001) oriented $Si_{0.5}Ge_{0.5}$/Ge/$Si_{0.5}Ge_{0.5}$ QW with dopants spread evenly through the well region at a concentration of $1 \times 10^{11}\,\mathrm{cm}^{-3}$. A lattice temperature of 4 K and an electron temperature of 24 K were used.

intervalley phonon scattering overtakes ionised impurity scattering as the dominant mechanism for $T > 45\,\mathrm{K}$.

5.7 Conclusion

Fermi's golden rule has been used to derive semi-classical expressions for scattering rates in SiGe based heterostructures. The mechanisms considered were categorised as either Coulombic interactions, structural imperfection interactions or electron–phonon interactions.

The computational challenges associated with Coulombic interactions were addressed, and substantial improvements in computational efficiency were developed without any loss of precision. It was shown that Coulombic interactions are fastest between energetically similar subbands in highly doped structures. Both electron–electron and ionised impurity scattering rates scale linearly with respect to sheet

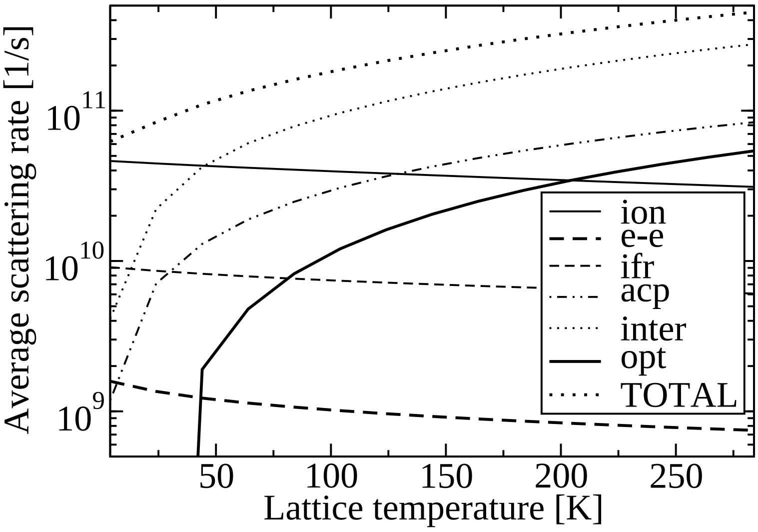


Figure 5.23: Scattering rates as a function of lattice temperature in a 20 nm wide (001) oriented $\mathrm{Si_{0.5}Ge_{0.5}/Ge/Si_{0.5}Ge_{0.5}}$ QW. An electron temperature of $T_e = T + 20\,\mathrm{K}$ was used. Dopants were spread evenly through the well region at a concentration of $n_{2\mathrm{D}} = 10^{11}\,\mathrm{cm}^{-2}$

doping concentration, but ionised impurity scattering is also heavily dependent upon the distribution of dopants. As a result, ionised impurity scattering is slower than electron–electron scattering in δ-doped structures, while the converse applies for evenly doped structures.

Alloy disorder scattering only occurs in SiGe alloys and it was shown that the scattering rates can be reduced significantly by using either pure Si or Ge in the well regions of a QCL. The situation was found to more complicated with respect to the barrier material. However, selecting a barrier alloy that is substantially different from the well material results in a large confining potential and the wavefunctions are less able to penetrate into the alloy region. This approach was shown to reduce the alloy disorder scattering.

A generalised interface roughness scattering model was developed, which allows scattering rates to be calculated for structures with arbitrary interface geometries. The rates were found to be fastest between energetically similar subbands in structures with high, thin barriers. Higher rates may exist in QCLs than in a single QW, as the wavefunctions may cross multiple interfaces.

Intravalley acoustic phonon scattering was shown to increase linearly with respect to the lattice temperature and to be approximately independent of the subband separation. Intervalley phonon scattering was considered in Si by fitting deformation potentials to branches of the phonon dispersion curves and was shown to be highly dependent upon subband separation. The rates were shown to drop rapidly if the subband separation decreases below the phonon energy, as transitions from low-wavevector initial states are forbidden due to the requirement for energy conservation. f-processes, which transfer electrons to perpendicular valleys in reciprocal space were shown to give rise to rapid scattering in (111) oriented Si Δ valleys. However, these processes were assumed to be negligible in (001) Si due to the large uniaxial strain splitting between the Δ_2 and Δ_4 valleys. In Ge, a phenomenological approach to intervalley and optical phonon scattering was taken, in which a single phonon energy and deformation potential were used to explain each process.

It was shown that ionised impurity scattering dominates in wide, evenly doped wells. In narrow wells, intervalley electron–phonon interactions dominate due to the large subband separation. A local minimum in the scattering rate was shown to exist for each material system. In (001) Si and Ge, the local minima are quite well defined at around L_w=9.2 nm (f_{21}=3.3 THz) and 20 nm (4.7 THz) respectively, whereas in (111) Si, the minimum extends between L_w=13 and 40 meV (f_{21}=3.1–9.7 THz). This implies that optimal QCL performance may be achieved around these frequencies as reduced non-radiative scattering rates may improve population inversion and hence increase the gain as discussed in chapter 7.

Chapter 6

Crystal growth related issues

In the work presented so far, a model has been developed to describe the band structure and charge transport characteristics of a QCL. It is now necessary to describe how this model applies to a realisable system. Despite the constant advancement of growth technology, there are fundamental and practical limitations on the thickness of layers and the step in alloy fraction across an interface.

Strain is unavoidable in SiGe heterostructures, and its effect upon band structure was shown to be a crucial factor in establishing a usable conduction band offset in chapter 2. However, strained systems are also susceptible to mechanical instability through the formation of dislocations. In this chapter, it is shown that this issue may be resolved by careful selection of the substrate material and the relevant technology for developing virtual substrates is reviewed.

Finally, experiments have shown that real interfaces between SiGe layers may not be perfectly abrupt. This has profound implications in terms of the band structure and scattering rates. A set of figures of merit are defined in this chapter to describe the tolerance of a system to diffuse interface geometries. The results are then compared for equivalent coupled QW systems in each material configuration.

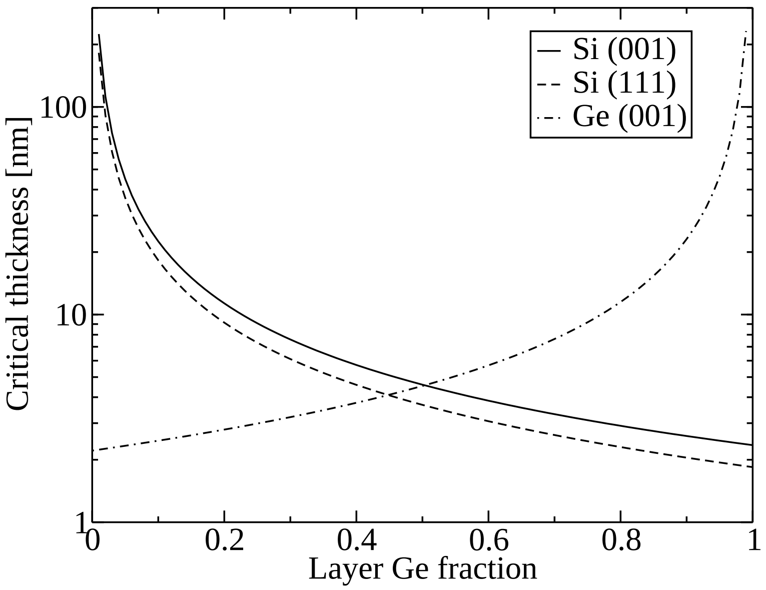


Figure 6.1: Critical thickness of layers with varying Ge fraction, when grown upon Si (001), Si (111) or Ge (001) substrates

6.1 Strain balancing

As shown in chapter 2, mechanical stresses due to the lattice mismatch between Si and Ge lead to unavoidable strain in SiGe heterostructures. In reality however, strained layers may only be grown up to a critical thickness h_c, beyond which the lattice relaxes via the formation of defects. This may have undesirable effects upon the electronic and optical properties of the structure and it is therefore necessary to limit layer widths to below their critical thicknesses [88].

A basic model of the critical thickness, derived by Van der Merwe [172] assumes that dislocations form in a periodic array. Here,

$$h_c \approx \frac{19}{16\pi^2}\left(\frac{1+\nu}{1-\nu}\right)\left(\frac{b}{\varepsilon_{\|}}\right), \tag{6.1}$$

where $b = a/\sqrt{2}$ and ν is the Poisson ratio[88]:

$$\nu = \frac{\varepsilon_{\perp}}{\varepsilon_{\perp} - 2\varepsilon_{\|}}. \tag{6.2}$$

$\varepsilon_\perp$ and $\varepsilon_{||}$ are defined as the strain tensor elements perpendicular and parallel to the interface respectively. Fig. 6.1 shows the critical thickness of varying alloy materials upon a range of substrates. All three curves show that an infinitely thick layer may be grown if the material is identical to that of the substrate. The worst case scenario occurs for Si grown upon Ge or vice versa, where the lattice mismatch is maximised and a limit of around 1 nm layer width is imposed.

A similar restriction occurs with multilayer structures. This may be illustrated by approximating the multilayer system as a single, thick layer containing the weighted mean of the layer Ge fractions:

$$\overline{x} = \frac{\sum_i l_i x_i}{\sum_i l_i}, \tag{6.3}$$

where l_i and x_i are the width and Ge fraction of each layer i respectively. A mismatch between this mean Ge fraction and that of the substrate leads to mechanical stress parallel to the interface, and a critical thickness exists for the entire stack in addition to the individual layers[97]. For a QCL, this limits the number of periods of the structure which may be grown.

A solution to this problem is to calculate an optimal Ge fraction for the substrate, which minimises the in-plane stress. The minimum strain energy corresponds to a substrate lattice constant

$$a_s = \frac{\sum_i A_i l_i / a_i}{\sum_i A_i l_i / a_i^2}, \tag{6.4}$$

where the elastic constants have been grouped together in a single property A [97]. The value of this property is

$$\begin{aligned} A^{(001)} &= 2\left(C_{11} + C_{12} - 2\frac{C_{12}^2}{C_{11}^2}\right) && (6.5) \\ A^{(111)} &= \frac{12C_{44}(4C_{11}C_{44} + 8C_{12}C_{44} + C_\alpha^2)}{c_\beta^2} && (6.6) \end{aligned}$$

for the (001) and (111) orientations respectively, where $C_\alpha = -(C_{11} + 2C_{12})$ and $C_\beta = C_{11} + 2C_{12} + 4C_{44}$ as defined in chapter 2.

6.2 Status of growth technology

Two techniques for growing SiGe heterostructures are in common use: molecular beam epitaxy (MBE) and chemical vapour deposition (CVD)[88]. In MBE growth, Si and Ge are heated in effusion chambers and beams of the atoms are selectively allowed to pass through shutters onto the growth medium. The growth rate is primarily determined by the flux of the source. The first demonstration of electroluminescence from Si/SiGe quantum cascade structures used a 12-period device grown using MBE. Absolute errors in layer thickness around ±2 Å were achieved[73]. Multiple QW stacks with 70–85% Ge wells and Si barriers have been grown upon $Si_{0.5}Ge_{0.5}$ virtual substrates using MBE. Total stack heights around 0.5 μm and interface roughness heights around 3–3.6 Å are achievable[173].

In CVD, a precursor such as silane or germane is heated to around 1000°C, which leads to Si or Ge being deposited on the growth medium by pyrolysis. The growth rate is a function of the gas flow rate, pressure and temperature. This means that calibration of CVD equipment is more complicated, and MBE is more flexible in a research environment. However, CVD tends to be favoured in industry due to the maturity of the technology and the lower achievable particulate densities.[88] The CVD growth rate on (111) oriented Si structures has been shown to be lower than on (001), although the ratio between Ge and Si deposition rates is greater.[174]

Doping in Si/SiGe heterostructures may present a challenge. Surface segregation occurs due to it being energetically favourable for dopants to rise to the surface during growth. The effect is quite weak in Ge[175], but has been shown to be sizeable for low temperature CVD[176] and gas-source MBE[177] growth of Si-rich structures. Segregation allows dopants to spread over large distances within devices, which makes modulation doping challenging. Zhang shows that the effect may be reduced, however, by removing the surface dopants with atomic hydrogen etching.[177]

6.2.1 Virtual substrates

Although it is impractical to obtain arbitrary Ge fractions in wafers, a "virtual substrate" (or pseudo-substrate) may be grown upon a standard Si wafer by depositing SiGe epitaxially, with gradually increasing Ge content[88]. This technique has been used with conventional CVD techniques to achieve a 24% Ge virtual substrate on Si[178] with a 5% Ge/μm composition gradient and 0.88 μm capping layer. In the same study, it was shown that the composition gradient could be increased to as much as 52% Ge/μm. However, the low temperature mobility dropped from 2.58×10^5 to $1.37\times10^5\,\mathrm{cm^2V^{-1}s^{-1}}$ due to the increased number of dislocations.

Using a low composition gradient and a thick capping layer gives a good quality surface, but SiGe alloys have a significantly lower thermal conductivity than pure Si.[179] In addition to this, a thick virtual substrate increases the dimensions of the QCL significantly. The optical mode of the device overlaps with the substrate to some extent, where it is unable to stimulate further emission. It is therefore desirable to develop thinner virtual substrates.

A thin film of the desired substrate material may be grown directly on top of a silicon-on-insulator (SOI) wafer. As long as the critical thickness of the SiGe film is greater than that of the Si on the wafer, the wafer can be annealed and dislocations thread down into the SiO_2 layer. The SiGe film therefore relaxes and may be used as a virtual substrate. Ge fractions up to 15% have been grown using this method.[180]

A similar approach involves growing a thin Si layer with many defects at low temperature. These defects act as nucleation sites for the dislocations when a substrate is annealed. Again, Ge fractions up to 15% have been grown, with threading dislocation density as low as $10^4\,\mathrm{cm^{-2}}$.[181] Defects below the surface can also be obtained by helium implantation, and 200 nm thick 30% Ge substrates have been developed.[182]

Virtual substrates have so far been difficult to obtain in (111) oriented systems using graded buffer regions, with threading dislocation density around ten times that in (001)[174]. In this orientation, significant relaxation of strain occurs by the formation of planar defects, and much gentler grading of buffer regions may be required[174].

Ge-rich virtual substrates may be obtained by using a graded SiGe buffer. However, great care must be taken to reduce threading dislocations and surface roughness. Samavedam showed that by using a miscut Si wafer and introducing a polishing step part way through the growth process, a threading dislocation density of $2.1\times10^{6}\,\mathrm{cm}^{-2}$ was achievable.[183] More recent work has reduced this to $1\times10^{5}\,\mathrm{cm}^{-2}$ in a 1μm Ge layer upon a $10\,\mu$m graded buffer.[184]

An alternative approach avoids the need for such a thick buffer by growing a 25 nm thick amorphous Ge seed layer directly upon Si using high temperature CVD. A crystalline Ge layer (up to 1.6μm thick) is then grown upon the seed layer at a lower temperature.[185] The dislocation density may be reduced to around $2.3\times10^{7}\,\mathrm{cm}^{-2}$ in such structures after annealing or $2.3\times10^{6}\,\mathrm{cm}^{-2}$ in limited areas[186].

6.3 Diffuse interfaces

Although the model has so far assumed that interfaces are perfectly abrupt, diffuse Ge profiles may result from processes such as surface segregation during growth,[177] or by annealing.[187] In this section, it is shown that this leads to significant changes in the subband spacing and scattering rates in a QW[7]. The issue of barrier degradation is then addressed by considering a single barrier separating a pair of QWs. A set of figures of merit for tolerance to Ge interdiffusion is defined and results are compared for each material configuration.

6.3.1 Diffusion model

Surface segregation yields asymmetric Ge interdiffusion profiles, as Ge preferentially rises to the surface during epitaxial growth.[177] In contrast, annealing leads to symmetric interdiffusion, as it depends only upon the concentration difference at interfaces.[187] Both processes change the electronic behaviour of a system by narrowing the QWs, degrading the barriers and reducing the purity of material within the QWs. Due to this intrinsic similarity between the processes, it is reasonable to approximate their combined effect as resulting from a single annealing-like process characterised by an effective diffusion length L_d.[187] Diffusion lengths around $L_d \approx 1$–2 nm have been observed experimentally.[188]

Annealing of an abrupt structure, with the Ge fraction x_I in layer I provides a simple model of a diffuse system. The abrupt-interface structure is embedded between infinitely thick barriers with composition x_0. The composition profile after annealing is[187]

$$\begin{aligned} x(z) &= \frac{1}{2}\sum_{I=1}^{N} x_I \left[\operatorname{erf}\left(\frac{z - z_{I-1}}{L_d}\right) - \operatorname{erf}\left(\frac{z - z_I}{L_d}\right)\right] \\ &+ \frac{x_0}{2}\left[\operatorname{erf}\left(\frac{z - z_N}{L_d}\right) - \operatorname{erf}\left(\frac{z - z_0}{L_d}\right)\right], \end{aligned} \tag{6.7}$$

where the I-th layer boundaries are z_{I-1} and z_I. For annealing at a constant temperature, the diffusion length is given by

$$L_d = 2(Dt)^{1/2}, \tag{6.8}$$

where D is a temperature dependent diffusion constant and t is the annealing time[187, 189].

6.3.2 Single quantum wells

A single QW with a pair of 5 nm $\mathrm{Si_{0.5}Ge_{0.5}}$ barriers was modelled in (001) Ge and in (001) and (111) Si. In each case, the width of the well was adjusted to give a separation of approximately 20 meV between the lowest pair of subbands. A virtual substrate was used in each case to minimise strain in the QW. For the (001) Si/SiGe QW, a 7.2 nm well width and 27% Ge virtual substrate composition were required. In the (111) Si/SiGe QW, the well width and virtual substrate Ge fraction were adjusted to 11.5 nm and 22% respectively. In the (001) Ge/GeSi QW, a 20 nm well width and 81% Ge virtual substrate were used.

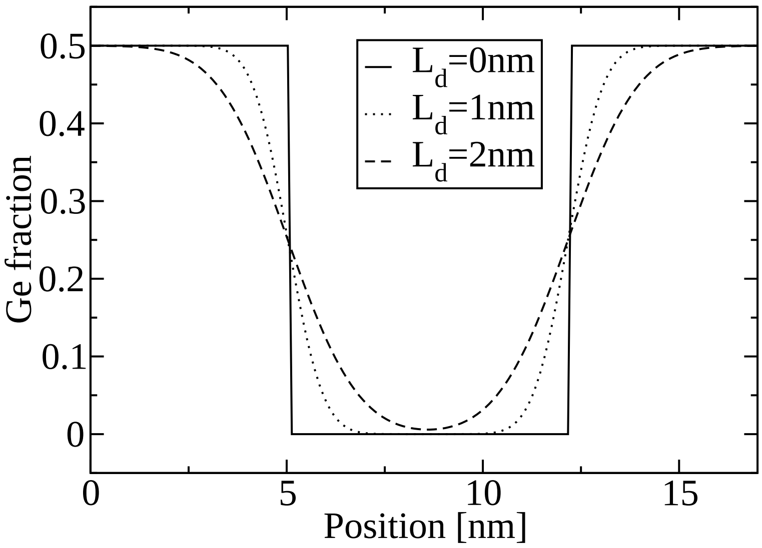


Figure 6.2: Interdiffusion in a single QW at diffusion lengths of 0, 1 and 2 nm.

The effect of interdiffusion upon the Ge profile in a QW is illustrated in fig. 6.2. As L_d increases, the bottom of the well narrows and the top widens. Subbands which are nominally near the bottom of the well are therefore pushed up in energy as interdiffusion increases, while those at the top drop in energy. Conversely, the effect is very small in subbands near the middle of the QW depth.

Fig 6.3 shows the effect of interdiffusion upon subband separation. The effect is quite complex, as it depends upon the depth of the QW and the nominal energies of the subbands. In the case of the (001) oriented structures, both subbands nominally lie well below the middle of the QW energy range and the well appears to shrink with increasing interdiffusion. The subband separation therefore increases. In the (111) Si system, the well is somewhat shallower and the upper subband lies near the middle of the energy range. The lower subband therefore increases in energy while the upper subband stays approximately unchanged, leading to a small reduction in subband separation.

In general, however, narrower layers are required to obtain a given subband separation in systems with larger effective mass. Small changes in layer width therefore

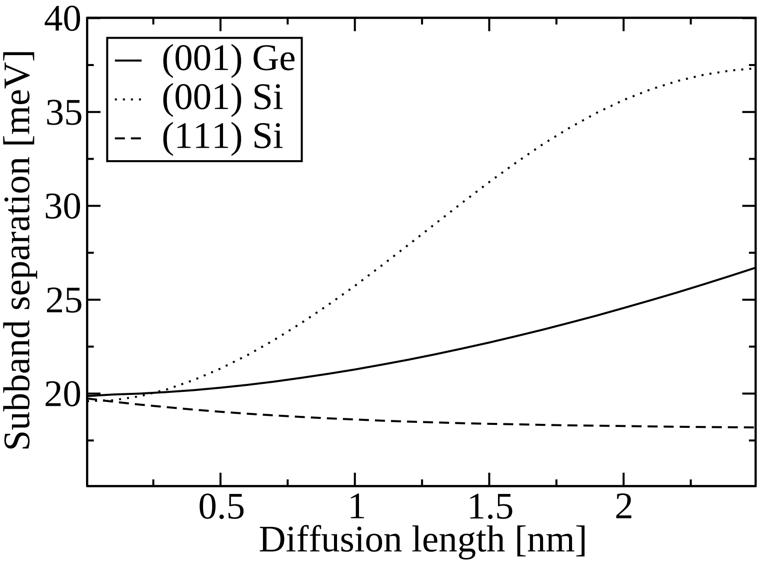


Figure 6.3: Separation between lowest pairs of subbands as a function of interdiffusion length in single QWs

have a greater effect and it can be seen in fig 6.3 that subband separation varies much more rapidly in (001) Si than in the other systems, due to its much greater effective mass.

The effect on scattering rates is also quite complex. The single QW systems described above were modelled with doping spread throughout the well layer at a sheet doping density of $10^{11}\,\mathrm{cm}^{-2}$ with $T = 4\,\mathrm{K}$ and T_e=24 K. Figure 6.4 shows the scattering rates as a function of interdiffusion in a single (001) oriented Si QW. Most of the rates only vary gently as the overlap between subbands changes. The interface roughness rate initially grows as the wavefunctions are allowed to penetrate deeper through the interfaces. At larger interdiffusion lengths, the degradation of the conduction band offset reduces the interface roughness scattering. The greatest change is observed in alloy disorder scattering, which grows by several orders of magnitude and becomes the fastest scattering mechanism at $L_d = 2.8\,\mathrm{nm}$. This is due to the combined effect of the decreasing purity of the Si in the well region and

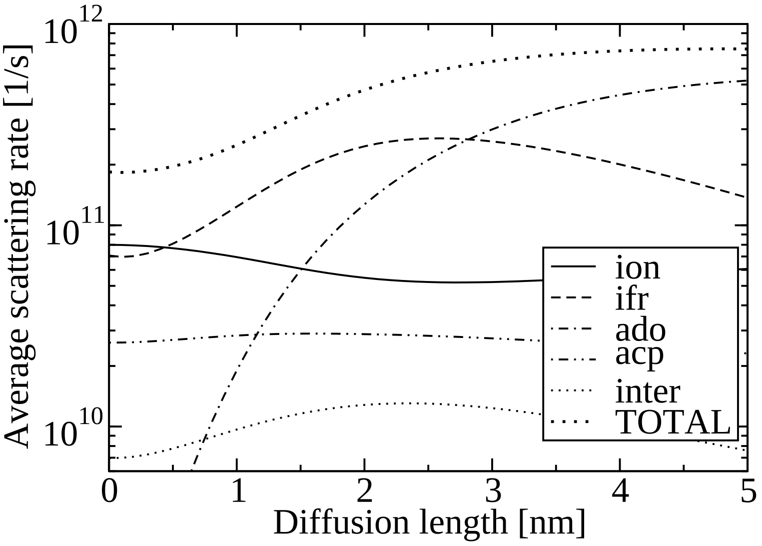


Figure 6.4: Average scattering rates from second to first subband in a (001) Si QW with 50% Ge barriers as a function of interdiffusion length.

the increasing leakage of the wavefunction into the barriers.

The total scattering rate in (001) Si rises appreciably for $0.4 < L_d < 3\,\text{nm}$ as the interface roughness and alloy disorder scattering both become faster. For $L_d \gtrsim 3\,\text{nm}$, the increase in alloy disorder scattering is approximately balanced by the decrease in interface roughness scattering and the total scattering stabilises 4.2 times faster than in the nominal structure.

Figure 6.5 shows that the situation is somewhat less dramatic in (111) oriented Si due to the greater well width. The Ge fraction in the centre of the well is therefore lower than in diffuse (001) oriented systems and the alloy disorder scattering increases more gently and only becomes the dominant mechanism around $L_d > 5\,\text{nm}$. The total scattering rate therefore remains fairly constant for smaller diffusion lengths.

Figure 6.6 shows the effect of interdiffusion upon scattering in Ge QWs. Although the previous observations about alloy disorder scattering rates remain qualitatively correct, the intervalley phonon scattering plays a much more important role.

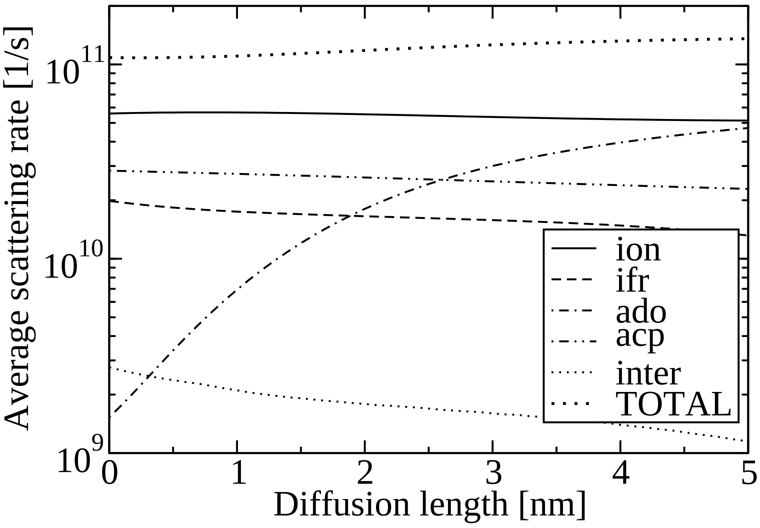


Figure 6.5: Average scattering rates from second to first subband in a (111) Si QW with 50% Ge barriers as a function of interdiffusion length.

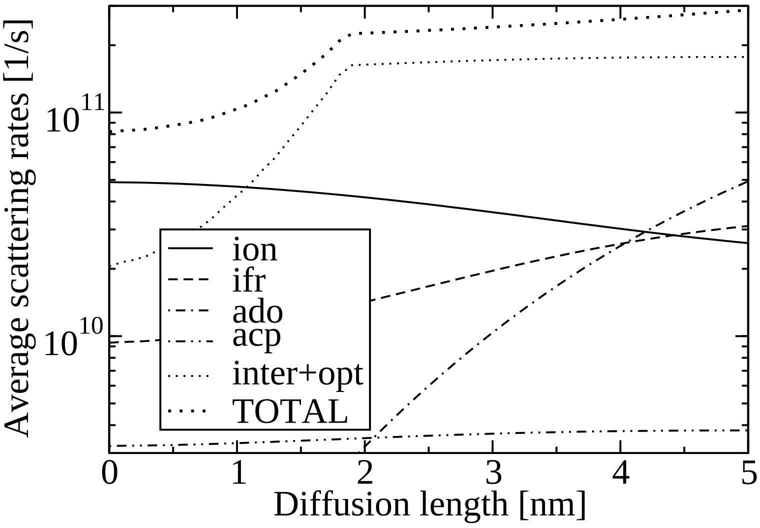


Figure 6.6: Average scattering rates from second to first subband in a (001) Ge QW with 50% Ge barriers as a function of interdiffusion length

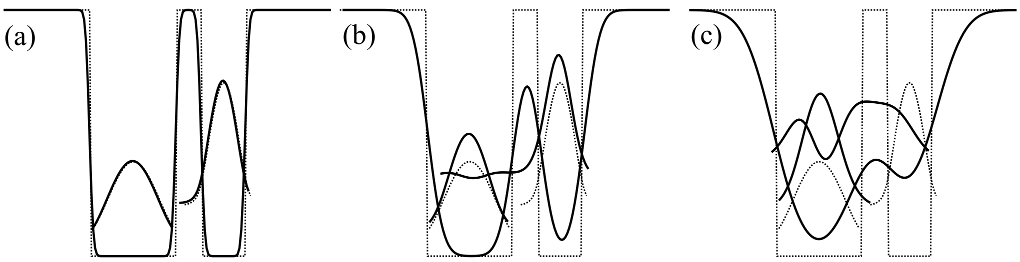


Figure 6.7: Conduction band potential and lowest two subband minima shown schematically (solid lines) for varying degrees of interdiffusion. The nominal values are shown as dotted lines in each plot. (a) shows the "uncoupled" regime, corresponding to low interdiffusion. (b) shows the weak coupling regime, in which interdiffusion degrades the barrier between wells. (c) shows the "single well" regime, in which large interdiffusion merges the wells.

In the nominal structure, the subband spacing is 20 meV, which is below the intervalley phonon energy. As interdiffusion increases, the subband separation grows and it can be seen in fig. 6.3 to exceed $\hbar\omega_q = 24\,\text{meV}$ for $L_d > 1.9\,\text{nm}$. The scattering rate therefore rises rapidly as interdiffusion increases, until the phonon scattering saturates at $L_d > 1.9\,\text{nm}$.

6.3.3 Barrier degradation

To estimate the effect of interdiffusion upon barrier degradation, pairs of coupled QWs separated by a thin barrier were considered. Initially, a (001) oriented Si-rich structure with well widths of 4.5 and 2.6 nm separated by a 1 nm barrier was modelled. A 28% Ge virtual substrate was used to minimise strain in the heterostructure. The resulting energy separation between the lowest pair of subbands is 20 meV. The lower energy electrons are strongly confined in the wider well, and the higher energy electrons in the narrower well as shown schematically in fig. 6.7(a).

Specific results for each material configuration are discussed shortly, but some general characteristics of diffuse double quantum well systems may be identified.

Figure. 6.7 shows the results for the (001) oriented n-type Si rich system described above. Three distinct regimes can be identified, as interdiffusion increases. For low interdiffusion [fig. 6.7(a)] the interfaces are almost abrupt, and the barrier is well defined. This effectively uncouples the wells, resulting in very small overlap between the lowest pair of subbands.

As interdiffusion increases, the barrier degrades [fig. 6.7(b)] and the wells become weakly coupled, leading to an increased overlap between subbands. The bottoms of the wells narrow, leading to an increased subband spacing. At very large interdiffusion lengths, [fig. 6.7(c)] the barrier potential is substantially reduced, and the system resembles a single quantum well with the nominal "barrier" region acting as a perturbation. The region of overlap between subbands now extends across the entire system, and the energy spacing between subbands is determined approximately by the width of the wide, single well and is hence lower than the nominal value.

Figures of merit

To provide a useful comparison between the material configurations, three figures of merit were defined for the tolerance to interdiffusion.

Figure 6.8 gives a schematic representation of the effect on subband spacing. A pair of figures of merit characterise such an effect. L_{pk} is the interdiffusion length which results in a peak shift in subband separation ΔE_{pk}. L_{pk} is therefore an indicator of the point at which the system moves into the "single well" regime. As this represents a large change in electronic behaviour, it is desirable to maximise the value of L_{pk}. Shifts in subband spacing are undesirable in a practical device, and a low value of ΔE_{pk} is preferable. A normalised value $\varepsilon = \frac{\Delta E_{\mathrm{pk}}}{E_0}$ gives the peak subband spacing relative to the nominal value.

As stated previously, the overlap between subband states increases as the barrier degrades and this causes an increase in scattering rates. A final figure of merit L_{w} is defined as the interdiffusion length which yields a "catastrophic" 50% increase in intersubband scattering rate. An ideal system would maximise this value.

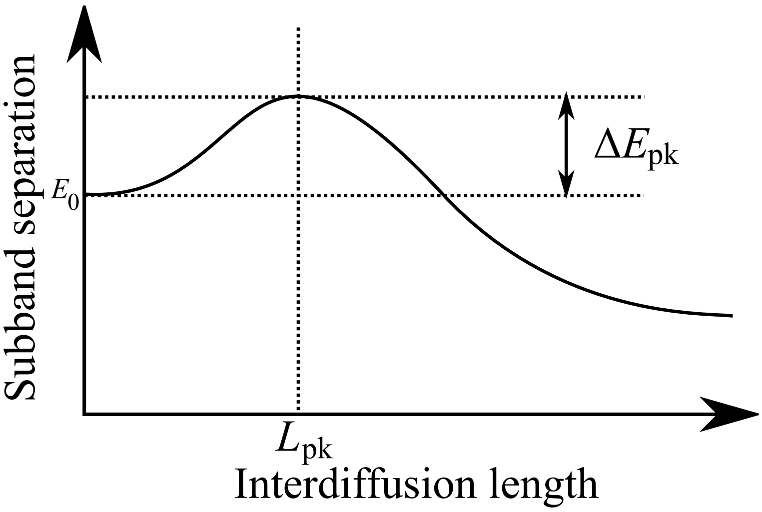


Figure 6.8: Schematic representation of subband spacing as a function of interdiffusion length. In the "weak coupling" regime ($L_d < L_{\mathrm{pk}}$), the subband spacing increases to a peak shift of ΔE_{pk}. At greater interdiffusion lengths, the subband spacing decreases.

(001) oriented Si-rich structures

Figure 6.9 shows the calculated intersubband scattering rates and subband separation for a (001) oriented Si-rich double QW. A large increase in subband spacing is visible even at small diffusion lengths, leading to a peak shift of $\varepsilon = 99.5\%$ at $L_{\mathrm{pk}} = 1.34\,\mathrm{nm}$.

At low interdiffusion lengths, interface roughness scattering dominates strongly. Increases in interdiffusion result in a greater overlap between wavefunctions, which causes all scattering rates to rise sharply. A 50% increase in total scattering rate is seen at $L_w = 0.87\,\mathrm{nm}$. In the "single well" regime, the large Ge fraction in the wells leads to the alloy disorder scattering becoming dominant. Simultaneously, the interface roughness scattering decreases as the central barrier decays and the total scattering rate becomes approximately constant.

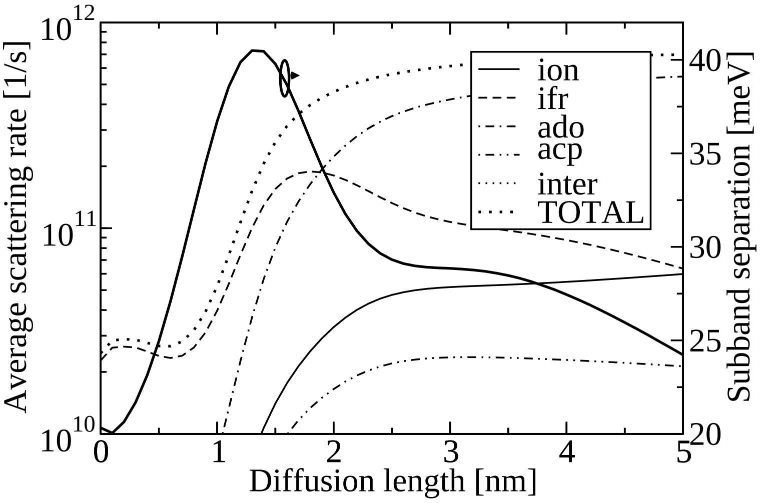


Figure 6.9: Average scattering rates and subband separation from second to first subband in a (001) Si double QW with 50% Ge barriers as a function of interdiffusion length.

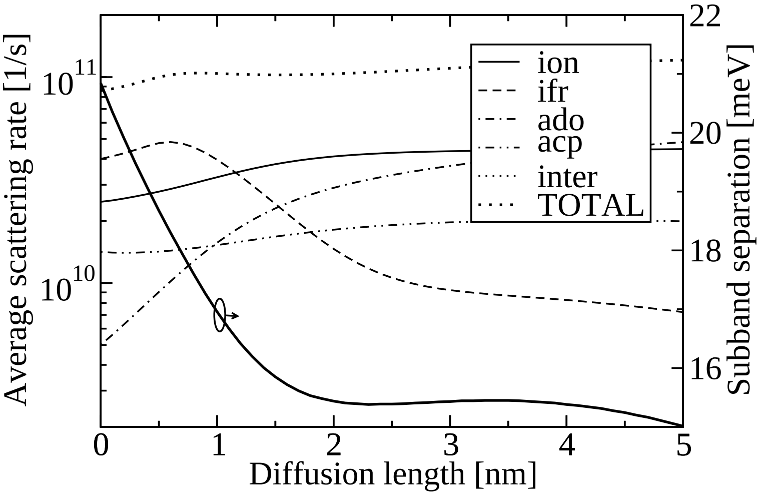


Figure 6.10: Average scattering rates and subband separation from second to first subband in a (111) Si double QW with 50% Ge barriers as a function of interdiffusion length.

(111) oriented Si-rich structures

In the (111) crystal orientation, the layer thicknesses were adjusted to 9, 3 and 2 nm for the two wells and the barrier respectively to preserve the nominal 20 meV subband separation. A 23% Ge virtual substrate was required to minimise strain. Figure 6.10 shows that the effect of interdiffusion on subband separation is relatively small due to the low effective mass. The lower subband energy only shifts downward by 6 meV, giving $\varepsilon = -26.0\%$ at $L_{\text{pk}} = 2.42$ nm.

As before, interdiffusion leads to a switch in dominance between interface roughness and alloy disorder scattering, although the large layer widths weaken the effects. The low conduction band offset allows some coupling between the wells, even in the absence of interdiffusion. The sudden appearance of electron–phonon and ionised impurity interactions seen in the (001) oriented system is therefore avoided. This fairly constant background level of scattering, combined with the smooth transition between ionised impurity and alloy disorder scattering gives a stable total scattering rate with respect to interdiffusion. The corresponding figure of merit $L_w > 5$ nm is far superior to that of the (001) orientation.

(001) oriented Ge-rich structures

The final set of results is shown in fig. 6.11, for a (001) oriented Ge rich system. Due to the low quantisation effective mass, the layer widths are increased to 12.5, 7.4 and 2.8 nm for the wells and the barrier respectively. A 78% Ge virtual substrate was used. The amount of interdiffusion required for single well behaviour is large, giving $L_{\text{pk}} = 3.81$ nm. Although the weak coupling regime persists over a large range of interdiffusion lengths, the subband separation is shown to rise more than in (111) oriented Si-rich systems, giving $\varepsilon = 167\%$. This occurs because the conduction band offset is larger than that of (111) oriented Si and the upper subband lies below the midpoint of the well and barrier potentials. The upper subband is therefore affected strongly by the narrowing of the bottom of the well.

As before, a transition between interface roughness and alloy disorder scattering

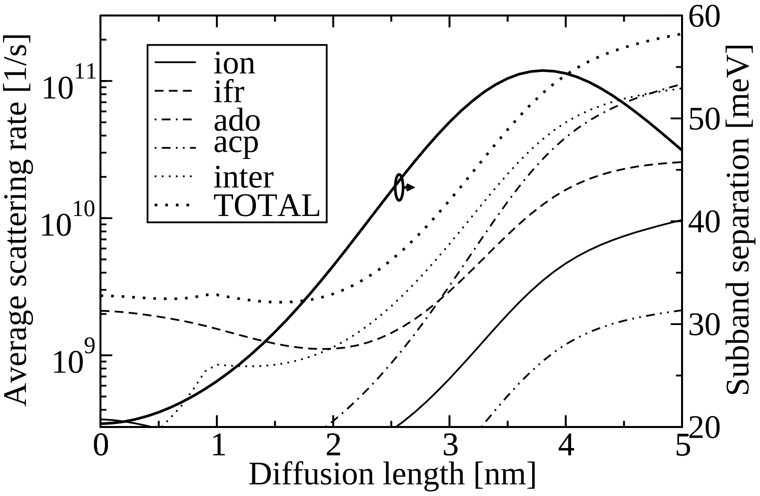


Figure 6.11: Average scattering rates and subband separation from second to first subband in a (001) Ge double QW with 50% Ge barriers as a function of interdiffusion length.

dominance is seen, although the rate of change of scattering rates is reduced due to the large layer widths. The effect of the gradual reduction in interface potential is exceeded by the increase in overlap between the subbands. This results in a slight increase in interface roughness scattering and the "balancing" of the interface roughness and alloy disorder scattering rates is not achieved.

Additionally, intervalley electron–phonon interactions are significant. The scattering rate increases for small amounts of interdiffusion, as the subband spacing approaches the phonon energy of 24 meV. For subband separations around this value, the electron–phonon interactions become the dominant scattering process. As the subband spacing increases above the phonon energy, the overlap between subbands leads to a further, gradual increase in electron–phonon scattering, although alloy disorder scattering regains dominance in the single well regime.

The combination of these effects leads to a faster increase in the total scattering than that seen in the (111) oriented Si-rich system. Consequently, the level of interdiffusion required for a 50% increase in scattering $L_{\mathrm{w}} = 2.36\,\mathrm{nm}$, is slightly lower.

6.4 Conclusion

By taking mechanical stability into account, an upper limit has been placed on the thickness of individual layers within a QCL. It has been shown, however, that strain balance may be achieved over the length of a period by selecting an appropriate substrate material. This implies that as long as all individual layers are kept below their critical thickness, an effectively infinite number of periods of the structure may be grown.

Epitaxial growth of SiGe heterostructures is an active area of research, but it has been shown that good quality growth of multi-layer (001) Si/SiGe structures is already achievable. Each of the material configurations considered in this work presents its own challenges, however. Modulation doping in (001) Si may be difficult due to surface segregation effects, although this is less of an issue in Ge. (111) Si has proved difficult to grow, due to the formation of planar dislocations, but a reduced composition gradient in the virtual substrate may reduce this problem. Finally, virtual substrate growth in (001) Ge may be challenging due to the large lattice mismatch between the Si wafer and the Ge-rich substrate material. Good progress is being made in each of these areas, however.

The effect of interdiffusion has been investigated in each of the material configurations. It was shown that as the diffusion length increases, the bottoms of QWs narrow and subband separations are affected. Alloy disorder scattering increases by several orders of magnitude as the alloy material diffuses into the nominally pure Si (or Ge) well region. (001) Si was found to be most strongly affected as its large quantisation effective mass led to the requirement for narrow wells. In Ge, a sudden jump in intervalley scattering was predicted as the subband separation switches from being below to above the relatively low phonon energy.

A set of figures of merit were defined to describe the tolerance of each material

System	L_{pk} [nm]	ε	L_{w} [nm]
(001) Si	1.34	99.5%	0.87
(111) Si	2.42	-26.0%	> 5.0
(001) Ge	3.81	167%	2.36

Table 6.1: Comparison between interdiffusion figures of merit for each material configuration.

configuration to barrier degradation by interdiffusion, with respect to subband separation and total scattering rate. The calculated values for these figures of merit are compared in table 6.1.

In general, it has been shown that, for all material configurations, interface roughness makes a greater contribution than alloy disorder to the total scattering rate in systems with abrupt interfaces. The converse is true for diffuse systems. As interdiffusion increases, the total scattering rate becomes larger due to increased overlap between subbands.

The subband spacing varies in a more complicated manner. Small levels of interdiffusion lead to a "weak coupling" regime, in which the bottoms of the wells become narrower and subband separation increases. Larger levels of interdiffusion cause the system to behave more like a single QW, and the subband separation decreases.

It has been shown that the tolerance of QCLs to interdiffusion may be improved by moving to the (111) orientation or by using Ge QWs. The low effective mass of these systems allows thicker layers to be used and the effect of interdiffusion on their geometry is less severe. The conduction band offset in (111) oriented systems is lower than that in Ge-rich systems and this leads to (111) oriented systems showing almost no variation in their subband spacing for small amounts of interdiffusion. An improvement in the stability of scattering processes is also achievable.

Ge-rich systems show a moderate improvement in the stability of subband spacing at low interdiffusion lengths. However, a significant barrier potential exists up to very large interdiffusion lengths. Care must be taken when designing Ge-rich QCLs due to the rapid electron–phonon interactions for subbands separated by more than 24 meV. By avoiding transitions just below the phonon energy in QC laser designs, much better stability of scattering rates may be achieved.

The results of this chapter are supported qualitatively by recent investigations of III-V heterostructures. A threefold increases in emission linewidth (and hence total scattering rate) has been observed experimentally in GaAs/AlAs QCLs as the diffusion length increases from 0 to 2.5 nm[190]. A blue shift of 5.6 meV in the interminiband emission frequency for GaAs/AlAs superlattices was also observed as the interdiffusion length increased from 0 to 2.7 nm[191]. Both sets of results lie within the same range of values as those predicted for SiGe-based heterostructures in this chapter.

Finally, it is important to revisit the key assumptions that were made in the previous chapter to handle the somewhat blurred distinction between interface roughness and alloy disorder scattering in diffuse systems. The ambiguity is due to the nonzero interface length and the continuous spatial variation of alloy composition. Spatial limits were introduced for each interface region and it was assumed that the interface roughness parameters are independent of interdiffusion. In reality, however, annealing-like diffusion processes may reduce the interface roughness height. Although the theoretical results agree qualitatively with III-V experimental data, TEM imaging of diffuse interfaces and FTIR spectroscopy of SiGe-based heterostructures would be extremely valuable in testing the validity of assumptions and tuning the roughness parameters.

Chapter 7

Transport and gain

In the work so far, the bandstructure and scattering rates have been determined for SiGe-based QCLs. This chapter describes methods for determining the current through a device and its optical gain.

At steady state, the rate of electrons entering each subband equals the rate of those leaving and the subband populations are found by solving a system of rate equations. Similarly, a steady state electron temperature may be found, for which the energy of electrons entering each subband balances that of the electrons leaving the subband.

Successful laser operation relies upon there being a net gain, or increase in the number of coherent photons as light passes through a QCL. The maximum operating temperature and threshold current are found by investigating the point at which the gain for SiGe QCLs equals the estimated threshold gain for SiGe waveguides.

7.1 Subband populations

Expressions for the intersubband scattering rates were found in chapter 5. The rate of change of subband population is given by the difference between the rate of electrons entering the subband and those leaving. It was shown in chapter 2 that the lowest conduction band potentials in SiGe alloys lie in multiple equivalent valleys in reciprocal space. The intersubband scattering processes described in chapter 5 are identical in equivalent valleys and the subband populations are also identical. It is therefore possible to gather the populations of subbands in equivalent valleys into a single term $n_{\mathrm{val}}^{i} n_i$, where n_i is the number of electrons in subband i and n_{val}^{i} is the number of equivalent valleys.

Charge neutrality requires that the total population of all subbands in all valleys must equal the number of donor ions, giving

$$N_{\mathrm{2D}} = \sum_i n_{\mathrm{val}}^{i} n_i. \tag{7.1}$$

In (001) oriented systems, the possible values of n_{val} are 2 (for subbands in Δ_2 valleys) or 4 (for subbands in Δ_4 or L valleys). If valley splitting effects are included, the split Δ_2 and L valley subbands must be considered independently and the number of equivalent valleys are halved. In (111) systems, n_{val} may be 6 (Δ valleys), 3 (L_3 valleys), or 1 (L_1 valley).

In this work, however, the calculations are greatly simplified by assuming that only the lowest valley set is populated. This is a reasonable approximation provided that subbands in only one valley set lie within the range of energy used in a QCL. In chapter 4, it was also shown that valley splitting becomes small (~1 meV) in relatively long structures with imperfect interfaces. It is assumed here, therefore, that the individual transitions due to valley splitting will be so close in energy that they simply appear as a linewidth broadening in the gain spectrum. The simplified expression for charge neutrality is given by

$$n_{\mathrm{val}} \sum_i n_i = N_{\mathrm{2D}}. \tag{7.2}$$

Periodicity requires that $n_j = n_{j+pN}$, where p is the index of the period and N is the number of subbands in a period. The rate of change of population is therefore[150]

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = \sum_{p=-1}^{1} \sum_{j=1}^{N} \left[n_j W_{(j+pN)i} - n_i W_{i(j+pN)} \right]. \tag{7.3}$$

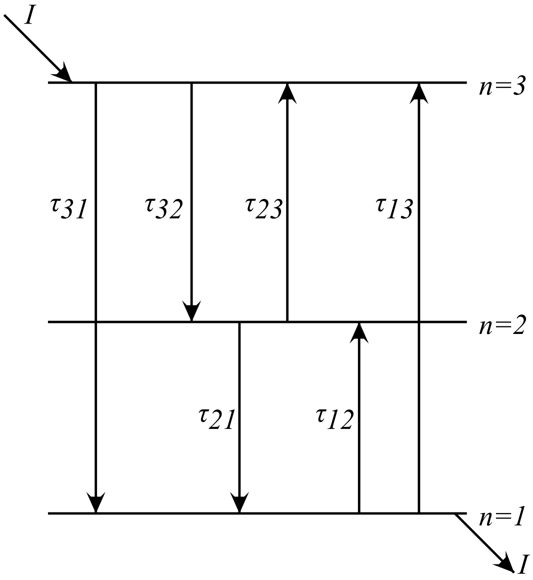


Figure 7.1: Schematic representation of subband energies and lifetimes in an idealised three-level QCL system. Current is injected only into the upper laser level and extracted from the extraction level.

At steady-state, $\frac{\mathrm{d}n_i}{\mathrm{d}t} = 0$ for all subbands and a set of rate equations may be solved to find the populations.

7.1.1 Three-level approximation

A real QCL may contain many subbands in each period, but an idealised three-level model is useful for developing an understanding of its behaviour. Figure 7.1 shows a schematic representation of such a model. Current I is injected into the upper laser level (n=3). Optical emission occurs between the upper laser level and lower laser level (n=2). Ultimately, electrons scatter to the depopulation level (n=1) and are extracted by an identical current.

The rate equations for the system are

$$\begin{aligned} 0 &= -n_1(W_{12}+W_{13}) + n_2W_{21} + n_3W_{31} \\ 0 &= n_1W_{12} - n_2(W_{21}+W_{23}) + n_3W_{32} \\ 0 &= n_1W_{13} + n_2W_{23} - n_3(W_{31}+W_{32}) \end{aligned} \tag{7.4}$$

These equations form an underdetermined set as the third equation is a linear combination of the first two. A fully determined set of equations may be obtained by replacing the third equation with the condition for charge neutrality, $N_{2D} = n_{\text{val}}(n_1 + n_2 + n_3)$.

Rewriting this as a matrix equation gives

$$\begin{array}{ccc} \begin{pmatrix} -(W_{12}+W_{13}) & W_{21} & W_{31} \\ W_{12} & -(W_{21}+W_{23}) & W_{32} \\ 1 & 1 & 1 \end{pmatrix} & \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} & = \begin{pmatrix} 0 \\ 0 \\ N_{2D}/n_{\text{val}} \end{pmatrix} \\ \mathsf{W} & \boldsymbol{n} & = \boldsymbol{N}. \end{array} \tag{7.5}$$

In the general case of N populated bands, the W matrix elements are

$$\begin{aligned} \mathsf{W}_{ii}|_{i\neq N} &= W_{ii} - \sum_j W_{ij}, \\ \mathsf{W}_{ij}|_{i\neq j\neq N} &= W_{ji}, \\ \mathsf{W}_{Nj} &= 1. \end{aligned} \tag{7.6}$$

The populations of the subbands are given by

$$\boldsymbol{n} = \mathsf{W}^{-1}\boldsymbol{N}. \tag{7.7}$$

For lasing, population inversion is required, meaning that more electrons exist in the upper laser level (ULL) than the lower laser level (LLL). Some insight into the effect of scattering rates upon populations may be gained from the three level model at low temperatures. Here, few carriers have sufficient kinetic energy to scatter to a higher subband and $W_{12} = W_{13} = W_{23} = 0$. At steady state, the carriers scattering

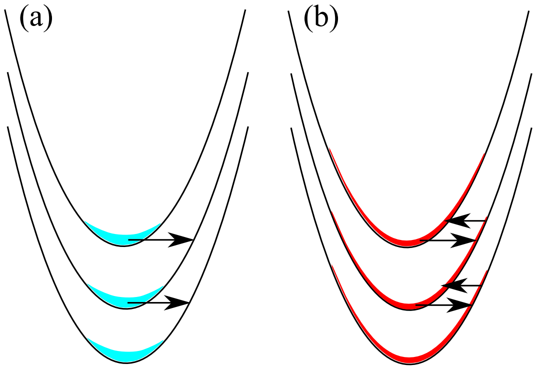


Figure 7.2: Schematic representation of elastic intersubband scattering. (a) In cold electron distributions, electrons may only scatter to higher kinetic energy states in lower subbands. (b) In hot electron distributions, electrons may also scatter to lower kinetic energy states in higher subbands.

from $|3\rangle$ to $|2\rangle$ replace the carriers scattering from $|2\rangle$ to $|1\rangle$. Thus, the rate equations reduce to

$$\frac{n_3}{n_2} = \frac{W_{21}}{W_{32}}. \tag{7.8}$$

A population inversion is therefore achieved between $|3\rangle$ and $|2\rangle$ if $W_{21} > W_{32}$.

7.2 Electron temperature

In chapter 5, electrons were assumed to take a quasi-thermal distribution within subbands due to the rapid intrasubband scattering, characterised by an electron temperature. In general, intersubband scattering changes the kinetic energy of carriers. The kinetic energy of carriers increases when they scatter elastically to a lower subband and decreases for scattering to higher subbands.

Figure 7.2(a) shows that at low electron temperatures, carriers may only scatter elastically to lower subbands. This leads to a net increase in kinetic energy in the destination subband. At high electron temperatures, electrons can scatter elastically into higher subbands [fig. 7.2 (b)]. For very high electron temperatures, this can cause a net decrease in kinetic energy.

It follows that a steady-state solution exists, for which the energy loss due to scattering to higher subbands equals the energy gain due to scattering to lower subbands[150]. The initial and final kinetic energies are related by

$$E_{k,f} = E_{if} + E_{k,i} \tag{7.9}$$

for elastic processes. An additional $\hbar\omega_q$ may be added to the right-hand side for phonon absorption processes or subtracted for emissions.

The net kinetic energy transfer rate in the final subband is

$$\frac{\mathrm{d}E_{k,f}}{\mathrm{d}t} = \sum_i \left(n_i \overline{w}_{if}^{+} - n_f \overline{w}_{fi}^{-} \right), \tag{7.10}$$

where $\overline{w}_{if}^{+}$ is the average kinetic energy gained by subband f when an electron scatters from subband i and $\overline{w}_{fi}^{-}$ is the average kinetic energy lost by subband f when an electron scatters to subband i. The kinetic energy gain in the final subband at a given initial wavevector is

$$w_{if}^{+}(k_i) = E_{k,f} W_{if}(k_i) \tag{7.11}$$

and the average value with respect to initial wavevector is

$$\overline{w}_{if}^{+} = \frac{\int k_i \mathrm{d}k_i E_{k,f} W_{if}(k_i, T_e) P_i(k_i, T_e)}{\pi n_i}. \tag{7.12}$$

$E_{if} + E_{k,i} = E_{k,f}$ for elastic processes and the average energy gain in the final subband may be rewritten as

$$\overline{w}_{if}^{+} = E_{if} \overline{W}_{if} + \frac{\int k_i \mathrm{d}k_i E_{k,i} W_{if}(k_i, T_e) P_i(k_i, T_e)}{\pi n_i}, \tag{7.13}$$

for elastic proceses. By deriving a similar expression for $\overline{w}_{if}^{-}$, it can be shown that

$$\overline{w}_{if}^{+} = \overline{w}_{if}^{-} + E_{if} \overline{W}_{if}. \tag{7.14}$$

For inelastic processes, $E_{if} \rightarrow E_{if} \pm \hbar\omega_q$ as before.

Substituting these results into eqn. 7.10 gives

$$\frac{\mathrm{d}E_{k,f}}{\mathrm{d}t} = \sum_i n_i \overline{w}_{if}^{-} + \sum_i n_i E_{if} \overline{W}_{if} - n_f \sum_i \overline{w}_{fi}^{-}. \tag{7.15}$$

Summing over all final subbands gives the net energy transfer rate:

$$\frac{\mathrm{d}E_k}{\mathrm{d}t} = \sum_f \sum_i \left(n_i \overline{w}^{-}_{if} - n_f \overline{w}^{-}_{fi} \right) + \sum_f \sum_i n_i E_{if} \overline{W}_{if}. \tag{7.16}$$

The first double summation equates to zero, leaving

$$\frac{\mathrm{d}E_k}{\mathrm{d}t} = \sum_f \sum_i n_i E_{if} \overline{W}_{if}. \tag{7.17}$$

A steady state condition may be located by calculating the energy transfer rate across a range of different electron temperatures and interpolating the results to find the temperature at which $\frac{\mathrm{d}E_k}{\mathrm{d}t} = 0$.

7.3 Current density

When an electric field is applied, the potential drop across a period of the QCL leads to a preferential direction of charge transport. The rate at which charge flows between two subbands is given by en_iW_{if} and the contribution to the current density is given by[192]

$$J_{if} = p_{if} e n_i W_{if}, \tag{7.18}$$

where p_{if} is the number of structural periods separating the final and initial subband.

In a simple model, electrons are treated as scattering over an integer number of periods, where $p_{if} = 1$ for left-to-right transitions or -1 for right-to-left transitions. Intraperiod scattering is neglected. A better model allows non-integer values of p_{if} by considering the difference in expectation positions $\langle z \rangle$ between the initial and final subbands, where

$$\langle z \rangle_i = \int \psi_i^*(z) z \psi_i(z) \mathrm{d}z. \tag{7.19}$$

The number of periods is then given by

$$p_{if} = \frac{\langle z \rangle_f - \langle z \rangle_i}{L_p} \tag{7.20}$$

and the total current density, assuming that only one valley set is populated, is given by summing over all intersubband interactions[3],

$$J = \frac{en_{\mathrm{val}}}{L_p}\sum_i n_i \sum_f \left(\langle z\rangle_f - \langle z\rangle_i\right) W_{if}. \tag{7.21}$$

It is useful to note that the scattering processes are identical for each valley in an equivalent set. The expression above, therefore, simply multiplies each initial subband's contribution to the current by the number of equivalent valleys.[1]

An approximate expression for the current may be obtained by returning to the three-level example in section 7.1.1. If current only enters the active region via the ULL and leaves via the LLL, and absorption to higher subbands is ignored then the steady state rate equations become

$$\begin{aligned} 0 &= -\frac{J}{en_{\mathrm{val}}} + n_2 W_{21} + n_3 W_{31} \qquad (7.22)\\ 0 &= n_3 W_{32} - n_2 W_{21}\\ 0 &= \frac{J}{en_{\mathrm{val}}} - n_3(W_{31} + W_{32}). \end{aligned}$$

It follows that $\frac{J}{e} = n_{\mathrm{val}} n_3 (W_{31} + W_{32})$ and therefore

$$J = \frac{n_{\mathrm{val}} n_3 e}{\tau_3}, \tag{7.23}$$

where the lifetime of the ULL is given by

$$\frac{1}{\tau_3} = W_{32} + W_{31}. \tag{7.24}$$

This simple approximation shows that the current density through the device depends largely upon the nonradiative depopulation rate of the ULL.

7.4 Gain

In order for lasing to occur, the number of coherent photons in a QCL must increase as light passes through the device. Photons may be absorbed as they interact with

---

[1] It is not necessary to include the number of destination valleys in the expression, as this was already accounted for in the scattering rate equations of chapter 5.

the structure, or they may stimulate the emission of other photons. The net gain is the number of additional photons introduced as an incident photon travels over a unit distance. A positive gain indicates that stimulated emission exceeds absorption, while the converse applies for negative gain.

The optical gain per unit length within a heterostructure is[90]

$$G(\omega) = \frac{\sigma(\omega)}{\varepsilon_0 c n_r}, \tag{7.25}$$

where n_r is the real part of the refractive index of the active region stack and σ is the real part of the optical conductivity. This is given by

$$\sigma(\omega) = \frac{\pi e^2}{2(m_q m_d^2)^{\frac{1}{3}} L_p} \sum_{i,j} f_{ji} n_i n_{\text{val}}^i \,\text{sgn}(E_{ij}) L_{ij}(\omega), \tag{7.26}$$

where f_{ji} is the oscillator strength and L_{ij} is a lineshape function, representing the spread of emission energy. The *signum* function determines whether the transition is an absorption or emission, and is defined by

$$\text{sgn}(x) = \begin{cases} -1; & x < 1 \\ 0; & x = 0 \\ 1; & x > 1. \end{cases} \tag{7.27}$$

It is important to note that optical transitions do not change the wavevector of electrons, and the subbands i and j must therefore be located in the same valley in the expression above. The optical conductivity is identical for each equivalent valley in a set, and the total value is found by simply multiplying the single valley result by n_{val}.

The oscillator strength is

$$f_{ji} = \frac{2(m_q m_d^2)^{\frac{1}{3}}}{\hbar} \omega_{ij} |D_{ij}|^2, \tag{7.28}$$

where $D_{ij} = \langle j|z|i\rangle$ is the optical matrix element. Substituting this into the expression for gain gives[150]

$$G(\omega) = \frac{e^2 \omega \pi}{\hbar c n_r \varepsilon_0 L_p} \sum_{ij} n_i n_{\text{val}}^i |D_{ij}|^2 \,\text{sgn}(E_{ij}) L_{ij}(\omega). \tag{7.29}$$

7.4.1 Lineshape

The lineshape may be treated as a Lorentzian distribution,

$$L_{ij}(\omega) = \frac{\gamma_{ij}}{\pi\left[(\hbar\omega - |E_{ij}|)^2 + \gamma_{ij}^2\right]}, \tag{7.30}$$

where γ_{ij} is the half-width at half-maximum of the energy distribution.

The non-zero linewidth has two main causes. Firstly, a QCL consists of many structural periods, and small variations in layer thickness and electric field lead to a slightly different optical emission energy from each period. This *inhomogeneous* broadening causes the spectrum to appear “smeared” in energy.

An additional *homogeneous* broadening effect is due to the finite lifetimes of the states involved in the transition. The corresponding broadening in energy may be derived from density matrix calculations of light-matter interactions as[193]

$$\gamma_{ij} \gtrsim \frac{\hbar}{2}\left(\frac{1}{\tau_i} + \frac{1}{\tau_j}\right). \tag{7.31}$$

The homogeneous broadening can be calculated by using the total scattering rates as described in chapter 5. This method has recently been shown to give good agreement with experimental data for III–V systems[194]. However, the inhomogeneous broadening is dependent upon the quality of construction of the device. A total linewidth of around $2\gamma \approx 2\,\mathrm{meV}$ is typical for the lasing transition in GaAs based THz QCLs[52, 195], and the inhomogeneous broadening is adjusted to give comparable low-temperature values in this work.

7.4.2 Gain coefficient

If a QCL is designed to emit light of energy $\hbar\omega = |E_{ij}|$, where i and j are the ULL and LLL respectively, the peak value of the lineshape function for the emission is

$$L_{ij}(\omega_{ij}) = \frac{1}{\pi\gamma_{ij}}, \tag{7.32}$$

and the peak contribution of the ULL→LLL transition to the gain is

$$\begin{aligned} G_{ij} &= \frac{e^2\omega_{ij}n_{\mathrm{val}}}{\hbar c n_r \varepsilon_0 L_p \gamma_{ij}}(n_i - n_j)|D_{ij}|^2 \\ &= \frac{2\pi e^2 n_{\mathrm{val}}}{n_r \varepsilon_0 L_p \gamma_{ij} \lambda_{ij}}(n_i - n_j)|D_{ij}|^2. \end{aligned} \tag{7.33}$$

By making the rather crude assumption that no other intersubband transitions contribute gain or loss at this frequency, an approximate expression for gain may be obtained from the three-level model of section 7.1.1. The population inversion between the ULL and the LLL is

$$\begin{aligned} n_3 - n_2 &= n_3\left(1 - \frac{W_{32}}{W_{21}}\right) \\ &= \frac{J\tau_3}{n_{\mathrm{val}}e}\left(1 - \frac{\tau_2}{\tau_{32}}\right). \end{aligned} \tag{7.34}$$

It follows that the peak gain is given by $G_{32} = g_{32}J$, where the gain coefficient g_{32} is defined as

$$g_{32} = \frac{2\pi e\tau_3|D_{32}|^2}{n_r\varepsilon_0 L_p\gamma_{32}\lambda_{32}}\left(1 - \frac{\tau_2}{\tau_{32}}\right). \tag{7.35}$$

The peak active region gain is therefore approximately proportional to the current density. However, there are very significant shortcomings in this approximation. Firstly, the current density is determined by the electric field, which also affects the subband separations, lifetimes and the optical matrix elements significantly. This breaks the linearity of the relationship.

Secondly, real QCLs may have many more than three subbands per period. Absorption of light may occur between any pair of subbands, yielding several absorption peaks at various energies across the gain spectrum. The nonzero linewidth of these absorptions means that they may significantly reduce the magnitude of the active region gain. A somewhat better approximation is therefore

$$G_{32} = g_{32}J - a_i, \tag{7.36}$$

where a_i represents the effect of intersubband absorption. Thirdly, the linear approximation suggests that an arbitrarily large gain may be achieved in virtually any device by simply driving it with a large enough current density. In reality, the maximum current density is finite, and in a poor design, it may be impossible to achieve a large enough gain to overcome losses.

7.4.3 Threshold current density

In the previous section, it was shown that the peak active region gain is a nonlinear function of current density. In this section, a threshold gain (*i.e.* the minimum required for lasing) is defined, which may be used to find the threshold (minimum) current density required to operate a QCL.

Stimulated emission is only achieved if a large photon flux is maintained within the optically active heterostructure. This is achieved by enclosing the QCL heterostructure within a waveguide. The *modal overlap* Γ is defined as the fraction of photons located within the active region of the QCL, and in an ideal waveguide Γ should be close to unity. Absorption of light within the waveguide is described by a waveguide loss α_w which ideally should be close to zero.

A second loss mechanism occurs due to the emission of light from the facets of the device. These *mirror losses* may be distributed over a round-trip through the waveguide and represented by an effective mirror loss

$$\alpha_m = \frac{1}{2L_w} \ln\left(\frac{1}{R_1 R_2}\right), \tag{7.37}$$

where R_1 and R_2 are the reflectivites of the front and back facets of the laser respectively. In simple dielectric waveguide structures, the reflectivity is given by the Fresnel reflection coefficient

$$R_1 = R_2 = \left(\frac{n-1}{n+1}\right)^2, \tag{7.38}$$

where n is the refractive index of the cavity material[196]. The refractive index is different for each layer of the QCL heterostructure and is given by $n \approx 3.42 + 0.37x + 0.22x^2$, where x is the Ge molar fraction in the layer[197]. However, the individual layers are much smaller than the optical wavelength and it is possible to approximate the entire structure as a bulk material with the average Ge fraction of the layers. The resulting mirror losses are 5.1 cm^{-1} for a 1 mm long Ge ridge waveguide structure and 6.0 cm^{-1} for Si. Although this approximation holds reasonably well for the semi-insulating surface plasmon waveguides described in the next section, it overestimates the mirror losses in a double–metal waveguide. The losses in such structures are dependent on the emission wavelength and the geometry of the waveguide and have values around $\alpha_m \approx 1\,\mathrm{cm}^{-1}$ in GaAs[198].

The *modal gain* is defined as

$$G_M = G_{ij}\Gamma \approx g_{ij}\Gamma J. \tag{7.39}$$

Net gain, and hence lasing, is only achievable when the modal gain exceeds the losses (*i.e.* $G_M > \alpha_w + \alpha_m$). A threshold gain G_{th} may be defined as the minimum active region gain required for lasing. It follows that

$$G_{th} = \frac{\alpha_w + \alpha_m}{\Gamma}. \tag{7.40}$$

Similarly, the threshold current density J_{th} is defined as the minimum current density required for lasing.

7.4.4 Waveguides

The active region of the waveguide[2] may consist of hundreds of periods of the QCL heterostructure and may be over 10 μm thick[195, 199]. In the case of THz QCLs, this thickness is substantially below the wavelength of the emitted light and standing-waves cannot form in a simple, dielectric-clad device. Alternative approaches to waveguiding are therefore required.

In recent collaborations with the present author[3, 18], Evans calculated the threshold gain for a 5 THz (111) Si-rich QCL using various waveguide configurations. The propagation constant for the electric field component of light perpendicular to the interfaces was calculated using a transfer matrix approach[200] and the bulk

---

[2] Not to be confused with the active region within a QCL period

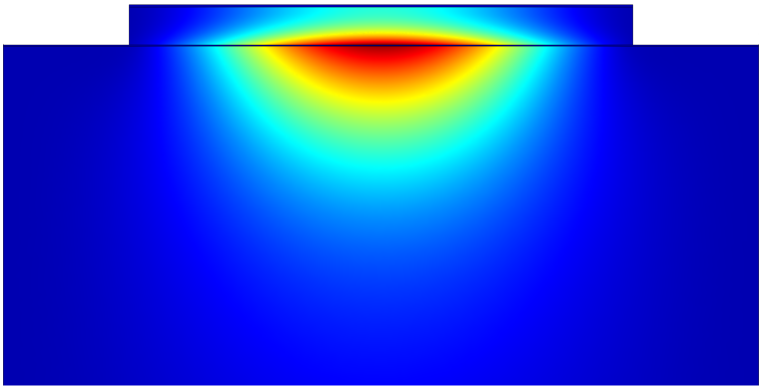


Figure 7.3: Optical mode profile in a single-metal Si waveguide upon a Si substrate. The active region is represented by the small rectangle at the top of the figure. Adapted from joint publication with C. A. Evans[18].

Drude model was used to estimate the permittivity of materials. Evans modelled the active region as bulk $Si_{0.9}Ge_{0.1}$, to match the virtual substrate in a (111) Si-rich QCL[3]. A doping concentration of $5 \times 10^{16}\,\mathrm{cm}^{-3}$ was used in the active region.

Surface plasmons (vibrations in electron gases near the electrical contacts in QCLs) have successfully been used to pin optical modes[201]. In a semi-insulating surface plasmon waveguide[49], the electrical contact at the top of the active region stack consists of a thin layer of highly doped semiconductor capped with metal. The bottom contact consists of a second highly doped layer between the substrate and the active region.

Figure 7.3 shows that the optical mode in such a structure spreads significantly into the substrate. Evans fixed the highly-doped top contact thickness to 20 nm, but used a much thicker bottom contact. Figure 7.4 shows that the mode is confined more effectively by increasing the thickness of the bottom contact. However, the waveguide also becomes more lossy, leading to a higher threshold gain.

Optimal bottom contact doping was found to be $2{\times}10^{19}\,\mathrm{cm}^{-3}$, as shown in fig. 7.5. This corresponds to a modal overlap of only 17% and losses of $10.2\,\mathrm{cm}^{-1}$. A fairly large threshold gain of $68\,\mathrm{cm}^{-1}$ is required to overcome the poor overlap.

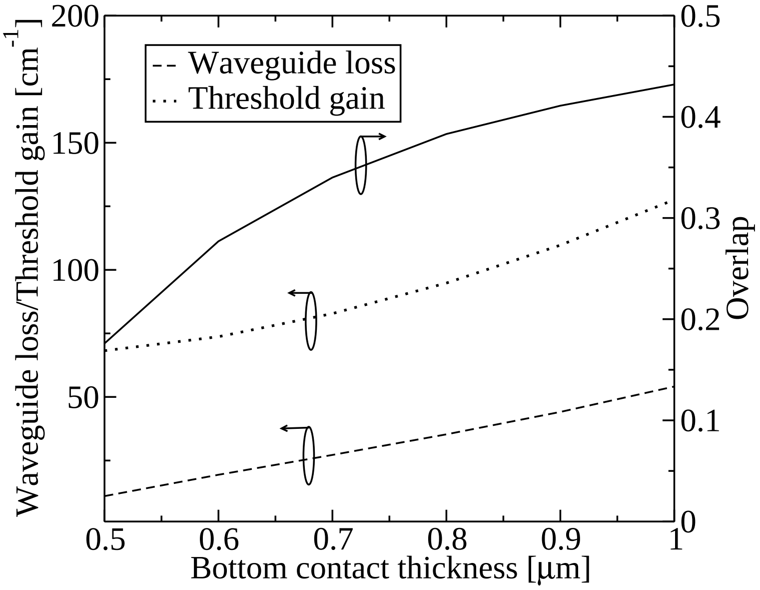


Figure 7.4: Modal overlap, waveguide loss and threshold gain as a function of bottom contact thickness in a Si waveguide. Doping was fixed at $2{\times}10^{19}\,\mathrm{cm}^{-3}$ in the bottom contact. Adapted from joint publication with C. A. Evans[18].

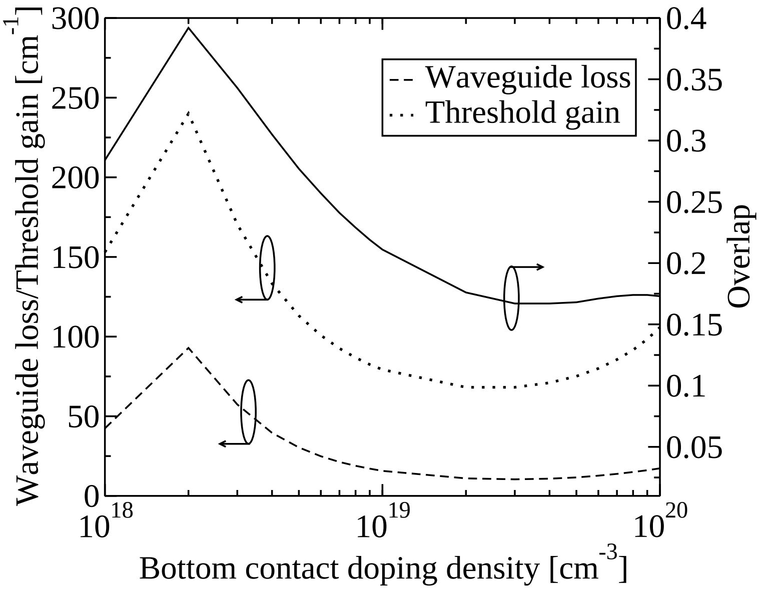


Figure 7.5: Modal overlap, waveguide loss and threshold gain as a function of bottom contact doping in a Si waveguide. The contact thickness was fixed at 0.5 μm. Adapted from joint publication with C. A. Evans[18].

Figure 7.6: Optical mode profile in a double-metal Si waveguide. The substrate has been replaced by a second metallic contact at the bottom of the figure. Adapted from joint publication with C. A. Evans[18].

The poor modal overlap may be avoided by using a double-metal waveguide, which has proved successful in GaAs THz QCLs[202]. Here, the bottom contact is formed by etching away the substrate and depositing a second metallic layer. The mode is now unable to escape the active region ($\Gamma = 99\%$) as shown in fig. 7.6. An initial design using gold contacts and a 10μm thick active region gave a waveguide loss of $\alpha_w = 50.7\,\mathrm{cm}^{-1}$. The threshold gain required to overcome these losses was $G_{th} = 51.7\,\mathrm{cm}^{-1}$.

High operating temperatures for GaAs-based THz QCLs have been achieved recently by using copper instead of gold in a double-metal waveguide[50, 203]. Evans showed that with a Cu-Cu Si waveguide, threshold gain could be reduced to $36.9\,\mathrm{cm}^{-1}$ for a $10\,\mu$m active region[3] or $31.0\,\mathrm{cm}^{-1}$ for $15\,\mu$m[18]. Evans's values for threshold gain were assumed to be invariant with respect to temperature and to remain approximately correct for Ge-rich structures.

The results of Evans for single-metal waveguides show that increasing the doping density improves the reflectivity of the bottom contact, but also leads to greater free carrier losses. A third option for waveguide design overcomes this problem by using a buried layer of a metal silicide as a reflector[204]. Buried tungsten[204] and cobalt[205] silicides have been grown successfully in SOI wafers. Although the losses have not been calculated for n-type buried silicide waveguides, the results for p-type structures show that for an $8\,\mu$m thick p-type SiGe active region emitting at $62\,\mu$m, $\Gamma \approx 90\%$ and $G_{th} = 31\,\mathrm{cm}^{-1}$. This represents a threshold gain around double that of a double-metal waveguide. Therefore, the best results are expected from double–metal and semi-insulating surface plasmon structures in n-type devices and these will be used in the remainder of this work.

7.5 Example simulation

It is useful to illustrate the theoretical methods discussed in this chapter by analysing a sample system from the literature. Driscoll and Paiella recently proposed a

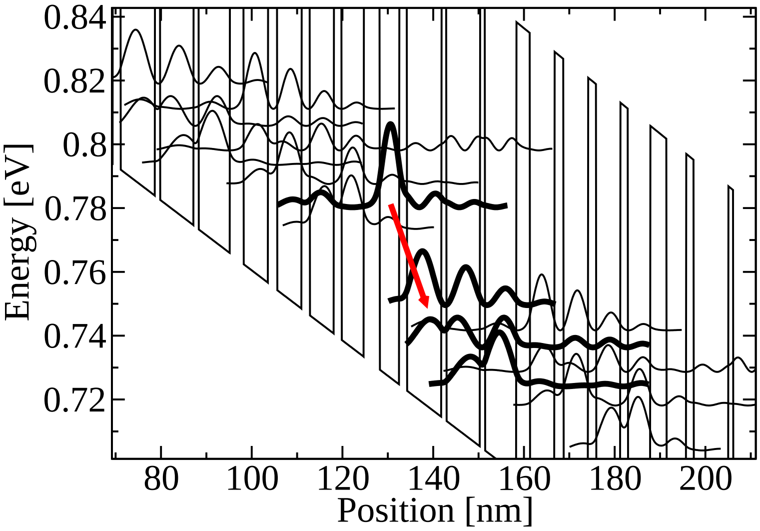


Figure 7.7: Conduction band profile and electron probability densities for Driscoll's 42 μm Ge QCL design[77]. The ULL, LLL and a pair of extraction levels are shown in bold and the nominal laser transition is denoted by the arrow. The layer widths were specified as **1.8**/4.3/**1.7**/7.6/**1.1**/7.4/**1.1**/6.9/**3.0**/5.4/**2.0**/5.4/**1.8**/5.3/**1.7**/4.9/**1.8** nm, where bold denotes 78% Ge barriers, and normal text denotes Ge wells. Dopants were spread evenly throughout the device at a concentration of $n_{2D} = 5 \times 10^{10}\,\mathrm{cm}^{-2}$.

Ge/GeSi QCL emitting near $42\,\mu$m[77], which is illustrated in fig. 7.7. Electrons are injected into the ULL from an *injector level*, which is just below the ULL in energy. The injector wavefunction extends slightly into the active region and electrons scatter into the ULL. The LLL overlaps strongly with the pair of extraction levels, which in turn overlap with the miniband states outside the active region, allowing rapid depopulation of the LLL.

The emission wavelength between the nominal laser subbands (arrow in fig. 7.7) was calculated as $40.4\,\mu$m, for an electric field of $11\,\mathrm{kV\,cm^{-1}}$, corresponding to a subband separation of 30.7 meV or 7.4 THz. This is slightly different from the value of $42\,\mu$m predicted by Driscoll as the original paper did not ensure strain balancing and used an 88% Ge substrate. Using the method of chapter 6, it can be shown that this limits the structure to a maximum thickness of 38 nm. In this chapter, a substrate composition of 94% Ge is used, which ensures mechanical stabililty, but changes the wavefunctions and subband energies slightly.

7.5.1 Electron temperature

Figures 7.8 and 7.9 show that the electron temperature is a function of both lattice temperature and electric field. Increasing the lattice temperature increases the rate of phonon absorption, while increasing the electric field increases the energy separation between subbands. Both processes increase the number of electrons scattering into high energy states and hence increase the electron temperature.

Figure 7.8 shows that when $T > 50\,\mathrm{K}$, the electron temperature rises linearly with respect to lattice temperature, with the functional form $T_e = T + T_0(E)$. T_0 increases from 25 to 87 K as the electric field increases from 5 to 15 kV/cm. Below 50 K, the inelastic scattering rate becomes very small, as shown in chapter 5, and the linear relationship breaks down.

Figure 7.9 shows that the relationship between electron temperature and electric field is quite complicated due to multiple subbands moving in and out of resonance. However, as the electric field increases, electrons are injected into higher kinetic

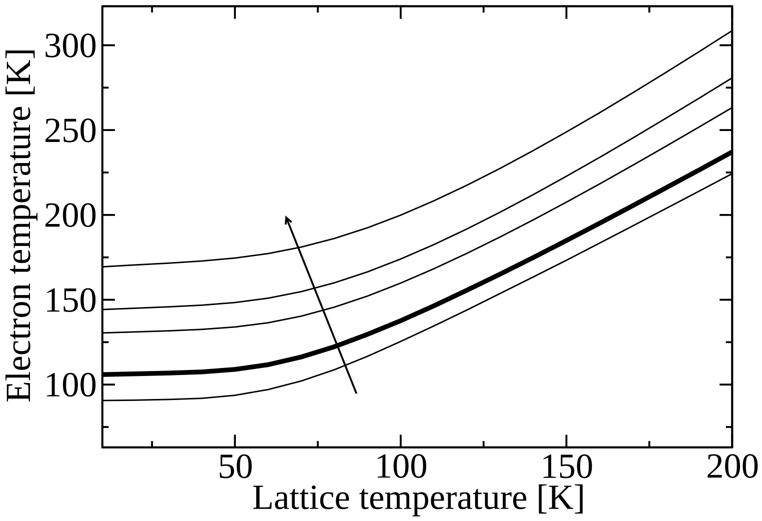


Figure 7.8: Variation of electron temperature with respect to lattice temperature in Driscoll's 42 μm design. Results are shown for electric fields linearly varying between 5 and 13 kV cm^{-1} (arrow denotes increasing electric field). The bold line denotes the design field of 11 kV cm^{-1}.

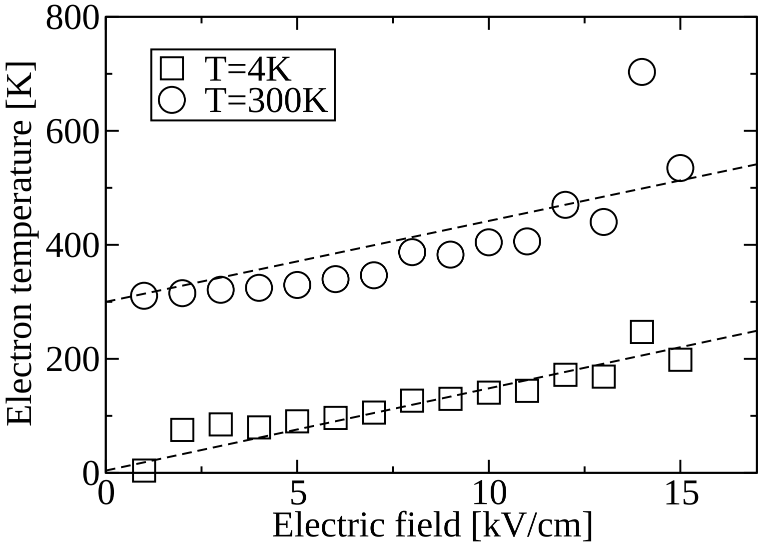


Figure 7.9: Variation of electron temperature with respect to electric field in Driscoll's 42 μm design.

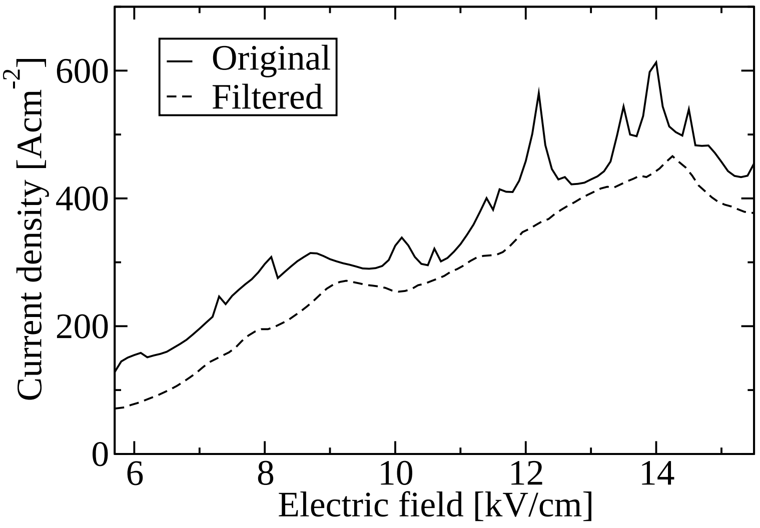


Figure 7.10: Current density as a function of electric field for Driscoll's $42\,\mu$m design at 4 K lattice temperature. The solid line shows the direct output of the simulation, including sharp spikes due to unrealistic parasitic current channels. The dashed line shows the results of the filtering algorithm, which removes the parasitic spikes.

energy states and the electron temperature generally rises. As the electric field shrinks toward zero, the device approaches thermal equilibrium and the electron temperature approaches the lattice temperature.

7.5.2 Current density

The current density is plotted against the electric field for Driscoll's design in fig. 7.10. Before commenting on the device-specific behaviour, it is important to note that the plot illustrates a fundamental limitation of the semi-classical approach used in the present work. At certain electric fields, pairs of miniband states move into close alignment, leading to hybridisation of wavefunctions which were previously localised on opposite sides of a barrier. The spatial extent of the wavefunctions becomes very large and electrons can apparently be transported instantaneously from one side of the barrier to the other once they enter the pair of hybridised states.

In reality, wavepackets oscillate coherently across barriers and decay via dephasing processes before such extended states can form.

A full treatment of this mechanism is only possible by moving to more detailed quantum mechanical approaches[148]. However, the following data analysis algorithm may be used to identify and remove the effects of the unrealistic parasitic current channels. Firstly, for each wavefunction ψ_i, the spatial extent is defined as the length $L_i = z^+ - z^-$ containing 98% of the electron probablility density such that

$$\int_{-\infty}^{z^-} \psi_i^2 \mathrm{d}z = \int_{z^+}^{\infty} \psi_i^2 \mathrm{d}z = 0.01. \tag{7.41}$$

Wavefunctions with spatial extent greater than a structural period $L_i > L_p$ are identified as sources of parasitic current and interminiband scattering originating from these states is disregarded. This completely eliminates the large current density spikes, but leaves small discontinuities (1–10 $\mathrm{A\,cm^{-2}}$) at the points where wavefunctions become long enough to meet the filtering criterion. Finally, these discontinuities are removed by applying a 3-sample moving average filter. The resulting line in fig. 7.10 shows that the parasitic spikes have been removed entirely without changing the overall shape of the plot. A 5–10% overall reduction in current density occurs away from the current spikes, indicating that interminiband scattering is normally only a minor contributor to the total current density.

Having removed the unrealistic features of the current density plot, it is now possible to observe several aspects of the behaviour of the QCL. Firstly, when there is no electric field, there is no preferential direction for charge transport and the current drops to zero. As the electric field increases, there is a drop in potential across each period of the structure and a current therefore flows. The predicted current densities are in the same range as those measured for bound-to-continuum THz QCLs[52, 206]. At the design field of 11 kV/cm, the injector level lies below the ULL. A local maximum in the current density occurs around an electric field of 14.2 kV/cm. This corresponds to the point at which the ULL and injector levels are in resonance. At higher electric fields, the injector and ULL move out of resonance

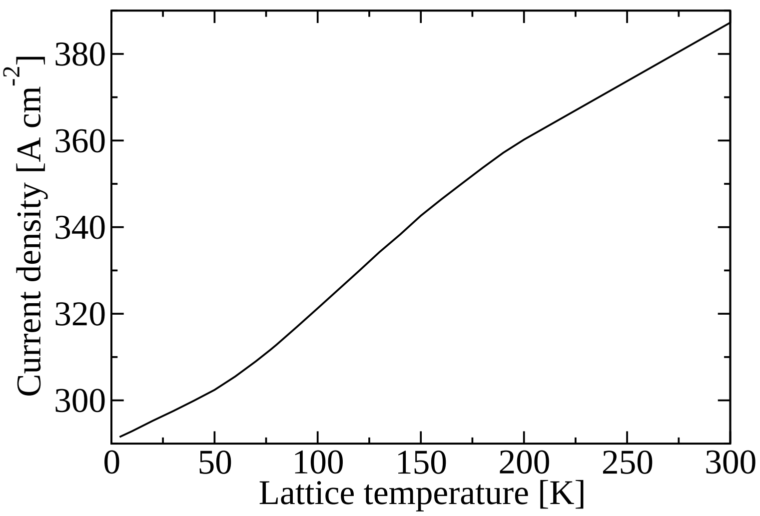


Figure 7.11: Current density as a function of electric field for Driscoll's 42 μm design as a function of lattice temperature.

and electrons are injected more slowly.

Finally, fig. 7.11 shows that current density increases as a function of lattice temperature, as observed experimentally in GaAs-based THz QCLs[36, 206]. This is due to the corresponding increases in electron temperature and phonon scattering. The relationship is approximately linear, with the form $J = 287 + 0.354T$ where J is in units of A cm^{-2} and T is in units of K.

7.5.3 Gain

The spectrum near the design field of 11 kV/cm is plotted in fig. 7.12. Figure 7.13 shows that the peak gain occurs at 7.26 THz at a field of 10.5 kV/cm rather than the intended value. The difference in value is due to the change in substrate composition, as explained previously. The gain never exceeds the estimated threshold for the waveguide structures discussed in this chapter, and the device is therefore incapable of lasing. As the electric field increases, the energy separation between the subbands increases and the emission frequency can be seen to increase. This is known as the

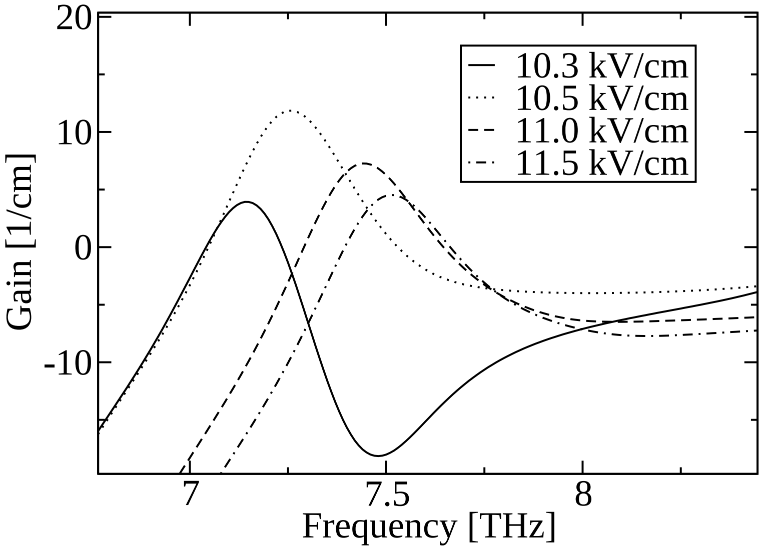


Figure 7.12: Active region gain spectrum for Driscoll's 42 μm design at electric fields near to the design field of 11 kV/cm. Results are shown at a lattice temperature of 4 K.

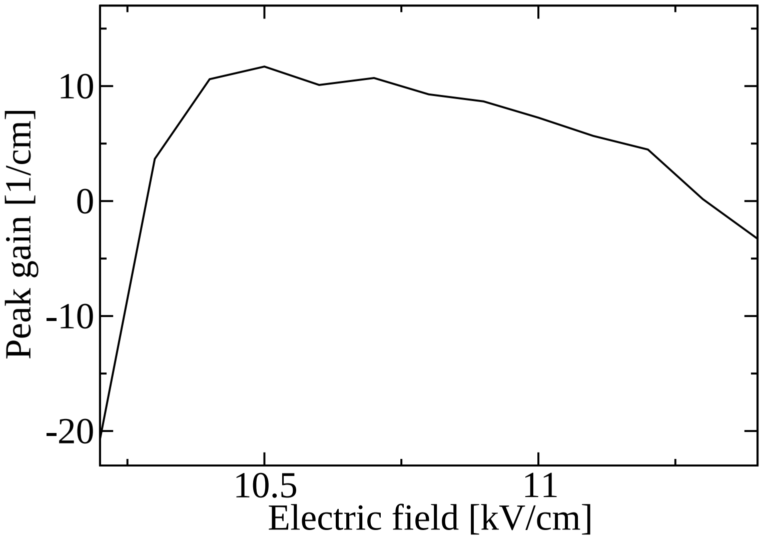


Figure 7.13: Peak gain as a function of electric field for Driscoll's 42 μm design at 4 K.

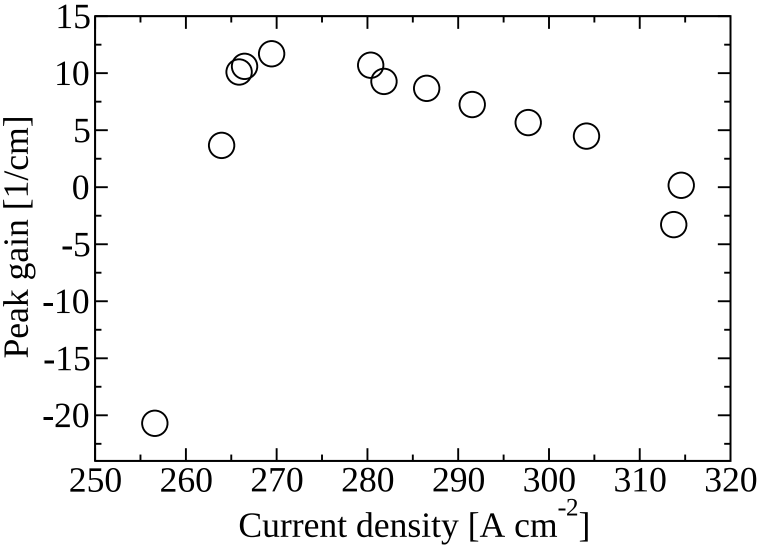


Figure 7.14: Peak gain as a function of current density for Driscoll's 42 μm design at 4 K.

quantum-confined Stark effect, and it allows the development of electrically-tunable QCLs[207]. Figure 7.14 shows that the relationship between gain and current density is highly nonlinear. The gain coefficient model is shown to be a poor approximation for the behaviour of THz QCLs as discussed previously.

Figure 7.15 shows that the peak gain rolls off for lattice temperatures above 105 K due to the LLL being repopulated by phonon absorption, which reduces the population inversion. Increasing the temperature from 4 to 105 K appears to improve to a maximum value of 14.7 cm^{-1}. This is due to enhanced injection of electrons into the ULL. The latter conclusion must be greeted with some caution, however, as experimental data for III-V devices suggests that the optical power output decreases monotonically with respect to lattice temperature[49, 195].

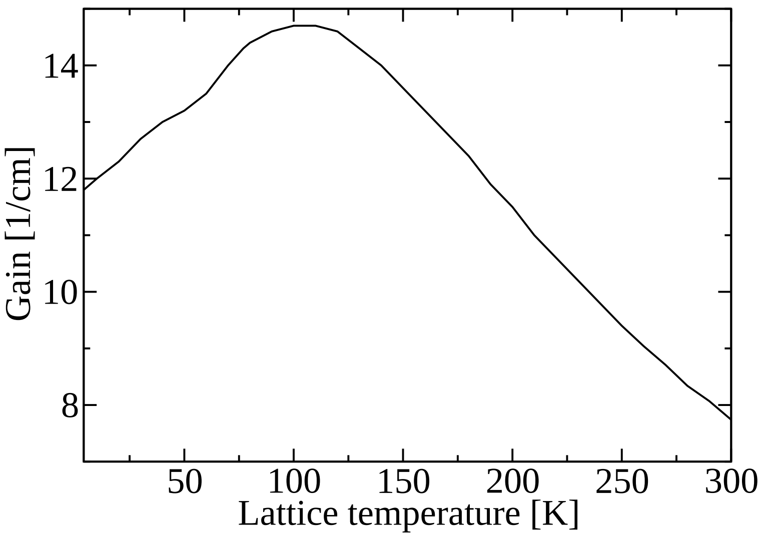


Figure 7.15: Active region gain as a function of temperature for Driscoll's $42\,\mu$m design in $10.5\,\mathrm{kVcm}^{-1}$ electric field as a function of lattice temperature.

7.6 Conclusion

A method has been described for calculating the gain and current density in a QCL, thus completing the QCL model. A flowchart for the full simulation is shown in fig. 7.16. The simulation now provides a fully self-consistent calculation of the Schrödinger, Poisson and energy-balance equations, using a semi-classical description of charge transport.

It has been shown that the population inversion required for lasing may be achieved by rapid depopulation of the LLL or rapid filling of the ULL such that $W_{21} > W_{32}$. The distribution of electrons within subbands was described with an effective electron temperature and a steady state value was found using an energy-balance approach. This affects the scattering rates significantly and hence the subband populations and current density within a QCL. Electron temperatures were shown to increase with respect to lattice temperature and electric field.

Current density was shown to vary in a complicated manner with respect to

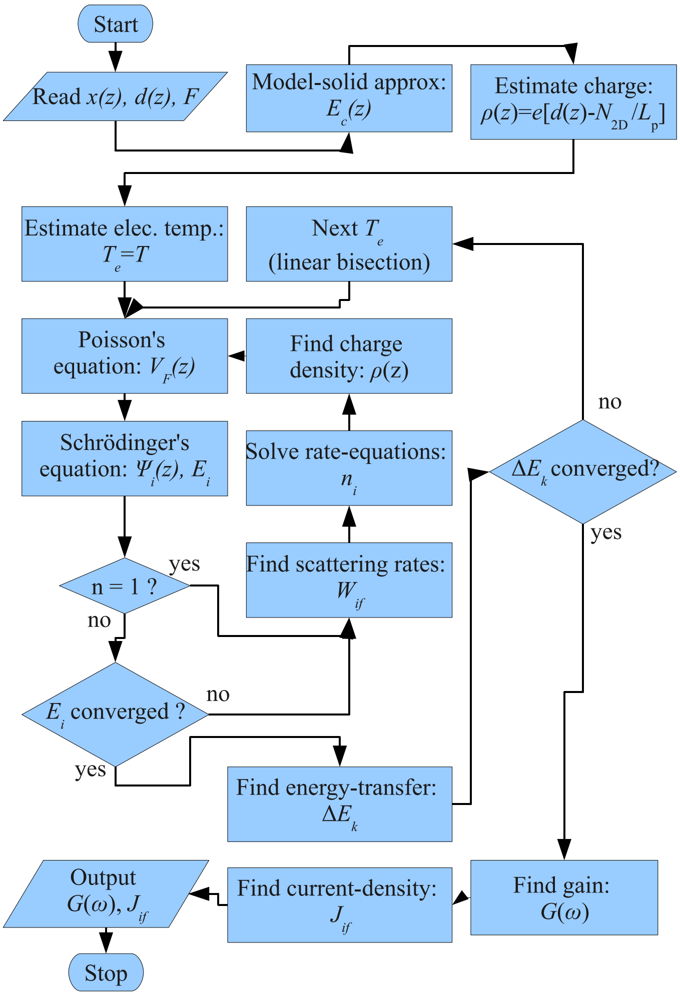


Figure 7.16: Complete flowchart for QCL simulation. “$n = 1$” denotes the first iteration of the self-consistent Poisson–Schrödinger calculation.

electric field. As the design field is approached, the injector level of a QCL comes into resonance with the ULL and a sharp increase in current density is seen. A negative differential resistance exists at higher fields, as the injector level exceeds the energy of the ULL.

It was shown that single-metal waveguides provide low losses and poor modal overlap giving an optimal threshold gain of $51.7\,\mathrm{cm}^{-1}$ for a $10\,\mu\mathrm{m}$ active region. The overlap was increased to almost 100% in double-metal waveguides at the expense of higher losses. The threshold gain was reduced to $36.9\,\mathrm{cm}^{-1}$. Buried silicide waveguides provide high overlap and moderate losses, giving a threshold gain of around $70\,\mathrm{cm}^{-1}$.

Finally, the full simulation was demonstrated by modelling the Ge QCL design of Driscoll. It was shown that unfortunately the design yields insufficient active region gain to overcome the waveguide losses. There are several reasons for this. Firstly, the emission energy exceeds the $L \rightarrow L$ phonon energy, meaning that the ULL population is reduced by rapid phonon emission. Secondly, the miniband is wide, meaning that strong optical absorption occurs near to the ULL→LLL frequency. Finally, the relatively large electric field yields high electron temperatures and electrons are able to backfill into the LLL, weakening the population inversion. In the next chapter, it is shown that each of these issues can be addressed and a large net gain is possible from Ge/GeSi QCLs.

Chapter 8

QCL design examples

The simulation tools presented in earlier chapters will now be used to investigate a set of sample QCL designs. An exhaustive exploration of the design space for THz QCLs is not feasible, due to the large set of possible material configurations, design schemes and doping profiles. Neither is it possible to give a true "like–for–like" comparison between QCL designs in different material configurations. This is because a design based on a given scheme must be altered significantly to tailor its performance to a particular material system. Therefore, rather than presenting a comparison of suboptimal designs in each material configuration, this chapter presents a set of distinct, optimised designs following the approximate chronological order in which they were investigated.

8.1 Design schemes

A superlattice is an infinitely long series of QWs (*i.e.* a periodic potential). Bloch's theorem shows that the electronic subbands are continuously distributed in energy within "minibands" separated by "minigaps". The superlattice serves as the basis for all QCL designs, however, when an external electric field is applied, the minibands are broken into discrete subbands. In a recent review paper[36], Williams noted that successful THz QCL designs have modified the biased superlattice

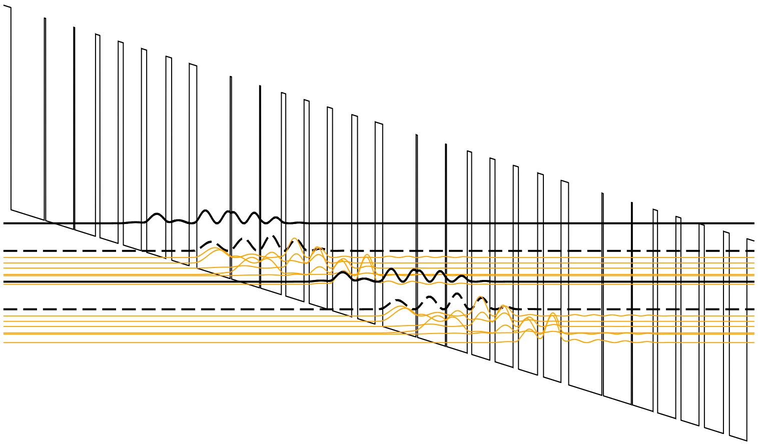


Figure 8.1: Wavefunctions in a chirped superlattice 4.4 THz GaAs/AlGaAs QCL design[49]. The optical transition occurs between the solid and dashed black lines at the edges of a pair of minibands. Image courtesy of Craig Evans, University of Leeds

structure to achieve population inversion according to two general design schemes.

The earliest approach to THz QCLs involves creating a chirped superlattice (CSL) structure[49] as shown in fig. 8.1. Here, QWs are made successively narrower across each period of the device, so that sets of subbands lie in narrow energy ranges when an electric field is applied. This effectively restores the miniband/minigap character of the unbiased superlattice. Due to the large overlap between subbands within a miniband, the intraminiband scattering is faster than scattering across a minigap, which leads to a population inversion between a pair of minibands. Nevertheless, a significant amount of nonradiative scattering still occurs across the minigap, limiting the population inversion.

According to Williams, the CSL design scheme has been effectively superseded by “bound-to-continuum” (BTC) devices, illustrated in fig. 8.2, in which a state bound in a narrow QW next to an injection barrier forms the upper laser level. The spatial extent of the upper laser level is reduced and hence the nonradiative scattering matrix element between the minibands is also reduced. Selective injection of

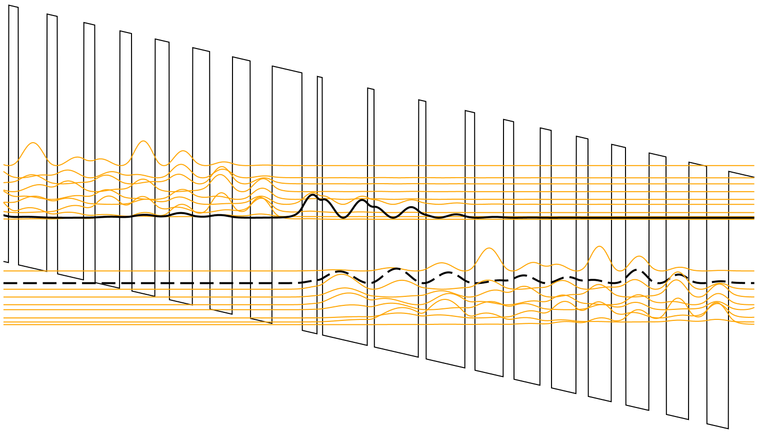


Figure 8.2: Wavefunctions in a bound–to–continuum mid-infrared GaAs/AlGaAs QCL[208]. The upper laser state is trapped in a narrow well next to the injection barrier. Image courtesy of Craig Evans, University of Leeds

electrons into the upper laser level is achieved by resonant tunneling through a thick injection barrier. GaAs-based QCLs in both the mid-infrared[208] and THz[209] regions have been successfully demonstrated. In principle, the BTC design scheme should not require any special modifications for SiGe designs. Indeed, a number of designs have been proposed for Ge/GeSi[77, 78, 210] and (111) Si/SiGe[3] BTC QCLs.

The second commonly used design scheme in III–V devices is the resonant-phonon (RP) QCL, illustrated in fig. 8.3. Rather than using intraminiband scattering to depopulate the lower laser level, these devices rely upon rapid inelastic scattering due to electron–phonon interactions. In III–V materials, the polarity of the chemical bonds in the material leads to rapid polar-LO phonon scattering between subbands separated by the phonon energy ($\hbar\omega_{\mathrm{LO}} = 36\,\mathrm{meV}$ in GaAs). The wavefunctions are engineered so that the lower laser level is in resonance with the second excited state of a quantum well, which in turn has a ground state $\sim \hbar\omega_{\mathrm{LO}}$ lower in energy[211]. The highest operating temperatures in III–V systems have been achieved using this approach[50].

In SiGe-based systems, there is no polar-LO phonon interaction, as no ionic

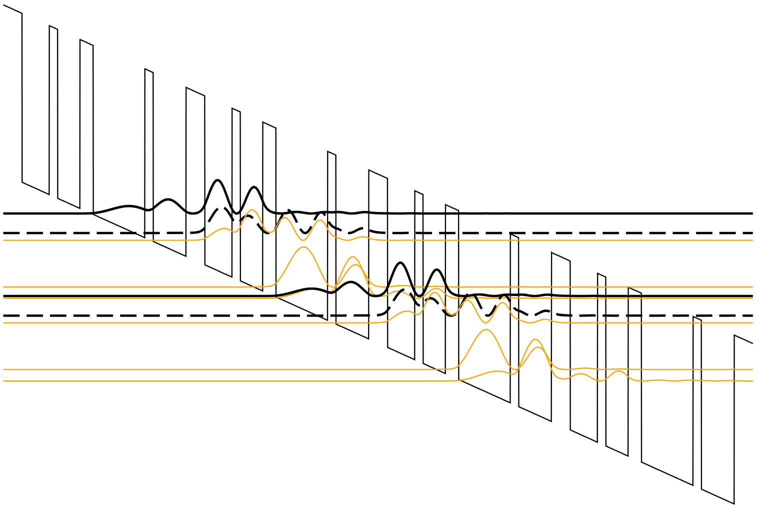


Figure 8.3: Wavefunctions in a resonant–phonon 3.4 THz GaAs/AlGaAs QCL[202]. The orange depopulation levels are separated by the LO-phonon energy. Image courtesy of Craig Evans, University of Leeds

bonds exist in the crystal. Instead, intervalley phonon interactions exist due to the presence of multiple equivalent valleys. The intervalley phonon scattering was shown in chapter 5 to saturate above the phonon energy, rather than to peak. This means that the overlap between the LLL and the depopulation state must be much greater than that between the ULL and the depopulation state to ensure selective depopulation of the LLL. Phonon depopulated designs are impractical in (001) oriented Si/SiGe, due to the very large g-LO phonon energy ($\hbar\omega = 63\,\text{meV}$), although the lower energy f-LA and $L \rightarrow L$ phonon processes are usable in (111) Si/SiGe and (001) Ge/GeSi respectively. Hybrid BTC/RP (or “interlaced”) devices have also been developed[212], which are fairly similar in design to a conventional BTC structure. They differ by having a depopulation miniband which is split in two by a minigap tuned to the phonon energy in order to exploit both depopulation mechanisms. The only phonon depopulated n-Si/SiGe QCL design published to date[4] employs such a design scheme.

During the course of the present work, four QCL designs were developed and simulated. The designs were based upon the well-established III–V design schemes described above, with the typical layer widths rescaled to account for the difference in quantisation effective mass via the transformation

$$L' = L\sqrt{\frac{m_q}{m'_q}}, \tag{8.1}$$

where L is the layer width, and the prime notation indicates the values for the modified system. The barrier material composition was selected to allow all the QCL subbands to lie within the bottom half of the confining potential and the electric field was scaled to account for the reduced period length. This approach avoids the need for tall, thin barriers which are likely to be challenging to grow. Preliminary designs were generated using this approach, and subsequently optimised by adjusting barrier composition, electric field and individual layer widths to maximise the gain. The following sections describe the devices which were investigated and their predicted performance.

8.2 (001) Si/SiGe QCLs

In chapter 5, the maximum intersubband transition lifetime for (001) oriented Si was shown to occur around 13.5 meV (3.3 THz), which implies that a QCL with a transition energy around this value will achieve maximal population inversion. Barbieri successfully demonstrated a 2.9 THz BTC GaAs/AlGaAs QCL operating up to 95 K with a threshold current density of 105 A/cm^2[206]. The sheet doping for Barbieri's design was 3.7×10^{10} cm^{-2}, the period length was 128.6 nm and the electric field was 2.1 kV/cm.

Figure 8.4 shows a plot of the electron densities and potential profile in a (001) Si/SiGe BTC QCL design based upon that of Barbieri. The population inversion was improved by trapping the ULL in the well adjacent to the injection barrier, and increasing the injection barrier thickness. Figure 8.5 shows that optimum gain occurs at an electric field of 5.2 kV/cm, with a frequency of 2.7 THz (11 meV). Strong intraminiband absorption occurs below 1.5 THz.

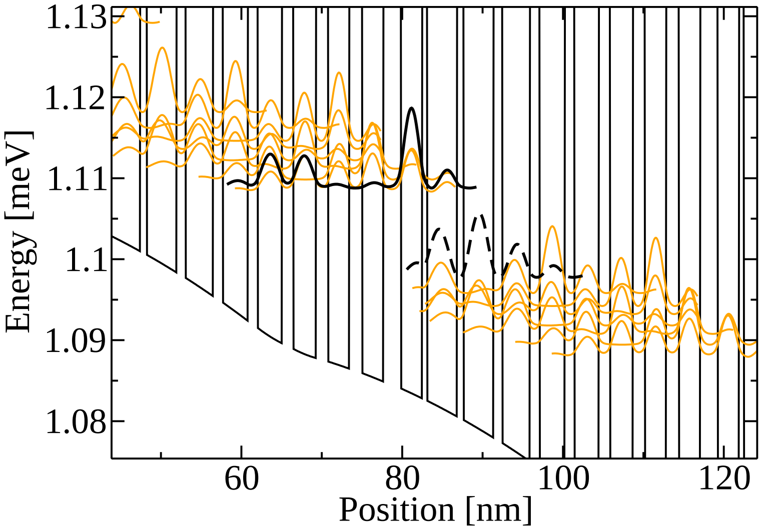


Figure 8.4: Potential profile and electron probability densities for a 2.7 THz (001) Si/SiGe BTC QCL. Layer widths of **2.2**/2.6/**0.6**/3.8/**0.8**/3.7/**1.1**/3.4/**1.2**/3.2/**1.2**/<u>3.0</u>/**1.4**/<u>2.8</u>/**1.5**/2.7/**1.6**/2.6 nm were used, where bold text denotes $\mathrm{Si_{0.8}Ge_{0.2}}$ barriers, normal text denotes Si wells and underlined text denotes *n*-doped layers at a concentration of $6.4\times10^{16}\,\mathrm{cm^{-3}}$. Results are shown at an electric field of $5.2\,\mathrm{kVcm^{-1}}$.

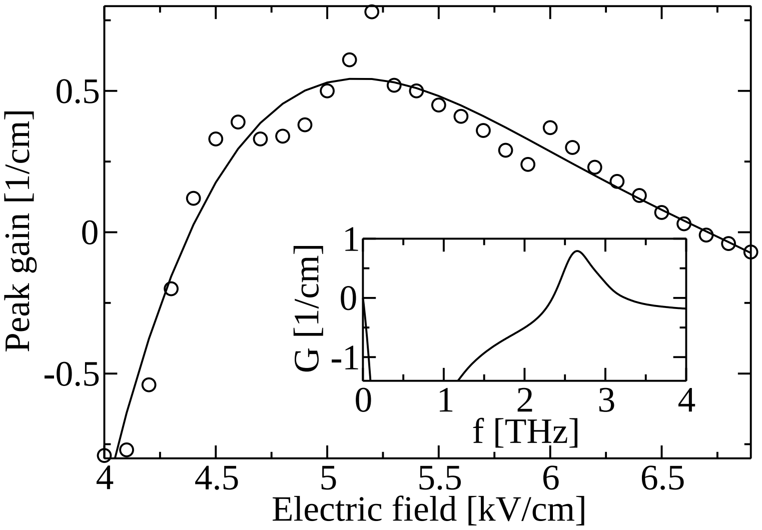


Figure 8.5: Peak gain as a function of electric field for a (001) Si/SiGe BTC QCL at a lattice temperature of 4 K. (Inset) gain spectrum at an electric field of 5.2 kV/cm.

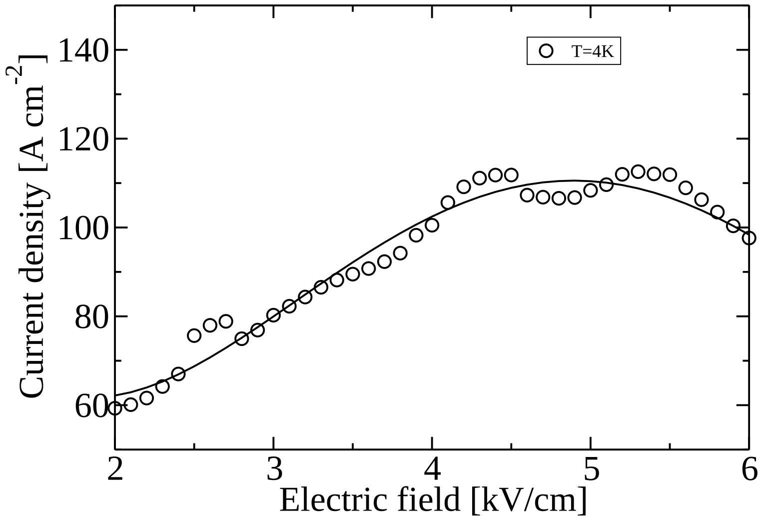


Figure 8.6: Current density as a function of applied electric field for a (001) Si/SiGe BTC QCL at a lattice temperature of 4 K.

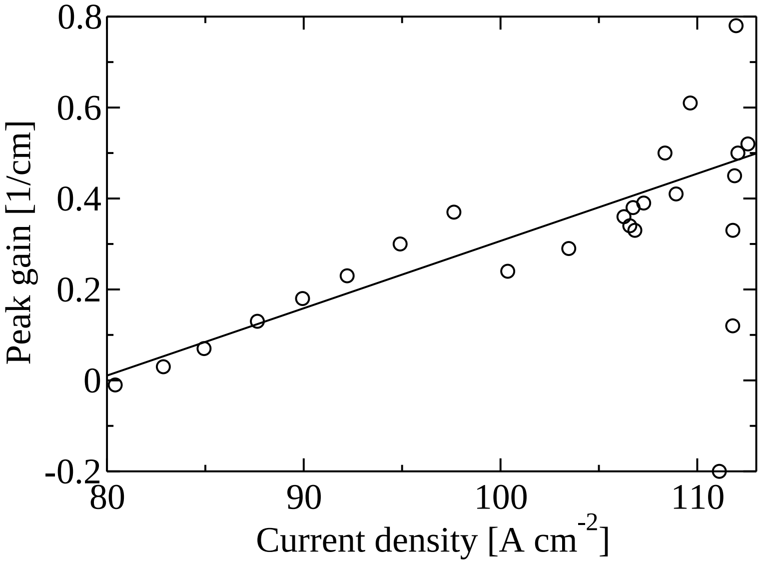


Figure 8.7: Peak gain as a function of current density for a (001) Si/SiGe BTC QCL at a lattice temperature of 4 K. A linear regression has been applied to data in the range $J < 110\,\mathrm{A\,cm^{-2}}$.

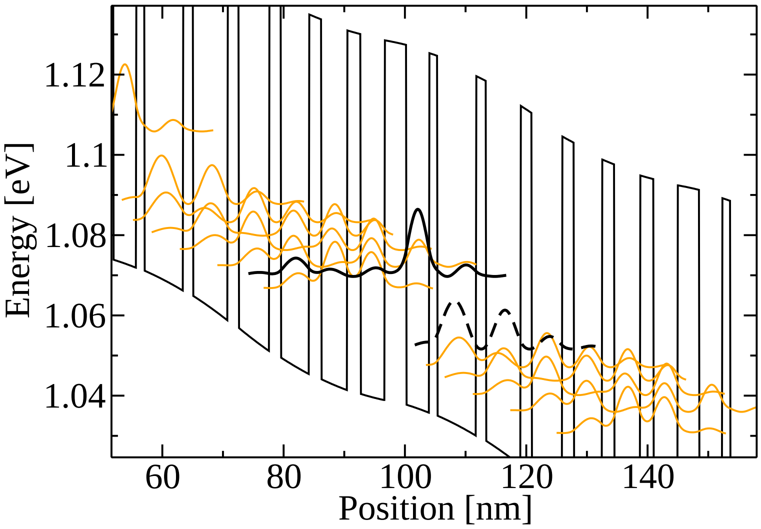


Figure 8.8: Potential profile and electron probability densities for a 4.2 THz (111) Si/SiGe BTC QCL. Layer widths of **3.6**/3.7/**1.4**/6.4/**1.6**/5.7/**1.8**/<u>5.1</u>/ <u>**1.9**</u>/<u>4.6</u>/<u>**2.1**</u>/<u>4.2</u>/ <u>**2.2**</u>/<u>4.0</u> nm were used, where bold text denotes $Si_{0.4}Ge_{0.6}$ barriers, normal text denotes Si wells and underlined text denotes n-doped layers at a concentration of 5×10^{16} cm^{-3}. Results are shown at an electric field of 7.1 $kVcm^{-1}$.

At this electric field, the injector subband resonates with the upper laser level, giving a peak current density of 113 Acm^{-2} at a lattice temperature of 4 K as shown in fig. 8.6. The maximum active region gain was calculated as 0.79 cm^{-1}, which is far below the threshold of 31 cm^{-1} estimated in chapter 7 for a 15 μm thick double-metal waveguide. Figure 8.7 shows the peak gain as a function of current density, along with a linear regression to the data in the range $J < 110\,\mathrm{A\,cm^{-2}}$. The relationship takes the approximate form $G = gJ - \alpha$, where $g = 15\,\mathrm{cm/kA}$ and $\alpha = 1.17\,\mathrm{cm^{-1}}$. The oscillator strength for the optical transition was calculated as 0.09 (*c.f.* 1.5–2.0 for GaAs[36]).

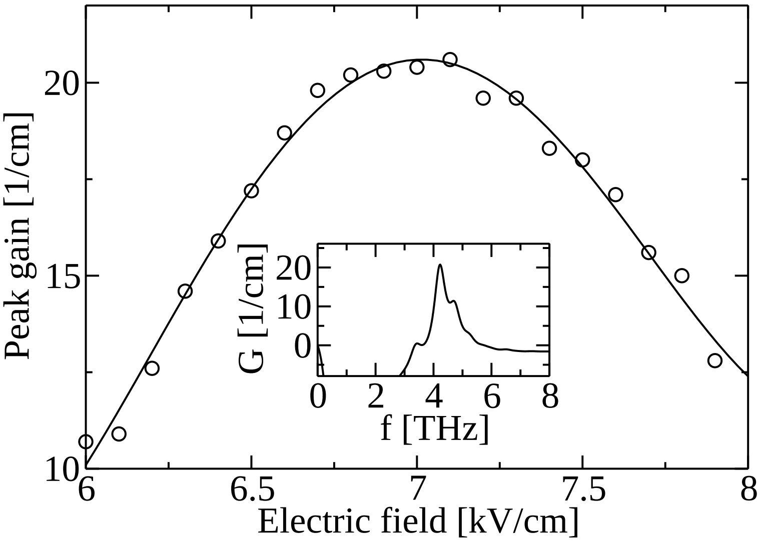


Figure 8.9: Peak gain as a function of electric field for 4.2 THz (111) Si/SiGe BTC QCL at a lattice temperature of 4 K. (Inset) Gain spectrum at a bias of 7.1 kV/cm.

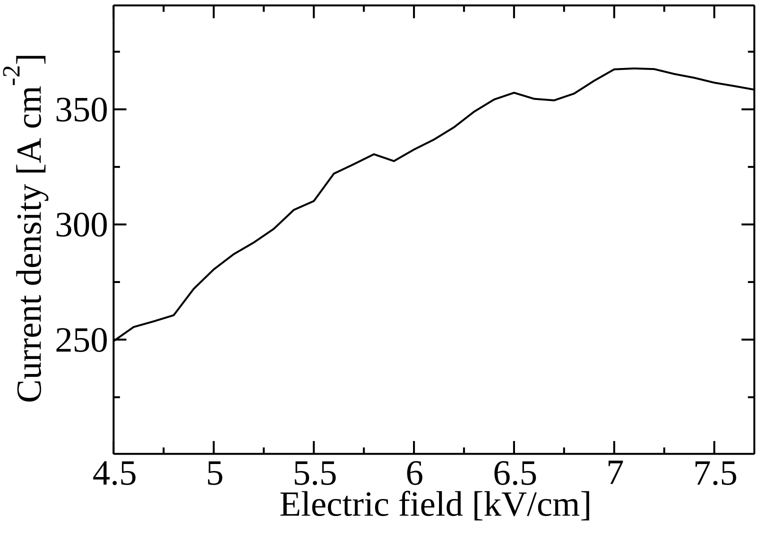


Figure 8.10: Current density as a function of applied electric field for a 4.2 THz (111) Si/SiGe BTC QCL at a lattice temperature of 4 K.

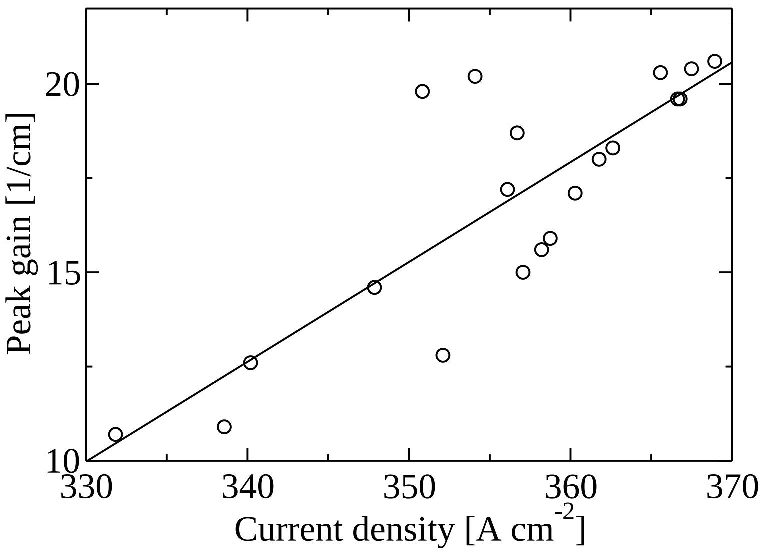


Figure 8.11: Peak gain as a function of current density for a 4.2 THz (111) Si/SiGe BTC QCL at a lattice temperature of 4 K. The solid line shows a linear regression to the data.

8.3 (111) Si/SiGe QCLs

Two (111) oriented Si/SiGe designs were investigated during the course of the present work. The first design is a BTC design similar to that of the previous section, but with seven QWs per period rather than nine. In chapter 5, the minimum scattering rate was shown to occur between 3.1 and 9.7 THz. The QCL shown in fig. 8.8 is designed to emit at 4.2 THz, thus minimising the nonradiative depopulation of the ULL. A virtual substrate Ge fraction of 16.1% is required for mechanical stability. Figure 8.9 shows that the peak gain occurs at an electric field of 7.1 kV/cm and from fig. 8.10, this field can be seen to give a current density of 367 A cm^{-2}. The largest achievable gain is 21 cm^{-1}, which is superior to both the (001) Si/SiGe design presented in the previous section and to Driscoll's Ge/GeSi design, but still insufficient to overcome the estimated gain threshold for a 15 μm thick Cu–Cu waveguide.

A linear regression to the relationship between gain and current density is shown in fig. 8.11 with the form $G = gJ - \alpha$, where $g = 265$ cm/kA and $\alpha = -77.5$ cm^{-1}.

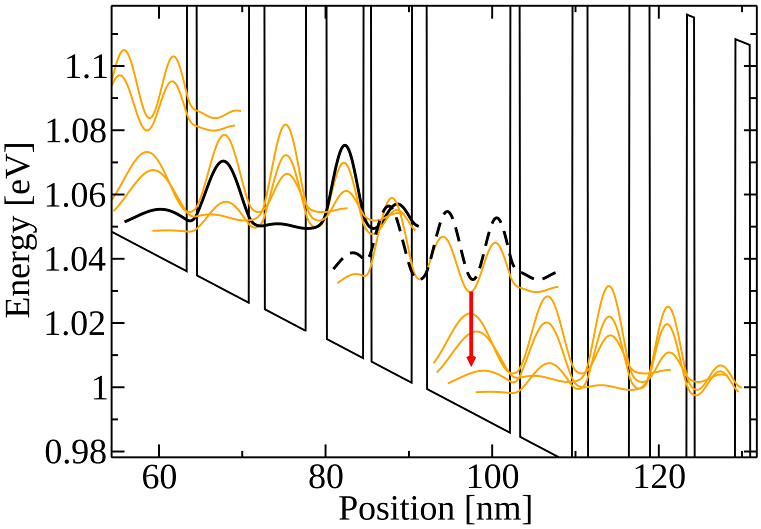


Figure 8.12: Potential profile and electron probability density plot for (111) Si/SiGe hybrid BTC/RP QCL at an electric field of 13 kV/cm. The phonon mediated depopulation is represented by the red arrow. Layer widths of 6.3/**1.8**/5.0/**2.5**/4.4/**0.9**/4.9/**1.8**/10.0/**1.2** were used, where bold text denotes $Si_{0.2}Ge_{0.8}$ barriers and normal weighted text denoted Si wells. A 14% Ge virtual substrate was used, and dopants were spread evenly through the structure at a density of $2 \times 10^{16}\,cm^{-3}$.

The oscillator strength for this design is calculated as 0.46; a factor of ~ 5 improvement over the previous design.

The second (111) Si/SiGe QCL design, produced in collaboration with Leon Lever at the University of Leeds, uses of a hybrid BTC/RP design scheme. Figure 8.12 shows the bandstructure for the design. The phonon depopulation requires a larger electric field than the conventional BTC designs considered so far (13 kV/cm in the present design, 7.1 kV/cm in the previous BTC design) to accommodate the extra phonon depopulation minigap. A larger confining potential is also required, and hence a larger step in Ge fraction is used at heterointerfaces (80% in the present design, 60% in the previous BTC design).

As shown in chapter 5, the intervalley $\Delta \rightarrow \Delta$ scattering rate in Si saturates for subband separations above the f-LA phonon energy (46.3 meV), giving optimal

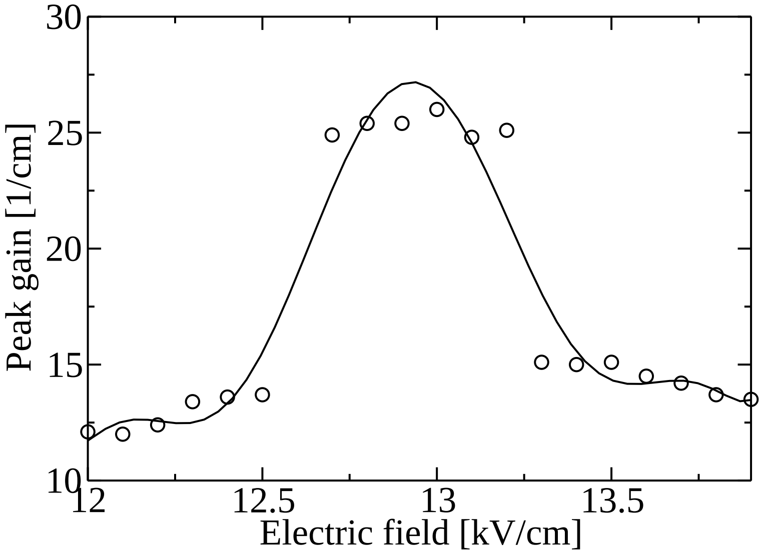


Figure 8.13: Peak gain as a function of applied electric field for a 3.8 THz (111) Si/SiGe resonant phonon QCL at a lattice temperature of 4 K.

depopulation of the lower laser level. However, a smaller subband separation of 25 meV was used in this design to limit the operating bias and hence the electron temperature. The optical emission occurs at a 16 meV (3.8 THz).

A population inversion between the upper and lower laser levels is achieved when the upper phonon depopulation level and the lower laser level move into resonance at a bias of 13 kV/cm as shown in fig. 8.13. The peak gain is 26 cm^{-1}, which represents an improvement over the previous BTC approaches, but is still below the estimated threshold for a Cu–Cu waveguide. In addition, the gain peak covers only a narrow range of electric fields (12.7 to 13.2 kV/cm). Figure 8.14 shows that this resonance corresponds to a peak in current density of 557 $\mathrm{A\,cm}^{-2}$, and fig. 8.15 shows that the relationship between gain and current density is approximately given by $G = gJ - \alpha$, where $g = 225\,\mathrm{cm/kA}$ and $\alpha = -99.3\,\mathrm{cm}^{-1}$. The oscillator strength for the optical transition was calculated as 0.39.

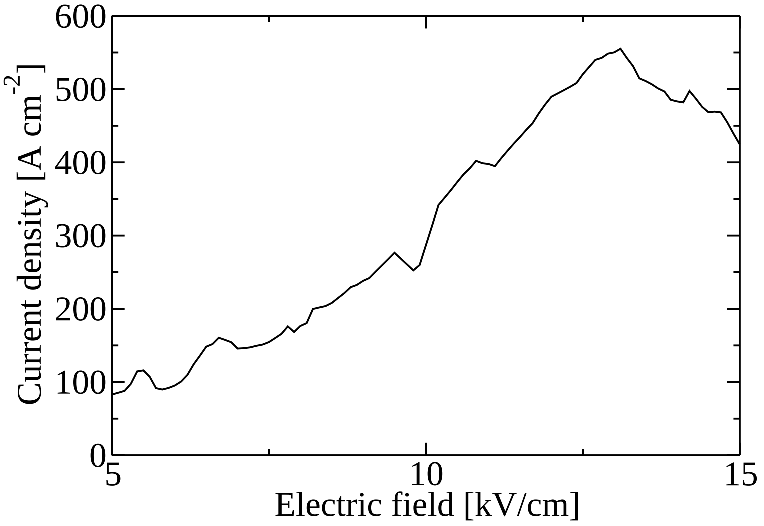


Figure 8.14: Current density as a function of applied electric field for a 3.8 THz (111) Si/SiGe resonant phonon QCL at a lattice temperature of 4 K.

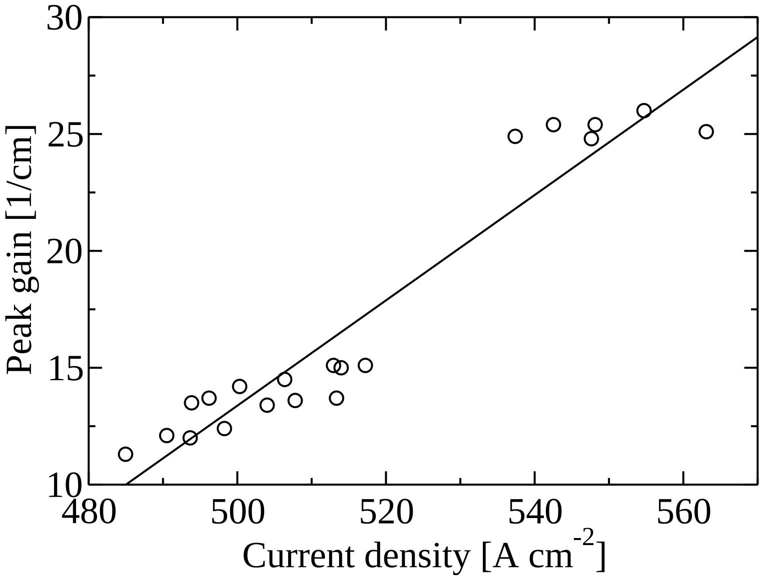


Figure 8.15: Peak gain as a function of current density in a 3.8 THz (111) Si/SiGe resonant phonon QCL at a lattice temperature of 4 K.

8.4 (001) Ge/GeSi QCLs

The results of previous chapters suggest that the (001) Ge/GeSi system is extremely promising for QCL development. The large conduction band offset allows good quantum confinement to be achieved without requiring large variations in Ge fraction across heterointerfaces. The small quantisation effective mass allows wider layers to be used than the other material configurations considered in this chapter, and enhances the oscillator strength of the device.

Despite the possible advantages of Ge/GeSi above the other Si-based material configurations, the last chapter showed that Driscoll's 42 μm BTC QCL design is unlikely to produce a net optical gain. The large emission energy, intraminiband absorption and high electron temperatures were all identified as likely causes. An alternative BTC design was developed in collaboration with Leon Lever at the University of Leeds to overcome these issues.

Figure 8.16 shows the bandstructure and electron probability densities in the new design. The emission energy has been reduced to 16 meV (3.8 THz) and the total width of the miniband to 14 meV, limiting the effect of intraminiband absorption. Nonradiative intervalley scattering between the laser levels is also lower than in Driscoll's design as the emission energy lies well below the $L \rightarrow L$ phonon energy (24 meV). A barrier composition of 85% Ge was selected in order to reduce interface roughness scattering and also because high quality Ge/Ge$_{0.85}$Si$_{0.15}$ multiple quantum well stacks have been reported[213]. A 97% Ge virtual substrate is required for mechanical stability. Dopants were spread only through the four QWs before the injection barrier to reduce the nonradiative ionised impurity scattering between the laser levels.

The applied electric field is reduced to 3.5 kV/cm in the present design, compared with 10.5 kV/cm in Driscoll's design. In the next section, it is shown that this causes a 45 K reduction in electron temperature in the current design, compared with that of Driscoll. Carriers are therefore located closer in energy to the subband minima,

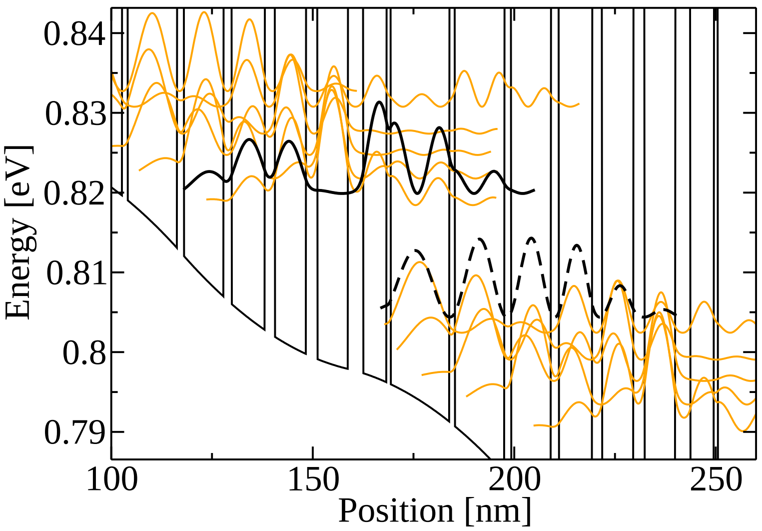


Figure 8.16: Potential profile and electron probability densities for 3.8 THz Ge/GeSi BTC QCL. Layer widths of 5.8/**1.0**/14.6/**1.4**/12.3/**1.6**/9.9/ **1.9**/8.3/**2.4**/ 7.8/**2.9**/7.6/**3.7** were used, where bold text denotes 85% Ge barriers and normal-weighted text denotes Ge wells. Underlined text denotes regions with a doping concentration of $2\times10^{16}\,\mathrm{cm}^{-3}$. Results are shown at an electric field of 3.5 kV/cm.

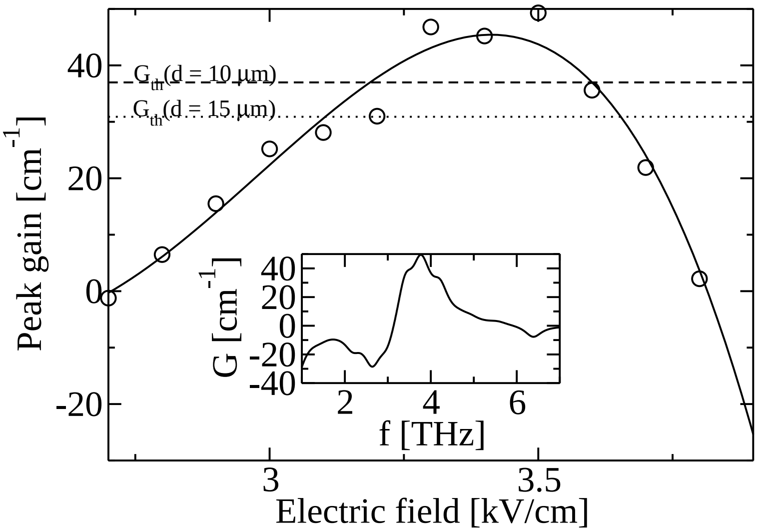


Figure 8.17: Peak gain as a function of applied electric field in 3.8 THz Ge/GeSi BTC QCL design at a lattice temperature of 4 K. The dashed and dotted lines show the estimated threshold gain for 10 μm and 15 μm thick Cu–Cu waveuides respectively. The gain spectrum at a frequency of 3.5 kV/cm is shown inset.

reducing the non-radiative depopulation rate of the upper laser level by intervalley scattering and reducing the thermal backfilling of the LLL.

Figure 8.17 shows that the optimised BTC design yields net gain at a lattice temperature of 4 K. For a 10 μm thick active region in a Cu–Cu waveguide, gain is predicted for applied electric fields between 3.19 and 3.60 kV/cm. By increasing the stack thickness to 15 μm, the dynamic range increases to 3.11–3.66 kV/cm. The peak gain is, however, insufficient to exceed the predicted threshold for a semi-insulating surface plasmon waveguide. Figure 8.18 shows that the peak gain increases as a function of the estimated current density to a maximum of 49.7 cm^{-1} at J=379 A cm^{-2}. For current densities below this peak, a roughly linear relationship exists, with the form $G = 315J - 70$, where G is given in units of cm^{-1} and J is in kA cm^{-2}. Interpolating this linear approximation gives an estimated threshold current density of 339 A cm^{-2} for a 10 μm thick active region. By increasing the active region thickness to 15 μm, the threshold current density is reduced to 319 A cm^{-2}.

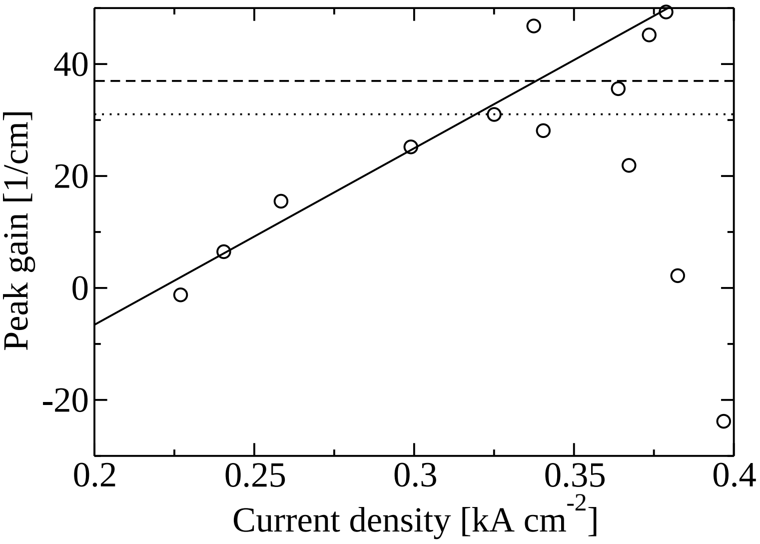


Figure 8.18: Peak gain as a function of current density in 3.8 THz Ge/GeSi BTC QCL design at a lattice temperature of 4 K. The dashed and dotted lines show the estimated threshold gain for 10 μm and 15 μm thick Cu–Cu waveuides respectively. The solid line shows an extrapolated linear regression to the data at currents below 350 A cm^{-2}.

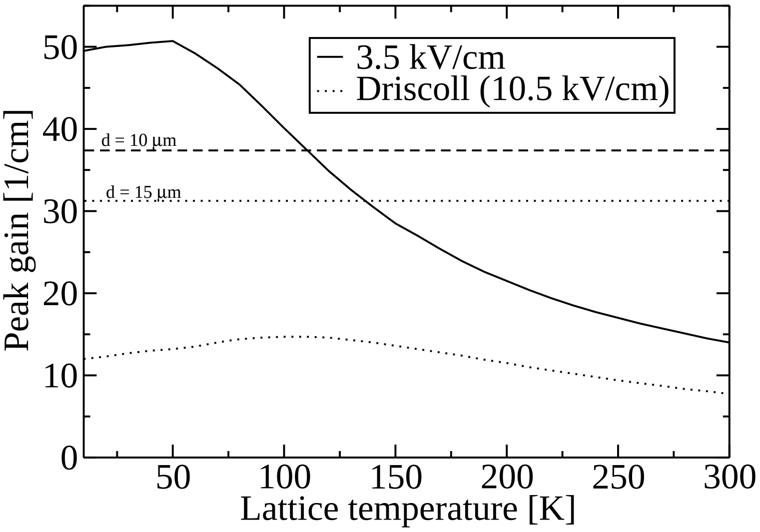


Figure 8.19: Peak gain as a function of lattice temperature for Ge/GeSi BTC QCL design at an electric field of 3.5 kV/cm. The results for Driscoll's 42 μm BTC QCL design are shown at the design field of 10.5 kV/cm for comparison. Estimated gain thresholds are shown for 10 and 15 μm thick active regions in a Cu–Cu waveguide.

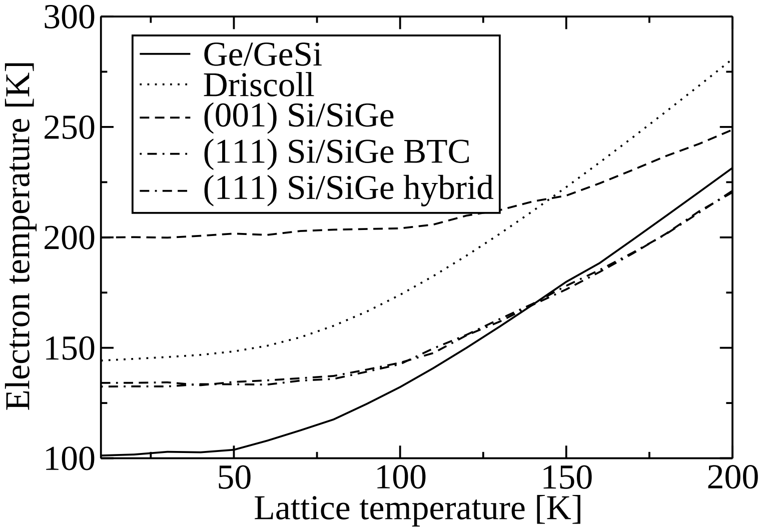


Figure 8.20: Electron temperature as a function of lattice temperature for each of the QCL designs considered in the present work. Each line is shown for the operating bias of the device.

The temperature dependence of the peak gain is shown in fig. 8.19. The gain is predicted to exceed the threshold for a 10 μm thick active region in a Cu–Cu waveguide structure up to a lattice temperature of 110 K. By increasing the active region thickness to 15 μm, the maximum operating temperature is increased to 136 K. By contrast, it can be seen that Drisoll's design does not yield sufficient gain to overcome the threshold even at 15 μm active region thickness.

8.5 Electron temperature

Electron temperatures were calculated self-consistently for each of the devices as described in chapter 7. Figure 8.20 shows that in each case the electron temperature rises with lattice temperature as expected, but varies considerably between the devices. The electron temperature has previously been shown to depend upon the electric field as well as the phonon energy and scattering rates. The (001) Si/SiGe structure has the highest electron temperature despite its relatively low electric field

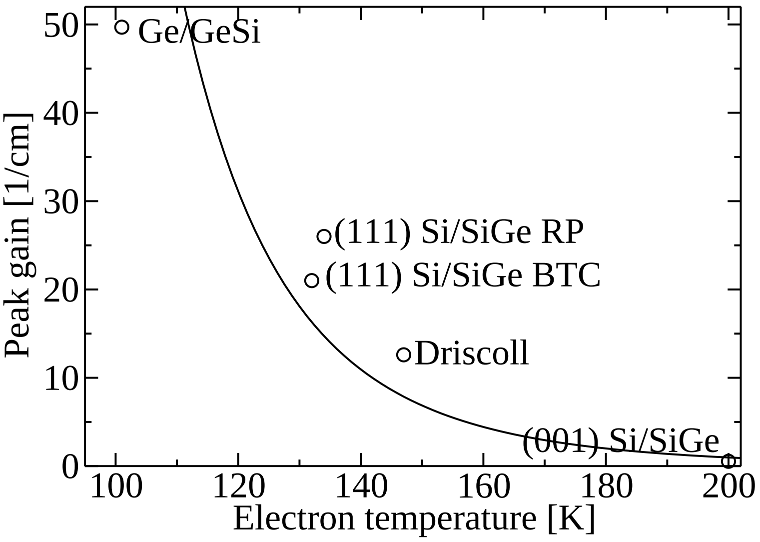


Figure 8.21: Peak gain as a function of electron temperature at $T = 4\,K$ for each of the devices simulated in the present work

of 5.2 kV/cm, due to the large g-LO phonon energy of 63.2 meV, given that the spacing between the subbands is much lower than this. This means that only electrons with very high in-plane kinetic energy can be removed from subbands by inelastic processes, resulting in a high effective carrier temperature.

In the (111) oriented structures, the electron temperature is lower due to the additional f-LA scattering with a phonon energy of 46.3 meV. Both (111) designs show similar electron temperature as the effect of the higher electric field in the hybrid design is compensated by the enhanced phonon emission rate. The Ge/GeSi BTC design presented in this chapter exhibits the lowest electron temperature due to the low intervalley phonon energy of 24 meV. The simulated electron temperature is 70 K lower than that of Driscoll's BTC design due to the reduction in applied electric field and reduced barrier height.

Figure 8.21 gives some insight into the difference between specific device designs by considering the electron temperature at $T = 4\,K$. A power function regression provides a good match to the data, with the form $G_{pk} = 10245\mathrm{e}^{-0.048T_e}$. High electron temperatures are therefore associated with low gain. This is because hot carriers are able to scatter into higher energy subbands and the resulting thermal backfilling of the lower laser level reduces the population inversion. Also, nonradiative phonon scattering competes with the optical transition in hot carrier distributions.

8.6 Conclusion

Four QCLs designed during the course of the present work have been simulated to illustrate the differences between the possible material configurations and design schemes. Net gain was predicted for the Ge/GeSi BTC design using a Cu–Cu waveguide, but none of the other designs were successful. The (001) oriented Si/SiGe device yielded negligible active region gain, while both (111) oriented devices exhibited active region gain approximately 5–10 cm^{-1} below the threshold.

The large variation in gain between designs can be explained by considering the effective masses of the materials and the simple three-level model from chapter 7, in which

$$G_{ij} \propto |D_{ij}|^2 \tau_i. \tag{8.2}$$

The dipole matrix element is proportional to the length of the device, and for a given transition energy, the layer widths are approximately inversely proportional to the quantisation effective mass. All the scattering rates considered in this work (chapter 5) are proportional to the density of states effective mass, and the lifetime of the upper laser level is therefore inversely proportional to m_d giving

$$G_{ij} \propto \frac{1}{m_d m_q^2}. \tag{8.3}$$

The figure of merit, $\mu_g = m_e^3/m_d m_q^2$ therefore gives an approximate measure of the influence of the effective mass upon the optical gain. Substituting the effective mass values from chapter 3 gives $\mu_g = 6.27$ for the Δ_2 valleys in (001) Si/SiGe, $\mu_g = 41.1$ for the Δ valleys in (111) Si/SiGe and $\mu_g = 231$ for the L valleys in (001) Ge/GeSi. This suggests that the maximum gain achievable with a (001) Ge/GeSi

Design	G_{pk} [1/cm]	T_e [K]	F [kV/cm]
(001) Si/SiGe BTC	0.54	200	5.2
(111) Si/SiGe BTC	21.0	132	7.1
(111) Si/SiGe hybrid	26.0	134	13.0
(001) Ge/GeSi BTC	49.7	101	3.5
(001) Ge/GeSi BTC (Driscoll)	12.6	147	10.5

Table 8.1: Device characteristics for each of the QCL designs simulated in this work. The simulated peak active region gain and electron temperature are given at the stated electric field, and at a lattice temperature of 4 K.

QCL should be larger than that of a (111) Si/SiGe device, which in turn is larger than that of a (001) Si/SiGe device. This is in qualitative agreement with the calculated peak gain presented in table 8.1. It was found that the gain is highest in structures with low electron temperature due to the reduced thermal backfilling of the lower laser level. (001) Si/SiGe devices are likely to have high electron temperatures due to the high g-LO phonon energy, while (111) Si/SiGe and (001) Ge/GeSi structures have lower phonon energies and hence lower electron temperature.

A substantial improvement in gain was achieved in a Ge/GeSi BTC design, compared with that of Driscoll, by performing a series of optimisations. Ionised impurity scattering between the laser levels was reduced by moving donor ions away from the active region and intervalley phonon scattering was reduced by selecting an emission energy below the $L \rightarrow L$ phonon energy. A 45 K reduction in electron temperature was achieved by reducing the bias from 10.5 kV/cm to 3.5 kV/cm. Figure 8.22 shows that a current density of 250–350 A cm^{-2} was calculated for all of the seven-well bound–to–continuum devices. The hybrid RP/BTC (111) Si/SiGe device has the largest current density due to the large electric field and rapid intervalley phonon scattering, while the (001) Si/SiGe BTC device has the lowest current density due to inefficient charge injection.

To summarise, the 3.8 THz Ge/GeSi BTC device has been shown to be the most

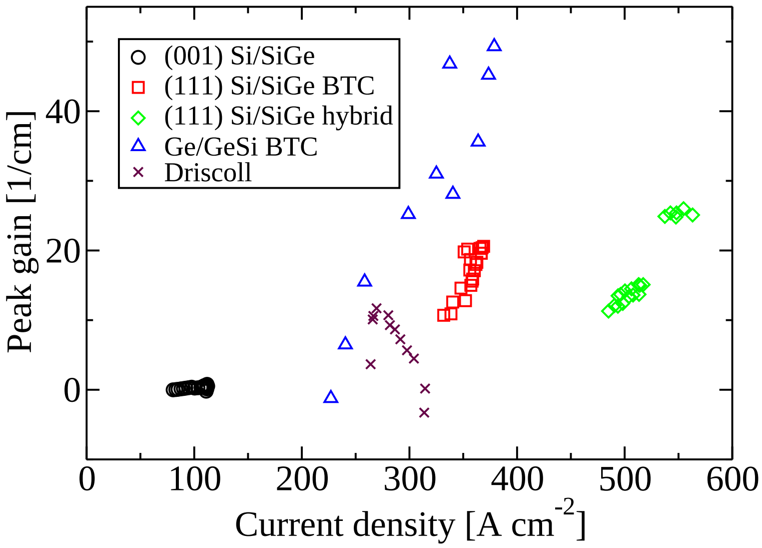


Figure 8.22: Comparison between gain–current relation for all QCL designs considered in the present work. Results are plotted at 4 K.

promising candidate for QCL development from the set of devices considered in the present work. The low effective mass and low electron temperature lead to high gain and a predicted maximum operating temperature of 136 K with a 15 μm thick Cu-Cu waveguide. By contrast, the (001) Si/SiGe structure was shown to exhibit negligible gain. The (111) oriented Si/SiGe structures were both predicted to exhibit gain just below the threshold for a 15 μm thick Cu-Cu waveguide. Although further optimisation of the (111) designs may possibly lead to a net gain, the discussion in chapter 6 indicates that high-quality epitaxial growth of (111) structures is not yet possible. (001) Ge/GeSi devices are also more attractive due to their potential compatibility with mainstream (001) oriented CMOS technology.

Chapter 9

Conclusion

A range of Si-based material configurations have been investigated for their suitability for THz QCL active regions, in particular, the ability to confine electrons effectively, to achieve a population inversion, to overcome waveguide losses and to be practically realisable. Comparisons between individual material properties have been made at the end of the relevant chapters, and this conclusion provides a cumulative summary of the findings. Proposals for further work are presented at the end of this chapter.

The conduction band offset in a QCL should be approximately twice the emission energy to allow for the emission of a photon and depopulation of the lower laser level within a period of the device. In chapter 2, strain and crystal orientation were shown to have an important effect upon the conduction band offset in Si/Ge/Si and Ge/Si/Ge QWs. The energy separation between the conduction band edge and the higher energy valleys must also be maximised to avoid intervalley scattering. The valley separation was found to be very small in (001) Ge/Si/Ge and (111) Si/Ge/Si QWs with realistic substrate compositions, limiting the maximum emission energy to around 30 meV (7.3 THz) in each system. Much better results were obtained for (111) Ge/Si/Ge and (001) Si/Ge/Si QWs. In the former, the conduction band offset provides the dominant constraint, giving a maximum emission energy around 117 meV (29 THz). In the latter system, valley splitting provides the dominant constraint, resulting in a maximum emission energy around 75 meV (19 THz).

A small quantisation effective mass is desirable at the conduction band edge in QCLs, as this allows wider heterolayers to be used, simplifying the crystal growth. A longer length scale may also give a larger dipole matrix element between the laser levels, and hence larger optical gain. In chapter 3, the quantisation effective mass at the conduction band edge is shown to be relatively low in (001) GeSi/Ge/GeSi ($m_q = 0.12m_e$) and (111) SiGe/Si/SiGe ($m_q = 0.26m_e$) QWs, but very high in (001) SiGe/Si/SiGe ($m_q = 0.92m_e$) and (111) GeSi/Ge/GeSi ($m_q = 1.64m_e$)[1]. The effect of various doping profiles upon the bandstructure was also investigated in chapter 3. It was shown that δ-doping results in much smaller changes to the bandstructure than wider doped regions.

The effect of valley splitting was considered in chapter 4. This occurs when two conduction band valleys have the same in-plane wavevector. (001) SiGe/Si/SiGe and (001) GeSi/Ge/GeSi are therefore affected, while (111) oriented structures are not. Interference between wavefunction components (from both valleys) and their reflections from interfaces leads to the formation of doublets of bound states. The single-valley Schrödinger equation solutions are effectively split into pairs, separated by a small energy (∼1–5 meV), which depends upon the width of the QW. The valley splitting was found to be smallest in wide structures with non-abrupt interfaces, meaning that spectral line doublets in QCLs may be too close in energy to be resolved. Nonetheless, it is likely to cause linewidth broadening of the emission spectrum.

In order to achieve a population inversion, rapid scattering is required to depopulate the lower laser level, while slow scattering is required within the active region to maximise the upper laser level lifetime. Expressions for electron scattering rates were derived in chapter 5 for interface roughness, alloy disorder, ionised impurity and phonon deformation potential interactions. It was shown that interface roughness scattering is usually the dominant process in Si-based QCLs, and is fastest in

---

[1] *c.f.* $m_q = 0.067m_e$ in GaAs

structures with narrow wells and high, thin barriers. Phonon scattering becomes significant when subbands are separated by an energy similar to or greater than the phonon energy. In (001) Si/SiGe THz QCLs, the phonon scattering rate is usually very low due to the high energy of the g-LO phonon (63 meV) compared with the subband separations in the device. The rate is much higher in (111) Si/SiGe and (001) Ge/GeSi systems due to the lower energy of the f-LA (46 meV) and $L \rightarrow L$ (28 meV) phonon interactions, respectively. Ionised impurity scattering was shown to be significant only in regions with widely-distributed doping.

The review of crystal growth technology in chapter 6 shows that (001) Si/SiGe and Ge/GeSi multiple-QW structures and virtual substrates have been realised, but very limited progress has been made with (111) oriented heterostructures. Interdiffusion of materials across interfaces (due to annealing or surface segregation) affects the subband separation and carrier lifetimes significantly. The effects were found to be most important in (001) Si/SiGe heterostructures due to the requirement for narrow layers. By contrast, (111) Si/SiGe and (001) Ge/GeSi heterostructures were predicted to tolerate interdiffusion lengths up to 2.3–5 nm without exhibiting significant changes in carrier dynamics or subband separation.

Electron temperature has an important effect upon the scattering rates in a QCL and in chapter 7, it was shown to depend upon the lattice temperature and electric field. A Ge/GeSi QCL design from the literature was simulated and predicted to yield insufficient active region gain to overcome the estimated losses in a double metal waveguide. This was due to depopulation of the upper laser level by phonon scattering, optical absorption within the wide miniband and thermal backfilling of the lower laser level.

Chapter 8 presented a range of QCL designs with simulated performance data. The optical gain of QCLs was shown to decrease as the quantisation effective mass increases (due to the change in dipole matrix element), and as the density-of-states effective mass increases (due to the decrease in ULL lifetime). As a result, the (001) Ge/GeSi and (111) Si/SiGe QCL designs were found to have much greater

	(001) Si/SiGe	(111) Si/SiGe	(001) Ge/GeSi
Large usable CB offset	✗	✓	✓
Low quantisation mass	✗	(✓)	✓
No valley splitting	✗	✓	✗
Epitaxial growth realised	✓	✗	✓
Interdiffusion tolerance	✗	✓	✓
Low electron temperature	✗	✓	✓

Table 9.1: Simplified criteria for a good THz QCL material. A ✓symbol means that a material configuration satisfies a criterion, a ✓in parentheses implies that the material configuration only just meets the criterion, while a ✗ symbol indicates failure to meet the criterion.

simulated peak gain than the (001) Si/SiGe design. The gain was also found to decay exponentially as the electron temperature increases. The highest simulated electron temperature occurs in the (001) Si/SiGe QCL design due to the large LO-phonon energy, while the lowest occurs in the Ge/GeSi device presented in this work, due to the low intervalley phonon energy and low electric field. The new Ge/GeSi device design was predicted to yield sufficient active region gain to overcome the estimated threshold for a Cu–Cu waveguide up to an operating temperature of 136 K.

Table 9.1 contains a simplified summary of desirable material properties for THz QCLs, and a comparison between the material configurations investigated in this work. (001) Si/SiGe fails all but one of the criteria in the list. Although large conduction band offsets exist, the small energy separation between conduction band valleys places a practical limit on the usable well depth. The large quantisation effective mass reduces the gain and results in the requirement for thin layers, leading to a poor tolerance to interdiffusion. The large g-LO phonon energy leads to high electron temperatures, further limiting the gain. Theoretically, (111) Si/SiGe appears to be far superior due to the lower effective mass and large separation between conduction band valley sets. Additionally, each of the six Δ conduction band valleys lies at a different in-plane wavevector, which eliminates valley splitting and potentially reduces the spectral linewidth. However, (111) oriented QCLs are far from realisable with existing crystal growth technologies. The best results are obtained for (001) Ge/GeSi, where the quantisation effective mass is by far the smallest. Relatively thick layer widths may be used, meaning that a lower electric field is possible (and hence less carrier heating) than in other material configurations. The small intervalley phonon energy allows effective cooling of carrier distributions, and hence a lower electron temperature and higher gain. Although valley splitting exists, the effect upon the peak gain is predicted to be quite small and may only be manifested as a slight spectral linewidth broadening.

In summary, a detailed study of n-type Si-based materials has shown that (001) oriented Ge/GeSi heterostructures offers the best opportunity to date for developing a Si-based THz QCL. The present theoretical work, combined with recent advances in crystal growth give rise to the possibility of experimental development of the first working device within the next few years.

9.1 Further work

A number of new avenues for research into the theory of Si-based QCLs have been identified. The most important of these, in the opinion of the author, are presented in this section. Additionally, a number of important areas for experimental investigation are proposed.

As described in chapter 5, coherent effects were assumed to be negligible in the present work. Although many useful conclusions about the charge transport characteristics have been reached, the fully incoherent model has yielded some unrealistic effects. The main issues arise from the apparent ability of electrons to tunnel instantaneously through barriers, giving rise to spatially extended wavefunctions and result in unrealistic parasitic current channels, as described in chapter 7. Previous investigations into coherent charge transport in III–V QCLs have also shown that fully incoherent models underestimate the effect of the injection barrier thickness as a charge transport bottleneck[148]. Although a full density matrix model of transport in the system would be too computationally expensive for use as a QCL design tool, investigations in III–V materials systems have obtained useful results by considering coherent transport between the injector and upper laser subbands on either side of the QCL injection barrier[53].

Further enhancements to the existing model could be implemented by considering the following effects. Intervalley mixing effects were considered only for (001) Si/SiGe QWs in this work. However, similar intervalley mixing effects are known to occur in (001) Ge/GeSi heterostructures[5, 8, 12]. The pseudopotential model used in chapter 4 could also be used to investigate valley splitting phenomena in Ge/GeSi QWs with widths appropriate for QCL designs. An estimate for the splitting between optical transitions could be incorporated in the gain spectrum as either a linewidth broadening or a spectral line doublet depending on its magnitude. Secondly, the assumption has been made that only one degenerate valley set is populated in Si-based QCLs. However, recent theoretical work suggests that intervalley phonon scattering may be an important process in GaAs/AlGaAs at high temperatures[214]. These effects can readily be incorporated into the rate equation model described in this work and used to predict leakage currents in Si-based QCLs at high temperatures.

The preliminary investigation of waveguide losses and modal overlap by Craig Evans (chapter 7) was conducted for (001) oriented structures with Ge fractions below 85%. Having now established that the best prospects for Si-based QCL development lie with (001) Ge/GeSi, it would be prudent to extend Evans' work to consider such structures. A more thorough investigation of waveguide losses would allow a minimum active region thickess to be determined for a given device. Additionally, according to the bulk Drude model, the free-carrier losses calculated for emission at a frequency of 5 THz should increase with the emission wavelength.

Having demonstrated the advantages of Ge/GeSi heterostructures above other

Si-based material configurations in chapter 8, a more thorough investigation of the QCL design space may yield higher gain and operating temperature. The non-polar nature of Si-based materials removes the "forbidden" Reststrahlen band of emission energies seen in III–V materials. Si-based QCLs could therefore be designed to emit THz radiation with frequencies above 5 THz. Additionally, superior device performance can be expected in this frequency range, due to the reduced free-carrier losses.

Recent theoretical and experimental work has shown that the active region temperature in III–V QCLs may be 20–100 K higher than the heatsink temperature for low duty-cycle pulsed mode operation[215, 216]. The geometry of the waveguide and the device mountings were found to have a significant effect upon active region temperature[217]. A potential advantage of Si-based structures is the large thermal conductivity ($1.3\,\mathrm{W\,cm^{-1}\,K^{-1}}$ in Si *c.f.* $0.55\,\mathrm{W\,cm^{-1}\,K^{-1}}$ in GaAs and $0.68\,\mathrm{W\,cm^{-1}\,K^{-1}}$ in InP[60]), which should allow heat to be efficiently dissipated away from the active region of the device, possibly allowing higher temperature operation. Modelling the thermal properties of Si-based QCLs would allow more accurate predictions of the maximum operating temperature, and improved design of waveguides and device mountings.

A theoretical model is only as good as its input parameters. Pump-probe spectroscopy has been used to determine the interface roughness scattering parameters in Si/SiGe QWs[7, 39]. However, the quality of material growth may differ somewhat in Ge/GeSi structures. Transmission electron microscopy can be used to measure the interface roughness height[218] at interfaces in Ge/GeSi double QWs and pump-probe spectroscopy can be used to find an effective correlation length. By investigating a number of different samples with varying Si content in the QWs, an effective alloy disorder potential may also be extracted. Measurements of the interdiffusion length as a function of time in samples annealed at various temperatures would allow a temperature dependent diffusion coefficient to be found. Carrier lifetime measurements in these structures would provide valuable data for experimental verification of the predictions in chapter 6.

The growth, fabrication and characterisation of a prototype Ge/GeSi QCL should now be carried out. The design proposed in chapter 8 makes use of a virtual substrate and heterostructure with alloy fractions and layer widths that have already been realised with existing CVD growth technology[213]. X-ray diffraction and transmission electron microscopy data may be used to analyse the accuracy and quality of crystal growth. Buried silicide waveguides have already been demonstrated for Si-based QCLs[219], but it would be highly advantageous to develop double-metal waveguides due to the large modal overlap and relatively low losses.

References

[1] A. Valavanis, Z. Ikonić, and R. W. Kelsall, "Growth variation effects in SiGe-based quantum cascade lasers," *J. Opt. A*, vol. 11, no. 5, p. 054012, 2009.

[2] A. Valavanis, Z. Ikonić, L. Lever, C. Evans, and R. Kelsall, "Substrate orientation and alloy composition effects in n-type SiGe quantum cascade structures," in *5^{th} IEEE Int. Conf. on group IV Photon.*, pp. 251–252, 2008.

[3] A. Valavanis, L. Lever, C. A. Evans, Z. Ikonić, and R. W. Kelsall, "Theory and design of quantum cascade lasers in (111) n-type Si/SiGe," *Phys. Rev. B*, vol. 78, no. 3, p. 035420, 2008.

[4] L. Lever, A. Valavanis, Z. Ikonić, and R. W. Kelsall, "Simulated [111] Si-SiGe terahertz quantum cascade laser," *Appl. Phys. Lett.*, vol. 92, no. 2, p. 021124, 2008.

[5] M. Virgilio and G. Grosso, "Valley splitting and optical intersubband transitions at parallel and normal incidence in [001]-Ge/SiGe quantum wells," *Phys. Rev. B*, vol. 79, no. 16, p. 165310, 2009.

[6] S. Pavlov *et al.*, "Stimulated terahertz emission due to electronic Raman scattering in silicon," *Appl. Phys. Lett.*, vol. 94, p. 171112, 2009.

[7] A. Valavanis, Z. Ikonić, and R. W. Kelsall, "Intersubband carrier scattering in n- and p-Si/SiGe quantum wells with diffuse interfaces," *Phys. Rev. B*, vol. 77, p. 075312, Feb. 2008.

[8] G. Ciasca *et al.*, "Terahertz intersubband absorption and conduction band alignment in n-type Si/SiGe multiple quantum wells," *Phys. Rev. B*, vol. 79, no. 8, p. 085302, 2009.

[9] A. Valavanis, Z. Ikonić, and R. W. Kelsall, "Intervalley splitting and intersubband transitions in n-type Si/SiGe quantum wells: Pseudopotential vs. effective mass calculation," *Phys. Rev. B*, vol. 75, no. 20, p. 205332, 2007.

[10] T. B. Boykin, N. Kharche, and G. Klimeck, "Valley splitting in finite barrier quantum wells," *Phys. Rev. B*, vol. 77, no. 24, p. 245320, 2008.

[11] M. O. Nestoklon, E. L. Ivchenko, J.-M. Jancu, and P. Voisin, "Electric field effect on electron spin splitting in SiGe/Si quantum wells," *Phys. Rev. B*, vol. 77, no. 15, p. 155328, 2008.

[12] M. Virgilio and G. Grosso, "Quantum-confined stark effect in Ge/SiGe quantum wells: A tight-binding description," *Phys. Rev. B*, vol. 77, no. 16, p. 165315, 2008.

[13] A. Valavanis, L. Lever, C. Evans, Z. Ikonić, and R. Kelsall, "Low electric field silicon-based THz quantum cascade laser employing L-valley intersubband transitions," in *UK Semicond.*, Inst. of phys., 2009.

[14] A. Valavanis, Z. Ikonić, L. Lever, C. Evans, and R. Kelsall, "Substrate orientation and alloy composition effects in n-type SiGe quantum cascade structures," in *The 5th Annual Conf. on Group IV Photon.*, IEEE LEOS, 2008.

[15] A. Valavanis, Z. Ikonić, and R. Kelsall, "Intersubband transitions in n-type group IV quantum cascade lasers," in *Int. quantum cascade lasers school and workshop*, (Monte Verita, Switzerland), 2008.

[16] A. Valavanis, L. Lever, C. Evans, Z. Ikonić, and R. Kelsall, "Theory and design of (111) oriented Si/SiGe quantum cascade lasers," in *QEP-18*, Inst. of phys., 2008.

[17] A. Valavanis, L. Lever, Z. Ikonić, and R. Kelsall, "Growth variation effects in Si/SiGe quantum cascade lasers," in *QEP-18*, Inst. of phys., 2008.

[18] A. Valavanis, L. Lever, C. Evans, Z. Ikonić, and R. Kelsall, "Crystal orientation and waveguide geometry effects in n-type Si/SiGe quantum cascade lasers," in *UK Semicond.*, Inst. of phys., 2008.

[19] A. Valavanis, L. Lever, Z. Ikonić, and R. Kelsall, "n-type Si/SiGe quantum cascade structures," in *The 9th Int. Conf. on Intersubband Transitions in Quantum Wells*, 2007.

[20] A. Valavanis, Z. Ikonić, and R. Kelsall, "Intervalley mixing in two-dimensional n-type Si/SiGe heterostructures," in *The 9th Int. Conf. on Intersubband Transitions in Quantum Wells*, 2007.

[21] A. Valavanis, Z. Ikonić, and R. Kelsall, "n-type Si/SiGe quantum cascade lasers," in *Mini-Symposium on Si-based optoelectronics*, The Rank Prize Funds, 2006.

[22] M. Tacke, "New developments and applications of tunable IR lead salt lasers," *Infrared Phys. Technol.*, vol. 36, no. 1, pp. 447–463, 1995. Proceedings of the Sixth International Conference on Infrared Physics.

[23] F. Capasso *et al.*, "New frontiers in quantum cascade lasers and applications," *IEEE J. Sel. Top. Quantum Electron.*, vol. 6, pp. 931–947, Nov. 2000.

[24] R. Kazarinov and R. Suris, "Amplification of electromagnetic waves in a semiconductor superlattice," *Sov. Phys., Semicond.*, vol. 5, pp. 707–709, 1971.

[25] C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho, "Recent progress in quantum cascade lasers and applications," *Rep. Prog. Phys.*, vol. 64, pp. 1533–1601, 2001.

[26] J. Faist *et al.*, "Quantum cascade laser," *Science*, vol. 264, pp. 553–556, Apr. 1994.

[27] J. Faist *et al.*, "Vertical transition quantum cascade laser with Bragg confined excited state," *Appl. Phys. Lett.*, vol. 66, no. 5, pp. 538–540, 1995.

[28] J. Faist *et al.*, "High power mid-infrared ($\lambda \sim 5\,\mu m$) quantum cascade lasers operating above room temperature," *Appl. Phys. Lett.*, vol. 68, pp. 3680–3682, June 1996.

[29] G. Scamarcio *et al.*, "High-Power Infrared (8-Micrometer Wavelength) Superlattice Lasers," *Science*, vol. 276, no. 5313, pp. 773–776, 1997.

[30] C. Sirtori *et al.*, "GaAs/$\mathrm{Al}_x\mathrm{Ga}_{1-x}$As quantum cascade lasers," *Appl. Phys. Lett.*, vol. 73, no. 24, pp. 3486–3488, 1998.

[31] J. Ulrich, J. Kreuter, W. Schrenck, G. Strasser, and K. Unterrainer, "Long wavelength (15 and 23 μm) GaAs/AlGaAs quantum cascade lasers," *Appl. Phys. Lett.*, vol. 80, pp. 3691–3693, May 2002.

[32] M. P. Semtsiv, M. Wienold, S. Dressler, and W. T. Masselink, "Short-wavelength ($\lambda \approx 3.05\,\mu$m) InP-based strain-compensated quantum-cascade laser," *Appl. Phys. Lett.*, vol. 90, p. 051111.

[33] L. Diehl *et al.*, "High-power quantum cascade lasers grown by low-pressure metal organic vapor-phase epitaxy operating in continuous wave above 400 K," *Appl. Phys. Lett.*, vol. 88, no. 20, p. 201115, 2006.

[34] H. Page, C. Becker, A. Robertson, G. Glastre, V. Ortiz, and C. Sirtori, "300 k operation of a GaAs-based quantum-cascade laser at $\lambda \approx 9\,\mu$m," *Appl. Phys. Lett.*, vol. 78, no. 22, pp. 3529–3531, 2001.

[35] J. Devenson, R. Teissier, O. Cathabard, and A. N. Baranov, "Inas/alsb quantum cascade lasers emitting below 3 mu m," *Appl. Phys. Lett.*, vol. 90, no. 11, p. 111118, 2007.

[36] B. S. Williams, "Terahertz quantum-cascade lasers," *Nat. Photonics*, vol. 1, pp. 517–525, Sept. 2007.

[37] W. S. Holland *et al.*, "Submillimetre images of dusty debris around nearby stars," *Nature*, vol. 392, pp. 788–791, Apr. 1998.

[38] D. Grischkowsky, S. Keiding, M. Exter, and C. Fattinger, "Far-infrared time-domain spectroscopy with terahertz beams of dielectrics and semiconductors," *J. Opt. Soc. Am. B*, vol. 7, no. 10, pp. 2006–2015, 1990.

[39] M. Califano *et al.*, "Interwell relaxation times in p-Si/SiGe asymmetric quantum well structures: Role of interface roughness," *Phys. Rev. B*, vol. 75, p. 045338, Jan. 2007.

[40] T. Löffler *et al.*, "Terahertz dark-field imaging of biomedical tissue," *Opt. Express*, vol. 9, no. 12, pp. 616–621, 2001.

[41] D. Graham-Rowe, "Terahertz takes to the stage," *Nat. Photonics*, vol. 1, no. 2, pp. 75–77, 2007.

[42] D. Woolard, R. Brown, M. Pepper, and M. Kemp, "Terahertz frequency sensing and imaging: A time of reckoning future applications?," *Proc. IEEE*, vol. 93, pp. 1722–1743, Oct. 2005.

[43] A. G. Davies, E. H. Linfield, and M. B. Johnston, "The development of terahertz sources and their applications," *Phys. Med. Biol.*, vol. 47, pp. 3679–3689, 2002.

[44] G. Mourou, C. V. Stancampiano, A. Antonetti, and A. Orszag, "Picosecond microwave pulses generated with a subpicosecond laser-driven semiconductor switch," *Appl. Phys. Lett.*, vol. 39, no. 4, pp. 295–296, 1981.

[45] A. Bonvalet, M. Joffre, J. L. Martin, and A. Migus, "Generation of ultrabroadband femtosecond pulses in the mid-infrared by optical rectification of 15 fs light pulses at 100 MHz repetition rate," *Appl. Phys. Lett.*, vol. 67, no. 20, pp. 2907–2909, 1995.

[46] B. Ferguson and X.-C. Zhang, “Materials for terahertz science and technology,” *Nat. Mater.*, vol. 1, pp. 26–33, Sept. 2002.

[47] P. Siegel, “Terahertz technology,” *IEEE Trans. Microwave Theory and Techniques*, vol. 50, pp. 910–928, Mar. 2002.

[48] G. P. Williams, “Far-IR/THz radiation from the Jefferson Laboratory, energy recovered linac, free electron laser,” vol. 73, pp. 1461–1463, AIP, 2002.

[49] R. Köhler *et al.*, “Terahertz semiconductor-heterostructure laser,” *Nature*, vol. 417, pp. 156–159, Feb. 2002.

[50] S. Kumar, Q. Hu, and J. L. Reno, “186 K operation of terahertz quantum-cascade lasers based on a diagonal design,” *Appl. Phys. Lett.*, vol. 94, no. 13, p. 131105, 2009.

[51] B. Williams, S. Kumar, Q. Hu, and J. Reno, “High-power terahertz quantum-cascade lasers,” *Electronics Letters*, vol. 42, pp. 89–91, Jan. 2006.

[52] C. Walther, G. Scalari, J. Faist, H. Beere, and D. Ritchie, “Low frequency terahertz quantum cascade laser operating from 1.6 to 1.8 THz,” *Appl. Phys. Lett.*, vol. 89, no. 23, p. 231121, 2006.

[53] G. Scalari *et al.*, “THz and sub-THz quantum cascade lasers,” *Laser and Photon. Rev.*, vol. 3, no. 1–2, pp. 45–66, 2009.

[54] A. W. M. Lee *et al.*, “Real-time terahertz imaging over a standoff distance ($>$ 25 meters),” *Appl. Phys. Lett.*, vol. 89, no. 14, p. 141125, 2006.

[55] L. Ajili, G. Scalari, N. Hoyler, M. Giovannini, and J. Faist, “InGaAs–AlInAs/InP terahertz quantum cascade laser,” *Appl. Phys. Lett.*, vol. 87, no. 14, p. 141107, 2005.

[56] D. J. Paul, “Silicon-germanium strained layer materials in microelectronics,” *Adv. Mater.*, vol. 11, no. 3, pp. 191–204, 1999.

[57] L. Pavesi, "Will silicon be the photonic material of the third millenium?," *J. Phys.-Condens. Mat.*, vol. 15, pp. R1169–R1196, June 2003.

[58] M. A. Foster *et al.*, "Broad-band optical parametric gain on a silicon photonic chip," *Nature*, vol. 441, pp. 960–963, June 2006.

[59] R. Baets *et al.*, "Integration of photonic functions in and with silicon," in *ESSDERC 2004: Proc. 34th Eur. Solid-State Device Research Conf.*, pp. 57–62, Sept. 2004.

[60] "`http://www.ioffe.ru/SVA/NSM/Semicond/index.html`."

[61] M. A. Green *et al.*, "Efficient silicon light-emitting diodes," *Nature*, vol. 412, pp. 805–808, Aug. 2001.

[62] S. Coffa, G. Franzò, F. Priolo, A. Pacelli, and A. Lacaita, "Direct evidence of impact excitation and spatial profiling of excited Er in light emitting Si diodes," *Appl. Phys. Lett.*, vol. 73, pp. 93–95, July 1998.

[63] F. Iacona *et al.*, "Electroluminescence at 1.54 μm in Er-doped Si nanocluster-based devices," *Appl. Phys. Lett.*, vol. 81, pp. 3242–3244, Oct. 2002.

[64] H.-S. Han, S.-Y. Seo, J. H. Shin, and N. Park, "Coefficient determination related to optical gain in erbium-doped silicon-rich silicon oxide waveguide amplifier," *Appl. Phys. Lett.*, vol. 81, pp. 3720–3722, Nov. 2002.

[65] S. Lazarouk, "Reverse biased porous silicon light emitting diodes," in *Conf. of the NATO Adv. Research Workshop on Towards the First Silicon Laser* [220], pp. 61–68.

[66] B. Garrido *et al.*, "Optical and electrical characteristics of LEDs fabricated from Si-nanocrystals embedded in SiO_2," in *Conf. of the NATO Adv. Research Workshop on Towards the First Silicon Laser* [220], pp. 45–54.

[67] L. Pavesi *et al.*, “Optical gain in silicon nanocrystals,” *Nature*, vol. 408, pp. 440–444, Nov. 2000.

[68] W. L. Ng *et al.*, “An efficient room-temperature silicon-based light-emitting diode,” *Nature*, vol. 410, pp. 192–194, Mar. 2001.

[69] M. A. Lourenço *et al.*, “Dislocation-based silicon light emitting devices,” in *Conf. of the NATO Adv. Research Workshop on Towards the First Silicon Laser* [220], pp. 11–20.

[70] O. Boyraz and B. Jalali, “Demonstration of a silicon Raman laser,” *Opt. Express*, vol. 12, pp. 5269–5273, Oct. 2004.

[71] H. Rong *et al.*, “A continuous-wave Raman silicon laser,” *Nature*, vol. 433, pp. 725–728, Feb. 2005.

[72] H. Rong *et al.*, “Monolithic integrated Raman silicon laser,” *Opt. Express*, vol. 14, pp. 6705–6712, July 2006.

[73] G. Dehlinger *et al.*, “Intersubband electroluminescence from silicon-based quantum cascade structures,” *Science*, vol. 290, pp. 2277–2280, Dec. 2000.

[74] S. A. Lynch *et al.*, “Intersubband electroluminescence from Si/SiGe cascade emitters at terahertz frequencies,” *Appl. Phys. Lett.*, vol. 81, no. 9, pp. 1543–1545, 2002.

[75] C. R. Pidgeon *et al.*, “THz intersubband dynamics in p-Si/SiGe quantum well emitter structures,” *Phys. Status Solidi B*, vol. 237, no. 1, pp. 381–385, 2003.

[76] S. Tsujino *et al.*, “Interface-roughness-induced broadening of intersubband electroluminescence in p-SiGe and n-GaInAs/AlInAs quantum-cascade structures,” *Appl. Phys. Lett.*, vol. 86, pp. 062113–1–3, Feb. 2005.

[77] K. Driscoll and R. Paiella, “Silicon-based injection lasers using electronic intersubband transitions in the L valleys,” *Appl. Phys. Lett.*, vol. 89, no. 19, p. 191110, 2006.

[78] K. Driscoll and R. Paiella, "Design of n-type silicon-based quantum cascade lasers for terahertz light emission," *J. Appl. Phys.*, vol. 102, no. 9, p. 093103, 2007.

[79] G. Sun, H. H. Cheng, J. Menéndez, J. B. Khurgin, and R. A. Soref, "Strain-free Ge/GeSiSn quantum cascade lasers based on L-valley intersubband transitions," *Appl. Phys. Lett.*, vol. 90, no. 25, p. 251105, 2007.

[80] G. Han and J. Yu, "An n-type SiGe/Ge QC structure utilizing the deep Ge quantum well for electron at the Γ point," *Semicond. Sci. Tech.*, vol. 22, pp. 769–773, June 2007.

[81] Adapted from image released under the GNU Free Document License, "File:DiamantEbene11.png" at Wikimedia Commons.

[82] C. G. Van de Walle, "Band lineups and deformation potentials in the model-solid theory," *Phys. Rev. B*, vol. 39, pp. 1871–1883, Jan. 1989.

[83] J. P. Dismukes, L. Ekstrom, and R. J. Paff, "Lattice parameter and density in germanium-silicon alloys," *J. Phys. Chem.*, vol. 68, pp. 3021–3027, Oct. 1964.

[84] S. Zwerdling, K. J. Button, B. Lax, and L. M. Roth, "Internal impurity levels in semiconductors: Experiments in p-type silicon," *Phys. Rev. Lett.*, vol. 4, pp. 173–176, Feb. 1960.

[85] E. O. Kane, "Energy band structure in p-type germanium and silicon," *J. Phys. Chem. Solids*, vol. 1, pp. 82–99, Apr. 1956.

[86] H. J. McSkimin and P. Andreatch Jr., "Measurement of third-order moduli of silicon and germanium," *J. Appl. Phys.*, vol. 35, pp. 3312–3319, Nov. 1964.

[87] C. G. Van de Walle and R. M. Martin, "Theoretical calculations of heterojunction discontinuities in the Si/Ge system," *Phys. Rev. B*, vol. 34, pp. 5621–5634, Oct. 1986.

[88] D. J. Paul, "Si/SiGe heterostructures: from material and physics to devices and circuits," *Semicond. Sci. Tech.*, vol. 19, pp. R75–R108, 2004.

[89] C. Kittel, *Introduction to Solid State Physics*, ch. 7, pp. 179–186. Wiley, Chichester, 7th edition ed., 1996.

[90] J. H. Davies, *The Physics of Low-Dimensional Semiconductors: An Introduction.* Cambridge University Press, Cambridge, 1998.

[91] Adapted from Creative Commons Attribution ShareAlike 2.5 licensed image, available as "File:Band_structure_Si_schematic.svg" at Wikimedia Commons.

[92] J. R. Chelikowsky and M. L. Cohen, "Electronic structure of silicon," *Phys. Rev. B*, vol. 10, pp. 5095–5107, Dec. 1974.

[93] M. M. Rieger and P. Vogl, "Electronic-band parameters in strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ alloys on $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ substrates," *Phys. Rev. B*, vol. 48, pp. 14276–14287, Nov. 1993.

[94] J. Weber and M. I. Alonso, "Near-band-gap photoluminescence of Si–Ge alloys," *Phys. Rev. B*, vol. 40, pp. 5683–5693, Sept. 1989.

[95] L. Yang *et al.*, "Si/SiGe heterostructure parameters for device simulations," *Semicond. Sci. Tech.*, vol. 19, pp. 1174–1182, Aug. 2004.

[96] S. Smirnov and H. Kosina, "Monte carlo modeling of the electron mobility in strained $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ layers on arbitrarily oriented $\mathrm{Si}_{1-y}\mathrm{Ge}_y$ substrates," *Solid State Electron.*, vol. 48, pp. 1325–1335, Apr. 2004.

[97] P. Harrison, *Quantum Wells, Wires and Dots.* Wiley, Chichester, 2nd edition ed., 2005.

[98] J. M. Hinckley and J. Singh, "Influence of substrate composition and crystallographic orientation on the band structure of pseudomorphic Si-Ge alloy films," *Phys. Rev. B*, vol. 42, pp. 3546–3566, Aug. 1990.

[99] G. S. Cargill III, J. Angilello, and K. L. Kavanagh, “Lattice compression from conduction electrons in heavily doped Si:As,” *Phys. Rev. Lett.*, vol. 61, pp. 1748–1751, Oct. 1988.

[100] J. Faist *et al.*, “Short wavelength ($\lambda \sim 3.4\,\mu$m) quantum cascade laser based on strained compensated InGaAs/AlInAs,” *Appl. Phys. Lett.*, vol. 72, no. 6, pp. 680–682, 1998.

[101] A. Wacker, “Gain in quantum cascade lasers and superlattices: A quantum transport theory,” *Phys. Rev. B*, vol. 66, p. 085326, Aug 2002.

[102] A. Rahman, M. S. Lundstrom, and A. W. Ghosh, “Generalized effective-mass approach for n-type metal-oxide-semiconductor field-effect transistors on arbitrarily oriented wafers,” *J. Appl. Phys.*, vol. 97, pp. 053702–1–11, Feb. 2005.

[103] G. Dresselhaus, A. F. Kip, and C. Kittel, “Cyclotron resonance of electrons and holes in silicon and germanium crystals,” *Phys. Rev.*, vol. 98, pp. 368–384, Apr 1955.

[104] J. Hensel, H. Hasegawa, and M. Nakayama, “Cyclotron resonance in uniaxially stressed silicon. II. nature of the covalent bond,” *Phys. Rev.*, vol. 138, pp. A225–A238, Apr. 1965.

[105] R. N. Dexter, H. J. Zeiger, and B. Lax, “Cyclotron resonance experiments in silicon and germanium,” *Phys. Rev.*, vol. 104, pp. 637–644, Nov 1956.

[106] A. Schenk, “Unified bulk mobility model for low- and high-field transport in silicon,” *J. Appl. Phys.*, vol. 79, no. 2, pp. 814–831, 1996.

[107] H. Mizuno, K. Taniguchi, and C. Hamaguchi, “Electron-transport simulation in silicon including anisotropic phonon scattering rate,” *Phys. Rev. B*, vol. 48, pp. 1512–1516, July 1993.

[108] M. G. Burt, “Fundamentals of envelope function theory for electronic states and photonic models in nanostructures,” *J. Phys. C Condens. Mat.*, vol. 11, pp. R53–R83, Mar. 1999.

[109] P. Harrison, D. Indjin, and R. W. Kelsall, “Electron temperature and mechanisms of hot carrier generation in quantum cascade lasers,” *J. Appl. Phys.*, vol. 92, no. 11, pp. 6921–6923, 2002.

[110] V. Spagnolo, G. Scamarcio, H. Page, and C. Sirtori, “Simultaneous measurement of the electronic and lattice temperatures in GaAs/$Al_{0.45}Ga_{0.55}As$ quantum-cascade lasers: Influence on the optical performance,” *Appl. Phys. Lett.*, vol. 84, no. 18, pp. 3690–3692, 2004.

[111] G. Curatola and G. Iannaccone, “Quantum confinement in silicon-germanium electron waveguides,” *Nanotechnology*, vol. 13, pp. 267–273, May 2002.

[112] L. R. Ram-Mohan, K. H. Yoo, and J. Moussa, “The Schrödinger-Poisson self-consistency in layered quantum semiconductor structures,” *J. Appl. Phys.*, vol. 95, pp. 3081–3092, Mar. 2004.

[113] D. G. Revin *et al.*, “InGaAs/AlAsSb quantum cascade lasers,” *Appl. Phys. Lett.*, vol. 85, no. 18, pp. 3992–3994, 2004.

[114] S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, “1.9 THz quantum-cascade lasers with one-well injector,” *Appl. Phys. Lett.*, vol. 88, no. 12, p. 121123, 2006.

[115] H. Luo *et al.*, “Terahertz quantum-cascade lasers based on a three-well active module,” *Appl. Phys. Lett.*, vol. 90, no. 4, p. 041112, 2007.

[116] B. S. Williams, S. Kumar, Q. Qin, Q. Hu, and J. L. Reno, “Terahertz quantum cascade lasers with double-resonant-phonon depopulation,” *Appl. Phys. Lett.*, vol. 88, no. 26, p. 261101, 2006.

[117] P. Weitz, R. Haug, K. von. Klitzing, and F. Schaflter, “Tilted magnetic field studies of spin- and valley-splittings in Si/Si$_{1-x}$Ge$_x$ heterostructures,” *Surf. Sci.*, vol. 361–362, p. 542, 1996.

[118] S. J. Koester, K. Ismail, and J. O. Chu, “Determination of spin- and valley-split energy levels in strained Si quantum wells,” *Semicond. Sci. Tech.*, vol. 12, p. 384, Apr. 1997.

[119] H. Koehler, M. Roos, and G. Landwehr, “Magnetic field dependence of the valley splitting in n-type inverted silicon MOSFET surfaces,” *Solid State Commun.*, vol. 27, pp. 955–959, 1978.

[120] R. Nicholas, K. von Klitzing, and T. Englert, “An investigation of the valley splitfing in n-channel silicon (100) inversion layers,” *Solid State Commun.*, vol. 34, pp. 51–55, 1980.

[121] F. F. Fang and P. J. Stiles, “Effects of a tilted magnetic field on a two-dimensional electron gas,” *Phys. Rev.*, vol. 174, pp. 823–828, Oct. 1968.

[122] G. Stöger *et al.*, “Shubnikov de Haas oscillations under hot-electron conditions in Si/Si$_{1-x}$Ge$_x$ heterostructures,” *Phys. Rev. B*, vol. 49, pp. 10417–10425, Apr. 1994.

[123] V. S. Khrapai, A. A. Shashkin, and V. T. Dolgopolov, “Strong enhancement of the valley splitting in a two-dimensional electron system in silicon,” *Phys. Rev. B*, vol. 67, p. 113305, Mar. 2003.

[124] K. Lai *et al.*, “Two-flux composite fermion series of the fractional quantum hall states in strained Si,” *Phys. Rev. Lett.*, vol. 93, p. 156805, Oct. 2004.

[125] T. B. Boykin *et al.*, “Valley splitting in strained silicon quantum wells,” *Appl. Phys. Lett.*, vol. 84, no. 1, pp. 115–117, 2004.

[126] T. B. Boykin and G. Klimeck, “The discretized Schrödinger equation for the finite square well and its relationship to solid-state physics,” *Eur. J. Phys.*, vol. 26, pp. 865–881, July 2005.

[127] T. B. Boykin *et al.*, “Valley splitting in low-density quantum-confined heterostructures studied using tight-binding models,” *Phys. Rev. B*, vol. 70, p. 165325, Oct. 2004.

[128] J.-C. Chiang, “Interference effects in Si/SiGe strained multivalley superlattice structures,” *Jpn. J. Appl. Phys.*, vol. 33, pp. L294–L296, Mar. 1994.

[129] M. O. Nestoklon, L. E. Golub, and E. L. Ivchenko, “Spin and valley-orbit splittings in SiGe/Si heterostructures,” *Phys. Rev. B*, vol. 73, p. 235334, June 2006.

[130] L. J. Sham and M. Nakayama, “Effective-mass approximation in the presence of an interface,” *Phys. Rev. B*, vol. 20, pp. 734–747, July 1979.

[131] T. B. Boykin *et al.*, “Valley splitting in V-shaped quantum wells,” *J. Appl. Phys.*, vol. 97, p. 113702, May 2005.

[132] G. Grosso, G. Pastori Parravicini, and C. Piermarocchi, “Valley splitting in triangular Si(001) quantum wells,” *Phys. Rev. B*, vol. 54, pp. 16393–16396, Dec. 1996.

[133] F. J. Ohkawa and Y. Uemura, “Theory of valley splitting in an n-channel (100) inversion layer of Si. I. Formulation by extended zone effective mass theory,” *J. Phys. Soc. Jpn.*, vol. 43, pp. 907–916, Sept. 1977.

[134] F. J. Ohkawa and Y. Uemura, “Theory of valley splitting in an n-channel (100) inversion layer of Si. II. Electric break through,” *J. Phys. Soc. Jpn.*, vol. 43, pp. 917–924, Sept. 1977.

[135] F. Ohkawa, “Electric break-through in an inversion layer: Exactly solvable model,” *Solid State Commun.*, vol. 26, pp. 69–71, 1978.

[136] A. Fang, Y.-C. Chang, and J. R. Tucker, "Simulation of Si:P spin-based quantum computer architecture," *Phys. Rev. B*, vol. 72, p. 075355, Aug. 2005.

[137] D. Z. Y. Ting and Y.-C. Chang, "L-valley-derived states in (001) GaSb/AlSb quantum wells and superlattices," *Phys. Rev. B*, vol. 38, pp. 3414–3422, Aug. 1988.

[138] M. Friesen, S. Chutia, C. Tahan, and S. N. Coppersmith, "Valley splitting theory of SiGe/Si/SiGe quantum wells," *Phys. Rev. B*, vol. 75, p. 115318, Mar. 2007.

[139] P. Friedel, M. S. Hybertsen, and M. Schülter, "Local empirical pseudopotential approach to the optical properties of Si/Ge superlattices," *Phys. Rev. B*, vol. 39, pp. 7974–7977, Apr. 1989.

[140] M. V. Fischetti and S. E. Laux, "Band structure, deformation potentials, and carrier mobility in strained Si, Ge, and SiGe alloys," *J. Appl. Phys.*, vol. 80, pp. 2234–2252, Aug. 1996.

[141] Z. Ikonić, R. W. Kelsall, and P. Harrison, "Virtual-crystal approximation and alloy broadening of intersubband transitions in p-type SiGe superlattices," *Phys. Rev. B*, vol. 64, p. 125308, Sept. 2001.

[142] Y. Shiraki and A. Sakai, "Fabrication technology of SiGe hetero-structures and their properties," *Surf. Sci. Rep.*, vol. 59, pp. 153–207, Aug. 2005.

[143] J. Zhang *et al.*, "SiGe quantum cascade structures for light emitting devices," *J. Cryst. Growth*, vol. 278, pp. 488–494, 2005.

[144] D. Paul *et al.*, "Si/SiGe quantum-cascade emitters for terahertz applications," *Physica E*, vol. 16, pp. 147–155, Jan. 2003.

[145] D. Paul *et al.*, "Electroluminescence from Si/SiGe quantum cascade emitters," *Physica E*, vol. 16, pp. 309–314, Mar. 2003.

[146] R. W. Kelsall *et al.*, “Intersubband lifetimes in p-Si/SiGe terahertz quantum cascade heterostructures,” *Phys. Rev. B*, vol. 71, p. 115326, Mar. 2005.

[147] M. Zhao *et al.*, “Low-temperature molecular beam epitaxy growth of Si/SiGe THz quantum cascade structures on virtual substrates,” *Thin Solid Films*, vol. 508, pp. 24–26, 2006.

[148] H. Callebaut and Q. Hu, “Importance of coherence for electron transport in terahertz quantum cascade lasers,” *J. Appl. Phys.*, vol. 98, no. 10, p. 104505, 2005.

[149] S.-C. Lee and A. Wacker, “Nonequilibrium Green’s function theory for transport and gain properties of quantum cascade structures,” *Phys. Rev. B*, vol. 66, p. 245314, Dec. 2002.

[150] V. D. Jovanovic *et al.*, “Influence of doping density on electron dynamics in GaAs/AlGaAs quantum cascade lasers,” *J. Appl. Phys.*, vol. 99, no. 10, p. 103106, 2006.

[151] F. T. Ulaby, *Fundamentals of applied electromagnetics*. Prentice Hall, 1999.

[152] M. Dür, S. M. Goodnick, and P. Lugli, “Monte carlo simulation of intersubband relaxation in wide, uniformly doped GaAs/Al$_x$Ga$_{1-x}$As quantum wells,” *Phys. Rev. B*, vol. 54, pp. 17794–17804, Dec 1996.

[153] F. Stern and W. E. Howard, “Properties of semiconductor surface inversion layers in the electric quantum limit,” *Phys. Rev.*, vol. 163, pp. 816–835, Nov 1967.

[154] J. H. Smet, C. G. Fonstad, and Q. Hu, “Intrawell and interwell intersubband transitions in multiple quantum wells for far-infrared sources,” *J. Appl. Phys.*, vol. 79, no. 12, pp. 9305–9320, 1996.

[155] T. Ando, A. B. Fowler, and F. Stern, “Electronic properties of two-dimensional systems,” *Rev. Mod. Phys.*, vol. 54, pp. 437–672, Apr. 1982.

[156] G. Bastard, *Wave mechanics applied to semiconductor heterostructures.* Wiley series in monographs of physics, John Wiley & Sons, Chichester, 1990.

[157] R. E. Prange and T.-W. Nee, "Quantum spectroscopy of the low-field oscillations in the surface impedance," *Phys. Rev.*, vol. 168, p. 779, Apr. 1968.

[158] U. Penner, H. Rücker, and I. N. Yassievich, "Theory of interface roughness scattering in quantum wells," *Semicond. Sci. Tech.*, vol. 13, pp. 709–713, 1998.

[159] H. Sakaki, T. Noda, K. Hirakawa, M. Tanaka, and T. Matsusue, "Interface roughness scattering in GaAs/AlAs quantum wells," *Appl. Phys. Lett.*, vol. 51, no. 23, pp. 1934–1936, 1987.

[160] T. Unuma, M. Yoshita, T. Noda, H. Sakaki, and H. Akiyama, "Intersubband absorption linewidth in GaAs quantum wells due to scattering by interface roughness, phonons, alloy disorder, and impurities," *J. Appl. Phys.*, vol. 93, pp. 1586–1597, Feb. 2003.

[161] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products.* Academic Press, San Diego, sixth ed., 2000.

[162] B. Mukhopadhyay and P. K. Basu, "Linewidth for interconduction subband transition in Si/Si$_{1-x}$Ge$_x$ quantum wells," *Phys. Status Solidi*, vol. 241, pp. 1859–1864, May 2004.

[163] R. M. Feenstra and M. A. Lutz, "Scattering from strain variations in high-mobility Si/SiGe heterostructures," *J. Appl. Phys.*, vol. 78, pp. 6091–6097, Nov. 1995.

[164] P. J. Price, "Two-dimensional electron transport in semiconductor layers. I. phonon scattering," *Ann. Phys.*, vol. 133, no. 2, pp. 217–239, 1981.

[165] M. V. Fischetti and S. E. Laux, "Monte Carlo study of electron transport in silicon inversion layers," *Phys. Rev. B*, vol. 48, pp. 2244–2274, July 1993.

[166] M. Fischetti, “Monte carlo simulation of transport in technologically significant semiconductors of the diamond and zinc-blende structures. I. homogeneous transport,” *IEEE Trans. Electron Dev.*, vol. 38, pp. 634–649, Mar. 1991.

[167] B. K. Ridley, *Electrons and phonons in semiconductor multilayers.* Cambridge University Press, Cambridge, 1997.

[168] C. Jacoboni, F. Nava, C. Canali, and G. Ottaviani, “Electron drift velocity and diffusivity in germanium,” *Phys. Rev. B*, vol. 24, pp. 1014–1026, July 1981.

[169] P. Dollfus, “Si/$\mathrm{Si}_{1-x}\mathrm{Ge}_x$ heterostructures: Electron transport and field-effect transistor operation using monte carlo simulation,” *J. Appl. Phys.*, vol. 82, pp. 3911–3916, Oct. 1997.

[170] D. K. Ferry, “First-order optical and intervalley scattering in semiconductors,” *Phys. Rev. B*, vol. 14, pp. 1605–1609, Aug. 1976.

[171] F. Monsef, P. Dollfus, S. Galdin, and A. Bournel, “First-order intervalley scattering in low-dimensional systems,” *Phys. Rev. B*, vol. 65, pp. 212304–1–4, June 2002.

[172] J. H. Van der Merwe, “Crystal interfaces. part II. finite overgrowths,” *J. Appl. Phys.*, vol. 34, no. 1, pp. 123–127, 1963.

[173] T. Fromherz, M. Meduna, G. Bauer, A. Borak, C. V. Falub, S. Tsujino, H. Sigg, and D. Grutzmacher, “Intersubband absorption of strain-compensated $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ valence-band quantum wells with $0.7 \leq x \leq 0.85$,” *J. Appl. Phys.*, vol. 98, no. 4, p. 044501, 2005.

[174] M. L. Lee, D. A. Antoniadis, and E. A. Fitzgerald, “Challenges in epitaxial growth of SiGe buffers on Si (111), (110), and (112),” *Thin Solid Films*, vol. 508, no. 1-2, pp. 136–139, 2006. Proc. Fourth Int. Conf. on Silicon Epitaxy and Heterostructures (ICSI-4) - ICSI-4.

[175] I. Yonenaga and T. Ayuzawa, “Segregation coefficients of various dopants in $\mathrm{Si}_x\mathrm{Ge}_{1-x}$ $(0.93 < x < 0.96)$ single crystals,” *J. Cryst. Growth*, vol. 297, no. 1, pp. 14 – 19, 2006.

[176] J. C. Bean, “Arbitrary doping profiles produced by Sb-doped Si MBE,” *Appl. Phys. Lett.*, vol. 33, no. 7, pp. 654–656, 1978.

[177] J. Zhang *et al.*, “Sharp n-type doping profiles in Si/SiGe heterostructures produced by atomic hydrogen etching,” *Surf. Sci.*, vol. 600, p. 2288, Apr. 2006.

[178] A. C. Churchill *et al.*, “Two-dimensional electron gas mobility as a function of virtual substrate quality in strained Si/SiGe heterojunctions,” in *Papers from the conference on silicon heterostructures: from physics to devices*, vol. 16, pp. 1634–1638, AVS, 1998.

[179] J. P. Dismukes, L. Ekstrom, E. F. Steigmeier, I. Kudman, and D. S. Beers, “Thermal and electrical properties of heavily doped Ge-Si alloys up to 1300°K,” *J. Appl. Phys.*, vol. 35, no. 10, pp. 2899–2907, 1964.

[180] K. Brunner, H. Dobler, G. Abstreiter, H. Schfer, and B. Lustig, “Molecular beam epitaxy growth and thermal stability of $\mathrm{Si}_{1-x}\mathrm{Ge}_x$ layers on extremely thin silicon-on-insulator substrates,” *Thin Solid Films*, vol. 321, no. 1-2, pp. 245 – 250, 1998.

[181] K. K. Linder *et al.*, “Reduction of dislocation density in mismatched SiGe/Si using a low-temperature Si buffer layer,” *Appl. Phys. Lett.*, vol. 70, no. 24, pp. 3224–3226, 1997.

[182] T. Hackbarth *et al.*, “High frequency n-type modfets on ultra-thin virtual SiGe substrates,” *Solid State Electron.*, vol. 47, no. 7, pp. 1179 – 1182, 2003. 3rd International Workshop on Ultimate Integration of Silicon March 7-8, 2002, Munich, Germany.

[183] S. B. Samavedam, M. T. Currie, T. A. Langdo, and E. A. Fitzgerald, “High-quality germanium photodiodes integrated on silicon substrates using optimized relaxed graded buffers,” *Appl. Phys. Lett.*, vol. 73, no. 15, pp. 2125–2127, 1998.

[184] J. Oh *et al.*, “Interdigitated Ge p-i-n photodetectors fabricated on a Si substrate using graded SiGe buffer layers,” *IEEE J. Quantum. Elect.*, vol. 38, pp. 1238–1241, Sept. 2002.

[185] J. M. Hartmann *et al.*, “Reduced pressure–chemical vapor deposition of Ge thick layers on Si(001) for 1.3–1.55 μm photodetection,” *J. Appl. Phys.*, vol. 95, no. 10, pp. 5905–5913, 2004.

[186] H.-C. Luan *et al.*, “High-quality Ge epilayers on Si with low threading-dislocation densities,” *Appl. Phys. Lett.*, vol. 75, no. 19, pp. 2909–2911, 1999.

[187] E. H. Li, B. L. Weiss, and K.-S. Chan, “Eigenstates and absorption spectra of interdiffused AlGaAs–GaAs multiple-quantum-well structures,” *IEEE J. Quantum. Elect.*, vol. 32, pp. 1399–1416, Aug. 1996.

[188] N. Ozguven and P. C. McIntyre, “Silicon-germanium interdiffusion in high-germanium-content epitaxial heterostructures,” *Appl. Phys. Lett.*, vol. 92, no. 18, p. 181907, 2008.

[189] K. Kash *et al.*, “Aluminum ion-implantation enhanced intermixing of GaAs–AlGaAs quantum-well structures,” *J. Appl. Phys.*, vol. 63, no. 1, pp. 190–194, 1988.

[190] S. Anders, W. Schrenk, T. Roch, C. Pflügl, and G. Strasser, “Tuning quantum-cascade lasers by postgrowth rapid thermal processing,” *Appl. Phys. Lett.*, vol. 84, no. 2, pp. 164–166, 2004.

[191] T. Roch, C. Pflugl, A. M. Andrews, W. Schrenk, and G. Strasser, “X-ray

investigation of quantum well intermixing after postgrowth rapid thermal processing," *J. Phys. D: Appl. Phys.*, vol. 38, no. 10A, pp. A132–A136, 2005.

[192] Z. Ikonić, P. Harrison, and R. W. Kelsall, "Self-consistent energy balance simulations of hole dynamics in SiGe/Si THz quantum cascade structures," *J. Appl. Phys.*, vol. 96, no. 11, pp. 6803–6811, 2004.

[193] A. Yariv, *Quantum Electronics.* John Wiley & Sons, Singapore, 3rd edition ed., 1989.

[194] A. Wittmann, Y. Bonetti, J. Faist, E. Gini, and M. Giovannini, "Intersubband linewidths in quantum cascade laser designs," *Applied Physics Letters*, vol. 93, no. 14, p. 141103, 2008.

[195] G. Scalari, N. Hoyler, M. Giovannini, and J. Faist, "Terahertz bound-to-continuum quantum-cascade lasers based on optical-phonon scattering extraction," *Appl. Phys. Lett.*, vol. 86, no. 18, p. 181101, 2005.

[196] P. Harrison, *Quantum Wells, Wires and Dots.* Wiley, Chichester, 3rd edition ed., 2009.

[197] F. Schaffler, *Properties of Advanced Semiconductor Materials GaN, AlN, InN, BN, SiC, SiGe.* John Wiley & Sons, 2001.

[198] S. Kohen, B. S. Williams, and Q. Hu, "Electromagnetic modeling of terahertz quantum cascade laser waveguides and resonators," *J. Appl. Phys.*, vol. 97, no. 5, p. 053106, 2005.

[199] J. Alton *et al.*, "Buried waveguides in terahertz quantum cascade lasers based on two-dimensional surface plasmon modes," *Appl. Phys. Lett.*, vol. 86, no. 7, p. 071109, 2005.

[200] E. Anemogiannis, E. Glytsis, and T. Gaylord, "Determination of guided and leaky modes in lossless and lossy planar multilayer optical waveguides: re-

flection pole method and wavevector density method," *J. Lightwave Technol.*, vol. 17, pp. 929–941, May 1999.

[201] C. Sirtori *et al.*, "Quantum cascade laser with plasmon-enhanced waveguide operating at 8.4 μm wavelength," *Appl. Phys. Lett.*, vol. 66, no. 24, pp. 3242–3244, 1995.

[202] B. S. Williams, S. Kumar, H. Callebaut, Q. Hu, and J. L. Reno, "Terahertz quantum-cascade laser at $\lambda \approx 100\,\mu$m using metal waveguide for mode confinement," *Appl. Phys. Lett.*, vol. 83, no. 11, pp. 2124–2126, 2003.

[203] M. A. Belkin *et al.*, "Terahertz quantum cascade lasers with copper metal–metal waveguides operating up to 178 K," *Opt. Express*, vol. 16, no. 5, pp. 3242–3248, 2008.

[204] R. Kelsall, Z. Ikonic, P. Harrison, S. Lynch, P. Townsend, D. Paul, D. Norris, S. Liew, A. Cullis, X. Li, J. Zhang, M. Bain, and H. Gamble, "Optical cavities for Si/SiGe tetrahertz quantum cascade emitters," *Opt. Mater.*, vol. 27, no. 5, pp. 851–854, 2005. Si-based Photonics: Towards True Monolithic Integration.

[205] D. Buca, S. Winnerl, S. Lenk, C. Buchal, and D.-X. Xu, "Fast time response from Si–SiGe undulating layer superlattices," *Appl. Phys. Lett.*, vol. 80, no. 22, pp. 4172–4174, 2002.

[206] S. Barbieri *et al.*, "2.9 THz quantum cascade lasers operating up to 70 K in continuous wave," *Appl. Phys. Lett.*, vol. 85, no. 10, pp. 1674–1676, 2004.

[207] A. Müller, M. Beck, J. Faist, U. Oesterle, and M. Ilegems, "Electrically tunable, room-temperature quantum-cascade lasers," *Appl. Phys. Lett.*, vol. 75, no. 11, pp. 1509–1511, 1999.

[208] J. Faist, M. Beck, T. Aellen, and E. Gini, "Quantum-cascade lasers based on a bound-to-continuum transition," *Appl. Phys. Lett.*, vol. 78, no. 2, pp. 147–149, 2001.

[209] G. Scalari *et al.*, "Far-infrared ($\lambda \approx 87\,\mu$m) bound-to-continuum quantum-cascade lasers operating up to 90 K," *Appl. Phys. Lett.*, vol. 82, no. 19, pp. 3165–3167, 2003.

[210] G. Han, J. Yu, and Y. Liu, "n-type Ge–SiGe quantum cascade structure utilizing quantum wells for electrons in the L and Γ valleys," *Photonics Technol. Lett., IEEE*, vol. 20, pp. 419–421, Mar. 2008.

[211] Q. Hu *et al.*, "Resonant-phonon-assisted THz quantum-cascade lasers with metal–metal waveguides," *Semicond. Sci. Tech.*, vol. 20, no. 7, pp. S228–S236, 2005.

[212] R. Köhler *et al.*, "Terahertz quantum-cascade lasers based on an interlaced photon–phonon cascade," *Appl. Phys. Lett.*, vol. 84, no. 8, pp. 1266–1268, 2004.

[213] C. Lange *et al.*, "Ultrafast nonlinear optical response of photoexcited Ge/SiGe quantum wells: Evidence for a femtosecond transient population inversion," *Phys. Rev. B*, vol. 79, no. 20, p. 201306, 2009.

[214] J. McTavish, Z. Ikonić, D. Indjin, and P. Harrison, "Intervalley scattering in GaAs/AlGaAs quantum wells and quantum cascade lasers," *Microelectronics Journal*, vol. 40, no. 3, pp. 577–580, 2009.

[215] C. Evans, V. Jovanovic, D. Indjin, Z. Ikonic, and P. Harrison, "Investigation of thermal effects in quantum-cascade lasers," *IEEE Journal of Quantum Electronics*, vol. 42, no. 9, pp. 859–867, 2006.

[216] A. Lops, V. Spagnolo, and G. Scamarcio, "Thermal modeling of GaInAs/ AlInAs quantum cascade lasers," *J. Appl. Phys.*, vol. 100, p. 043109, 2006.

[217] C. Evans *et al.*, "Thermal modeling of terahertz quantum-cascade lasers: comparison of optical waveguides," *IEEE J. Quantum Electron.*, vol. 44, no. 7, pp. 680–685, 2008.

[218] A. Carim and R. Sinclair, “The Evolution of Si/SiO$_2$ Interface Roughness,” *J. Electrochem. Soc.*, vol. 134, p. 741, 1987.

[219] M. Bain *et al.*, “SiGe HBTs on bonded SOI incorporating buried silicide layers,” *IEEE Trans. Electron Dev.*, vol. 52, pp. 317–324, Mar. 2005.

[220] *Conf. of the NATO Adv. Research Workshop on Towards the First Silicon Laser*, Sept. 2003.