THE UNIVERSITY
of ADELAIDE

Quantum Transport and Spin-Orbit Coupling in Semiconductor Nanostructures

Tyler Whittaker

Supervisors:
Dr. Giuseppe Tettamanzi
Assoc. Prof. Rajib Rahman

A thesis submitted towards the degree of Master of Philosophy

The Faculty of Sciences
The University of Adelaide

July 7, 2023

Abstract

Semiconductor nanostructures present a promising path to enhance current computing technology and develop quantum information technologies. The electronic structure of these nanostructures has direct implications on the transport of electrons through them and their potential device applications. This thesis explores the electronic structures of various semiconducting nanostructures and their dependence on electromagnetic fields, dopant atoms, spin-orbit coupling and atomistic features through experiment and modelling. A form of quantum transport, known as single electron charging, is demonstrated through the electron-bound states of dopant atoms in a silicon nanostructure. A portion of a recently published work is presented which models the response of valley states in a gate-defined quantum dot within a silicon quantum well to an interface step and applied electric fields. The main work of this thesis employs an atomistic tight-binding model to determine the effective g-factor anisotropy of InAs nanowires under various atomistic and electromagnetic conditions. The spin-orbit interactions present in the nanowires are extracted from the effective g-factor anisotropy with an effective model. The modelling results provide insights for InAs nanowire applications in fields such as Majorana zero mode research, spintronic devices and quantum information technology.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree. The author acknowledges that copyright of published works contained within the thesis resides with the copyright holder(s) of those works. I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time. I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Tyler Whittaker

Acknowledgements

Firstly, I would like to thank my supervisor Dr. Giuseppe Carlo Tettamanzi for introducing me to the field, giving me the opportunity to work on this project and for his constant support and encouragement. I would like to acknowledge my co-supervisor Dr. Rajib Rahman for providing me with a direction for the project when difficulties arose and for guidance throughout my candidature. I would like to extend many thanks to Dr. Yuling Hsueh for her insights and time spent checking calculations.

Additionally, I would like to thank my colleagues Ross, Alan, Tom, Daniel and Jonathan for countless stimulating conversations over a coffee or a croissant. I would like to thank my friends for providing an escape from my thesis, and never failing to ask when it will be finished. I would like to extend a special thanks to Declan for his efforts in proof reading this manuscript and for always being willing to listen to my thesis related rambles. Finally, I'd like to thank my family, especially my parents Ben and Trina, for their continual support.

Contents

1. Introduction & Outline **1**

1.1. Outline 2

2. Background Theory **3**

2.1. Quantum confinement and nanostructures 3

2.1.1. Preliminaries: The single electron hamiltonian 4

2.1.2. Subbands in nanostructures 5

2.2. Quantised conductance 7

2.3. Single electron tunnelling and Coulomb blockade 9

2.3.1. Coulomb blockade 9

2.3.2. Single electron tunnelling 13

2.3.3. Single atom transistor 17

2.4. Spin-orbit interaction 18

2.4.1. Atomic physics origins 18

2.4.2. Derivation from Dirac's equation 19

2.4.3. Spin-orbit interaction in semiconductors 22

2.4.4. Spin-orbit interactions in nanowires 25

3. Investigating quantum transport through dopants in silicon **33**

3.1. Preliminaries 34

3.1.1. The devices 34

3.1.2. Device preparation 35

3.1.3. FinFET operation 36

3.1.4. Room temperature testing 38

3.2. FinFET operation at low temperatures 42

3.3. Identifying dopant states from Coulomb blockade 46

3.3.1. Physical interpretation of Coulomb peaks 46

3.3.2. Determining the coupling constant 48

3.3.3. Binding energies of donor states 51

3.3.4. Difficulties in identifying the dopants 53
3.4. Summary 55

4. Atomistic tight-binding modelling tool - NEMO3D 57
4.1. Nanoelectronic modelling 58
4.1.1. Common assumptions 58
4.1.2. Simplified models 59
4.1.3. NEMO3D 60
4.2. Atomistic tight-binding 61
4.2.1. Construction of the tight-binding Hamiltonian 61
4.2.2. The spin-orbit interaction 65
4.2.3. Electromagnetic coupling 66
4.2.4. Solving the Hamiltonian 67
4.3. Example of bulk band structures: Si and InAs 68
4.3.1. Silicon bulk band structure 68
4.3.2. Indium arsenide bulk band structure 70
4.4. Valley splitting of quantum dot in Si/SiGe quantum well 72

5. Effective g-factor anisotropy of indium arsenide nanowires 77
5.1. Indium arsenide nanowire in NEMO3D 78
5.1.1. Nanowire growth direction 78
5.1.2. Nanowire geometry 79
5.1.3. Nanowire band structure 82
5.2. Effective g-factor dependence on nanowire confinement 83
5.3. Effective model of the nanowire g-factor anisotropy 86
5.3.1. The nanowire geometry for g-factor anisotropy 87
5.3.2. The effective Hamiltonian 87
5.3.3. Deriving the effective g-factor expression 89
5.4. Effective g-factor anisotropy of square cross-section nanowires 94
5.4.1. Electric field dependence 94
5.4.2. Mixed-ion termination 98
5.4.3. Larger square cross-section nanowires 102
5.5. Rectangular cross-section nanowires 107
5.5.1. Effective g-factor anisotropy 108
5.5.2. Mixed-ion termination and applied electric fields 111
5.6. Discussion of results and applications 113
5.7. Summary 118

6. Conclusion **119**

A. Useful Derivations **123**

A.1. A review: The 1D infinite potential well . . . 123
A.2. An infinite box: Periodic boundary conditions . . . 124
A.3. The Pauli matrices and identities . . . 125
A.4. Unitary transformations . . . 126
A.4.1. The Baker-Campbell-Hausdorff formula . . . 126
A.4.2. Rotation of the spin basis . . . 127
A.5. Time-independent perturbation theory . . . 128

B. Measurement techniques **131**

B.1. Measurement setup . . . 132
B.1.1. Computer . . . 132
B.1.2. Instruments . . . 133
B.1.3. Control Box . . . 135
B.1.4. Cryostat . . . 136
B.1.5. Measurement philosophy . . . 137
B.2. Testing the measurement setup . . . 138
B.2.1. Transistor IV curve setup . . . 139

C. Cryogenic cooling **141**

C.1. Cooling with liquid helium . . . 141
C.1.1. Cooling to 1K . . . 142
C.2. Thermometry . . . 142
C.3. The cooling procedure . . . 143
C.3.1. Required equipment . . . 144
C.3.2. Safety . . . 146
C.3.3. The cool down procedure . . . 147
C.3.4. The warm-up procedure . . . 151

D. Quantum transport through dopants in silicon supplementary **153**

D.1. Determining the energy barrier and conduction band minimum . . . 153
D.2. Limits of the measurement setup . . . 156
D.2.1. Current amplifier gain and offset . . . 156
D.2.2. Problems encountered with cooling . . . 158

E. Effective g-factor anisotropy of InAs nanowires supplementary 161
E.1. Magnetic field dependence of the effective g-factor anisotropy 161

Bibliography 163

"There are two possible outcomes: if the result confirms the hypothesis, then you've made a measurement. If the result is contrary to the hypothesis, then you've made a discovery"
— Enrico Fermi

Chapter 1.

Introduction & Outline

The field of condensed matter physics, particularly the subfields of mesoscopic and solid-state physics, have inspired the development of modern computing technology. The field boasts key technological advancements from the first working transistor in 1947 by James Bardeen, Walter Brattain and William Shockley at Bell labs [1, 2] to the CMOS integrated circuits [3] that power modern day smart phones and computers. The continual advancement of technology is built upon the principle of increasing the complexity of circuitry on a limited space and to date has roughly followed Moore's law [4]. State-of-the-art CMOS integrated circuit technology is 5 nm node [5, 6, 7] and soon to reach 3 nm node technology [8]. It is not long before this technology reaches the atomic limit. As such, condensed matter physicists are constantly searching for new ways to improve on current integrated circuit technology.

The demand for the miniaturisation of integrated circuits led to the development of devices built from small semiconductor nanostructures. These semiconducting nanostructures were found to exhibit interesting quantum properties which inspired the start of new fields of research such as spintronics and quantum computing. Spintronics aims to utilise the spin degree of freedom to improve computing performance. This is because less power is required to control a single carrier spin [9]. Quantum computers describe information in the form of states that obey quantum principles such as superposition, entanglement and decoherence. As a result they can solve specific problems that classical computers cannot and solve other problems more efficiently [10, 11]. The implementation of a quantum computer involves constructing a robust two level system called a qubit. Semiconductor nanostructures, particularly in silicon, have become an attractive host for these qubits because they allow for the quantum control of electron states and are compatible with commercial semiconductor processing [12, 13]. There is also a significant

research effort focused on constructing topologically protected qubits from semiconductors in proximity with superconductors [14, 15, 16]. The realisation of these systems will require the observation of novel electron-hole excitations, which are the condensed matter analogs of Majorana fermions – particles which are their own anti-particles [17].

This thesis explores properties of various semiconducting nanostructures, particularly their electronic band structure, quantum transport and spin-orbit coupling which are important for applications in modern electronics/spintronics and quantum information. Coulomb blockade and single electron tunnelling through dopant atoms in a silicon nanostructure is observed. Previous experiments of this kind initiated the development of dopant in silicon qubits. From our recently published work, the coupling of two valley states in a gate-defined quantum dot within a silicon quantum well with a step interface is presented. The coupling is found to be electrically tunable which is ideal for creating valley qubit gates. The main work of this thesis is modelling the effective g-factor anisotropy of the conduction band in InAs nanowires to gain insight on the spin-orbit interactions present in the nanowires for applications in topological qubits hosting Majorana excitations, spin-orbit qubits and spintronic devices.

1.1. Outline

The structure of this thesis is as follows. Chapter 2 explains the background theory relevant to this thesis, including quantum confinement, quantised conductance, Coulomb blockade, spin-orbit interaction and Majorana zero bias peaks. Chapter 3 discusses the measurement of Coulomb blockade and single electron tunnelling through the electron bound states of dopants in silicon. Chapter 4 introduces the tight-binding formalism and the numerical modelling techniques of the three-dimensional Nano-Electronic MOdelling (NEMO3D) tool. The modelling of the valley state coupling in the gated quantum dot of a silicon quantum well is presented as an example of NEMO3D's modelling capabilities. Chapter 5 discusses a method for modelling the effective g-factor anisotropy and spin-orbit interactions present in a [100] InAs nanowire with different applied electric fields, surface-ion terminations and confinement conditions using NEMO3D and an analytic continuum model. Based on the results of this modelling, possible applications of the nanowires in spintronics and quantum information technology are suggested. Chapter 6 summarises the work of this thesis, highlighting significant findings and potential directions for future research.

Chapter 2.

Background Theory

This chapter outlines the background physics that will be referenced throughout this thesis. Quantum nanostructures and the field of quantum transport are introduced, specifically Coulomb blockade and single electron transport, which are fundamental to modern mesoscopic research. These quantum transport phenomena are employed in this thesis to determine the bound states of dopants in a silicon channel and can be used to determine the electronic structure of InAs nanowires modelled in later chapters. In addition, the spin-orbit interaction is introduced because it has important implications on the InAs electron spin structure e.g. g-factor anisotropy which in turn influences the quantum transport properties of nanowires. The last section of this chapter provides a brief introduction to one of the largest areas of research employing III-V nanowires with large spin-orbit coupling, the search for Majorana zero modes.

2.1. Quantum confinement and nanostructures

The field of mesoscopic physics – the study of micro- to nano-scale materials – is built upon the concept of quantum confinement of carriers (typically electrons) in complex many body materials e.g. semiconductors. Quantum confinement of a bulk semiconductor to nano-scales creates a semiconductor *nanostructure*. The semiconductor nanostructure allows for the control of an electron's state and energy on a quantum level which leads to interesting observable phenomena. These nanostructures have vast applications in modern nano-electronics such as the transistors used in our computers and phones. Currently, a large portion of mesoscopic research focuses on nanostructures for quantum computing applications, including quantum dot spin qubits [18], dopant in silicon spin

qubits [19, 20, 21], spin-orbit qubits [22, 23, 24] and topological qubits [14, 15, 16, 17]. Additionally, nanostructures are employed for spintronics applications, such as the spin field effect transistor [25], and various quantum optics applications, such as single-photon sources [26] and detectors [27].

From introductory quantum mechanics, an electron confined to a box, with dimensions much larger than its wavelength, behaves as a free particle. However, when one confines an electron in a box with dimensions comparable to or smaller than the de Broglie wavelength of the electron λ_B, typically the nanoscale, the properties of the electron change significantly from that of the free electron i.e. the electron gains quantised energy modes. This is called quantum confinement. Quantum confinement in one dimension creates a *quantum well*, confinement in two dimensions a quantum wire or *nanowire* and confining in all three dimension creates a *quantum dot*. If the thermal energy of the system is small in comparison to the energy spacing of the quantised energy modes, often requiring cryogenic temperatures, the transport of the electrons through such a semiconducting nanostructure is no longer classical but follows the laws of quantum mechanics. This is known as quantum transport [28, 29]. Quantum transport manifests as various observable phenomena depending on the system. In quantum point contacts and nanowires it results in a quantisation of the conductance [30, 31, 32], while in quantum dots coupled to three-terminals it can result in single electron tunnelling and Coulomb blockade [33, 34, 35].

In this section, the single electron Hamiltonian is introduced along with the main effect of quantum confinement on the electronic band structure of a semiconductor. Namely, that quantum confinement introduces quantised subbands into the band structure of semiconductor nanostructures. These subbands can facilitate quantum transport.

2.1.1. Preliminaries: The single electron hamiltonian

A semiconductor is made up of the interactions between many ion cores organised in a crystal lattice and the electrons surrounding these ion cores. Typically, these ion cores are assumed to consist of the atom nuclei and their inner-shell electrons and only the valence electrons are considered to be free. The semiconductor system can be described by a many particle hamiltonian

$$H = H_e + H_i + H_{ei} \tag{2.1}$$

where H_e describes the electrons including the electron-electron Coulombic interactions, H_i describes the ions and the Coulombic interactions between them and H_{ei} describes the interaction of electrons with the ions. This Hamiltonian will have eigenfunctions and eigenvalues in the form of many particle wavefunctions and their corresponding energies.

For semiconductor systems it is reasonable to assume that the ion cores are stationary with respect to the electron's motion, the Born-Oppenheimer approximation [36]. This allows one to neglect the dynamics of the ions H_i and focus on the dynamics of the electrons in the potential of a (stationary) ion core lattice. Typically, the many electron Hamiltonian, H_e, is simplified to a single electron Hamiltonian by a mean field approximation. That is the assumption that every electron experiences the same potential due to the ion core lattice V_i. Additionally, an effective potential describing the electron-electron interactions that each electron feels, V_{ee}, can be included. The resulting single electron Hamiltonian is

$$H_{1e} = \frac{\mathbf{p}^2}{2m^*} + V_i + V_{ee} \tag{2.2}$$

with $\mathbf{p} = -i\hbar\boldsymbol{\nabla}$ the electron momentum operator and m^* the effective mass of the electron in the semiconductor. The energy eigenvalues of this Hamiltonian, $E_n(\mathbf{k})$ (where $\mathbf{k}$ is the wave vector), make up the semiconductor's electronic band structure with the bands indexed by n. The different methods of modelling the dynamics of electrons in semiconductors, specifically electron-electron interactions, is discussed further in Chapter 4. However, for the majority of this thesis the focus is on the single electron Hamiltonian without V_{ee}. The following sections will build upon equation (2.2) to explain quantum transport, electromagnetic coupling and spin-orbit interactions.

2.1.2. Subbands in nanostructures

From the elementary quantum mechanics problem of the infinite potential well[1], the confinement of an electron to a 1D infinite potential well produces a quantised wavevector and resonant frequencies for the electron wavefunction. In a semiconductor nanostructure, where one or more dimensions has been confined on nanoscales comparable to the de Broglie wavelength, the quantum confinement produces quantised subbands in the semiconductor band structure.

[1] Refer to Appendix A.1.

Following an example from the work of Datta [29], consider electrons in a homogeneous solid moving freely in three dimensions. The energy levels describing the motion of electrons are $E_n(k_x, k_y, k_z)$ where n denotes the different bands. To create a quantum well confined in z by length L_z one can introduce an infinite potential like equation (A.1) into the single electron Hamiltonian. This would quantise the wavevector in the z direction to $k_z = p\pi/L_z$ and energy bands would become

$$E_{n,p}(k_x, k_y) \approx E_n(k_x, k_y, k_z = p\pi/L_z) \tag{2.3}$$

where p is the subband index and takes integer values. The quantum well has 1D subbands each with a 2D dispersion relation. However, keep in mind this is only an approximation because the solid is assumed to be homogeneous and the surface is approximated as an infinite potential. For this to be classified as a quantum well L_z must be small enough so that quantum effects are experimentally observable. This occurs when the discrete energy levels. $\epsilon_p = \hbar^2 k_z^2/2m = \hbar^2 p^2 \pi^2/2mL_z^2$, are larger than the thermal energy k_BT [29]. Alternatively, L_z must be on the order of the Fermi wavelength λ_F [37]. The Fermi wavelength is the de Broglie wavelength of an electron at the Fermi energy (the highest occupied electron energy level at absolute zero).

A rough approximation of the energy levels of an electron in a quantum well, can be made by first assuming a parabolic dispersion relation for the bulk semiconductor, the effective mass approximation. For an electron in the conduction band of the bulk semiconductor with electron effective mass m^* and conduction band minimum E_c, the energy levels are given by [29]

$$E(k_x, k_y, k_z) \approx E_c + \frac{\hbar^2(k_x^2 + k_y^2 + k_z^2)}{2m^*}. \tag{2.4}$$

For a quantum well confined in the z direction by length L_z, just as above, the z-component of momentum is quantised with subband index p. The dispersion relationship becomes

$$E_p(k_x, k_y) \approx E_c + p^2\varepsilon_z + \frac{\hbar^2\left(k_x^2 + k_y^2\right)}{2m^*} \tag{2.5}$$

with

$$\varepsilon_z = \frac{\hbar^2\pi^2}{2m^* L_z^2} = \frac{m_0}{m^*}\left(\frac{10\,\mathrm{nm}}{L_z}\right)^2 \times 3.8\,\mathrm{meV} \tag{2.6}$$

where m_0 is the free electron mass and $\hbar^2\pi^2/2m_0(10\,\text{nm})^2 \approx 3.8\,\text{meV}$ [29]. A quantum well of width 10 nm and conduction band electron effective mass equivalent to the free electron mass has subband energy splitting of $\sim 4\,\text{meV}$. The smaller the effective mass of the conduction band, the larger the subband energy splitting and the easier it is to observe quantum phenomena[2]. The smaller the confinement length L_z the larger this subband energy splitting is.

Following a similar procedure one can determine an approximation for the subbands of a nanowire confined in two dimensions and a quantum dot confined in three dimensions. The key take away is that the extent of confinement determines the magnitude of the splitting of the electronic band structure into subbands. These subbands can lead to important quantum phenomena as will be seen in the following sections.

2.2. Quantised conductance

One example of quantum confinement influencing the transport of electrons through a semiconducting nanostructure is quantised conductance. It was first observed in 1988 in a GaAs 2 dimensional electron gas (2DEG) [30, 31] and has since been commonly demonstrated in a structure called a quantum point contact (QPC) [32].

Consider, as in Figure 2.1, a 2DEG in the xy plane with a constriction of width d in the y direction that restricts the flow of electrons in the x direction. This device is known as a QPC. Only electrons travelling in the x direction with specific quantised energies are transmitted through the constriction. All other electrons are reflected. This is analogous to a quantum wire in which the electron can be considered as travelling plane waves in the x direction with quantised subbands in the y direction. The dispersion relation of the QPC is then given by

$$E_n(k_x) = \frac{\hbar k_x^2}{2m^*} + \varepsilon_n \tag{2.7}$$

where ε_n are the energy subbands introduced due to confinement in the y direction with subband index n. Each electron state can be given a transmission coefficient $T_n(E)$ depending on whether it passes through the constriction at a given energy ($T = 1$) or is

[2] Since the effective mass is inversely proportional to the curvature of the conduction bands parabolic minimum, this is analogous to a steeper curvature of the conduction band minimum giving a larger energy splitting.

reflected ($T = 0$), assuming no scattering. There will only be a finite number of states that are transmitted through the constriction.

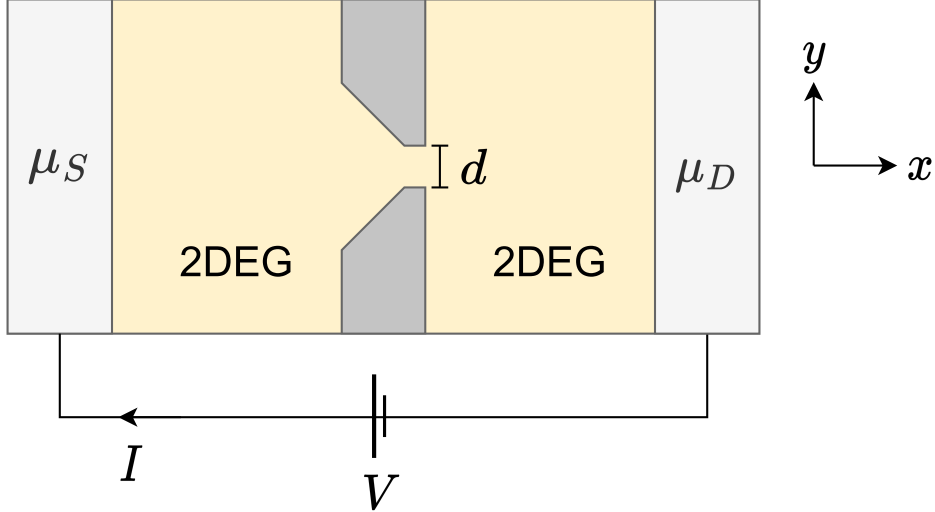


Figure 2.1. A schematic of a quantum point contact (QPC). A 2DEG between two reservoirs (light grey) with a gate defined constriction of width d in the middle. Applying a positive gate to the dark grey region depletes the 2DEG in these regions so that electrons can only flow through the constriction. A voltage bias V can be applied to the reservoirs to produce a current I.

If there is no external voltage bias applied the states of positive and negative momentum are equally occupied. Applying a small voltage bias $V = (\mu_S - \mu_D)/e$, where $\mu_{S,D}$ are the chemical potentials of the source and the drain reservoirs, creates an imbalance around the Fermi level E_F and a current begins to flow from source to drain (assuming $\mu_S < \mu_D$). The current passing through the constriction can be written as [32]

$$I = e \sum_n \int_{\mu_D}^{\mu_S} \frac{1}{2} \rho_n(E) v_n(E) T_n(E) dE \tag{2.8}$$

where $v_n = (dE_n/dk_x)/\hbar$ is the effective electron velocity and $\rho_n = \frac{2}{\pi}(dE_n/dk_x)^{-1}$ is the 1D density of states. Using the definition of the voltage bias above and assuming for small bias ($eV \ll E_F$) that $T_n(E)$ is constant, this leads to the expression for the conductance

$$G = \frac{I}{V} = \frac{2e^2}{h} \sum_n T_n(E). \tag{2.9}$$

Every subband below the Fermi level contributes exactly one conductance quantum $G_0 = 2e^2/h$ to the overall conductance. The factor of 2 comes from the spin degeneracy

of the subbands. If the Fermi level in the channel is altered using an external gate, there is a stepwise increase of the conductance by G_0 every time the Fermi level crosses a new subband. This is an example of discrete subbands, created by quantum confinement, governing the quantum transport of electrons.

2.3. Single electron tunnelling and Coulomb blockade

Another example of quantum transport through semiconductor nanostructures is single electron tunnelling through states on a zero dimensional charge island. The single electron charging theory, including single electron tunnelling and Coulomb blockade, was first introduced by Van Houten and Beenakker to describe conductance oscillations in silicon quantum dot systems [33, 34]. Experiments on Coulomb blockade and single electron transport through Si/SiO_2 and Si/SiGe quantum dots dates back to the early 1990s, shortly after the discovery of Coulomb blockade [38, 39]. In Chapter 3, single electron tunnelling and Coulomb blockade will be demonstrated with the electron bound states of dopant atoms in a silicon nanostructure.

2.3.1. Coulomb blockade

From Section 2.1, when a bulk material is confined in all three dimensions, to the order of the Fermi wavelength λ_F, a quantum dot is formed. The quantum dot is small enough in size that a finite number of electrons reside on it. Due to confinement the dot also has discrete energy states that the electrons can reside in, similar to the electron states of an atom.

Consider now a quantum dot coupled to two electron reservoirs, referred to as 'source' and 'drain', via tunnel junctions with capacitances C_S and C_D, respectively, and capacitively coupled to a gate V_G with capacitance C_G, as in Figure 2.2. This configuration is often called the single electron transistor. The number of electrons on the quantum dot is discrete, with total charge given by $Q = Ne$ where N is an integer and e is the charge of the electron. The phenomena of Coulomb blockade is caused by the Coulomb repulsion between the electrons on the dot blocking any flow of electrons from the leads to the dot [35]. For an electron to be added to the dot the Coulomb interaction of the electrons on the dot must be overcome. The charging energy required for an electron to be added to the dot is given by $\epsilon_C = e^2/2C$ where $C = C_G + C_S + C_D$

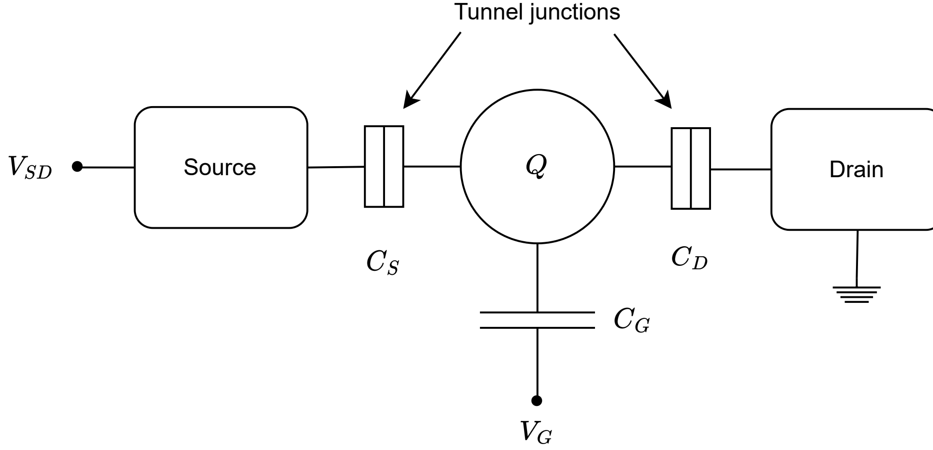


Figure 2.2. A schematic of the single electron transistor circuit. A quantum dot with a total isolated charge of Q coupled to two electron reservoirs, 'source' and 'drain', via tunnel junctions with capacitance C_S and C_D, respectively. A voltage gate V_G is also capacitively coupled to the quantum dot with capacitance C_G. A voltage bias is applied across the transistor by applying a voltage V_D at the drain and grounding the source.

is the total capacitance between the dot and the rest of the system [35, 37, 38, 40]. An electron will only flow from the lead to the dot if the charging energy ϵ_C is overcome, otherwise the system is in Coulomb blockade.

In addition to the charging energy, a semiconducting quantum dot has discrete energy levels spaced by the energy splitting ΔE. From Section 2.1.2, a dot with length L in all dimensions has a level splitting of $\Delta E \propto \frac{\hbar^2}{m^* L^2}$ where m^* is the electron effective mass of the material. Coulomb blockade adds a distinction between one and two electrons occupying an energy level [35]. The separation of the energy levels is increased from just ΔE to an addition energy of $\Delta E + e/C$ which will be proven in the next section.[3]

Coulomb blockade can only be observed under specific conditions. The first condition is that the dot's coupling to the electrodes must be weak compared to ϵ_C. The coupling Γ is used to define the tunnelling rate $\Gamma/\hbar$ and determines the width of the energy levels on the dot. Hence, to resolve the quantised charge states of the dot the condition $\Gamma \ll \epsilon_C$ must be satisfied [35]. This condition also ensures that the isolated charge on the dot has a large enough lifetime to be considered fixed and avoids effects such as co-tunnelling (virtual tunnelling of electrons through the barriers) which are detrimental to the orthodox physics of single-electron charging [35, 37, 41]. The second condition is

[3] For a metallic dot the splitting of the quantised energy levels ΔE is negligible and only the charging energy needs to be considered.

that the temperature must be low, $k_BT \ll \epsilon_C$, as temperature also determines the width of the energy states [35, 37, 42]. The same conditions apply for resolving the discrete energy levels of the dot but with the energy scale ΔE: $\Gamma \ll \Delta E$ and $k_BT \ll \Delta E$. If the coupling to the electrodes, Γ, is too strong or the temperature too high the electrons in the leads would easily overcome the charging energy and Coulomb blockade would not be observed. To ensure these conditions are satisfied, ϵ_C must be large which requires the capacitance of the dot C to be small and hence the dot must be small since $C \propto R$, the radius of the dot [35]. The dimensions of the dot must be on the order of 100 nm for observation of Coulomb blockade at temperatures less than 4 K and 1 nm–5 nm for observation near room temperature [41, 42].

Constant interaction model

For a clearer understanding of Coulomb blockade and single electron transport it is useful to think about it in terms of a simple model, the constant interaction model [33, 34, 35, 43]. This model makes two assumptions. The first is that interactions between electrons on the dot and between electrons on the dot and the leads are parameterised by a single constant capacitance C, the total capacitance of the system. The second is that the discrete energy levels are independent of these interactions and hence independent of the number of electrons on the dot N. In the Coulomb blockade regime, the Coulomb interaction contributes to the energy of the electrons on the dot. From basic electrostatics of the single electron transistor setup, Figure 2.2, the charge on the island can be written as

$$Q = CV_{QD} - C_SV_S - C_DV_D - C_GV_G \tag{2.10}$$

where V_{QD} is the potential of the quantum dot and C is the total capacitance. By rearranging this equation, one can see that the potential of the quantum dot is made up of a potential from the charge already on the dot and and an external potential from the leads

$$V_{QD} = \frac{Q}{C} + V_{\text{ext}} \tag{2.11}$$

where

$$V_{\text{ext}} = (C_SV_S + C_DV_D + C_GV_G)/C. \tag{2.12}$$

Now taking the reference energy of the system to be the configuration where the charge Q and all potentials are zero. The total electrostatic energy required to place N electrons on the dot is given by

$$U(N) = \frac{(Ne)^2}{2C} - NeV_{\text{ext}} + \sum_{n=1}^{N} E_n \tag{2.13}$$

where E_n are the single particle energy levels due to quantum confinement [35]. These are the states that the electrons on the dot can occupy which are independent of the number of electrons on the dot. Adding or removing an electron changes the Coulomb energy which is taken into account by the first term. The second term is the energy contribution of the external potential from the source, drain and gate.

From statistical mechanics, the transport of electrons is driven by the difference in chemical potential [35, 43]. Hence, it is useful to determine the chemical potential of the dot for comparison with that of the source and drain. The chemical potential in the independent particle picture is identical to the difference between the total energy of a system with $N-1$ particles and one with N particles, the single particle energy. The chemical potential of the quantum dot with N electrons is [35, 43]

$$\mu(N) = U(N) - U(N-1) = \left(N - \frac{1}{2}\right)\frac{e^2}{C} - eV_{\text{ext}} + E_N. \tag{2.14}$$

The notable feature of this chemical potential is that it can be tuned by the gate voltage V_G, with the source and drain voltage held constant. The gate voltage is related to the chemical potential by the gate coupling factor $\alpha = C_G/C$. The chemical potentials for different numbers of electrons on the dot N create a 'ladder' of levels separated by an addition energy

$$E_{\text{add}} = \mu(N+1) - \mu(N) = \frac{e^2}{C} + E_{N+1} - E_N \tag{2.15}$$

where the last two terms make up the energy level splitting introduced earlier, $\Delta E = E_{N+1} - E_N$. In semiconductors, this level splitting is not necessarily constant [40]. For large semiconducting dots this level splitting becomes negligible and the addition energy is just the first term due to the Coulomb interaction. With this picture of a ladder of chemical potential levels corresponding to the number of electrons on the quantum dot the transport of the electrons across the dot can be considered.

For comparison with the rest of this thesis, it may be helpful to consider the Hamiltonian that describes the electrons on the quantum dot. From the second quantisation formulation and the arguments above the Hamiltonian describing N electrons on the dot is

$$H_{\text{dot}} = \sum_{n}^{N} E_n \hat{c}_n^\dagger \hat{c}_n + \frac{e^2 \hat{N}^2}{2C} - \hat{N} e V_{\text{ext}} \tag{2.16}$$

where $\hat{c}_n^\dagger$ ($\hat{c}_n$) is the creation (annihilation) operator of an electron and $\hat{N} = \sum_n \hat{c}_n^\dagger \hat{c}_n$ is the number operator. The Hamiltonian acting on an electron state will reproduce energies similar to that of equation (2.13). The first term is due to the occupied quantised energy levels of the dot, similar to that of an atom. Compared to the QPC or nanowire there is no free dimension so the wave vector is completely quantised. The second term is the charging energy due to the Coulomb repulsion of electrons on the dot and the last term is the electrostatic energy due to the combined external potential of the source/drain electrodes and the gate.

2.3.2. Single electron tunnelling

The transport of electrons through the dot is dependent on the position of the dot's chemical potential with respect to the chemical potential of the source μ_S and the drain μ_D. Consider the ladder diagrams in Figure 2.3 of the single electron transistor setup from Figure 2.2. A bias applied across the source and drain, $V_{SD} = (\mu_S - \mu_D)/e$, creates a difference in the source and drain chemical potentials called the bias window. The chemical potential ladder of the quantum dot can be tuned by the gate voltage as described above. Electrons travel from higher chemical potential to lower chemical potential. Hence, the tunnelling of an electron from the source to the dot and then from the dot to the drain – *sequential electron tunnelling* – will only occur if one or more of the chemical potentials of the dot is within the bias window. If there are no chemical potential levels of the dot between the bias window as in Figure 2.3a, no electrons can flow to the dot (the charge on the dot remains constant) and current is suppressed. If the dominant energy scale is the charging energy, ϵ_C, this is called Coulomb blockade. The Coulomb blockade can be lifted by increasing the gate voltage until a chemical potential level $\mu(N)$ is within the bias window as in Figure 2.3b. In this state a current flows and the charge on the quantum dot fluctuates between N and $N - 1$.

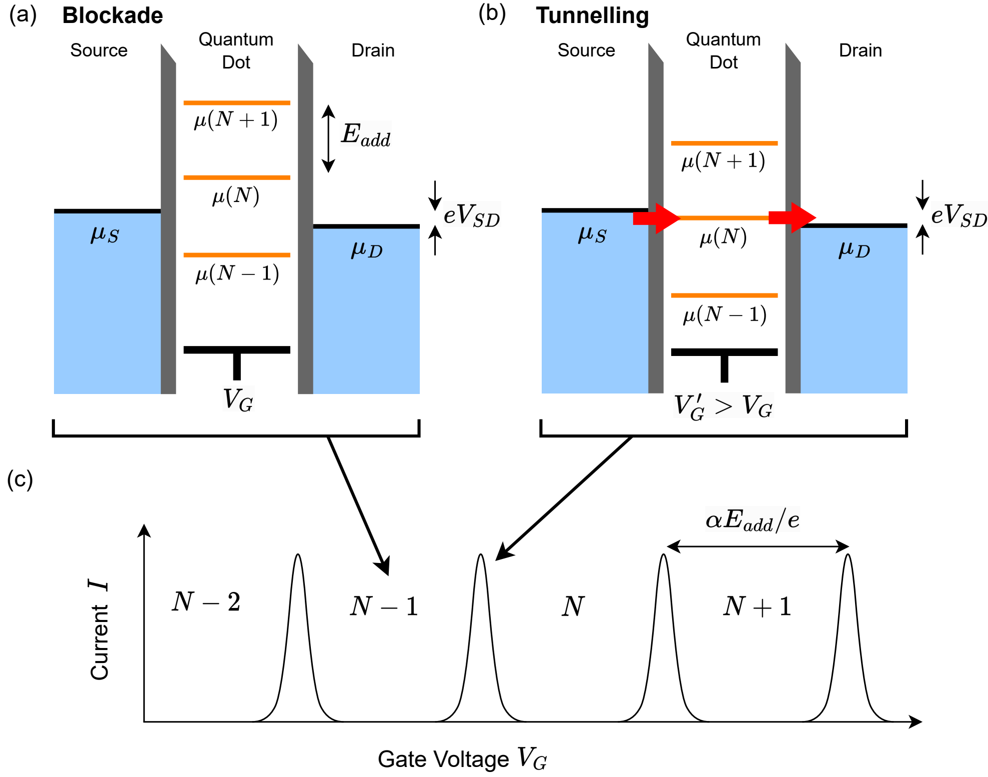


Figure 2.3. Single electron transport. (a) Ladder diagram of Coulomb blockade. The current flow is blocked by the Coulombic repulsion in E_{add} and the number of charges on the dot is constant, $N-1$. (b) Ladder diagram of single electron tunnelling. The gate voltage applied is increased so that a chemical potential level is within the bias window. This allows a single electron to tunnel from source to the dot and then from the dot to drain. (c) The Coulomb peaks that form in a current versus gate voltage plot. The peaks occur due to the tunnelling in (b) and the current suppression occurs due to the blockade (a). Each suppression corresponds to an integer number of electrons on the dot.

For a constant small bias, Coulomb blockade can be observed as Coulomb peaks in a current versus gate voltage plot as in Figure 2.3c. The regions of current suppression are due to Coulomb blockade, when the chemical potential states are outside the bias region, and correspond to a finite number of charges on the quantum dot. The Coulomb peaks occur when a single chemical potential level is within the bias window and the sequential tunnelling of electrons through the quantum dot creates a current. The spacing of the Coulomb peaks is proportional to the addition energy such that $e\Delta V_G = \alpha E_{\text{add}}$. Note that the spacing between all peaks is approximately the same in the regime where $\epsilon_C \gg \Delta E$, while spacing may vary if ΔE is dominant [35]. This dramatic sensitivity of the current to changes at the single electron level gives this system its name, the single electron transistor.

Now consider varying the bias voltage with the gate voltage held constant. Let the gate voltage be set such that the system is in Coulomb blockade, Figure 2.3a If the bias voltage is increased, the bias window increases and eventually a chemical potential level enters the window and current flows. The sign of the current depends on the sign of the bias applied[4]. As the bias window increases further a second chemical potential level may enter the bias window allowing the charge on the dot to fluctuate between three integer numbers of electrons ($N-1$, N and $N+1$) which allows for a double electron tunnelling current [43].

Plotting the current as a function of the bias voltage and gate voltage produces the stability diagram depicted in Figure 2.4. The white diamonds are Coulomb blockade regions corresponding to a constant number of electrons on the dot. These are called *Coulomb diamonds*. The height of the Coulomb diamonds measured from zero bias corresponds to the addition energy, E_{add}, because this is the largest voltage bias that can be applied before one of the dot's chemical potential levels enters the bias window. The largest width of the diamonds on the gate voltage axis is the same as for the Coulomb peaks. The coloured regions of finite current (orange: positive current and blue: negative current) correspond to the sequential tunnelling of electrons when at least one chemical potential level is between the source drain bias. The current increases stepwise as more chemical potential levels of the dot enter the bias window, not shown in this diagram.

The Coulomb peaks from Figure 2.3c appear along the centre of the stability diagram for almost zero small bias and are often referred to as degenerate points. In the stability diagram, each degenerate point is the intersection of two straight resonance lines separating the regions of current suppression and finite current. These lines correspond to the chemical potential of either the source or drain being aligned with a chemical potential level of the dot. Assuming the drain is grounded as in Figure 2.2, from the chemical potential of the dot (2.14), an expression for each of these resonance lines can be derived. For the dot chemical potential aligned with the source chemical potential [35]

$$V_{SD} = \kappa(V_G - V_C) \tag{2.17}$$

with $\kappa = C_G/(C_G + C_D)$ and $V_C = (N - 1/2)e/C_G + CE_N/(eC_G)$ the chemical potential on the dot without the external potential. For the chemical potential aligned with the

[4] If the bias is negative $\mu_D > \mu_S$, so the electrons will flow to the left in the ladder diagrams producing negative current.

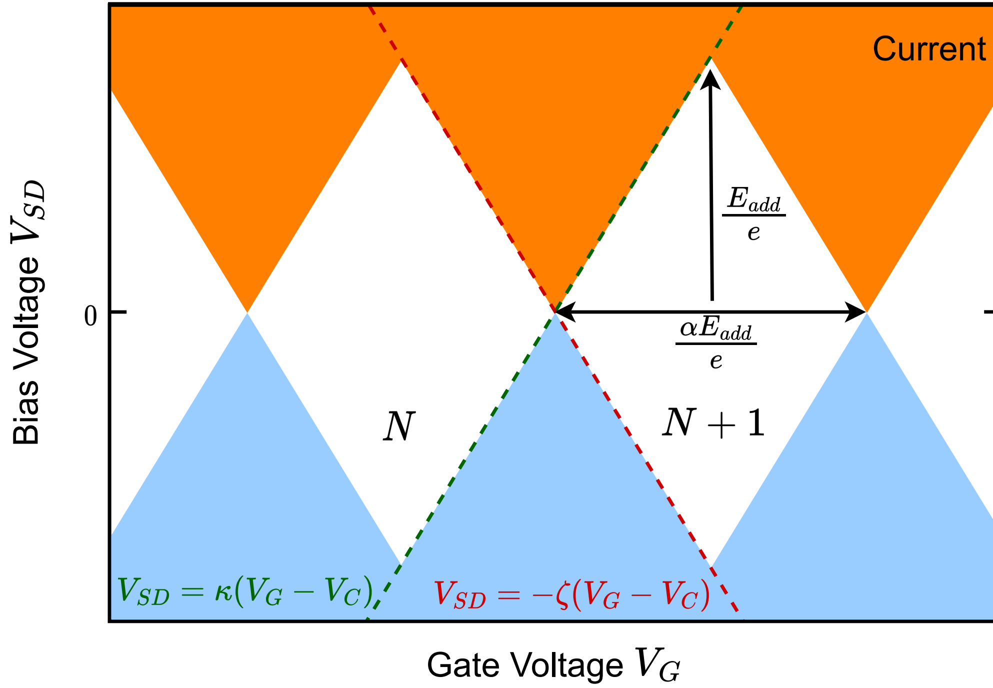


Figure 2.4. The current stability diagram. Current plotted against the bias and gate voltage. The orange indicates a region of positive current, the blue indicates a region of negative current and the white a region of zero current or Coulomb diamonds. At small bias, current only flows in the Coulomb peaks (or degeneracy points) as in Figure 2.3c. The green (red) dashed resonance line corresponds to when the source (drain) aligns with the chemical potential level $\mu(N+1)$ on the dot.

drain

$$V_{SD} = -\zeta(V_G - V_C) \tag{2.18}$$

with $\zeta = C_G/C_S$. The slopes of these lines relate to the gate coupling parameter by $1/\alpha = 1/\kappa + 1/\zeta$. Hence, the stability diagram can be used to determine the capacitances of the system as well as the addition energy as depicted in Figure 2.4. This technique is used in the analysis of Chapter 3.

The single electron transistor is the fundamental single electron device for demonstrating Coulomb blockade and single electron tunnelling. Since its discovery, smaller and more efficient versions of the device have been investigated. One of the popular designs under investigation is the single atom transistor whose island is made from a single donor atom in a semiconductor instead of a quantum dot.

2.3.3. Single atom transistor

The single atom transistor follows the single electron transistor setup but instead of confining electrons in a quantum dot the electrons are bound to the radial Coulomb potential of a donor atom in a semiconductor. Single electron tunnelling and Coulomb blockade through the electron bound states of a donor atom in a silicon nanostructure was demonstrated in the works [44, 45, 46]. Since then there have been many other manifestations of the single atom transistor [19, 47, 48].

The most common donors in silicon are phosphorous, P, and arsenic, As. Silicon is a valence four element and P and As are both valence five elements. When a P or As atom replaces a Si atom in the silicon lattice there is a left over free electron that is not used in the covalent bounds to silicon atoms. The single dopant atom creates a steep radially symmetric electrostatic potential which attracts the free electron. The mean radius of the orbit that an electron takes around a nucleus atom in its ground state is given by the effective Bohr radius $a_B = a_0 \epsilon_r m_0 / m^*$ where ϵ_r and m^* are the static dielectric constant and effective mass of the material, and $a_0 = \frac{4\pi\epsilon_0 \hbar^2}{m_0 e^2} \approx 0.053\,\mathrm{nm}$ is the Bohr radius of a hydrogen atom in a vacuum. For a phosphorus donor in silicon the effective Bohr radius is approximately 2.5 nm [49] and for an arsenic donor in silicon it is closer to 3 nm [50].

The dopant atom in bulk silicon has two charge states: the ionised state D^+ with no electrons bound to the donor atom and the neutral state D^0 with one electron bound to the dopant atom. However, the work of [44, 45], observed that an As dopant in a silicon nanostructure capacitively coupled to source, drain and gate electrodes allows for another bound state, the negative state D^- with two electrons bound to the dopant atom.

In Chapter 3, the low temperature measurement setup described in Appendices B and C is used to reproduce the Coulomb blockade of these dopant bound states with similar devices to that used in [44, 45]. This experiment was performed as a preliminary test of the low temperature measurement setup and also to demonstrate Coulomb blockade, a fundamental form of quantum transport, in a semiconducting nanostructure.

2.4. Spin-orbit interaction

In this section, the spin-orbit interaction is introduced. First in the more familiar atomic physics framework and then in terms of the semiconductor band structure. Chapter 5 of this thesis models the effective g-factor anisotropy and spin-orbit interactions in InAs nanowires for various atomistic and electromagnetic conditions. As such, the common spin-orbit interactions present in III-V semiconductors and their nanostructures will be introduced. Lastly, some applications of InAs nanowires with large spin-orbit interaction will be discussed, in particular the search for Majorana zero modes.

2.4.1. Atomic physics origins

The spin-orbit interaction couples the orbital momentum of the electron with the orientation of its intrinsic spin. It originates from relativistic corrections to the electron energy levels of the hydrogen atom in atomic physics. The spin-orbit interaction can be thought of as an effective Zeeman splitting due to the apparent magnetic field, felt by an electron in its rest frame, acting on its spin magnetic dipole moment.

An electron moving in the electric field $\mathbf{E}$ produced by an atom's nucleus with velocity $\mathbf{v}$, experiences an apparent magnetic field in its rest frame

$$\mathbf{B} = -\frac{1}{c^2}\mathbf{v} \times \mathbf{E} \tag{2.19}$$

where c is the speed of light and the non-relativistic limit is assumed, $\frac{1}{\sqrt{1-(v/c)^2}} \approx 1$. It is also well known that an electron in a magnetic field possesses a Zeeman energy $-\boldsymbol{\mu}_S \cdot \mathbf{B}$ where $\boldsymbol{\mu}_S = -\frac{1}{2}g_S\mu_B\boldsymbol{\sigma}$ is the spin magnetic dipole moment of the electron, with $g_S \approx 2$ the electron spin g-factor, μ_B the Bohr magneton, $\hbar$ the reduced Planck's constant and $\boldsymbol{\sigma}$ the vector of Pauli matrices. The apparent magnetic field in the electron's rest frame acts on the spin magnetic dipole momentum of the electron which leads to the spin-orbit interaction term [51]

$$H_{\mathrm{SO}} = -\frac{\hbar}{4m_0^2c^2}\boldsymbol{\sigma} \cdot (\mathbf{p} \times \boldsymbol{\nabla} V) \tag{2.20}$$

where m_0 is the rest mass of the electron, $\mathbf{p}$ is the electron's momentum and V is the Coulomb potential of the nucleus atom acting on the electron such that $\mathbf{E} = \frac{1}{e}\boldsymbol{\nabla} V$ (central potential approximation [52]). The extra factor of 1/2 comes from the Thomas correction that takes into account the precession of the electron's spin [53]. This spin-orbit

interaction term, commonly called the Pauli spin-orbit term, can also be derived by taking a non-relativistic approximation of Dirac's equation.

2.4.2. Derivation from Dirac's equation

This section presents a common derivation of the spin-orbit interaction from Dirac's equation which can be found in [51, 52]. The derivation demonstrates that the spin-orbit interaction is a relativistic correction to the single electron Hamiltonian.

In relativistic quantum mechanics, fermions are described by Dirac's equation (with electromagnetic coupling)

$$\gamma^{\mu}\left(i\hbar\partial_{\mu} + eA_{\mu}\right)\psi = m_0 c\psi \tag{2.21}$$

where $A^{\mu} = (V/c, \mathbf{A})$ is the gauge field with electric potential V and vector potential $\mathbf{A}$, $\partial_{\mu} = (\frac{1}{c}\frac{\partial}{\partial t}, \boldsymbol{\nabla})$ is the covariant four-derivative, ψ is a four-component spinor, e and m_0 are the charge and mass of the electron respectively and γ^{μ} are the 4×4 gamma matrices

$$\gamma^0 = \begin{pmatrix} I_{2\times 2} & 0 \\ 0 & -I_{2\times 2} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \tag{2.22}$$

with the Pauli matrices σ_i for $i = x,\ y,\ z$ which are given in Appendix A.3.

Assuming that A_{μ} is time independent, let the time dependence of ψ be

$$\psi = \psi(\mathbf{x}, t=0)e^{-iEt/\hbar} \tag{2.23}$$

where ψ is an eigenfunction of $\frac{1}{c}\frac{\partial}{\partial t}$ with eigenvalue E. The spatial part of the four component spinor can be separated into upper and lower two component spinors $\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix}$. The Dirac equation can then be written as two coupled equations for ψ_A and ψ_B

$$\left(E + eV - m_0c^2\right)\psi_A = c(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\psi_B \tag{2.24}$$

$$\left(E + eV + m_0c^2\right)\psi_B = c(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\psi_A \tag{2.25}$$

where $\boldsymbol{\pi} = \mathbf{p} + e\mathbf{A} = -i\hbar\boldsymbol{\nabla} + e\mathbf{A}$.

The lower component ψ_B can be eliminated by substituting equation (2.25) into equation (2.24)

$$\left(E + eV - m_0c^2\right)\psi_A = (\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\,\frac{c^2}{E + eV + m_0c^2}\,(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\,\psi_A. \tag{2.26}$$

No approximations have been made up until this point. Now in the non-relativistic limit, assume that $E \approx m_0c^2$ and $eV \ll m_0c^2$. Define the non-relativistic energy $E^{(\mathrm{NR})} = E - m_0c^2$ as the energy difference from the electron rest mass. As a result, the quantity $\left(E^{(\mathrm{NR})} + eV\right)/2m_0c^2 \approx (\mathbf{p} + e\mathbf{A})^2/2m_0c^2 \approx (v/c)^2 \ll 1$ and the following power expansion can be performed up to first order in $(v/c)^2$

$$\frac{c^2}{E + eV + m_0c^2} = \frac{1}{2m_0}\left[\frac{2m_0c^2}{2m_0c^2 + E^{(\mathrm{NR})} + eV}\right] \approx \frac{1}{2m_0}\left[1 - \frac{E^{(\mathrm{NR})} + eV}{2m_0c^2} + \ldots\right].$$

Keeping terms up to first order in $(v/c)^2$ equation (2.26) can be written

$$E^{(\mathrm{NR})}\psi_A = \left\{(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\,\frac{1}{2m_0}\left[1 - \frac{E^{(\mathrm{NR})} + eV}{2m_0c^2}\right](\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) - eV\right\}\psi_A. \tag{2.27}$$

This expression may look like a form of the Schrödinger equation but this is not the case for two reasons. Firstly, this is not an eigenvalue problem because $E^{(\mathrm{NR})}$ is present on both sides of the equation. Secondly, ψ_A is not normalised to unity because the normalisation condition of the Dirac equation states that

$$\int \psi^\dagger\psi d^3x = \int (\psi_A^\dagger\psi_A + \psi_B^\dagger\psi_B)d^3x = 1, \tag{2.28}$$

and in the non-relativistic limit, from equation (2.25) the lower component ψ_B is given by

$$\psi_B \approx \frac{(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})}{2m_0c}\psi_A. \tag{2.29}$$

The coefficient of ψ_A is of the order (v/c) so in the non-relativistic limit the lower spinor ψ_B is negligible and the spin components ψ_A and ψ_B are called the *large* and *small* spinor components of ψ respectively.

Using equation (2.29) the normalisation condition equation (2.28) can be written to first order in $(v/c)^2$ as

$$\int \psi_A^\dagger \left(1 + \frac{p^2 + e\hbar\boldsymbol{\sigma}\cdot\boldsymbol{B}}{4m_0c^2}\right)\psi_A d^3x \approx 1 \tag{2.30}$$

this comes about by using the identity (A.12) and noting that $\mathbf{p}\times\mathbf{A} = -i\hbar(\boldsymbol{\nabla}\times\mathbf{A})+\mathbf{A}\times\mathbf{p}$ (where $\mathbf{p}$ acts on everything to the right and $\boldsymbol{\nabla}$ only acts on $\mathbf{A}$). The $\boldsymbol{\pi}^2$ is reduced to $\mathbf{p}^2$ by assuming for simplicity that $\mathbf{A} = 0$.

From equation (2.30), it is clear that the normalisation to unity condition can be satisfied by introducing a new two component spinor

$$\tilde{\psi} = \Omega\psi_A \quad \text{where} \quad \Omega = \left(1 + \frac{p^2 + e\hbar\boldsymbol{\sigma}\cdot\mathbf{B}}{8m_0^2c^2}\right) \tag{2.31}$$

so that to first order in $(v/c)^2$

$$\int \tilde{\psi}^\dagger\tilde{\psi}d^3x \approx \int \psi_A^\dagger\left(1 + \frac{p^2 + e\hbar\,(\boldsymbol{\sigma}\cdot\mathbf{B})}{4m_0^2c^2}\right)\psi_A\, d^3x \approx 1,$$

making use of the fact that Ω is Hermitian and that $(1+x)^2 \approx 1+2x$ for $x \ll 1$. Substituting this new spinor $\tilde{\psi}$ into equation (2.27) and multiplying from the left by $\Omega^{-1} \approx \left(1 - \frac{p^2+e\hbar\boldsymbol{\sigma}\cdot\mathbf{B}}{8m_0^2c^2}\right)$, after some rearranging, to first order in $(v/c)^2$ the equation becomes

$$\left[\frac{p^2}{2m_0} - eV + \frac{e\hbar}{2m_0}(\boldsymbol{\sigma}\cdot\mathbf{B}) - \frac{e\hbar}{4m_0^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{p}\times\mathbf{E}) + \frac{e\hbar^2}{8m_0^2c^2}(\boldsymbol{\nabla}\cdot\mathbf{E}) \right.$$
$$\left. -\frac{p^4}{8m_0^3c^2} - \frac{e\hbar p^2}{4m_0^3c^2}\boldsymbol{\sigma}\cdot\mathbf{B} - \frac{(e\hbar B)^2}{8m_0^3c^2}\right]\tilde{\psi} = E^{(\mathrm{NR})}\tilde{\psi} \tag{2.32}$$

where $\mathbf{E} = -\boldsymbol{\nabla}V$. This is now a true eigenvalue problem that can be treated as the Schrödinger equation for the two component spinor $\tilde{\psi}$. Most of the terms fall out after some simple algebra and using similar steps to that for deriving equation (2.30) [51, 52]. The derivation of the fourth and fifth terms requires writing $E^{(\mathrm{NR})}p^2 = \frac{1}{2}\{p^2, E^{(\mathrm{NR})}\}$ and using the identity $\{A^2, B\} - 2ABA = [A, [A, B]]$ where $[.,.]$ ($\{.,.\}$) stands for the commutator (anti-commutator) [52]. The third term on the left of equation (2.32) is the Zeeman term. The fourth term is the Pauli spin-orbit term which matches equation (2.20) by substituting $\mathbf{E} = \frac{1}{e}\boldsymbol{\nabla}V$. The fifth term is the Darwin term and the terms on the second line are higher order corrections to the kinetic energy and Zeeman term.

2.4.3. Spin-orbit interaction in semiconductors

Similar to the atomic case, an electron in a semiconductor crystal lattice can also experience a spin-orbit interaction. In the semiconductor, the potential V in equation (2.20) is no longer the atomic potential but instead it is a combination of the periodic potential of the ion-core lattice, any external potentials and disturbances of the periodic potential produced by defects and phonons.

Consider the energy bands that an electron can occupy in a semiconductor, $E_{n,s}(\mathbf{k})$ where $\mathbf{k}$ is the wavevector, n is the band index and $s =\downarrow, \uparrow$ is the spin index. Note that s is an index and not the orientation of the spin up or down because the spin polarisation can vary as a function of $\mathbf{k}$ [51]. When the orbital momentum and spin orientation are independent and there are no applied magnetic fields, the bands are spin degenerate

$$E_{n\uparrow}(\mathbf{k}) = E_{n\downarrow}(\mathbf{k}). \tag{2.33}$$

This spin degeneracy is the result of conserved spatial inversion symmetry and time reversal symmetry [51]. Spatial inversion flips the sign of the wavevector $\mathbf{k} \to -\mathbf{k}$ while time reversal flips the sign of the wavevector and the spin

$$\text{Spatial inversion symmetry}: \quad E_{n\uparrow}(\mathbf{k}) = E_{n\uparrow}(-\mathbf{k}) \tag{2.34}$$

$$\text{Time reversal symmetry}: \quad E_{n\uparrow}(\mathbf{k}) = E_{n\downarrow}(-\mathbf{k}) \tag{2.35}$$

The degeneracy present in time reversal symmetric systems is also known as Kramer's degeneracy. The combination of these two symmetries gives us the spin degeneracy in equation (2.33). The spin degeneracy can be lifted by breaking either spatial inversion symmetry or time reversal symmetry.

Zeeman splitting

The time reversal symmetry can be broken by applying a magnetic field $\mathbf{B}$ which introduces the Zeeman interaction term in the single electron Hamiltonian (2.2)

$$H_Z = \frac{1}{2} g^* \mu_B \mathbf{B} \cdot \boldsymbol{\sigma} \tag{2.36}$$

where g^* is the material dependent effective g-factor of the electron. This Zeeman term lifts Kramer's degeneracy and splits the spin bands by a Zeeman energy $E_Z = g^* \mu_B |\mathbf{B}|$.

The spins in semiconductor structures are not *pure* spins but rather they are a mixture of the electron's intrinsic spin and orbital angular momentums due to the spin-orbit interaction. As a result, the g-factor in the Zeeman term is no longer that of the pure electron spin, $g_0 = 2.0023$ but rather an effective g-factor g^* [54].

Spatial inversion asymmetry and spin-orbit interaction

Analogous to the Zeeman term breaking time reversal symmetry, the spatial inversion symmetry must be broken by introducing a term into the single electron Hamiltonian. This term must satisfy time reversal symmetry but break spatial inversion symmetry. Hence, it should be odd in both $\mathbf{k}$ and $\boldsymbol{\sigma}$. This term is the spin-orbit interaction and it takes the form

$$H_{\mathrm{SO}} = \mathbf{B}_{\mathrm{SO}}(\mathbf{k}) \cdot \boldsymbol{\sigma} \tag{2.37}$$

where $\mathbf{B}_{\mathrm{SO}}(\mathbf{k})$ is a function odd in $\mathbf{k}$ that acts like an effective magnetic field in the electron's rest frame due to the non-zero electric field caused by a spatial inversion asymmetry [54]. Examples of different spatial inversion asymmetries are described below.

There are various spatial inversion asymmetries that can occur in semiconductors which induce different spin-orbit interaction terms. The spatial inversion asymmetries of interest are bulk inversion asymmetry (BIA) and structural inversion asymmetry (SIA) [54]. BIA occurs in bulk zincblende III-V semiconductors due to the lack of an inversion centre in the crystal structure. SIA occurs due to an electric field, either external or built-in. Explicit forms of the spin-orbit interaction induced by these asymmetries can be derived from $\boldsymbol{k} \cdot \mathbf{p}$ theory [51].

The spin-orbit interaction that occurs due to BIA is called the Dresselhaus spin-orbit interaction [55]. If x, y, z are chosen to lie along the [100], [010], [001] crystallographic axes respectively, the Dresselhaus term for the lowest conduction subband in a bulk III-V semiconductor is [43, 54]

$$H_D^{\mathrm{3D}} = \gamma_D \{[k_x(k_y^2 - k_z^2)]\sigma_x + [k_y(k_z^2 - k_x^2)]\sigma_y + [k_z(k_x^2 - k_y^2)]\sigma_z\} \tag{2.38}$$

where γ_D is the Dresselhaus coefficient which is material dependent, $\mathbf{k} = -i\boldsymbol{\nabla}$ is the momentum operator and σ_i for $i = x, y, z$ are the Pauli matrices defined in Appendix A.3.

The SIA due to the total electric field $\mathbf{E}$ consisting of external gates, built-in electrostatic potentials and asymmetric interfaces induces a Rashba spin-orbit interaction [56]. For the coordinates as above the bulk Rashba term can be written [43, 54]

$$\begin{aligned} H_R^{3\mathrm{D}} &= -\alpha_0(\mathbf{E}\times\mathbf{k})\cdot\boldsymbol{\sigma} \\ &= -\alpha_0\left[(E_y k_z - E_z k_y)\sigma_x + (E_z k_x - E_x k_z)\sigma_y + (E_x k_y - E_y k_x)\sigma_z\right] \end{aligned} \tag{2.39}$$

where α_0 is the material dependent Rashba coefficient, $\mathbf{E} = -\boldsymbol{\nabla}V$ where V is the total electric potential of the SIA. The strength of the Rashba and Dresselhaus constants can be predicted using tight-binding or $\mathbf{k}\cdot\mathbf{p}$ theory. Typically, semiconductors with heavier elements, like InAs and InSb, have a stronger spin-orbit interaction [51].

Effective g-factor anisotropy

The spin-orbit interaction is responsible for electrons in the semiconductor having an effective g-factor, g^*, different to the free electron g-factor, g_0. Spin-orbit interactions and the confinement in III-V semiconductor nanostructures can result in an anisotropy of this effective g-factor in i.e. different values of g^* for different orientations of the external magnetic field $\mathbf{B}$ [51, 54]. This is best understood by considering the spin-orbit interaction as an effective magnetic field $\mathbf{B}_{\mathrm{SO}}$ as in equation (2.37). When an external magnetic field $\mathbf{B}$ is applied there are two competing fields, $\mathbf{B}$ and $\mathbf{B}_{\mathrm{SO}}$, acting on spins and depending on their orientation with respect to each other and the quantisation of $\mathbf{k}$ due to confinement, the spin splitting and hence the value of g^* is different. The effective g-factor anisotropy of InAs nanowires will be explored thoroughly in Chapter 5.

Spin-orbit interaction in a quantum well

In this thesis, the spin-orbit interaction occurring in semiconductor nanostructures is of particular interest. Naturally, the first nanostructure to consider is the quantum well which is created by quantum confinement of a semiconductor in one dimension. The effective spin-orbit interaction terms of a quantum well will be derived from the bulk spin-orbit interaction terms equations (2.38) and (2.39) using the envelope function approximation [43, 51, 54]. In Chapter 5, these terms will be modified again to model the spin-orbit interaction in a 1D nanowire by confining a second dimension.

The Dresselhaus term depends on the crystallographic axes chosen for x, y, z and the confinement direction. To obtain the Dresselhaus spin-orbit interaction for a quantum well confined in the [001] direction, start with the bulk Dresselhaus spin-orbit interaction equation (2.38) and average over the lowest conduction subband envelope function in the growth direction [001]. From the infinite potential well problem, assume that the envelope function of the lowest subband wavefunction, ψ_{1z}, is given by equation (A.5). As a result, $\langle k_z \rangle = \langle \psi_{1z} | k_z | \psi_{1z} \rangle = 0$ and $\langle k_z^2 \rangle \propto (\pi/L_z)^2$ is a nanostructure dependent constant. Generally, $\langle k_z^2 \rangle \gg k_x^2, k_y^2$ due to strong confinement in z, so only the linear in k_x and k_y terms are kept. The resulting Dresselhaus term for a [001] quantum well is then [43, 54]

$$H_D^{\mathrm{2D}} = \beta_D (k_x \sigma_x - k_y \sigma_y) \tag{2.40}$$

where $\beta_D = -\gamma_D \langle k_z^2 \rangle$. If the axes are along different crystallographic directions a different bulk Dresselhaus term would be derived for the given axes and the same envelop function approximation method can be used to derive quantum well terms.

In a quantum well grown along [001], the external electric field due to a gate potential or the substrate is typically in the z direction. The Rashba term equation (2.39) then simplifies to [43, 54]

$$H_R^{\mathrm{2D}} = -\alpha_R (k_x \sigma_y - k_y \sigma_x) \tag{2.41}$$

where $\alpha_R = \alpha_0 \langle E_z \rangle$ and $\langle E_z \rangle$ is the electric field averaged over the lowest subband in the growth direction [001].

2.4.4. Spin-orbit interactions in nanowires

The spin-orbit interaction in III-V semiconducting nanowires has become an increasingly popular area of research since the turn of the century. The spin-orbit interaction in nanowires has shown promise for various applications such as spin transistors [25], spin qubits [23], and – arguably the main driving force – Majorana fermion research for topological qubits [14, 15, 16, 17].

Majorana fermions, fermions that are their own anti-particle, were first theoretically predicted to exist in high energy particle physics by Ettore Majorana in 1937 [57]. In recent years, condensed matter physicists have discovered that these Majorana fermions

may be realised as charge neutral excitations of electrons and holes at the edges of a topological superconducting state, known as Majorana Zero Modes (MZMs). A topological superconducting state is one whose band structure cannot be deformed to that of an insulator (or more generally the vacuum) without closing the energy band gap [58]. Interestingly, this topological superconducting state can be formed in a III-V semiconductor nanowire when it is interfaced with an s-wave superconductor as in Figure 2.5 [15, 59, 60]. The MZMs are theoretically predicted to occur at the ends of a semiconductor nanowire in this topological state. If one has a network of nanowires, with each nanowire hosting MZMs at its ends, it is possible store information non-locally in the MZMs and exchange the information according to non-abelian statistics [14, 61]. The non-local nature of MZMs protects the information against corruption by the external environment, resulting in a topologically protected system [17]. These topologically protected systems are expected to revolutionise the world of quantum computing. However, there is still significant controversy surrounding the experimental observation of MZMs.

One of the key ingredients for creating this topological superconducting state is the large spin-orbit coupling in the III-V nanowire, typically InAs. When combined with the application of a magnetic field, the spin-orbit interaction in the nanowire can create helical states [62, 63] – a pair of states who have both opposite momentum and opposite spin – which can be paired by the superconducting proximity effect to create a topological superconducting state [17]. The caveat is that the application of a magnetic field can suppress the superconductivity of the system. The work around is to ensure that the semiconductor nanowire has a large g-factor which allows for a large Zeeman splitting with only a small applied magnetic field [15, 17]. The large spin-orbit interaction and the large effective g-factor present in InAs nanowires is what makes them an attractive structure for realising MZMs and manufacturing spin qubits or transistors. The work presented in Chapter 5, aims to model the band structure of InAs nanowires, namely the spin splitting of the lowest conduction subband, for different atomistic conditions and electromagnetic conditions to explore the possible spin-orbit interactions that can be induced in a III-V nanowire.

The observation of MZMs and control of spin qubits in these nanowires typically requires quantum transport measurements through the nanowire in a setup similar to that of the single electron transistors in Figure 2.2, with the nanowire taking the place of the dot. The Coulomb blockade measurements discussed in Section 2.3.1 can identify the spectrum of energy states in the nanowire [22, 23] and conductance measurements

can be performed to transport electrons through these states and measure excitations such as MZMs [15, 62, 64, 65, 66]. There was intention to perform quantum transport experiments on InAs nanowires for this thesis, with the low temperature measurement setup described in Appendix B, and compare results with the modelling of the InAs nanowire in Chapter 5. However, due to uncontrollable circumstances these experiments were not completed within candidature.

Rashba and Dresselhaus spin-orbit interaction

The common Rashba and Dresselhaus spin-orbit interaction terms for the nanowire can be derived from the spin-orbit interaction terms of the quantum well. Starting from equations (2.40) and (2.41), confine in the y direction to create a nanowire along x and apply the envelop function approximation. The resulting nanowire Dresselhaus and Rashba spin-orbit interaction are [54]

$$H_D^{1D} = \beta_D k_x \sigma_x, \tag{2.42}$$

$$H_R^{1D} = -\alpha_R k_x \sigma_y. \tag{2.43}$$

In the $k_x = 0$ case, the k_y terms of (2.40) and (2.41) become important, however, for non-zero k_x these terms are sufficient. The Rashba spin-orbit interaction is the term used in the standard derivation of the helical gap and topological superconducting state that hosts MZMs. These models typically ignore Dresselhaus spin-orbit interaction because they assume the nanowire growth direction of [111], which is a high-symmetry direction assumed to have negligible Dresselhaus spin-orbit interaction [51, 67].

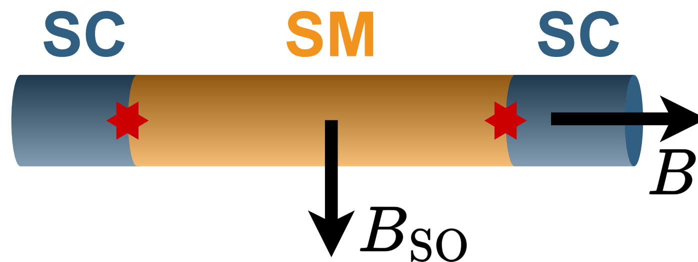


Figure 2.5. The 1D system for creating Majorana zero modes. The semiconductor (SM) nanowire with a Rashba spin-orbit field, B_{SO}, is sandwiched between two s-wave superconductors with a magnetic field B applied along the wire axis. The Majorana zero modes (red stars) occur at the boundary of SM and SC when the SM is in a topologically superconducting state.

Majorana zero modes in 1D nanowires

The most studied architecture in Majorana zero mode research is the semiconductor-superconductor nanowire heterostructure [15, 59, 60]. The system consists of a III-V semiconductor nanowire with a large spin-orbit field $\mathbf{B}_{\text{SO}}$, usually InAs or InSb, in proximity with s-wave superconducting leads and a magnetic field $\mathbf{B}$ applied along the nanowire, as depicted in Figure 2.5. The typical Hamiltonian that is used to describe this system if the nanowire is lying along the x axis and the spin-orbit field is perpendicular to the nanowire axis is [17, 59, 60]

$$\begin{aligned} H &= H_{SM} + H_{SC}, \\ H_{SM} &= \int dx \psi^\dagger \left(-\frac{\partial_x^2}{2m^*} - \mu + i\alpha_R \sigma_y \partial_x + E_Z \sigma_x \right) \psi, \\ H_{SC} &= \int dx \Delta (\psi_\uparrow \psi_\downarrow + \text{h.c.}) \end{aligned} \tag{2.44}$$

where $\psi_\sigma^\dagger$ creates an electron with effective mass m^*, chemical potential μ and spin σ in the nanowire. α_R is the strength of the Rashba spin-orbit interaction that aligns the spin in the y-direction due to the breaking of SIA in the z direction and $E_Z = g^* \mu_B B_x$ is the Zeeman energy due to the applied magnetic field. H_{SM} describes the semiconductor nanowire with Rashba spin-orbit interaction in the limit where the confinement of the nanowire is strong enough that only the lowest transverse subband is relevant. H_{SC} is the superconducting proximity effect of the superconducting leads on the nanowire with Δ, the proximity induced pairing potential in the nanowire. The nature of the MZMs requires charge neutral, spinless excitations. The charge neutrality is induced by the Cooper pairing of electrons and holes in the superconductor and the spinless nature of the excitation is created by the helical gap.

The Helical Gap

Focusing on H_{SM} first, the corresponding dispersion relation is

$$E_\pm(k) = \frac{k^2}{2m^*} - \mu \pm \sqrt{(\alpha_R k)^2 + E_Z^2}. \tag{2.45}$$

with k the wave vector along the nanowire. If $\alpha_R = 0$ and E_Z is non-zero the magnetic field lifts the spin degeneracy of the lowest subband by the Zeeman energy E_Z. If $E_Z = 0$ and α_R is non-zero, the spin-orbit coupling lifts the spin degeneracy of the lowest subband by shifting the spin bands in k. This creates two shifted parabolas, one for each spin,

represented in Figure 2.6a by the red and blue parabolas. If a magnetic field is then applied so that both α_R and E_Z are non-zero an anti-crossing forms between the two shifted parabolas at $k = 0$ as seen in Figure 2.6b. This is called a helical gap [63, 68].

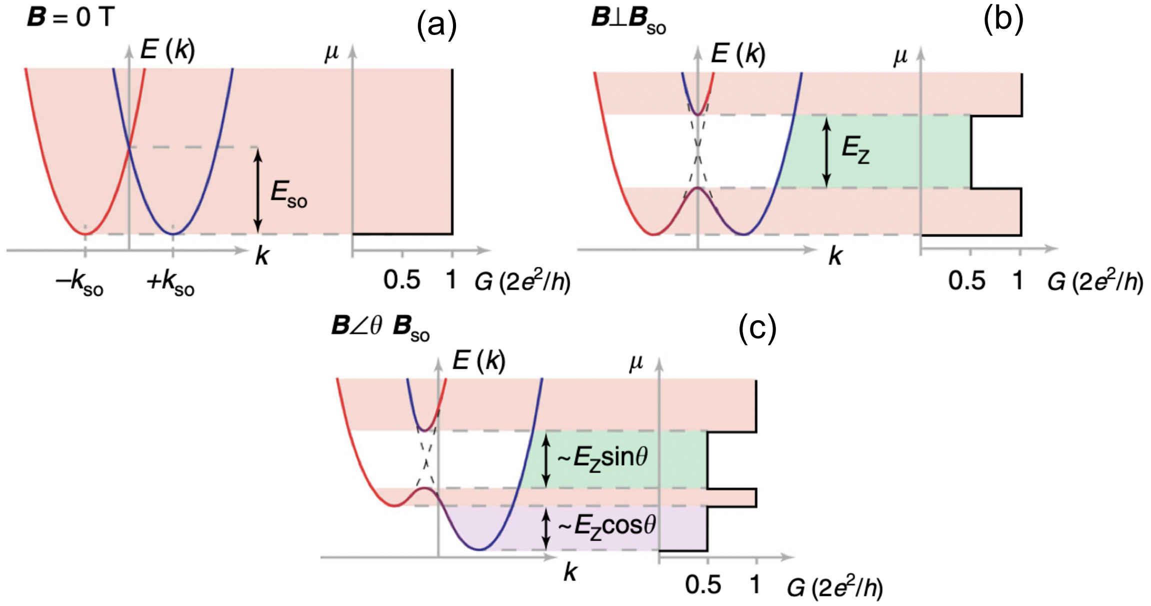


Figure 2.6. Left: The dispersion relation of the lowest subband spin states in the presence of Rashba spin-orbit field, $\mathbf{B}_{\mathrm{SO}}$, and different magnetic field, $\mathbf{B}$, orientations. Right: The corresponding conductance output. (a) With no external magnetic field the spins are momentum shifted by $k_{\mathrm{SO}} = m^*\alpha_R/\hbar^2$ so that opposite spins have opposite momentum and the states cross at $E_{\mathrm{SO}} = m^*\alpha_R^2/2\hbar^2$. (b) In the presence of the non-zero magnetic field $\mathbf{B}$ perpendicular to $\mathbf{B}_{\mathrm{SO}}$, the spin bands hybridise and a helical gap of size E_Z opens (green). The conductance drops to $G_0/2$ in the gap. (c) When the magnetic field is orientated at an angle θ to $\mathbf{B}_{\mathrm{SO}}$ the size of the helical gap is proportional to the component of $\mathbf{B}$ perpendicular to $\mathbf{B}_{\mathrm{SO}}$ and a Zeeman energy shift (purple) is introduced due to the component of $\mathbf{B}$ parallel to $\mathbf{B}_{\mathrm{SO}}$. Figures were adapted from the work of Kammhuber et al. [62] which is licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

The origin of the helical gap can be understood physically as follows. Consider the spin-orbit field, $\mathbf{B}_{\mathrm{SO}} = i\alpha_R\partial_x\hat{y}$, polarising the electron spins so that opposite spins have opposite momentum. Applying a magnetic field, $\mathbf{B}$, causes the spins to precess about the component of the magnetic field perpendicular to the spin-orbit field. The precession (spin-mixing) is maximised if $\mathbf{B}$ is completely perpendicular to $\mathbf{B}_{\mathrm{SO}}$ and minimised if the two are parallel. Similarly, the size of the helical gap is maximised when the applied magnetic field $\mathbf{B}$ is perpendicular to the spin-orbit field $\mathbf{B}_{\mathrm{SO}}$ and minimised when they are parallel [62, 63, 68]. The dispersion of the spin states for different orientations of $\mathbf{B}$ with respected to $\mathbf{B}_{\mathrm{SO}}$ are depicted in Figure 2.6 along with quantised conductance plots

for each dispersion. When μ is within the helical gap the electron spin-states are two-fold degenerate corresponding to half a conductance quantum, $G_0 = 2e/h$. This can be seen as one 'spinless' band due to the helical edge states having opposite spins at opposite k values. Above and below this gap there is a four-fold degeneracy corresponding to a full conductance quantum.

Topological criterion

In the spinless regime, ignoring the upper unoccupied band which is allowed if $\Delta \ll E_Z$, when superconducting proximity Δ is introduced weakly, the opposite spin and momentum carriers are p-wave paired. At this point the wire enters a topological superconducting state. From equation (2.44) one can derive the topological criterion [17]

$$E_Z > \sqrt{\Delta^2 + \mu^2}. \tag{2.46}$$

When this condition is satisfied the nanowire is in a topological state. If the system changes to the point where $E_Z = \sqrt{\Delta^2 + \mu^2}$ the gap closes and a topological phase transition occurs. When $E_Z < \sqrt{\Delta^2 + \mu^2}$, the gap is dominated by pairing and the nanowire is in a topological trivial state. If the nanowire is in its topologically non-trivial phase it naturally creates a boundary with a trivial state, the bulk s-wave superconductor, at its ends. These boundaries can host a pair of MZMs, one at each end of the nanowire. The non-local nature of the MZMs means that they are topologically protected from environmental effects and ideal for creating topological quantum gates [14, 17].

Conditions for experimental realisation of MZMs

Although, there have been various experiments looking for MZMs [64, 66, 69, 70], no experiment has conclusive evidence of an observed Majorana bound state. This is mainly due to disorder that is introduced into the device by the current limitations of fabrication [71, 72]. Disorder can be pair-breaking and destroy the topological gap and any MZMs along with it [71]. Part of the work of Chapter 5 is to determine if the spin-orbit interaction of InAs nanowires can be optimised to protect against this disorder.

The topological state's sensitivity to disorder is dependent on the extent of time-reversal symmetry breaking in the semiconductor [73]. The BCS theory of superconductivity [74] requires the pairing of particles which are the time reverse of each other. The breaking of time reversal symmetry reduces the pairing of the superconductor and

hence the width of the energy gap. The smaller the energy gap, the more susceptible it is to disorder in the material. The extent of time-reversal symmetry breaking can be quantified by the ratio $R = E_Z/E_{\mathrm{SO}}$ where E_Z is the Zeeman splitting as in Section 2.4.3 and $E_{\mathrm{SO}} = \frac{\hbar^2 k_{\mathrm{SO}}}{2m^*} = \frac{m^*\alpha^2}{2\hbar}$ is the spin-orbit energy [17, 63, 72]. If R is large, time reversal symmetry breaking is strong and the topological gap is more susceptible to disorder suppression. On the other hand, if R is small time-reversal symmetry is only weakly broken and the disorder suppression of the topological gap is reduced. For one to achieve $R < 1$, it is ideal to either have control over the strength of spin-orbit coupling real-time with gating or possess the knowledge to manufacture III-V semiconducting nanowires with a desired spin-orbit field magnitude and orientation.

A secondary problem results from the fact that the Zeeman field must be large enough to meet the topological criterion (2.46). Unfortunately, a large magnetic field can also suppress superconductivity in the bulk superconductor and reduce the superconducting pairing in the semiconductor. The suppression of superconductivity is typically minimised by ensuring that the effective g-factor of the semiconducting nanowire is significantly larger than that of the superconductor [17]. The anisotropic effective g-factor in III-V nanowires is dependent on the spin-orbit interaction present. As such knowing the g-factor anisotropy of a nanowire and orienting the magnetic field along the maximum of this anisotropy is important for minimising superconductivity suppression.

In chapter 5, an atomistic 20 band tight-binding model is used to determine the effective g-factor anisotropy and spin-orbit interactions present in zincblende InAs nanowire. The multi-band atomistic model is expected to give a more comprehensive understanding of the spin-orbit interaction and effective g-factor anisotropy in InAs nanowires. These simulations will provide insight into how atomistic features such as surface-ion termination and nanowire cross-section, as well as applied electric fields, influence the g-factor anisotropy and the spin-orbit interactions in InAs nanowires. In particular, this work will be useful for designing nanowires with maximised spin-orbit coupling and effective g-factor for MZMs, as well as spintronic devices and spin-orbit qubits.

Chapter 3.

Investigating quantum transport through dopants in silicon

In this chapter, single electron tunnelling and Coulomb blockade is demonstrated through a few dopants in a silicon nanostructure. The experiment both demonstrates said form of quantum transport and validates our low temperature measurement setup[1]. The original plan was to use this setup to perform quantum transport measurements on InAs nanowires and compare results with the theoretical work of Chapter 5. However, due to delays experienced by our collaborators in the fabrication of the nanowires, the experiments could not be performed during candidature.

Dopant atoms are an important ingredient in current semiconducting technology. They are typically a passive element implanted in semiconductors, like silicon, to increase the n or p carrier concentration in regions of the semiconductor. With the reduction in transistor size it was found that a single dopant atom could act as the Coulomb island of a single electron transistor, a device now known as the single atom transistor. The original discovery was made by Sellier et al. in 2006 [44], where a Coulomb blockade spectrum was observed due to the electron bound states of an isolated As dopant in the silicon channel of an n-p-n Fin Field Effect Transistor (FinFET). Since then the single donor in silicon architecture has been employed to build single atom transistors from intentionally implanted dopants [47, 48] and to design high fidelity donor in silicon qubits [19, 20, 21]. At low temperatures, the transport of electrons through these bound donor states can be observed experimentally as Coulomb peaks and Coulomb diamonds. This Coulomb blockade measurement is a standard technique used to map the spectrum of states in quantum devices.

[1]The measurement setup is outlined in Appendices B and C

The experiment performed in this chapter, aims to reproduce the dopant in silicon Coulomb blockade measurements performed by Sellier et al. [44] and investigated further in [45, 75]. Two silicon n-p-n FinFETs similar to those used by Sellier et al. [44, 45, 46, 75], were cooled to cryogenic temperatures for current-voltage measurements. The results and subsequent analysis indicate single electron tunnelling and Coulomb blockade in the devices, due to the electron bound states of arsenic dopants, in agreement with the work of Sellier et al.

3.1. Preliminaries

This section gives a brief description of the design and conventional operation of the devices being tested. It also details the preparation of the devices for measurement and the results of room temperature testing. If the devices operate conventionally at room temperature, they are good candidates for low temperature measurements.

3.1.1. The devices

The n-p-n FinFETs are transistors whose current flow is dictated by the field effect of a three-sided gate capacitively coupled to a tall, narrow, semiconducting channel. The semiconducting channel on top of the substrate resembles a 'fin,' giving the device its name. This architecture was introduced to succeed the planar Metal-Oxide-Silicon Field Effect Transistor (MOSFET) which experiences undesirable short channel effects when reduced below 32 nm node technology [76]. Interestingly, these short-channel effects are what allow us to observe the Coulomb blockade of dopants in the FinFET.

The in-depth fabrication process of these wires is presented in previous works [44, 45]. The FinFET consists of a p-doped channel with an n-doped silicon wire deposited over the channel to form the gate, see Figure 3.1. The n-doped source and drain are separated from the gate by nitride spacers as depicted in the bottom panes of Figure 3.1a and b. The doping configuration is depicted in Figure 3.1b. The height of the channel is $H = 60\,\text{nm}$ while its width W and length L can vary: $W = 35\,\text{nm}$ to $1\,\mu\text{m}$ and $L = 50\,\text{nm}$ to $1\,\mu\text{m}$. The dimensions of the channel influence the capacitive coupling to the electrodes. The shorter the length of the channel the more likely it is to find an isolated dopant in the channel [75].

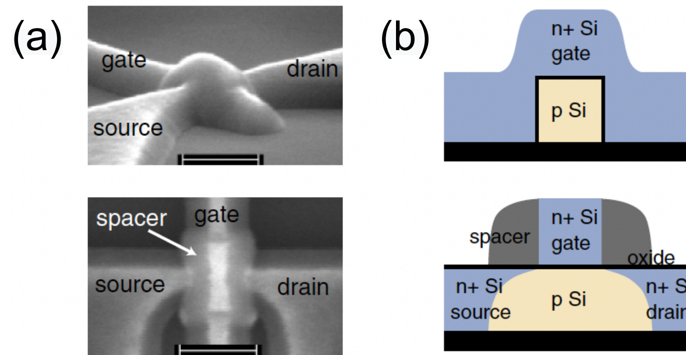


Figure 3.1. (a) Scanning electron micrographs of the FinFET with a 200 nm scale bar. (b) Cross-sections of the FinFET showing the n- and p-type regions. The top is a cut along the gate and the bottom a cut along the silicon fin. Reprinted (a) and (b) with permission from [44] copyright 2006 by the American Physical Society.

3.1.2. Device preparation

The FinFET devices are distributed on chips each holding approximately 80 devices. Each chip is glued in the centre of a male 16 pin dual-inline chip holder with gold plates along the inside edge of the chip holder. The gold plates of the chip holder are then connected to the source, drain and gate pads of FinFETs on the chip via aluminium wire bonds. Figure 3.2 is an image of Chip C2 in the chip holder and placed in the Variable Temperature Insert (VTI) sample area with two devices wire bonded. The center blow-up is an optical microscope image of the chip holder's gold plates wire bonded to the pads of the FinFETs on chip. Two sets of empty FinFET pads are pictured in the right blow-up. The two devices measured in this cooldown are labelled FinFET1 and FinFET2. Table 3.1 lists the chip holder pin numbers corresponding to the source, drain and gate pads of each FinFET on Chip C2. It is ideal to have multiple working devices on the chip to be cooled. This maximises data collection per cooldown and ensures there are backup devices to measure if a device or its wire bonds break while the chip is in the cryostat.

Table 3.1. The pin numbers corresponding to the source, drain and gate pads of Chip C2 devices, FinFET1 and FinFET2.

Device	Source	Drain	Gate
FinFET1	1	2	16
FinFET2	3	4	15

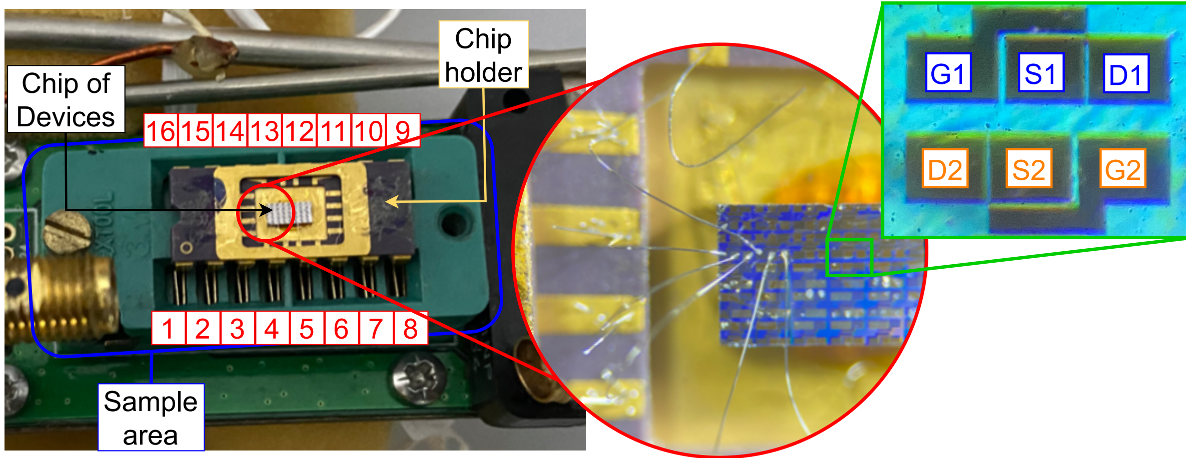


Figure 3.2. Chip C2 in its 16 pin dual-inline chip holder which is locked into the VTI sample area. The pin numbers are labelled in red. The center blow-up is an optical microscope image of the wire bonds between device pads and the gold plates of the chip holder, refer to Table 3.1. The right blow-up is a zoomed optical microscope image of two sets of FinFET pads on the chip which have not been wire bonded. The source, drain and gate are labelled as S, D and G respectively.

3.1.3. FinFET operation

This section will outline the conventional room temperature operation of FinFETs. The *n*-*p*-*n* FinFET conventionally operates by tuning the channel resistance with the gate, similar to the MOSFET depicted in Figure 3.3. The difference between the two systems is that the MOSFET is a planar structure with a metal oxide silicon (MOS) capacitor for the gate while the FinFET is a three-dimensional structure with a polysilicon oxide silicon capacitor for its gate. For simplicity, the field effect will be described for the MOSFET but the same principles can be universally applied to other capacitively gated devices [77, 78] such as the FinFET and gated nanowires [23, 79] studied in Chapter 5.

When a positive voltage is applied to the gate it creates a build up of charge at the oxide surface near the channel. The positive potential bends the conduction and valence bands, E_C and E_V, of the semiconductor channel near the oxide surface. This causes holes to populate the valence band near the surface as seen in Figure 3.3a. In the charge picture, this corresponds to the positive oxide charge repelling the positive carriers in the *p*-doped channel just below the oxide. The positive carriers leave behind a layer of immobile, negative acceptor ions called the depletion region as depicted in Figure 3.3b. The charge neutrality of the system is conserved by the negative acceptor ions balancing the positive charge at the oxide surface.

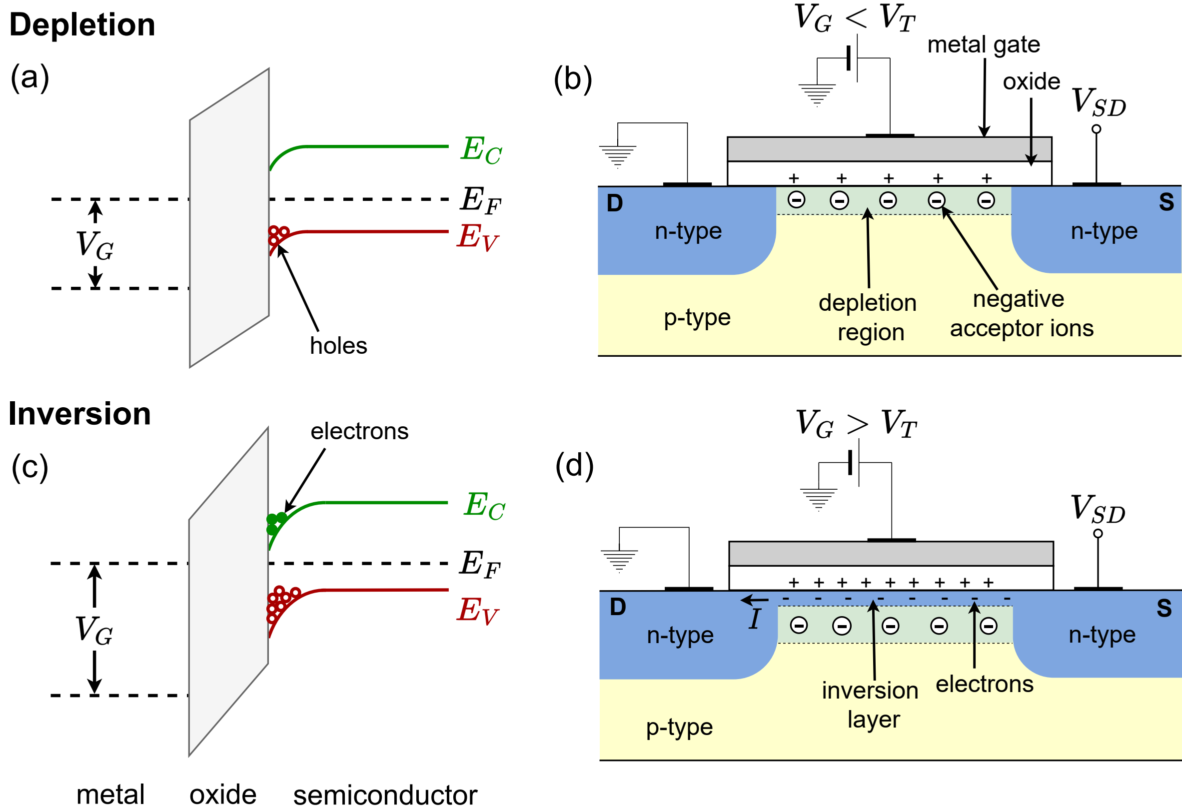


Figure 3.3. The field effect of a MOS capacitor and the operation of a MOSFET. Left: The band diagram of the MOS capacitor. Right: The MOSFET diagram of the carriers. The top diagrams are when the MOSFET is in depletion and the bottom diagrams are when the MOSFET is in inversion.

Increasing the positive gate voltage bends the bands further which at first increases the width of the depletion region or the number of holes in the valence band. If a large enough positive voltage is applied, the conduction band E_C is bent close enough to the Fermi level E_F that it is populated with electrons as shown in Figure 3.3c. The region of the channel near the oxide is populated by a layer of n carriers. This layer is called an inversion layer because the region of the channel near the oxide changes from p-type to n-type. Charge neutrality is now conserved by increasing the number of electrons below the oxide as the gate voltage increases. This inversion layer allows n carriers to travel from the drain to the source and creates a current as depicted in Figure 3.3d. The gate voltage at which the inversion of the channel begins is called the threshold voltage. After this threshold voltage is reached, current will flow in the device.

3.1.4. Room temperature testing

In this section, the devices are tested and a few important features are extracted including the threshold voltage, sub-threshold swing and estimates of the FinFETs channel dimensions. These devices go through many stages of fabrication and preparation in the lab and as a result there are many points at which the devices may become defective. Hence, it is imperative to ensure that the devices work correctly at room temperature before they are cooled.

The basic circuit diagram of the transistor setup is depicted in Figure 3.4 and resembles the single electron transistor in Figure 2.2. The SRS SIM928 modules output DC voltages for the gate V_G and the source-drain bias V_{SD}. The Femto DLPCA200 current amplifier takes the drain current I_D as an input and outputs a gained DC voltage which is measured by the Keithley 2000 multimeter. The temperature of the device is measured by sensing the cryostats' resistor with the Lakeshore 340 temperature controller as described in Appendix C. All instruments are controlled on the lab PC using a Python 3 package called Qcodes [80] and the data is saved in accessible databases. The negative terminals of the SIM928 modules, the current amplifier ground and the cryostat are all connected to the control box chassis which acts as the common ground. The gate is capacitively coupled to the channel with capacitance C_G. The source and drain are coupled to the channel via tunnel junctions with capacitances, C_S and C_D.

FinFET1 and FinFET2 from Chip C2 where tested at $T = 291.3 \pm 0.5\,\mathrm{K}$ by measuring drain current I_D against gate voltage V_G for various source-drain bias V_{SD}. The results are given in Figure 3.5. Both devices displayed the gate action of a field effect transistor. The current is initially zero and turns on after the gate voltage reaches a threshold voltage. The current continues to increase as the inversion layer increases. The sign and magnitude of the current in the on state is dependent on the source-drain bias.

The room temperature threshold voltage, V_T, for each FinFET can be determined by extrapolating from the steepest slope of these current traces to zero drain current [77]. The intercept at zero drain current is the threshold voltage. Figure 3.6a and c are the current traces of FinFET1 and FinFET2, respectively, with linear extrapolations to determine values for V_T. The threshold voltages extracted from the intercepts are $266 \pm 6\,\mathrm{mV}$ and $249 \pm 3\,\mathrm{mV}$ for FinFET1 and FinFET2, respectively.

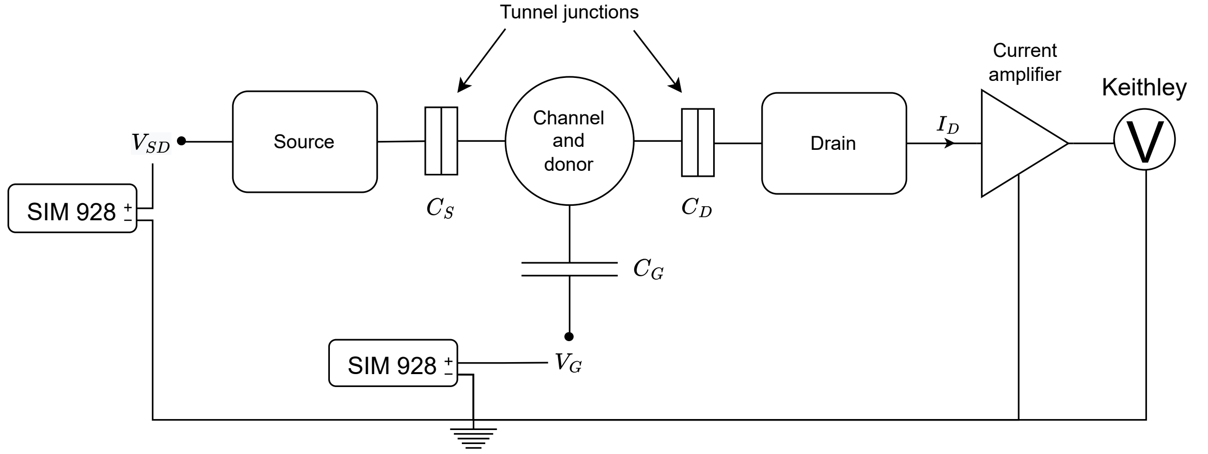


Figure 3.4. Schematic of the measurement setup for transistors including the FinFET devices and single electron transistors. The gate voltage is capacitively coupled to the channel or any donors by a capacitance C_G. The source and drain electrodes are coupled to the channel via tunnel junctions with capacitances C_S and C_D. A source-drain bias V_{SD} is applied at the source with the drain grounded and a drain current I_D is measured at the drain.

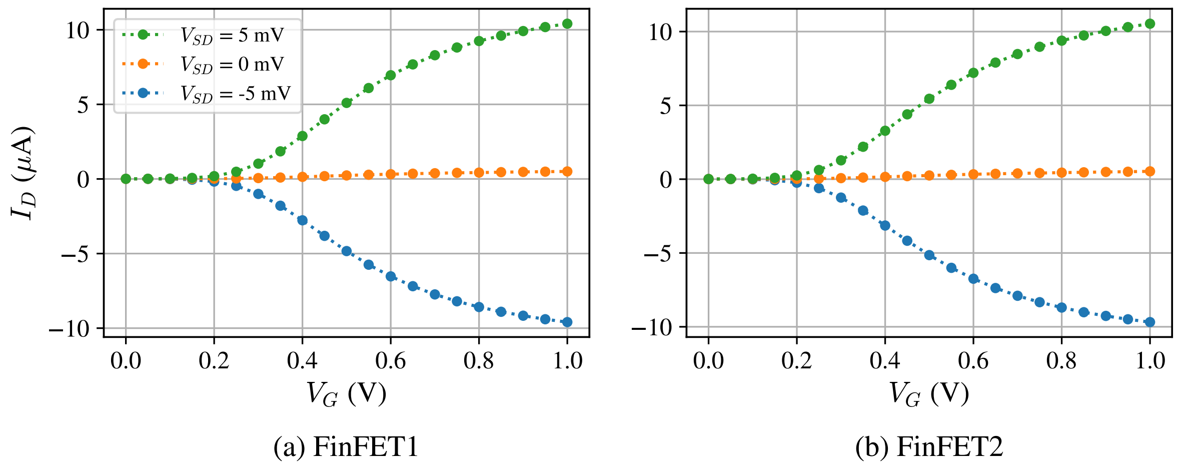


Figure 3.5. The gate action of the FinFETs at room temperature. Drain current I_D versus gate voltage V_G for 5 mV steps of source drain bias V_{SD} of devices on chip C2.

Another common definition of the threshold voltage is the gate voltage at which the drain current takes the value [77]

$$I_D = 0.1\,\mu\text{A} \times \frac{W}{L}, \tag{3.1}$$

this is known as the threshold current.

From equation (3.1), the threshold voltages V_T determined above and the corresponding threshold current I_D determined from a linear interpolation of the FinFET current traces, one can estimate the ratio of channel dimensions, $r_{ch} = W/L$, for each device. For FinFET1 the ratio is $r_{ch} = 6.6$ and for FinFET2 $r_{ch} = 6.1$. These ratios provide insight into the actual dimensions of the channel. The average of these ratios is $\bar{r}_{ch} = 6.4$. A similar batch of FinFET devices used by Sellier et al. [44, 75] have dimensions $L \times W = 60\,\text{nm} \times 385\,\text{nm}$, corresponding to $r_{ch} = 6.4$. The agreement between this value and $\bar{r}_{ch}$ is an indication that the tested devices are good candidates for observing the Coulomb blockade spectrum of isolated dopant atoms.

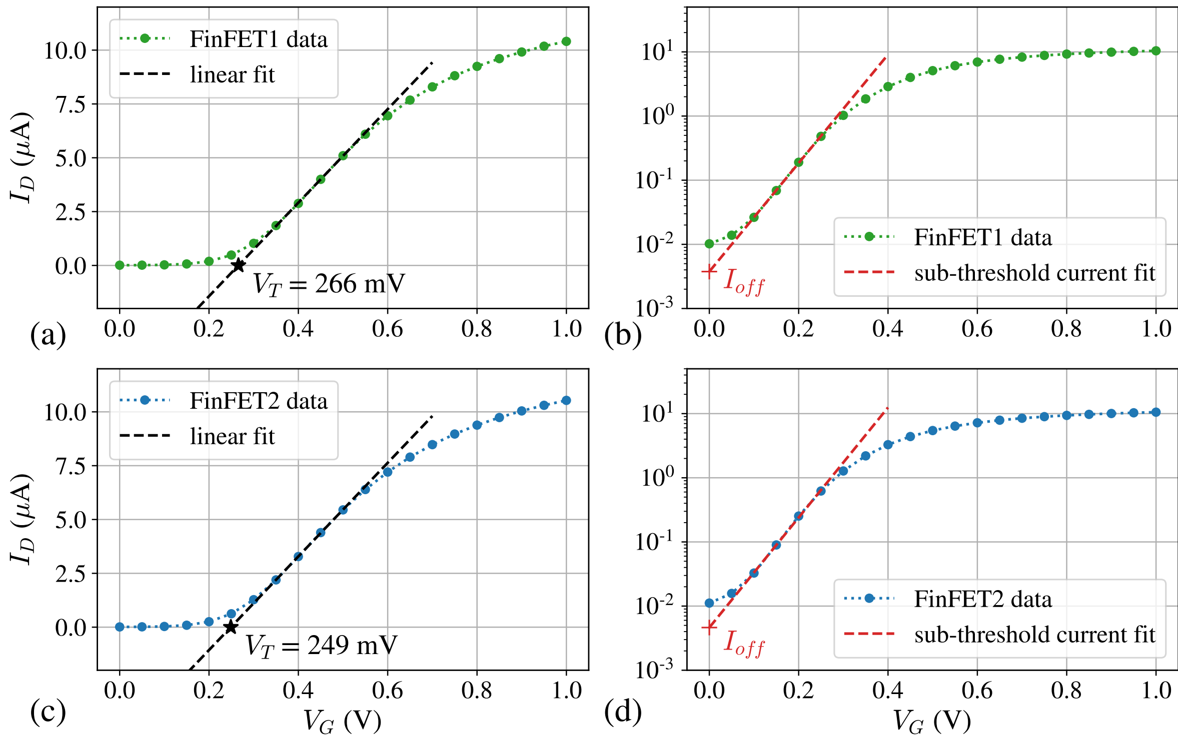


Figure 3.6. Analysis of the FinFETs' room temperature current traces. The gate was swept at a source drain bias of 5 mV. (a) and (c) are the extrapolation of the steepest point in the current trace to zero current (black dashed line) to determine the threshold voltage (black star) for FinFET1 (green) and FinFET2 (blue) respectively. (b) and (d) are linear regression fits (red dashed line) to the sub-threshold region of the current traces for FinFET1 and FinFET2, respectively. The plus symbol marks the off current, I_{off}.

Figure 3.6b and d are the log scale of the drain current against gate voltage for FinFET1 and FinFET2, respectively. The log scale is useful for analysing the sub-threshold characteristics and off current, I_{off}, of the FinFETs. The sub-threshold current is the current output when the gate voltage is below the threshold voltage. The off

current is the current output of the device at zero gate voltage. For both FinFETs the drain current departs from linear near zero gate voltage and approaches 10^{-2} µA. This is due to the floor of the current amplifier output at 10^6 V/A gain with an offset of 10 mV as discussed in Appendix D.2.1. When $V_G < V_T$, the channel is either in depletion or weak inversion. The current in this sub-threshold region is dominated by diffusion and operates similar to a bipolar transistor. The sub-threshold current can be written as [77, 78]

$$I_D = I_{\text{off}} \exp\left(\frac{qV_G}{\eta k_B T}\right) \tag{3.2}$$

where the off current is given by

$$I_{\text{off}} = 100\,\text{nA} \times \frac{W}{L} \exp\left(-\frac{qV_T}{\eta k_B T}\right) \tag{3.3}$$

and $\eta = 1 + \frac{C_{\text{dep}}}{C_{\text{ox}}}$ with C_{dep} the depletion region capacitance per unit area[2] and C_{ox} the oxide capacitance per unit area. T is temperature and k_B is the Boltzmann constant. A common parameter used to describe the sub-threshold region is the sub-threshold swing, SS, which is defined as the gate voltage required to change the drain current by a decade. It can be written as [78]

$$SS = \ln(10)\left(\frac{\eta k_B T}{q}\right). \tag{3.4}$$

The base 10 logarithm of the sub-threshold current (3.2) can be written in terms of this sub-threshold swing and the off current

$$\log(I_D) = \log(I_{\text{off}}) + V_G/SS. \tag{3.5}$$

The sub-threshold swing of the FinFETs can be determined from the slope of a linear regression fit to the linear sub-threshold region in Figure 3.6b and d. The sub-threshold swing for FinFET1 is 118 ± 1 mV/decade and for FinFET2 117 ± 2 mV/decade. These sub-threshold swings can give an estimate of η, 2.05 ± 0.02 and 2.02 ± 0.04, respectively. From the intercepts of the linear regression fits, the off current for FinFET1 is 3.8 ± 0.2 nA and for FinFET2 4.6 ± 0.4 nA. Typically, η is between 1.1 and 1.8 for MOSFETs and should be close to 1.1 for FinFETs [77]. That is the depletion region capacitance C_{dep} is

[2] The depletion region is the layer of negative acceptor charges left by the repelled p carriers as indicated in Figure 3.3. In the MOSFET with a depletion layer of width w and permittivity ϵ_{dep} the depletion capacitance per unit area is $C_{dep} = \epsilon_{dep}/w$.

typically smaller than the oxide capacitance C_{ox}. The values of η are larger than expected as are the values of I_{off}. The main reason for this is likely short channel effects. As the channel length decreases there is a point where the threshold voltage drops significantly because carriers can easily diffuse across the short channel. From equation (3.3), if the threshold voltage is too small ($V_T \ll \eta k_B T/q$) the off current can be large. A shorter channel would reach inversion faster to create current flow, leaving a depletion region of smaller width. If this width is small compared to the width of the oxide (assuming the oxide and depletion region have a similar permittivity) then $C_{\text{dep}} > C_{\text{ox}}$ which would explain a larger η. Another possibility is that the source-drain bias is too small to determine an accurate off current and η. Typically, a 50 mV bias is applied to determine the off current, however, here a small drain bias was applied because our setup was optimised for smaller current measurements.

From equation (3.3) and the values for the off current and sub-threshold swing determined above, another estimate of the channel dimension ratio r_{ch} can be made. Assuming the room temperature threshold voltages V_T determined earlier, for FinFET1 $r_{ch} = 6.7 \pm 0.9$ and for FinFET2 $r_{ch} = 6.3 \pm 0.7$. These values agree with the previously determined ratios of the channel dimensions. In conclusion, FinFET1 and 2 have similar dimensions to the FinFETs tested by Sellier et al. [44, 75] and operate conventionally at room temperature with observed short channel effects in the form of a large off current.

3.2. FinFET operation at low temperatures

In this section, the evolution of the FinFET operation during cooldown is discussed in detail. There are clear changes in the threshold voltage of the devices and near cryogenic temperatures peaks appear due to dopant atoms in the channel and a quantum dot formed by the channel itself.

After identifying multiple working devices from the same chip (C2) it was time to cooldown. With chip C2 in the sample area of the VTI and all the box pins grounded, the cooldown procedure outlined in Appendix C.3.3 was followed to cool the devices. After the VTI was inserted into the dewar of liquid helium, an infinite loop of gate voltage sweeps from 0 V to 0.8 V was initiated at a small source drain bias of 3 mV to observe the evolution of FinFET2's operation as it cooled. The threshold voltage of the FinFET increased as the device was cooled. This is due to the suppression of thermal fluctuations which aid in the transport of electrons through the device at higher temperatures.

Once FinFET2 reached temperatures less than 15 K, peaks began to appear in the current trace. Figure 3.7a compares a high temperature trace at 252.1 K to a trace with clear peaks at 11.1 K. Of particular interest, are the peaks that appear in the sub-threshold region of the current trace clearly plotted on the log scale in Figure 3.7b. These peaks are associated with the electron bound states of donors in the silicon channel [44, 45, 75].

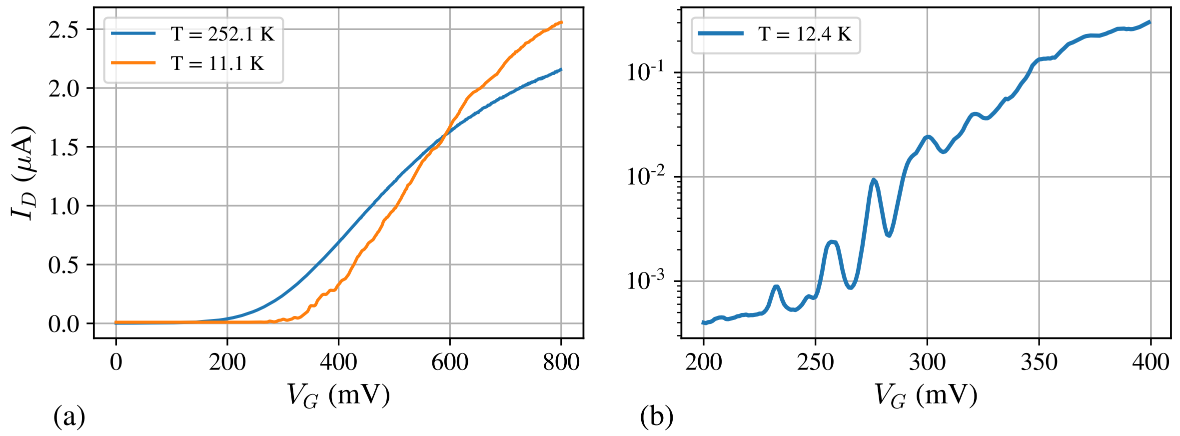


Figure 3.7. Gate voltage V_G sweeps of the drain current I_D of FinFET2 at source-drain bias 1 mV. (a) Compares the sweep at high temperature to the sweep at low temperatures. (b) A zoom in on a gate voltage sweep where the Coulomb peaks first appear with the current on a log scale.

From Section 2.3.1, there are two main requirements for the observation of Coulomb blockade: an isolated island (atom-like quantum dot) with weak capacitive coupling to the source, drain and gate electrodes, $\Gamma \ll \epsilon_C$, and temperatures below that of the charging energy, $k_B T \ll \epsilon_C$. The latter condition is easily satisfied since the measurements are taken at approximately 11 K–12 K corresponding to thermal energy $k_B T \approx 1$ meV and the charging energies of the electrons are typically on the order of 30 meV [44]. The first condition is only satisfied due to two fortunate features of the FinFET architecture.

The first feature is an artefact of the all-round gate geometry of the FinFETs. The combined potential of the three perpendicular gate faces is strongest at the top corners of the silicon channel. Hence, when the silicon channel is in a regime of weak inversion, there is a higher density of n carriers (electrons) at these top corner regions. At sub-threshold voltages electron transport through the channel is restricted to these corner regions. This is known as the corner effect [81]. These corner regions are depicted in the centre blow-up of Figure 3.8a. The cross-sectional area of the corner regions is reported to be $\sim 4\,\text{nm}^2$ in these devices [75]. Due to their small size, these corner regions may host a

finite number of localised As (donor) atoms which have diffused from the highly-doped source and drain contacts into the channel. Since these donors reside just below the gate they are well coupled to the gate. The gate voltage bends the conduction band to create a two dimensional triangular potential in the silicon channel as depicted in the right blow-up of Figure 3.8a. Lansbergen et al. [45] report that the gate voltage can tune the hybridisation of a donor bound electron wavefunction between the donor potential and this triangular potential, red line in Figure 3.8a. The ability to electronically manipulate the dopant bound electron in a gated silicon channel is a demonstration of quantum control useful for engineering quantum logic devices.

The second fortunate feature of the FinFET is due to the silicon nitride spacers either side of the gate, depicted in Figure 3.1. The poorly n-doped regions of silicon below the spacers form energy barriers E_B between the channel and the source/drain contacts. Figure 3.8b is a diagram of the conduction band energy along the silicon channel. The highly n-doped source and drain are well below the Fermi level E_F due to the introduction of positive dopants and associated free electrons. In the regions below the spacers, there is the presence of some doping depicted by small sharp potential wells but overall these regions are higher in energy than the source and drain and form energy barriers. When no gate voltage is applied the conduction band edge E_C of the p-doped channel sits above the source, drain and barriers (the black dotted line). The channel is a barrier between the source and drain.

At high temperatures, this barrier can be overcome by thermionic emission of electrons to create current. At low temperatures thermally activated transport is suppressed and no current can flow via thermionic emission. When a gate voltage is applied the conduction band of the channel is lowered. If a dopant is present in the channel, depicted as a sharp potential well, it is isolated from the source and drain by the energy barriers and the rest of the channel. By applying enough gate voltage so that the dopant is at the Fermi level electrons may tunnel from the source to a bound state of the dopant and then to the drain. The dopant is isolated from the source and drain by tunnel barriers (weak coupling) and capacitively coupled to the gate satisfying the conditions for a Coulomb island. Hence, Coulomb peaks of the two electron bound dopant states D^0 and D^- can explain the peaks observed at low temperatures and sub-threshold gate voltages.

The energy barriers and the corner region of the channel also create a quantum dot that can act as a Coulomb island and form Coulomb peaks. This quantum dot will reach the Fermi level at higher gate voltages and will have smaller energy barriers. As the gate voltage increases past the conduction band, the inversion in the FinFET channel will

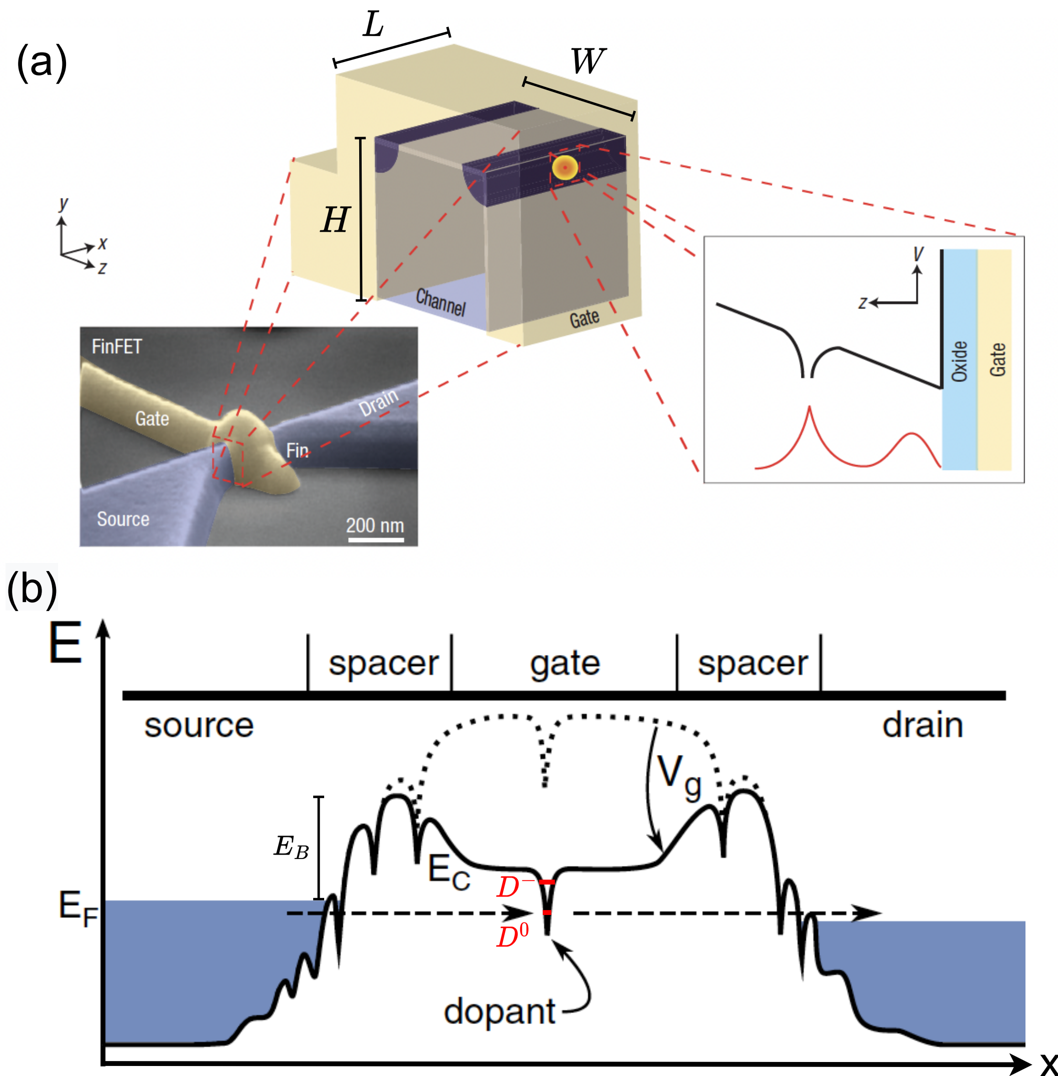


Figure 3.8. Low temperature FinFET operation and conditions for Coulomb blockade. (a) The bottom left is a scanning electron micrograph of the FinFET. The centre blow-up is a diagram of the silicon channel with three sides covered by the gate. The dark blue regions are the corner regions where the donor atoms of interest are located. The right blow-up is the potential along z due to the donor and the gate (black line) and the hybridised electron wavefunction (red line). (b) The energy band diagram of the FinFET along the silicon channel. The spacers create energy barriers E_B in the conduction band that reduce as the gate voltage is increased. The dopant bound states sit below the conduction band minimum E_C and reaches the Fermi level E_F first. Reprinted (a) from [45] under RightsLink License 5582420204090. Reprinted (b) with permission from [44] copyright 2006 by the American Physical Society.

increase allowing current to flow like the room temperature case. The electron transport

is more sensitive to dopants and impurities in the channel at low temperatures which may explain why there are some fluctuations in the drain current at high gate voltages.

In order to differentiate between peaks formed by the dopants and the quantum dot formed by the channel, it is useful to estimate the gate voltage at which the conduction band edge aligns with the Fermi level. From now on this will be referred to as the edge of the conduction band. The edge of the conduction band for FinFET2 is determined in Appendix D.1. This was done using an analysis of the thermionic emission of carriers over the energy barrier E_B for various gate voltages as done in [75]. The conduction band edge was found to be $E_C = 330 \pm 10\,\mathrm{mV}$ from the high temperature data of FinFET2. The results also gave an estimate for the coupling of the gate to the device, $\alpha = 0.42 \pm 0.04$, and the cross-sectional area of current flow, $S = 4.5 \pm 0.5\,\mathrm{nm}^2$. These results are similar to that of Sellier et al. [44, 75] and confirm the corner effect.

3.3. Identifying dopant states from Coulomb blockade

The collected low temperature data has signatures of Coulomb blockade and single electron quantum transport through the FinFETs. In this section, the electron states causing the Coulomb peaks are identified by comparing binding energies, capacitive coupling to the leads determined from stability diagrams and the expected Bohr radius of As dopants in silicon.

3.3.1. Physical interpretation of Coulomb peaks

Due to some mishaps in the cooling procedure and limitations of the low temperature setup discussed in Appendix D.2.2, the lowest temperatures reached in this run (11 K–12 K) were above the desired temperature of 4 K. Despite this, the current traces obtained at these temperatures still produced Coulomb peaks. However, there was also the presence of a baseline current of the conventional FinFET operation due to the warmer temperatures. By removing this baseline it was possible to clearly identify the Coulomb peaks and determine the electron states causing them.

At the lowest temperatures (11 K–12 K), multiple gate voltage sweeps were taken in the region below the conduction band E_C. The current resolution used for these measurements was 100 pA, with the current amplifier gain set to $10^7\,\mathrm{V/A}$, refer to

Appendix D.2.1. Figure 3.9a and c are the average of these gate voltage sweeps for each FinFET, ten sweeps for FinFET1 and six sweeps for FinFET2. The current traces are made up of a Coulomb peak spectrum and a baseline current. This baseline current is the standard FinFET operation after the threshold voltage but at sub-threshold voltages this is due to a combination of diffusion of electrons from source to drain and thermal fluctuations. The baseline current, the orange dot-dash line in Figure 3.9a and c, was determined using an iterative asymmetric least squares smoother developed by P. Eilers and H. Boelens in 2005 [82] and discussed further in [83]. The baseline was removed from the current data leaving the Coulomb peak spectrums in Figure 3.9b and d.

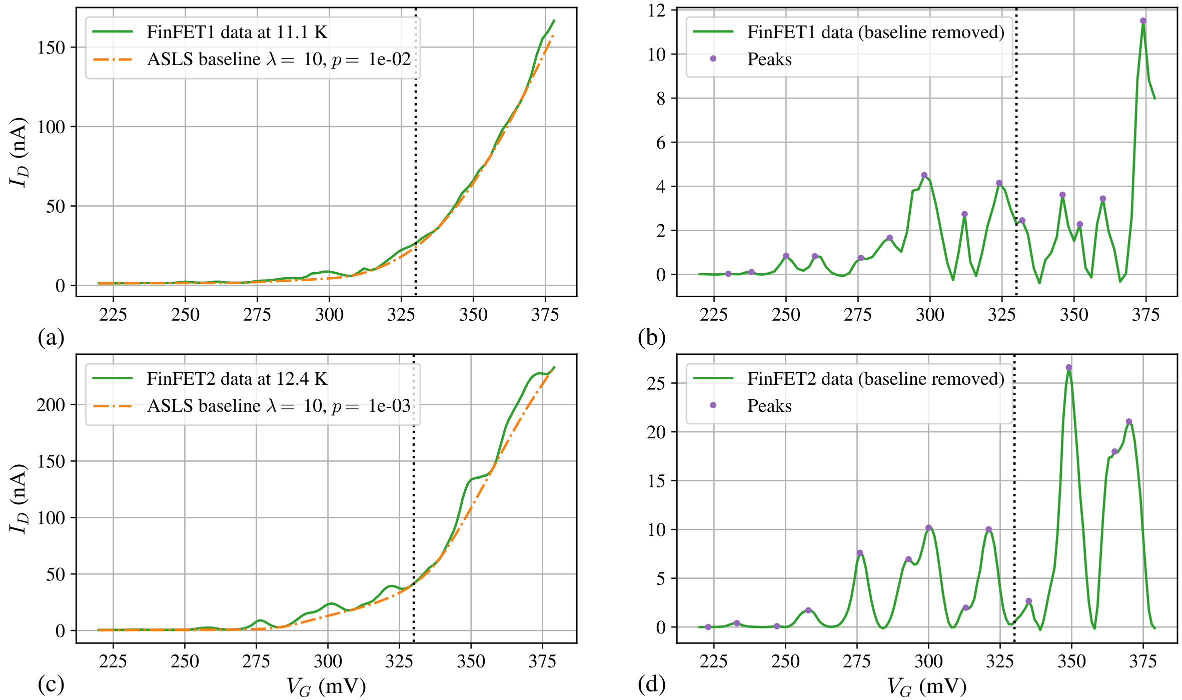


Figure 3.9. The gate voltage sweeps of FinFET1 and FinFET2 with and without the baseline current. (left) The solid green lines are the average of multiple sweeps for (a) FinFET1, ten sweeps, and (c) FinFET2, six sweeps, with the current amplifier gain at 10^7 V/A. The gate voltage is in steps of 2 mV and 1 mV respectively. The dot-dash orange line is the baseline of the drain current determined using the asymmetric least squares smoother (ASLS) with the smoothing parameters λ and p specified [82]. (right) The Coulomb blockade spectrum without the baseline for (b) FinFET1 and (d) FinFET2. The peak maximums are indicated by purple dots. The dotted black line is the conduction band determined in Appendix D.1.

From Section 2.3.1, the Coulomb peaks correspond to degeneracy points where the total charge on the Coulomb island is increased by a single electron. In this system the Coulomb island is made up of dopants in the corners of the silicon channel and

the quantum dot formed by the corner region of the channel. The peaks before the conduction band, the black-dotted line in Figure 3.9b and d, are due to the degenerate points between the electron bound states of donors in the channel: the D^+ zero electron state, the neutral D^0 one electron state and D^- two electron state. The D^+ state is unoccupied by an electron so it does not appear in the Coulomb spectrum. For energies above the conduction band, larger spectral peaks are observed which are suspected to be due to quantised charge states in the quantum dot. The quantum dot peaks are larger because the coupling of the leads to the quantum dot is stronger than that of the donor atoms in the channel [35]. The conduction band minimum for FinFET1 is assumed from the data of FinFET2 but may occur at larger gate voltages given that larger peaks do not occur until well after the black dotted line in Figure 3.9b.

Both FinFETs have a complicated spectrum of more than eight clear peaks below the determined conduction band. This indicates the presence of multiple donors in the channel corners each with two electron bound states, similar to the results of [44]. The donor states are filled in the order of most energetically favourable. For instance, starting from the lowest gate voltage, the first peak would correspond to the degeneracy point between the first dopant's D^+ state and D^0 state and the blockade following that peak corresponds to the D^0 state. The next peak could be the first degeneracy point for the next donor and so on until all donors are in the D^0 state. The next peaks would then be the degeneracy points between the dopants' D^0 and D^- states in the same order. The energy required to bind an electron to a dopant is dependent on the dopant's position in the channel and capacitive coupling to the source, drain and gate contacts and hence the filling of the dopant states may not follow this exact order; one dopant may reach the second bound state before another has reached its first. For simplicity, the model of filling all the D^0 states first and then all the D^- states is used in this analysis.

3.3.2. Determining the coupling constant

From Section 2.3.1, the capacitive coupling of the donor states to the source, drain and gate can be determined from Coulomb blockade stability diagrams. With the drain grounded as in Figure 3.4, the positive sloped resonance line of the Coulomb blockade state corresponds to $\kappa = C_G/(C_G + C_D)$ and the negative sloped resonance line corresponds to $\zeta = C_G/C_S$. The gate coupling constant α of a dopant state can be written in terms of the capacitances to the source, drain and gate contacts, $\alpha = C_G/(C_G + C_S + C_D)$.

The stability diagrams for each device were taken by measuring the drain current against the source drain bias V_{SD} and the gate voltage V_G. The absolute value drain current stability diagram for FinFET2 measured at current amplifier gain of 10^6 V/A is depicted in Figure 3.10. At lower gate voltages one can make out Coulomb diamonds. The coloured dot-dash lines are rough guides of the resonance lines through three degenerate points which match the Coulomb peaks of Figure 3.9d. Poorly resolved diamonds can also be seen at higher gate voltages. Each diamond represents a blockade where an electron has been bound to a donor in the channel. Figure 3.11 is the absolute value drain current stability diagram for FinFET1 with the current amplifier gain set to 10^7 V/A.

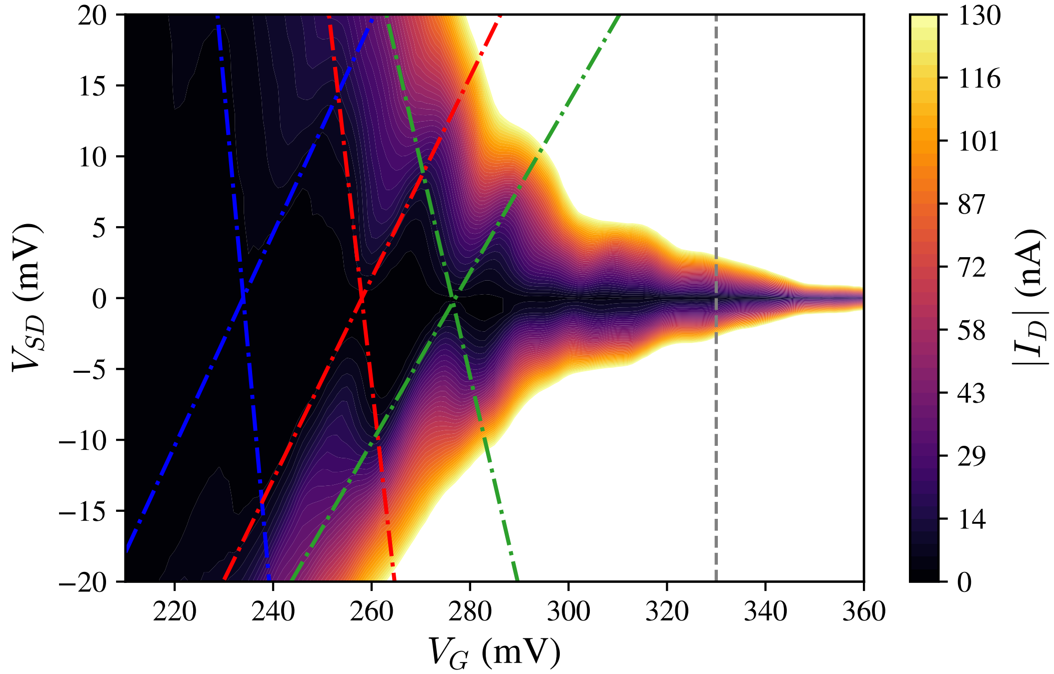


Figure 3.10. Stability diagram for FinFET2 at 12.7 K and current amplifier gain 10^6 V/A. The absolute value of the drain current against source drain voltage and gate voltage. The coloured dot-dash lines are the resonance lines of three degenerate points that form Coulomb diamonds. The dashed grey line is the conduction band edge determined in Appendix D.1.

For this analysis, the focus will be on the FinFET2 data since it has fewer Coulomb peaks and the correct conduction band minimum. The ratios κ and ζ were determined from the slope of the resonance lines in Figure 3.10 and used to determine the source and drain capacitances and the gate coupling constant α for the degeneracy points highlighted in the stability. The capacitance ratios and capacitances for each degeneracy point are presented in Table 3.2. The similar values of C_S and C_D for each state indicate that

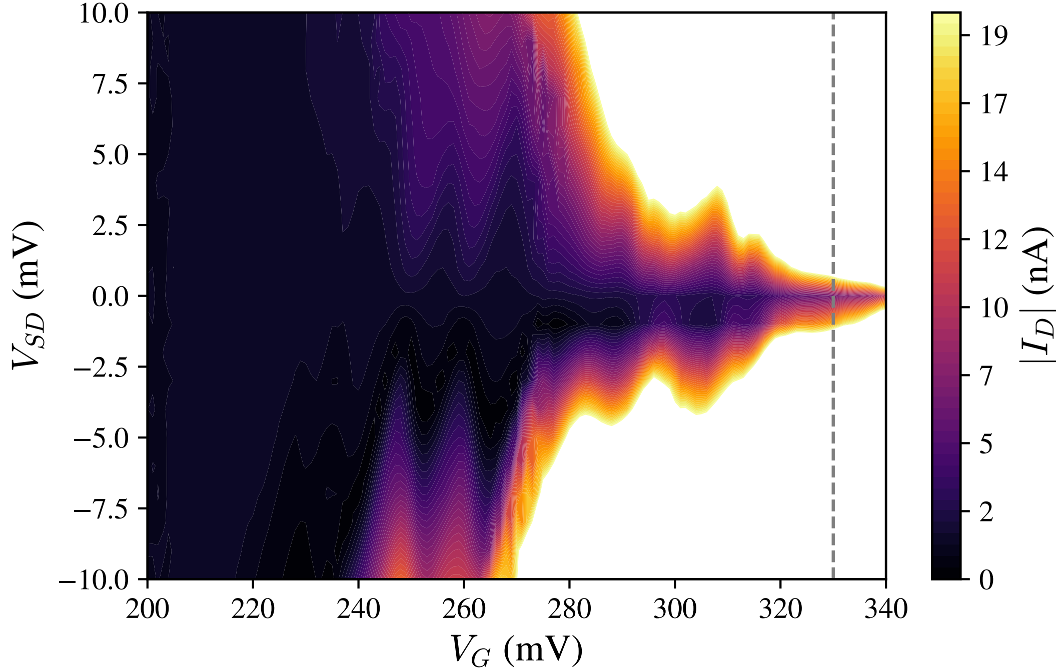


Figure 3.11. Stability diagram for FinFET1 at 11.3 K. The absolute value of the drain current against source drain voltage and gate voltage. The dashed grey line is the conduction band determined in Appendix D.1.

the dopants of these states are well centred in the channel. The fact that $C_G > C_{S,D}$ indicates that the donor states are more strongly coupled to the gate.

Table 3.2. Capacitance ratios and capacitances for the degeneracy points of the FinFET2 stability diagram Figure 3.10.

Degeneracy point	κ	ζ	C_S (C_G)	C_D (C_G)	α
Blue	0.75	3.8	0.26	0.33	0.63
Red	0.71	3.0	0.33	0.41	0.57
Green	0.60	1.6	0.63	0.67	0.44

The degeneracy points outlined in the stability diagram correspond to three of the peaks in Figure 3.9d. The peaks around these that are not as prominent in the stability are likely due to donors that are not centralised between the source and drain. The non-centralised states will have contrasting C_S and C_D values and overall weaker coupling to the leads Γ which would produce a dull peak in the current trace [35]. These states clutter the stability diagram's Coulomb diamonds and broaden peaks in the Coulomb peak traces.

The coupling constants α in Table 3.2, determined for each degeneracy point are all larger than that found in the thermionic emission analysis in Appendix D.1, $\alpha = 0.42$. For the remainder of this chapter we take the rounded average of these four values as the gate coupling constant, $\alpha = 0.5$.

3.3.3. Binding energies of donor states

The electron bound states of dopants in silicon are identifiable by their binding energies. In this section, the spacing of the Coulomb peaks of FinFET2 and the refined coupling constant, $\alpha = 0.5$, are used to identify the donor bound states.

Figure 3.12 is a zoomed plot of FinFET2's Coulomb peak spectrum below the conduction band. There are ten clear Coulomb peaks. A likely physical explanation of these states is that there are five dopants (labelled a to e) each with a D^0 (a1 to e1) and a D^+ (a2 to e2) bound state. Each peak represents the transition into the blockade of the corresponding donor bound state. The binding energies of the dopant states are determined by the distance of the corresponding Coulomb peak from the conduction band minimum. The distance of each Coulomb peak from the conduction band, $V_G - E_C$, and the corresponding binding energies of the peaks, calculated using the average conversion factor $\alpha = 0.5$, are given in Table 3.3.

The binding energy of an arsenic atom in bulk silicon is 46 meV [84]. The prominent states in Figure 3.10 the blue dot-dash state (c1) and red dot-dash state (e1) are most likely D^0 states given their binding energies: 48.5 meV and 36 meV respectively. Peaks a1, b1 and d1 are also likely D^0 donor states but may not be as centralised between source and drain lending to binding energies different to that of bulk donors and less prominent diamonds in the stability diagram. It is worth noting that a1 and b1 peaks are at currents of the order of 100 pA and the stability diagram has a resolution floor of 1 nA with the current amplifier gain set to 10^6 V/A, refer to Appendix D.2.1 for more detail. Hence, these peaks do not appear clearly in the stability diagram but their small current peaks suggest a weaker coupling to the leads.

In bulk silicon, the binding energy of the D^- state, two electrons bound to the donor, is only about 2 meV because of the strong Coulomb repulsion between the electrons [85]. In a FinFET, the D^- state is negatively charged and capacitively coupled to the source, drain and gate electrodes reducing the charging energy to e^2/C where C is the

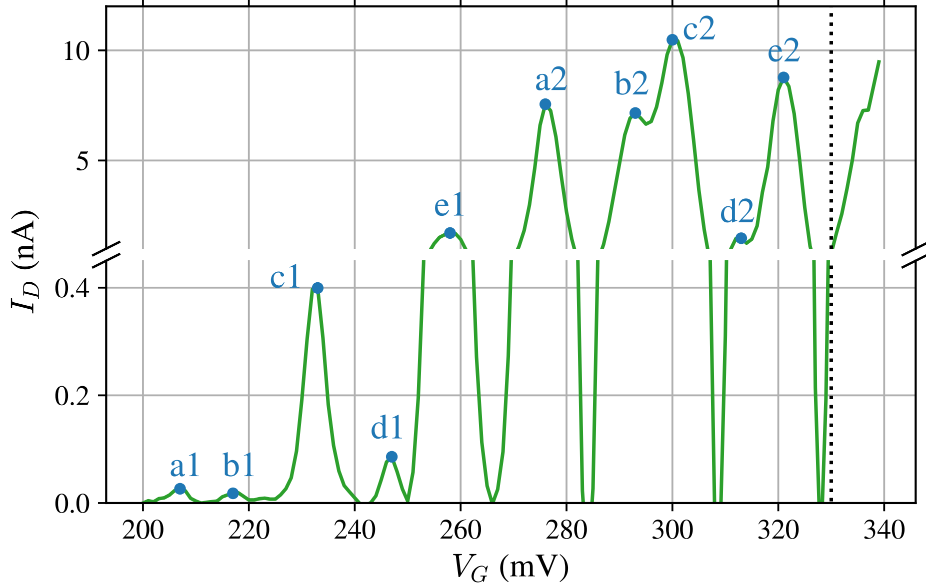


Figure 3.12. The Coulomb peaks of FinFET2 with 1 mV source drain bias at 12.4 K. The blue dots and labels identify the peaks and their corresponding dopant state. a to e are the dopants and 1 and 2 correspond to the D^0 and D^- state, respectively. The black dotted line is the conduction band.

Table 3.3. The binding energies and distance from conduction band to peak of the dopant states of FinFET2, $V_G - E_C$, which differ by a factor of the coupling constant $\alpha = 0.5$. Peaks are labelled as in Figure 3.12. Note that ΔV_G has an error of about 10 mV based on the determined conduction band minimum.

Peak No.	$V_G - E_C$ (mV)	Binding energy (meV)
a1	123	61.5
b1	113	56.5
c1	97	48.5
d1	83	41.5
e1	72	36
a2	54	27
b2	37	18.5
c2	30	15
d2	17	8.5
e2	9	4.5

total capacitance. This reduction of charging energy is of great interest for dopants in silicon quantum gates, since it could provide longer lifetimes of the two-electron donor state [44, 46]. The charging energy should be fairly consistent across all dopants and

a charging energy of 30 meV is common for donors in these devices [45]. The charging energies between the identified single electron and two electron states of the donors are: a1 $\rightarrow$ a2 $= 34.5$ meV, b1 $\rightarrow$ b2 $= 38.0$ meV, c1 $\rightarrow$ c2 $= 33.5$ meV, d1 $\rightarrow$ d2 $= 33.5$ meV and e1 $\rightarrow$ e2 $= 31.5$ meV. The charging energies are consistently around 30 meV supporting the hypothesis of five donors. The variations in these charging energies e.g. b1 $\rightarrow$ b2 and c1 $\rightarrow$ c2, could be due to the dopants being located in different locations along the corner of the channel and hence different capacitive coupling to the leads. Alternatively, some dopants may be located deeper in the channel and less coupled to the leads.

Another check to confirm that these states are due to dopants is to determine the gate capacitance of the bound states. The gate capacitance of the dopants can be determined from the charging energy between D^0 and D^-. From the capacitance ratios of e1 (the red dot-dash stability state) and the charging energy of donor e (e^2/C), the gate capacitance is $C_G = 2.9 \times 10^{-18}$ F which agrees with the value reported in [44]. This capacitance corresponds to a sphere of radius 2.2 nm in bulk silicon which agrees with the 2.5 nm Bohr radius of a neutral D^0 bound state of an As donor in bulk silicon [50, 86].

FinFET1 also has multiple Coulomb peaks which would correspond to a handful of dopants in the corners of the silicon channel. The stability diagram Figure 3.11a has two very clear states between gate voltages 240 mV and 260 mV that would likely be single electron D^0 states of dopants. There is also the traces of a state before these two corresponding to the small Coulomb peak at around 238 mV in Figure 3.9b. This could also correspond to a D^0 state. The binding energies of these three peaks using FinFET2's conversion factor α and conduction band value are: 48 meV, 42 meV and 37 meV respectively. These values are similar to the binding energies of As donors in bulk silicon. Following the same arguments as above the D^- states corresponding to these donors could be determined.

3.3.4. Difficulties in identifying the dopants

The Coulomb peak spectrums in Figure 3.9 and the stability diagrams Figures 3.10 and 3.11 all have broad and overlapping neighbouring peaks. This makes it difficult to identify Coulomb peaks and results in poorly resolved Coulomb diamonds. In both stability diagrams, the extent of the Coulomb blockade regions are much smaller than that observed in previous works with similar devices [44, 45, 75]. These works boast Coulomb diamonds extending up to ± 30 mV sourced drain bias compared to the diamonds

presented here extending from ±15 mV for FinFET1 and only ±5 mV for FinFET2. The broadening of the Coulomb peaks in both the stability and Coulomb spectrums is a result of two factors.

The first is that the device is too warm, 11 K–12 K. This causes thermal broadening of the peaks. When k_BT is greater than the coupling to the leads, Γ, the width of the peaks is proportional to temperature, however, for $\Gamma > k_BT$, the peak widths are independent of temperature and proportional to Γ [35]. The thermal energy for temperatures of $T = 12.4$ K and $T = 11.1$ K is $k_BT \approx 1$ meV. Typical values of Γ are 0.1 meV–1 meV [35] and from similar FinFETs in [44] the width of the peaks is around 0.5 mV at 0.1 K corresponding to a coupling similar to 0.34 meV. The thermal energy of this experiment is larger than 0.34 meV which suggests these results are in the $\Gamma < k_BT$ regime. The average full width half maximum of the peaks below the sub-threshold region in Figure 3.9b and d are 2.2 mV (1.1 meV) and 5.5 mV (2.75 meV), respectively. These are on the order of the thermal energy. If temperatures less than 4 K are achieved, $k_BT < 0.086$ meV, the width of these peaks will be significantly reduced and will approach the lead coupling Γ. This would resolve the neighbouring peaks and Coulomb diamonds more clearly. A discussion of the problems encountered that restricted the cryostat from reaching sub 4 K temperatures in this run is found in Appendix D.2.2.

The second is that the spectrum is complicated. There are multiple neighbouring states that are overlapping so that Coulomb peaks are crowded and the Coulomb diamond regions are small. For instance, the D^0 Coulomb diamond of the dopant c is filled with transitions to the D^0 states of dopants d and e and D^- states of dopants a and b, in Figure 3.10. If it weren't for these states, there would be a full Coulomb diamond extending to the source-drain bias of approximately ±30 mV. This problem is amplified when combined with the thermal broadening. For a more accurate analysis of the Coulomb spectrum in the FinFETs it is ideal to find a device with only a single dopant in the corners of the channel. Then there would be only two clear donor bound states to analyse. However, finding these single dopant devices is a time consuming game of chance. Lansbergen et al. [45] reports a one in seven probability of finding a working FinFET with two resonances from a single dopant, the D^0 and D^- states. The devices measured here are similar to the devices tested by Lansbergen et al. and the sample area allows for only two or three devices to be cooled down at a time. Hence, testing at least seven working devices to find a FinFET with a single dopant would require cooling down at least three different sets of working devices. Due to time constraints and a limited

helium supply it was not possible for us to conduct multiple cool downs and search for such a device.

3.4. Summary

In this chapter, Coulomb blockade spectrums and stability diagrams corresponding to transitions between the one- and two-electron bound states of dopants in the silicon channels of two FinFETs were observed. The sub-threshold Coulomb peaks for FinFET2 were identified to be the D^0 and D^- states of five dopants. A similar spectrum of Coulomb peaks due to multiple dopants was observed in FinFET1. Quantum dot states due to the confinement of the corner effect and the energy barriers were observed above the conduction band. Coulomb diamonds corresponding to the Coulomb blockade between D^0 and D^- dopant states were observed in the stability diagrams of FinFET1 and FinFET2. The size of the Coulomb diamonds and resolution of Coulomb peaks was limited due to the presence of thermal fluctuations and multiple dopants. These are not novel results but a good example of quantum transport due to the sequential tunnelling of single electrons. This experiment doubled as a validation of the low temperature measurement setup, although there are still improvements to be made. An accurate analysis of the D^0 and D^- states (including excited states) of a dopant in silicon could be performed if a device with a single dopant in the channel is found and cooled below 4 K.

Chapter 4.

Atomistic tight-binding modelling tool - NEMO3D

In recent years, there has been an overwhelming interest in novel nanoscale semiconductor devices which are expected to revolutionise computing and quantum information technologies. The charge carriers in these nanoscale devices are governed by quantum mechanics which determines their electronic band structure and ultimately the properties and operation of the device. Understandably, there is an increasing need for tools which can accurately model the electronic band structures of nanoscale systems built with various geometries, materials, dopants and applied electromagnetic fields. One such tool developed for this purpose is the three dimensional Nanoelectronic Modelling tool (NEMO3D) developed by Klimeck et al. [87, 88]. NEMO3D uses an atomistic nearest neighbour tight-binding $sp^3d^5s^*$ model to compute the electronic structure of up to approximately 64 million atoms, corresponding to a volume of $(110\,\mathrm{nm})^3$. The tool has been optimised to scale from a single computer processing unit to a large number of processors in a supercomputer cluster.

The NEMO3D tool is different to continuum methods of calculating band-structure e.g. effective mass and $k \cdot p$ theory in that it accounts for crystal symmetry on an atomic resolution. The tight-binding model also has the advantage over the common pseudopotentials method as far as it is less expensive on storage and processing power [89]. The main criticism of the tight-binding method is that it is semi-empirical i.e. it requires input from experiment to obtain the material parameters for a given semiconductor. Regardless, the tool has successful modelled various nanoelectronic structures including but not limited to self-assembled quantum dots [90], gated donors in silicon [91, 92], Si/SiGe quantum wells [93] and SiGe nanowires [94].

NEMO3D will be used for the remainder of this thesis to model the electronic band structure of InAs nanowires. In particular, to investigate how different internal and external properties of the nanowire influence its effective g-factor anisotropy and spin-orbit interactions. In this chapter, the atomistic tight-binding method underpinning NEMO3D is introduced, with specific emphasis placed on how the model incorporates electromagnetic coupling and the spin-orbit interaction. The InAs and Si bulk band structures produced by NEMO3D are compared to experimental data. An example of NEMO3D's ability to model electronic spectra is given by modelling the valley splitting of a quantum dot in a Si/SiGe quantum well with an interface step. These valley splitting results were incorporated in a recent publication [95].

4.1. Nanoelectronic modelling

In this section, the various nanoelectronic modelling techniques are briefly introduced and compared to the approach of NEMO3D. The modelling of nanoscale semiconductor devices is generally approached from the quantum mechanical level. The focus is typically on a few periodic unit cells of a crystal structure containing a few thousand atoms. The main goal of the modelling is to determine the electronic structure e.g. eigenenergies and eigenfunctions of the electrons in a semiconductor. From first principles, one would typically start from the time-independent Schrödinger equation with a many body Hamiltonian containing interactions between the ion cores of the crystal lattice, the ion cores and electrons and between the electrons themselves, see equation (2.1). However, due to the complexity of the interactions between a large number of atoms and electrons all models make approximations to simplify calculations. These approximations reduce the computational cost of the model at the expense of the model's realism.

4.1.1. Common assumptions

The first assumption typically made for semiconductors is to consider only the electrons in the outer-most shell of the atoms in the crystal lattice, the valence electrons. The core shell electrons bound to each atom in the lattice can be lumped together as ion cores. The next approximation made by most models is that the ion cores of the crystal are effectively stationary as seen by the free electrons. This is called the Born-Oppenheimer approximation [36]. This essentially means that the electron and ion core wavefunctions

may be treated separately. The problem can then be simplified to a Hamiltonian considering the motion of electrons through a potential of stationary ion cores. The ion motion is negligible and typically ignored but can be used to derive the electron-phonon interactions.

The Born-Oppenheimer approximation greatly simplifies the many particle model, however, there are still complex electron-electron interactions present that are difficult to solve. There are computationally expensive approaches that can solve full electron models by explicitly including electron-electron interactions such as quantum Monte Carlo [96]. However, typically the problem is reduced to a single electron Schrödinger equation by making a mean field approximation: that every electron experiences the same crystal potential [97]. The electron-electron interactions are then introduced by including an appropriate local or non-local potential in this single electron Schrödinger equation. Sophisticated *ab initio* methods using this approach include the Hartree-Fock approximation [98, 99] and density functional theory [100, 101]. At a high level these methods solve the all-electron problem by including all the electrons on the atoms in a self consistent solution. These ab initio methods can be accurate but they are computationally intense and time consuming. Typically, further assumptions are made to determine the electronic structure of semiconducting nanostructures.

4.1.2. Simplified models

One simplified model assumes that the valence electron's are mostly free and do not feel the full potential of the semiconductor nuclei due to the screening of the core electrons. The true potential felt by the valence electrons is approximated by including a repulsive potential in the crystal potential which represents the screening interaction between core electrons and valence electrons. The resulting effective potential is called a *pseudopotential* [97]. Pseudopotential methods can be empirical or derived from first principles. Typically, pseudopotential methods are restricted to small systems and are computationally expensive because they require full matrix manipulations to transform between real and momentum space [87, 89].

The simplest and least computationally intensive models are non-atomistic models which approximate the atoms as a continuum rather than modelling them individually. The most common model is the effective mass approximation which retains only the information of the effective masses and band edges from the electronic band structure [51, 97]. Another common model is $k \cdot p$ theory [51], which is an extension to this

approximation by including the coupling of multiple bands. The parameters for these continuum models are determined from experimental data of semiconductors e.g. band gaps. These non-atomistic approximations are computationally inexpensive compared to *ab initio* methods and have found great agreement with multiple experiments [97]. However, these models do not contain information on the atomic level which is crucial for the modelling of nanoscale features such as growth direction, interfaces, and disorder which depend on the geometric constraints induced by the crystal lattice symmetry.

An intermediate model that is not extremely computationally expensive and also considers the atomistic resolution and crystal symmetry is the atomistic nearest-neighbour tight-binding approach. This is the approach used by the Nanoelectronic Modelling tool (NEMO3D) [89]. The atomic information is included by considering the atomic orbital wavefunctions (s, p and d) as a basis for the valence electron wavefunctions. The orbital wavefunctions on the nearest neighbouring atoms overlap and build up the electronic band structure of the semiconductor [97].

4.1.3. NEMO3D

The semi-empirical nearest-neighbour tight-binding method used in NEMO3D was introduced by Slater and Koster in 1954 [102] using an sp^3s^* atomic basis. The atomic basis was then extended to include all five d orbitals, $sp^3d^5s^*$, by Jancu et al. in 1998 [103]. The method selects a basis of orbitals for each atom in the crystal and builds a single electron Hamiltonian representing the electronic properties of the material. The interactions between same-atom orbitals and nearest neighbour atom orbitals are treated as empirical fitting parameters. These parameters are determined from a genetic algorithm that fits a large parameter space to the bulk band structure of the desired semiconducting material [89]. Once these parameters are determined the tight-binding model is transferable to any nanostructure of the semiconductor. For this work, the $sp^3d^5s^*$ with spin basis of 20 atomic orbital states is used.

The tight-binding method stores information in sparse matrices making it less computationally expensive compared to pseudopotential calculations. The computational cost of this method scales linearly with the number of atoms in the system and can easily be performed on supercomputing clusters [89]. . There is some controversy over the validity of the atomistic tight-binding method given its empirical nature [87]. However, there is now an extensive catalogue of tight-binding parameters and results that have found agreement with experiments for various nanoelectronic devices [87, 89, 91, 92, 93].

4.2. Atomistic tight-binding

This section discusses in detail the atomistic nearest-neighbour tight-binding method including the implementation of electromagnetic coupling and spin-orbit interactions.

4.2.1. Construction of the tight-binding Hamiltonian

Consider a semiconductor nanostructure consisting of N primitive unit cells each located at Bravais lattice points

$$R_{n_1,n_2,n_3} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \tag{4.1}$$

where a_i for $i = 1, 2, 3$ are the primitive direct lattice vectors and n_i are integers. For simplicity, denote the location of the jth unit cell or Bravais lattice point as R_j where j captures the nature of the integers n_i. In the case of zincblende and diamond lattices the primitive cell contains two atoms. Let the atoms be located at equilibrium atom positions $\mathbf{R}_{\nu j} = \mathbf{v}_\nu + \mathbf{R}_j$ where $\mathbf{v}_\nu$ is the offset of the νth atom of the unit cell from the unit cell location. For each atom in the primitive cell let there be a basis of orthogonal localised atomic-like orbitals $|\mu, \nu, \mathbf{R}_{j,\mu}\rangle$, called Löwden orbitals [104], where μ is the index over the orbitals including spin. To simplify notation ν is absorbed into μ so that the basis of the whole primitive unit cell is $|\mu, \mathbf{R}_{\mu j}\rangle$ with μ the index over the atomic orbitals on both atoms of the unit cell and the lattice vectors $\mathbf{R}_{\mu j} = \mathbf{R}_{\nu j}$.

From these localised atomic orbitals a basis of Bloch states can be created by summing over all unit cells in the crystal structure

$$|\mu, \mathbf{k}\rangle = \frac{1}{\sqrt{N}} \sum_{j}^{N} \exp(i\mathbf{k} \cdot \mathbf{R}_{\mu j})\, |\mu, \mathbf{R}_j\rangle \tag{4.2}$$

where $\mathbf{k}$ is the three dimensional wave vector restricted to the first Brillouin zone. The wavefunctions that diagonalise the tight-binding Hamiltonian and make up the bands of the semiconductor can be expressed as linear combinations of these atomic orbital Bloch states

$$|\Psi_{n,\mathbf{k}}\rangle = \sum_{\mu} C_{n,\mu}(\mathbf{k})\, |\mu, \mathbf{k}\rangle \tag{4.3}$$

where $C_{n,\mu}(\mathbf{k})$ are the expansion coefficients and n is the band index. These wavefunctions are eigenfunctions of a single electron tight-binding Hamiltonian H with eigenvalues $E_n(\mathbf{k})$

$$H\left|\Psi_{n,\mathbf{k}}\right\rangle = E_n(\mathbf{k})\left|\Psi_{n,\mathbf{k}}\right\rangle . \tag{4.4}$$

From equations (4.3) and (4.4), the tight-binding Hamiltonian can be represented as a matrix whose order is the total number of atomic orbitals in the unit cell. For the zincblende primitive cell with a 20 orbital basis this would be a 40×40 matrix. Assuming the Hamiltonian conserves translational symmetry due to a periodic crystal potential, the matrix elements of the Hamiltonian take the form [102, 105]

$$\begin{aligned} H_{\mu',\mu}(\mathbf{k}) = \left\langle \mu', \mathbf{k}\right| H \left|\mu, \mathbf{k}\right\rangle &= \frac{1}{N}\sum_{j',j} \exp(i\mathbf{k}\cdot(\mathbf{R}_{\mu j} - \mathbf{R}_{\mu' j'}))\left\langle \mu', \mathbf{R}_{\mu' j'}\right| H \left|\mu, \mathbf{R}_{\mu j}\right\rangle \\ &= \sum_{j} \exp(i\mathbf{k}\cdot(\mathbf{R}_{\mu j} - \mathbf{R}_{\mu'}))\left\langle \mu', \mathbf{R}_{\mu'}\right| H \left|\mu, \mathbf{R}_{\mu j}\right\rangle \end{aligned} \tag{4.5}$$

where $\mathbf{R}_{\mu'} = \mathbf{R}_{\mu' 0}$. In the last step, the summation over all of the unit cells indexed by j' is removed which corresponds to multiplying the expression by N and the value for $\mathbf{R}_{\mu' j'}$ is set to the constant $\mathbf{R}_{\mu'}$ [102]. The distance $(\mathbf{R}_{\mu j} - \mathbf{R}_{\mu'})$ corresponds to the distance between atoms in unit cell j and the atoms in the 0th unit cell. Typically, in nearest-neighbour tight-binding j is restricted to the sum over the nearest neighbour in the current unit cell and the three nearest-neighbours in adjacent unit cells. The elements of the Hamiltonian matrix are dependent on $\mathbf{k}$ by the plane waves due to the displacement between the four nearest-neighbouring atoms. The integrals $\left\langle \mu', \mathbf{R}_{\mu'}\right| H \left|\mu, \mathbf{R}_{\mu j}\right\rangle$ are determined from fitting to the experimentally determined bulk band structure.

In NEMO3D, a geometry can be modelled with periodic boundary conditions or closed boundary conditions. These two conditions can be explained simply by considering a one-dimensional chain of n atoms along x with positions x_i for $i = 1, ..., n$. Periodic boundary conditions are the requirement that $x_{n+1} = x_1$ so that the one-dimensional chain can be thought of as a ring or an infinite chain. Closed boundary conditions requires that the chain is finite, with x_1 and x_n the atoms at the end of the chain. The typical form of the electron wavefunction for periodic and closed boundary conditions are derived in Appendices A.1 and A.2.

The atomic basis of NEMO3D takes the Bloch form (4.5) only if the nanostructure has periodic boundary conditions. If the crystal geometry is chosen to have closed boundary

conditions there is no longer translational symmetry and $\mathbf{k} = \mathbf{0}$. The eigenfunctions of the Hamiltonian are then a linear combination of all the atomic orbitals localised on each of the N atoms in the lattice with positions $\mathbf{R}_I$. The basis for this linear combination will be denoted $|\mu, I\rangle$, where μ runs over the atomic orbitals of an atom and I runs over all the atom positions in the crystal lattice. The Hamiltonian can then be written as a $NM \times NM$ matrix where M is the number of atomic orbitals on each atom with elements $\langle \mu', I'| H |\mu, I\rangle$.

Regardless of whether the Hamiltonian elements are constructed from the basis of Bloch states $\{|\mu, \mathbf{k}\rangle\}$ or the basis of all atomic orbitals in the crystal lattice $\{|\mu, I\rangle\}$, the tight-binding Hamiltonian has the same general form. First consider a 1D chain of N atoms each with a single atomic orbital, the Hamiltonian can be expressed as an $N \times N$ matrix that acts on a column vector Ψ of length N,

$$[H]_{nn}\Psi_m = [\varepsilon_n \delta_{n,m} - t\delta_{n,m+1} - t\delta_{n,m-1}]\Psi_m. \tag{4.6}$$

where n is the row index of the Hamiltonian matrix, and m is the column index of the matrix and the index of the column vector. The diagonal terms of the matrix ε_n are the on-site energies for the atomic orbital on the nth atom and the non-zero off-diagonal terms t are the overlap/interactions between the nearest-neighbour atomic orbitals, which are assumed to be the same for each atom here.

When multiple orbitals are localised at each atom, say M orbitals, the Hamiltonian matrix becomes an $NM \times NM$ matrix. The restriction to only nearest-neighbour atom interactions means that the majority of the off-diagonal terms in this matrix are zero and the Hamiltonian is a sparse block tridiagonal matrix. This matrix can be written as

$$H = \begin{pmatrix} h_1 & t_{12} & 0 & \dots & 0 & t_{bc} \\ t_{21} & h_2 & t_{23} & 0 & \ddots & 0 \\ 0 & t_{32} & h_3 & t_{34} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & 0 & t_{N-1,N-2} & h_{N-1} & t_{N-1,N} \\ t_{bc} & 0 & \dots & 0 & t_{N,N-1} & h_N \end{pmatrix}. \tag{4.7}$$

. The matrix consists of $M \times M$ diagonal blocks h_n for each atom $n \in [1, N]$, whose diagonals are the on-site orbital energies, similar to ϵ_n above, and whose off-diagonals

are the intra-atomic orbital interactions between orbitals on the nth atom. The off-diagonal blocks $t_{n,n\pm1}$ are made up of the inter-atomic orbital interactions between orbitals on nearest neighbour atoms. The top corners of the matrix t_{bc} are set depending on the desired boundary conditions; zero for closed boundary conditions and t_{N1} or t_{1N} for periodic boundary conditions. These sparse matrices simplify the computation of eigenvalues which is one of the great advantages of the tight-binding method.

The general tight-binding Hamiltonian used in NEMO3D for three dimensional crystal structures can be written in terms of the three different interactions mentioned above [87, 106]

$$H = \sum_{i}^{N} \sum_{\mu,\upsilon} (\varepsilon_{\mu,i}\delta_{\mu\upsilon} + \Delta_{\mu\upsilon,i}) c^{\dagger}_{\mu i} c_{\upsilon i} + \sum_{i}^{N} \sum_{j\neq i}^{4\text{nn}} \sum_{\mu,\upsilon} t_{\mu i,\upsilon j} c^{\dagger}_{\mu i} c_{\upsilon j} \tag{4.8}$$

where i is the index over the atoms of the lattice, N is the total number of atoms, 4nn denotes that the sum is over the four nearest-neighbouring atoms, and μ and υ index the atomic orbitals with spin degree of freedom. $c^{\dagger}_{\mu i}$ ($c_{\mu i}$) is the creation (annihilation) operator of an electron on the orbital μ localised on atom i. The first term $\epsilon_{\mu,i}$ is the on-site energies of atomic orbitals on each atom and the second term $\Delta_{\mu\upsilon,i}$ is the intra-atomic orbital interactions. The last term is the interaction between the orbital μ on the ith atom with the atomic orbital υ on the four the nearest-neighbouring atoms indexed by j. The on-site energies, $\varepsilon_{\mu,i}$, the intra-atomic interactions, $\Delta_{\mu\upsilon,I}$, and the inter-atomic interactions, $t_{\mu i,\upsilon j}$, in equation (4.8) correspond to the Hamiltonian integrals introduced earlier which are all empirical fitting parameters. In the atomic orbital basis notation, $|\mu, I\rangle$, the Hamiltonian matrix elements can be written as [105]

$$\langle \mu, I| H |\mu, I\rangle = \varepsilon_{\mu,I} \tag{4.9}$$

$$\langle \mu', I| H |\mu, I\rangle = \Delta_{\mu'\mu,I} \text{ for } \mu' \neq \mu \tag{4.10}$$

$$\langle \mu', I'| H |\mu, I\rangle = t_{\mu' I',\mu I} \text{ for } I' \neq I. \tag{4.11}$$

Similar results could also be obtained for the Hamiltonian written in terms of the Bloch states $|\mu, \boldsymbol{k}\rangle$. These integrals are determined by fitting to the band structure of a bulk semiconductor with the genetic algorithm described in [89].

4.2.2. The spin-orbit interaction

The spin-orbit interaction of an electron bound to an atom is given in equation (2.20)

$$H_{\mathrm{SO}} = -\frac{\hbar}{4m_0^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{p}\times\boldsymbol{\nabla}V) \tag{4.12}$$

where V is the total potential felt by an electron on the atom and $\boldsymbol{\sigma}$ is the vector of Pauli matrices.

In NEMO3D, the potential V is split into the intrinsic potential of the atom and an external potential applied to electrons on the atom. The spin-orbit interaction of the external potential is calculated by symmetrising it and using the electromagnetic coupling methods discussed in the next section. For the intrinsic atom potential, the result of the spin-orbit interaction is the coupling of the electron's orbital angular momentum $\mathbf{l}$ and the electron's intrinsic spin momentum $\mathbf{s}$. The Hamiltonian interaction term could be written $H_{SO} = \lambda\mathbf{l}\cdot\mathbf{s}$ with λ the spin-orbit coupling constant. The spin-orbit interaction couples p orbitals on the same atom [97, 107]. The atomic orbitals are eigenstates of the total angular momentum $\mathbf{j} = \mathbf{l} + \mathbf{s}$. For the p orbital, $l = 1$ and $s = 1/2$. The total angular momentum of a p orbital can take two values $j = l + s = 3/2$ or $j = l - s = 1/2$. The result of the spin-orbit interaction is to split the $j = 3/2$ and $j = 1/2$, p orbital states by a spin-orbit splitting, $\Delta_0 = \frac{3}{2}\lambda$. There are four $j = 3/2$ states and two $j = 1/2$ states. At the Γ point ($\mathbf{k} = (0,0,0)$) of a III-V semiconductor's valence band, the spin-orbit interaction creates a four-fold degeneracy of two heavy-hole and two light-hole bands and a double degenerate split-off band separated by Δ_0 since these three valence bands (heavy-hole, light-hole and split-off) are mostly composed of p-orbitals. The s orbitals are not influenced by the spin-orbit coupling.

The spin-orbit coupling enters the NEMO3D Hamiltonian through the intra-atomic orbital constants $\Delta_{\mu v,i}$ in equation (4.8). In NEMO3D, only the coupling of the p orbitals is considered following the work of Chadi [107]. The coupling of the p orbitals can either be determined from first principles or experimental data of the valence band splitting, Δ_0, at the Γ point of a bulk semiconductor. For a III-V zincblende semiconductor, the atomic spin-orbit splitting of the anion and cation, Δ_a and Δ_c, is required. Typically the atomic spin-orbit splitting is renormalised to obtain the correct spin-orbit splitting of the valence bands in the semiconductor [107].

4.2.3. Electromagnetic coupling

Time-dependent electromagnetic fields are included in the tight-binding model by a form of minimal coupling described in [105, 108, 109]. This method maintains gauge invariance, conservation of charge and does not introduce any extra free parameters.

In the presence of an applied vector potential, $\mathbf{A}(\mathbf{r}, t)$, and scalar potential, $\Phi(\mathbf{r}, t)$, extra terms are introduced into the single electron Hamiltonian. The momentum operator becomes $\mathbf{p} \to \mathbf{p} + e\mathbf{A}$ and the scalar potential is added. By substituting this momentum operator and scalar potential into the non-relativistic limit of Dirac's equation, see Section 2.4.2, the Pauli spin-orbit term and a Zeeman splitting are introduced into the single electron Hamiltonian. The spin-orbit term is accounted for as described in Section 4.2.2. The Zeeman splitting, $\mu_B \boldsymbol{\sigma} \cdot \mathbf{B}$, is added to the diagonal on-site elements of the matrix depending on the spin of the orbitals. The scalar potential is also added to the diagonal on-site elements of the matrix in the form $e\Phi$ and will influence the spin-orbit interaction by changing the total potential of the system.

The inclusion of these terms to the on-site elements of the tight-binding Hamiltonian may have introduced the electric field and magnetic field but the Hamiltonian is not yet gauge invariant. To enforce gauge invariance a phase factor must be introduced to the interactions of nearest-neighbouring atoms. This is the phase that an electron gains when hopping between nearest neighbours in the presence of a vector potential. The Hamiltonian with electromagnetic coupling, ignoring the spin-orbit and Zeeman interactions for now, is

$$H_{\mathrm{EM}} = \frac{1}{2m} \left(-i\hbar\boldsymbol{\nabla} + e\mathbf{A}\right)^2 + V \tag{4.13}$$

where V contains the atom potential and any external potentials. According to [52] the solution to this Schrödinger equation is

$$\Psi(\mathbf{r}, t) = \Psi^{(0)}(\mathbf{r}, t) \exp\left(-\frac{ie}{\hbar} \int^{\mathbf{r}} \mathbf{A}(\mathbf{r}', t) \cdot d\mathbf{r}'\right) \tag{4.14}$$

where $\Psi^{(0)}$ is the solution to the Schrödinger equation with $\mathbf{A} = 0$, the corresponding Hamiltonian will be denoted H_0. It follows that H_{EM} and H_0 can be related by the unitary transformation

$$H_{\mathrm{EM}} = \exp\left(-\frac{ie}{\hbar} \int^{\mathbf{r}} \mathbf{A}(\mathbf{r}', t) \cdot d\mathbf{r}'\right) H_0 \exp\left(\frac{ie}{\hbar} \int^{\mathbf{r}} \mathbf{A}(\mathbf{r}', t) \cdot d\mathbf{r}'\right). \tag{4.15}$$

If H_0 now includes the spin-orbit term, equation (4.12), the electromagnetic Hamiltonian H_{EM} can still be written in this form and satisfies gauge invariance as described in [108].

Letting $H_0 = H_{1e} + H_{\mathrm{SO}}$, the matrix elements of the tight-binding Hamiltonian in the presence of a non-zero vector potential can be written

$$\langle \mu', I' | H_{\mathrm{EM}} | \mu, I \rangle \tag{4.16}$$

$$= \exp\left(-\frac{ie}{\hbar}\int_0^{\mathbf{R}_{I'}} \mathbf{A}(\mathbf{l},t)\cdot d\mathbf{l}\right) \langle \mu', I' | H_0 | \mu, I \rangle \exp\left(\frac{ie}{\hbar}\int_0^{\mathbf{R}_I} \mathbf{A}(\mathbf{l},t)\cdot d\mathbf{l}\right) \tag{4.17}$$

$$= \exp\left(-\frac{ie}{\hbar}\int_{\mathbf{R}_I}^{\mathbf{R}_{I'}} \mathbf{A}(\mathbf{l},t)\cdot d\mathbf{l}\right) \langle \mu', I' | H_0 | \mu, I \rangle \tag{4.18}$$

where the integral in the last line is the straight line connecting two nearest-neighbours. The vector potential is incorporated into the off-diagonal elements of the Hamiltonian by introducing a phase which is the contour integral of the vector potential along the line between nearest-neighbouring atoms. The introduction of this phase is known as the Peierls substitution [110]. The phase factor influences interactions between orbitals on different atoms, $I' \neq I$, the off-diagonal elements and has no influence on the intra-atomic orbital interactions, the block diagonal elements.

In summary, the introduction of the electromagnetic scalar and vector potentials changes the elements of the NEMO3D tight-binding Hamiltonian equation (4.8) to

$$\varepsilon_{\mu,I} = \varepsilon^0_{\mu,I} - e\Phi(\mathbf{R}_I, t) \tag{4.19}$$

$$t_{\mu' I', \mu I} = t^0_{\mu' I', \mu I} \exp\left(-\frac{ie}{\hbar}\int_{\mathbf{R}_I}^{\mathbf{R}_{I'}} \mathbf{A}(\mathbf{l},t)\cdot d\mathbf{l}\right) \quad \text{for } I' \neq I \tag{4.20}$$

where $\varepsilon^0_{\mu,I}$ and $t^0_{\mu' I', \mu I}$ are the zero field matrix elements.

The scalar potential Φ is added to the on-site elements of the matrix, while the vector potential introduces a phase to the off-diagonal elements of the tight-binding Hamiltonian. If the spin degree of freedom is considered the corresponding Zeeman energy is included on the diagonal spin elements of the tight-binding Hamiltonian.

4.2.4. Solving the Hamiltonian

In order to solve the Hamiltonian, first the desired geometry and electromagnetic conditions of a semiconducting nanostructure are given as an input to NEMO3D. Then the

tight-binding Hamiltonian is constructed from the input values and the parameters of the semiconductor's bulk band structure. The desired eigenvectors and eigenvalues of the tight-binding Hamiltonian are solved for by running a parallel Lanczos algorithm on a supercomputing cluster [89]. Typically, in a closed structure only the first few conduction bands or valence bands are required and not the whole band structure. If the full Hamiltonian for 1 million atoms is constructed with the $sp^3d^5s^*$ atomic basis (20 orbitals), then the order of the Hamiltonian would be 20 million which is unrealistic to solve completely. Instead, algorithms such as the Lanczos algorithm approximate the solution on a small subspace that is gradually increased until a certain tolerance is achieved. This is a computationally feasible method for solving the tight-binding Hamiltonian because it requires only a few eigenvectors to be held in memory. The ability to solve the tight-binding sparse matrices with parallel algorithms is the reason for NEMO3D's superior computational efficiency when compared to other modelling methods.

4.3. Example of bulk band structures: Si and InAs

Before one begins modelling nanostructures in NEMO3D, it is important to check that the empirical tight-binding parameters for the semiconductor of interest reproduce the bulk band structure. Once the bulk tight-binding parameters are correct they are transferable to any nanostructure of the semiconductor. In this section, the bulk band structures of two semiconducting materials are produced, namely Si and InAs.

4.3.1. Silicon bulk band structure

Silicon (Si) is a group IV element that forms a diamond lattice, the same structure as zincblende lattice except all atoms are the same element. The empirical tight-binding parameters for bulk silicon are detailed in [111]. Figure 4.1a is the bulk band structure of silicon from NEMO3D. NEMO3D uses an eight atom unit cell rather than the diamond (or zincblende) primitive cell of two atoms. This results in Brillouin zone folding, where the true primitive cell band structure is folded due to the use of a larger supercell comprising multiple primitive cells [112]. The NEMO3D unit cell is made up of four primitive cells which correspond to a cube with the side length of two primitive cells in each of the three dimensions. If we consider the band structure in one dimension, the

NEMO3D Bloch state is summed over half as many unit cells as in the primitive cell case. As a result, the Brillouin zone in this dimension is half the size of the primitive Brillouin zone and each band is folded in half. For clarity, the high symmetry points of the NEMO3D Brillioun zone will be denoted L', Γ' and X' to differentiate them from those of the primitive cell Brillouin zone, L, Γ and X. Similarly, $\Gamma' \to X'$ corresponds to the [100] direction or Δ' and $\Gamma' \to L'$ corresponds to the [111] direction or Λ'.

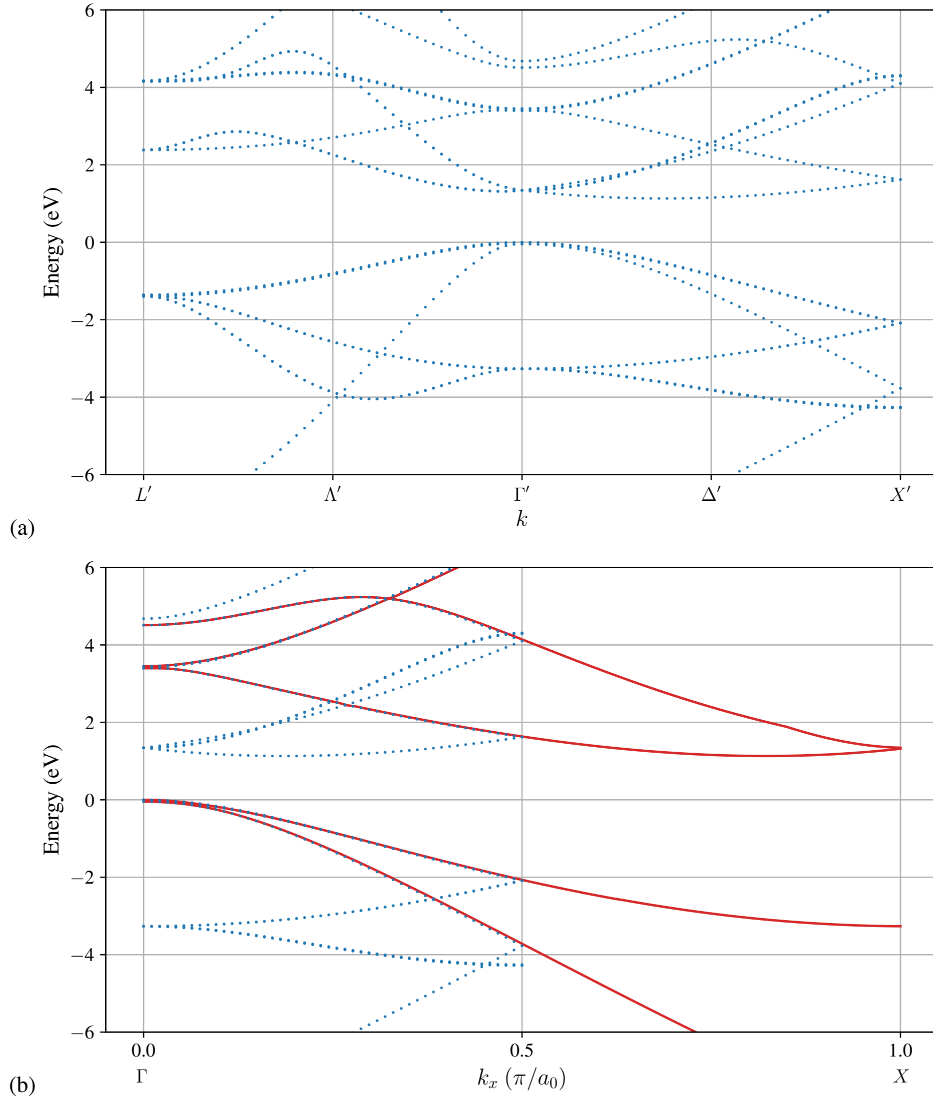


Figure 4.1. The bulk band structure of silicon from NEMO3D. (a) the zone folded band structure from $L' \to \Gamma' \to X'$ in the NEMO3D Brillouin zone. (b) The unfolded band structure in the standard primitive cell Brillouin zone. The solid red lines are the unfolded band structure and the blue dots are the NEMO3D eigenvalues.

An example of unfolding the NEMO3D band structure along the [100] direction is shown by the solid red lines in Figure 4.1b. There are some folded bands (blue dots)

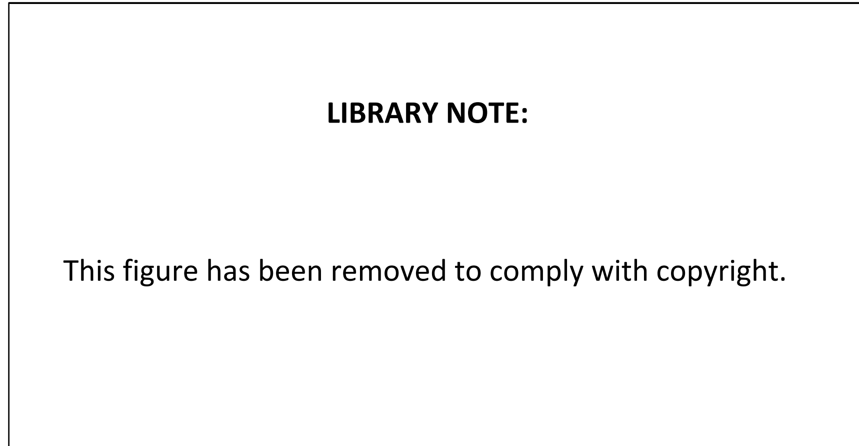


Figure 4.2. The bulk silicon band structure determined from experimental data. Figure adapted from [113].

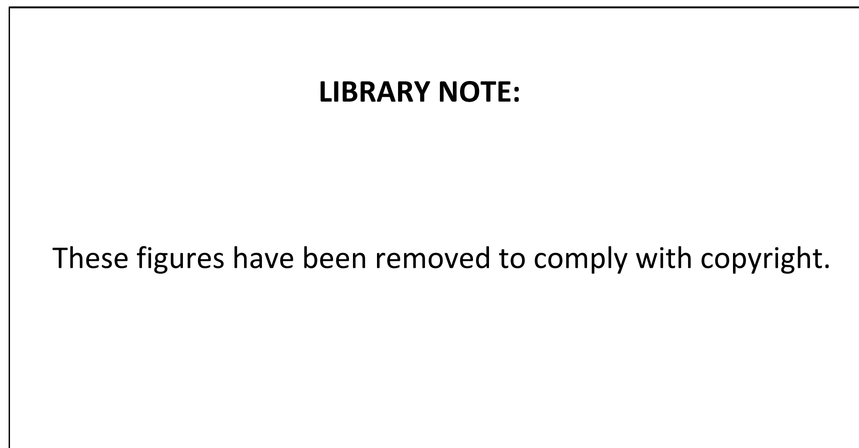
which are not accounted for in the unfolded bands (red lines). This is because we have only considered the unfolding of bands in one dimension. The zone folding of the full three dimensional band structure can lead to bands running in other directions being projected onto the [100] direction. The unfolding of the full three dimensional band structure requires extensive algorithms [112] which are outside the scope of this thesis. Silicon is an indirect band gap semiconductor, meaning that the conduction band minimum is not directly above the valence band maximum. The conduction band minimum occurs at around $k_x = 0.85\pi/a_0$ with an energy of 1 meV. The valence band maximum occurs at the Γ point with an energy of around 0 meV. These features can be seen more clearly in the unfolded band structure. The shape of the band structure agrees with the band structure formed from experimental results in Figure 4.2 and the band structure from [111]. The band gap of silicon from NEMO3D is $E_g = 1.13$ eV in good agreement with the experimental value of $E_g = 1.12$ eV [113].

4.3.2. Indium arsenide bulk band structure

Indium Arsenide (InAs) is a III-V semiconductor that forms either a zincblende or wurtzite lattice. In this thesis, InAs will be considered as a zincblende crystal structure. The empirical tight-binding parameters for the bulk InAs band structure are taken from [89]. Figure 4.3a is the bulk band structure of InAs from NEMO3D. Notice again the

presence of zone folding and the use of the prime notation to distinguish the symmetry points of the NEMO3D Brillouin zone from the standard Brillouin zone. The bands could be unfolded as done for silicon, however this is a time consuming process and the main features of the InAs band structure can be seen in the zone folded band structure.

The valence band maximum at the Γ' point is clearly shown near $0\,\text{eV}$ with a degeneracy of the light-hole and heavy-hole bands. The light- and heavy-hole bands split as k moves away from Γ' as expected and there is a clear split-off band below these. Unlike silicon, InAs is a direct band semiconductor so the conduction band minimum is located directly above the valence band maximum at around $0.8\,\text{eV}$. The NEMO3D band structure of InAs agrees with the experimental data band structure Figure 4.3b. The NEMO3D data finds a band gap of $E_g = 0.37\,\text{eV}$ and spin splitting $\Delta_0 = 0.39\,\text{eV}$ in agreement with the experimental values of $E_g = 0.35\,\text{eV}$ and $\Delta_0 = 0.38\,\text{eV}$ from [113].



(a) (b)

Figure 4.3. (a) The bulk band structure of indium arsenide from NEMO3D. (b) The bulk InAs band structure from experimental data [113].

The tight-binding parameters for silicon and indium arsenide accurately reproduce their respective bulk band structures. These parameters can now be applied to various nanostructure geometries. The next section is a brief demonstration of NEMO3D's ability to model the band structure of semiconductor nanostructures. The silicon tight-binding parameters are employed to determine the valley splitting of a quantum dot potential in a Si/SiGe quantum well under various electrical conditions. In Chapter 5, there is a more extensive analysis of the effective g-factor and spin-orbit coupling in an InAs nanowire using the InAs tight-binding parameters.

4.4. Valley splitting of quantum dot in Si/SiGe quantum well

The results of this section were included in a recent publication [95].

This section concludes the chapter and provides an example of modelling the electronic states of a semiconductor nanostructure with NEMO3D. Boross et al. [114] found that the two lowest valley states of an electron trapped in a Si/SiGe quantum well by a gate defined quantum dot experience an anti-crossing in the presence of a single-atomic layer step at the interface. The valley states of the electron were modelled by a hybrid model; an envelope-function approximation for the plane of the quantum well and a two band tight-binding model along the quantum well axis. Boykin et al. [115] verified the two band tight-binding model for the modelling of valley splitting in a silicon quantum well by comparing to NEMO3D simulations. The work of Gardin et al. [95], published during the author's candidature, found that the magnitude of this anti-crossing due to the interface step could be tuned by an out-of-plane electric field. The ability to control the anti-crossing between a two level system is an important criteria for building quantum logic gates [114]. The NEMO3D $sp^3d^5s^*$ nearest-neighbour tight-binding model was used to verify this novel result of the simpler two band model.

Unstrained silicon has sixfold degenerate valley states corresponding to the conduction band minima of the three p orbitals. These are known as valleys. The sixfold valley degeneracy can be lifted to a four-fold valley degeneracy and two-fold valley degeneracy by the strain in a Si/SiGe quantum well. In a quantum well, confined along the [001] direction the two lowest valleys are the z valleys. The two-fold valley degeneracy can be lifted by applying an electric field perpendicular to the plane of the quantum well. The splitting of these two valleys creates a two level system or valley qubit [36].

Figure 4.4 shows the Si/SiGe quantum well geometry. The width of the quantum well is N_z atomic layers before the step (green circle) and after the step the width of the quantum well is reduced by an atomic layer, $N_z - 1$. This step geometry is assumed to extend along the y direction, hence the geometry does not change in the y direction and it can be neglected. The quantum dot is defined by two top gates which can manipulate the electrons position in x. An electric field is applied along z confining the electron to the top interface where the step is located.

The hybrid model uncovered two features of this system. First, upon varying the position of the quantum dot with respect to the interface step an anti-crossing in the

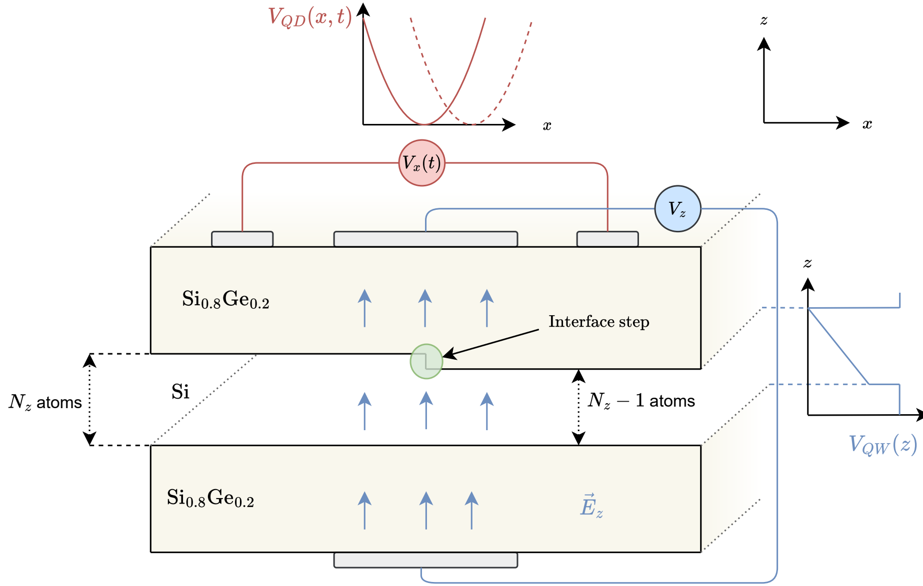


Figure 4.4. Geometry of the two-dimensional Si/SiGe heterostructure model with the y axis neglected. A gate controlled harmonic quantum dot potential V_{QD} is applied in x, and an electric field in z. The quantum well in the z axis, originating from the Si/SiGe interfaces, has a width of N_z atoms and decreases by one atomic layer at the step (green circle). Reprinted figure with permission from [95] copyright 2022 by the American Physical Society.

two lowest valley states can appear before or after the step due to valley orbit coupling. This was discovered and explained by Boross et al. [114]. Second, the magnitude of the anti-crossing between these valley states can be tuned by an electric field. An interesting aside is that the electric field dependence of the anti-crossing gap seems to change depending on the width of the quantum well, however, the reason for this is not completely understood. Neither of these features had been verified by a more realistic model like the $sp^3d^5s^*$ tight-binding model of NEMO3D. Hence, a key aspect of the work by Gardin et al. [95] was to reproduce these phenomena with NEMO3D.

The geometry modelled in NEMO3D was essentially the same as the hybrid model except it was three dimensional; constructed by extending the 2D geometry of Figure 4.4 along y. The hybrid model used hard wall boundary conditions for the quantum well interface, which corresponds to an infinite barrier height. This was modelled in NEMO3D by using a SiO2 interface, which has a large barrier height of 3 eV. To reduce simulation time, a tighter confinement was adopted for the quantum dot. The curvature of the quantum dot for NEMO3D simulations was set to 10^{-1} meV/nm^2 corresponding to an orbital level spacing of 9 meV while the hybrid model uses a parabola with an orbital

level spacing of 2 meV. The approach taken to validate the hybrid model follows that used by Boykin et al. [115]. Boykin et al. found the two-band tight-binding model to produce results in qualitative agreement with NEMO3D simulations of the valley splitting as a function of quantum well width. The valley splitting was of the correct order of magnitude but differed from the NEMO3D calculations. Similarly, in this work differences between the hybrid model and NEMO3D were expected, especially given the slight discrepancies between the models mentioned above. The focus of the validation was to find trends of NEMO3D data that agreed with the hybrid model and supported the conclusions about the anti-crossing and the control of the energy gap.

Figure 4.5 is the evolution of the two lowest valley states with respect to the quantum dot's distance from the step. The solid lines are the hybrid model and the filled circles are from NEMO3D with dashed lines as guides to the eye. The energy levels are labelled $\epsilon_{0,v}$ where 0 denotes the lowest orbital and the valley states are indexed by $v = \pm$. The spectrum shows that the step influences one valley state more than the other. Boross et al. [114] attributed this to the spatial oscillations that occur in the wavefunctions of the two valley states. The lower valley state has a maxima at the N_zth atomic layer of the quantum well while the higher valley state has a minima. Hence, when the dot is moved passed the step the lower valley state 'feels' the step and increases in energy while the other valley state does not 'feel' the step. When the two states meet near the step, an anti-crossing opens due to valley-orbital coupling induced by the step potential. The NEMO3D and hybrid model both demonstrate an anti-crossing between the valley states which can occur before the step, in the quantum well of width N_z, Figure 4.5b or after the interface step, in the region of width $N_z - 1$, Figure 4.5a. As expected there are slight differences between the models. In Figure 4.5a, the quantum well width is 16.5 monolayers for the hybrid model and 18 monolayers for NEMO3D, where a monolayer is two atomic layers or half a unit cell. In Figure 4.5b, the width is 21 monolayers for the hybrid model and 20 monolayers for NEMO3D.

Figure 4.6 is the anti-crossing gap plotted as a function of the out-of-plane electric field for the hybrid model (blue) and NEMO3D (red). The NEMO3D and hybrid model show similar trends. Both models find a minimum in the anti-crossing gap which varies with the width of the quantum well. The minimum for (a) is at around 7 mV/nm and (b) at 2.5 mV/nm. This agreement suggests that the anti-crossing gap can indeed be tuned by an out-of-plane electric field and the hybrid model is correct.

In conclusion, NEMO3D supports the results of the hybrid model. The valley states of a gate defined quantum dot in a Si/SiGe quantum well can be coupled when passing

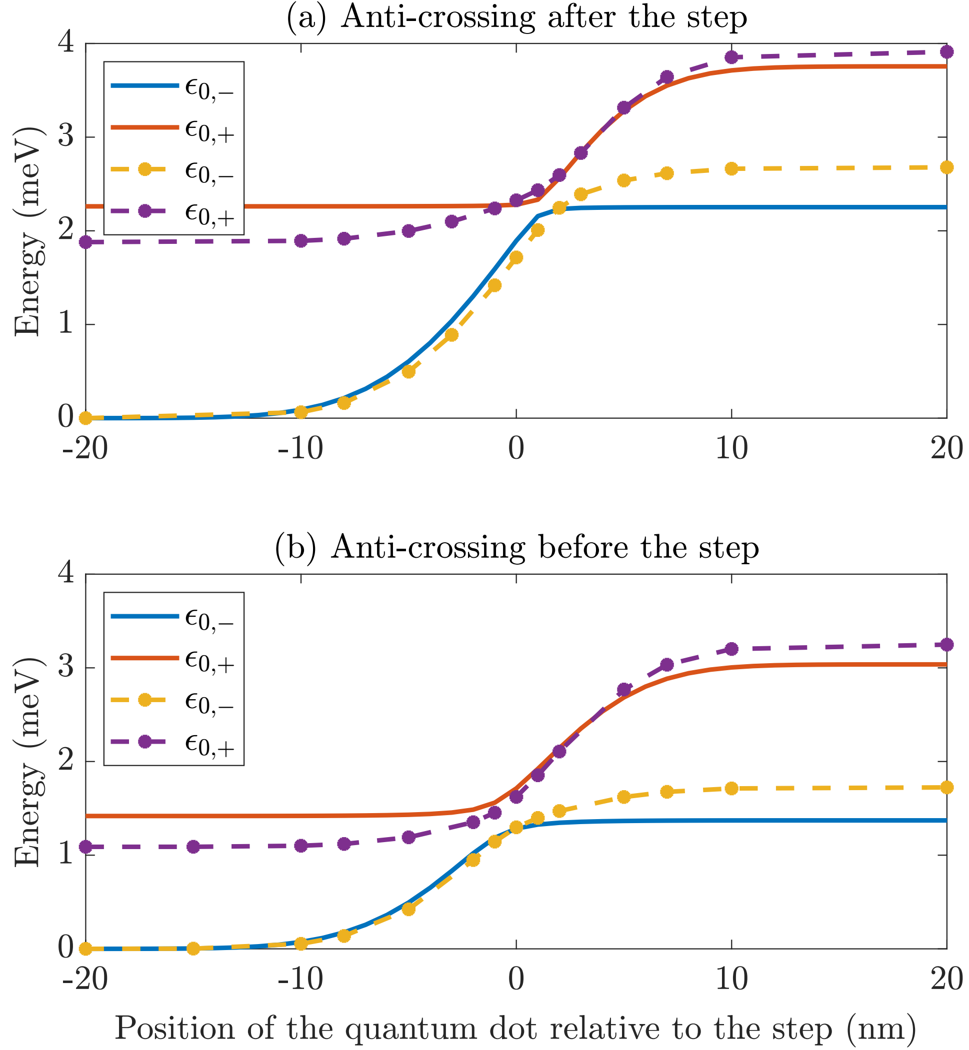


Figure 4.5. Evolution of the spectrum with the location of the quantum dot relative to the step at 0 nm. The anti-crossing can occur either after (a) or before (b) the step; a fact captured in both the hybrid model (solid lines) and NEMO3D's 20 band $sp^3d^5s^*$ model with spin (filled circles). The dashed lines are guides for the eye. The applied electric field is $E_z = 10\,\mathrm{mV/nm}$ for all curves in the figure. Reprinted figure with permission from [95] copyright 2022 by the American Physical Society.

through an interface step, producing an electrically tunable anti-crossing between the valley states. The control of the interaction between valley states is an important step towards making quantum gates from valley qubits. This work also demonstrates that the atomistic tight-binding NEMO3D can model the valley states in silicon quantum wells and validate simplified models.

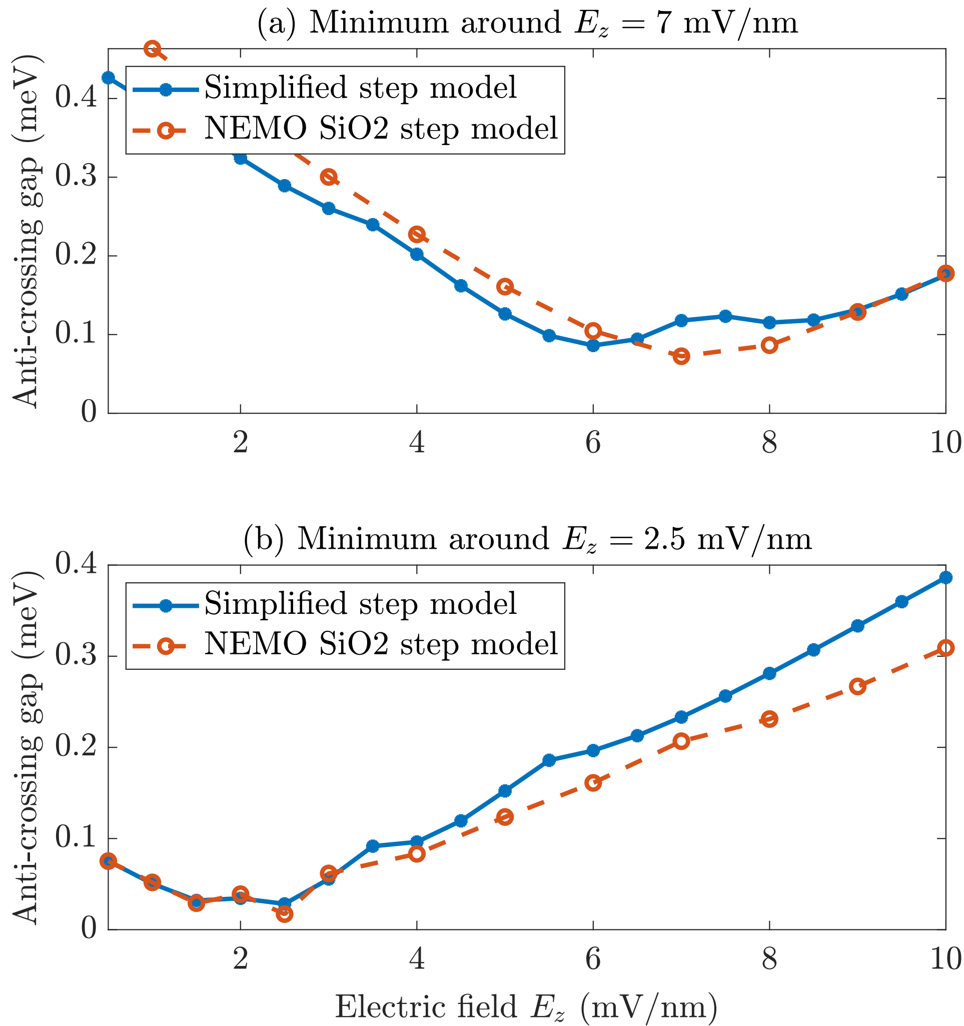


Figure 4.6. Evolution of the anti-crossing energy gap with applied electric field in the step model (solid-dotted blue) and NEMO3D simulations (dashed-circled red). The magnitude and location of the minimum of the gap can change depending on the quantum well width. (a) The hybrid model quantum well width was 16.5 monolayers, while the width for NEMO3D was 19 monolayers. (b) The hybrid model width was 19 monolayers while that of NEMO3D was 20 monolayers. Reprinted figure with permission from [95] copyright 2022 by the American Physical Society.

Chapter 5.

Effective g-factor anisotropy of indium arsenide nanowires

In this chapter, the atomistic nearest neighbour tight-binding model of NEMO3D is employed to model the effective g-factor anisotropy of an indium arsenide (InAs) nanowire (with a zincblende crystal structure) under various atomistic and electromagnetic conditions. From this g-factor anisotropy data it is possible to determine the spin-orbit interactions of the nanowire by fitting a one band effective model, which incorporates the well established Rashba and Dresselhaus spin-orbit interactions. This modelling approach provides insight for designing InAs nanowire devices with a desired spin-orbit interaction and effective g-factor anisotropy. The chapter begins by selecting a nanowire geometry and modelling the band structure of the nanowire to confirm the presence of subbands and the spin-orbit interaction. The majority of the chapter focuses on modelling the effective g-factor anisotropy and corresponding spin-orbit interaction of the nanowire's lowest conduction subband for different applied electric fields, surface ion terminations and nanowire confinement geometries. The chapter closes with a discussion of potential applications for these nanowires in the fields of spintronics and quantum information including spin-transistors, spin-orbit qubits and topologically protected qubits with Majorana zero modes.

5.1. Indium arsenide nanowire in NEMO3D

InAs is a III-V semiconductor typically of the zincblende crystal structure with a lattice constant $a_0 = 0.605\,83\,\mathrm{nm}$. The electron effective mass of InAs is $m^* = 0.023m_0$ where m_0 is the free electron mass. Other notable properties of InAs are its high carrier mobility (2–3$\times 10^4\,\mathrm{cm}^2/\mathrm{Vs}$ for electrons), narrow band gap of 0.35 eV, large effective g-factor of -15.3 and large spin-orbit coupling [113]. The electronic band structure of InAs nanostructures can be calculated with the NEMO3D $sp^3d^5s^*$ tight-binding model with the InAs empirical parameters from [89]. These empirical parameters were validated in Section 4.3 by reproducing the InAs bulk band structure as determined by experiments. In this section, two geometries of the InAs nanowire are introduced for modelling: periodic and closed boundary condition rectangular nanowires. Then an important atomistic feature of the III-V nanowires, surface ion termination, is introduced which will be helpful for interpreting the modelling results. The electronic band structure of the nanowire is then modelled with NEMO3D to check that the band structure produces quantised subbands and to identify the conduction subband of interest for this chapter.

5.1.1. Nanowire growth direction

InAs nanowires with a zincblende crystal structure are typically grown vertically from a substrate via molecular beam epitaxy or molecular organic crystal vapour epitaxy [116, 117]. The most common growth direction is the [111] crystal direction which naturally forms nanowires with a hexagonal cross-section. The hexagonal [111] nanowires often form a zincblende crystal structure with stacking faults; layers of the wurtzite crystal structure, which are undesirable for most applications [116]. III-V nanowires can also be grown along the [100] direction with a square/rectangular cross section [117], although this is less common because it requires a less energetically favourable lateral growth. The least common growth direction is the [110] direction with a rhombus/diamond shaped cross-section [118, 119].

For this work only nanowires grown along the [100] crystal growth direction are considered because it is the simplest crystal construction to model in NEMO3D. The [100] nanowire may also be advantageous over the [111] nanowire which is susceptible to stacking faults. It is of interest and left to future work to model the effective g-factor anisotropy and spin-orbit interactions present in the [110] and [111] nanowire.

5.1.2. Nanowire geometry

The [100] nanowires will be modelled as rectangular prisms with a square or rectangular cross-section, as shown in Figure 5.1, to match the shape of real [100] grown nanowires. The main geometry modelled will be a square cross-section nanowire of side length $L_y = L_z = L$. This L can be thought of as an effective diameter for comparison with the diameters of hexagonal cross-section [111] nanowires that are typically measured in experiments. Of course, this is an approximation and hence one cannot expect exact agreement unless a [111] hexagonal nanowire is modelled. The nanowire will also be modelled with a rectangular cross-section, $L_y > L_z$, to consider how the change in the shape of confinement influences the g-factor anisotropy. The nanowire can be modelled with periodic boundary conditions to determine the band structure of the nanowire, or closed boundary conditions to calculate the eigenstates of the system at the Γ point ($\mathbf{k} = \mathbf{0}$) as discussed in Section 4.2.

The periodic nanowire is created from a rectangular prism of InAs with a yz cross-section of dimensions L_y nm by L_z nm and the length of a unit cell, a_0, in x. The boundary conditions are closed in the y and z dimensions and the surface atoms are passivated with hydrogen atoms to avoid any surface effects. The periodic boundary conditions are applied in the x direction so that the box of unit cell length and cross-section $L_y \times L_z$ nm^2 is repeated along x to create a nanowire which is effectively infinite in the x direction. The periodic nanowire geometry is depicted in Figure 5.1a. This nanowire geometry can be solved efficiently due to the small number of atoms. However, it does encounter problems when applying magnetic fields perpendicular to the nanowire's periodic axis. A magnetic field perpendicular to the nanowire axis creates a vector potential component along the nanowire but, due to the periodic boundary conditions, the phase gained by the electrons from this vector potential is discontinuous.

The closed nanowire is a rectangular prism with the same cross-section as the periodic nanowire but with a length of 100 nm in the x direction. The boundary conditions are closed in all dimensions with the surface atoms passivated by hydrogen. The closed nanowire geometry is depicted in Figure 5.1b. For this work $L_z \leq L_y \leq 100$ nm. The nanowire length of 100 nm was chosen because it was one of the smallest lengths that captures the bulk properties of InAs along the nanowire, see Section 5.2. The length of the nanowire cannot be extended to realistic nanowire lengths, on the order of 1 µm, without the NEMO3D simulations becoming too computationally intensive to complete. However, a nanowire with a finite length of 100 nm can be solved within a day assuming L_y and L_z are not too large.

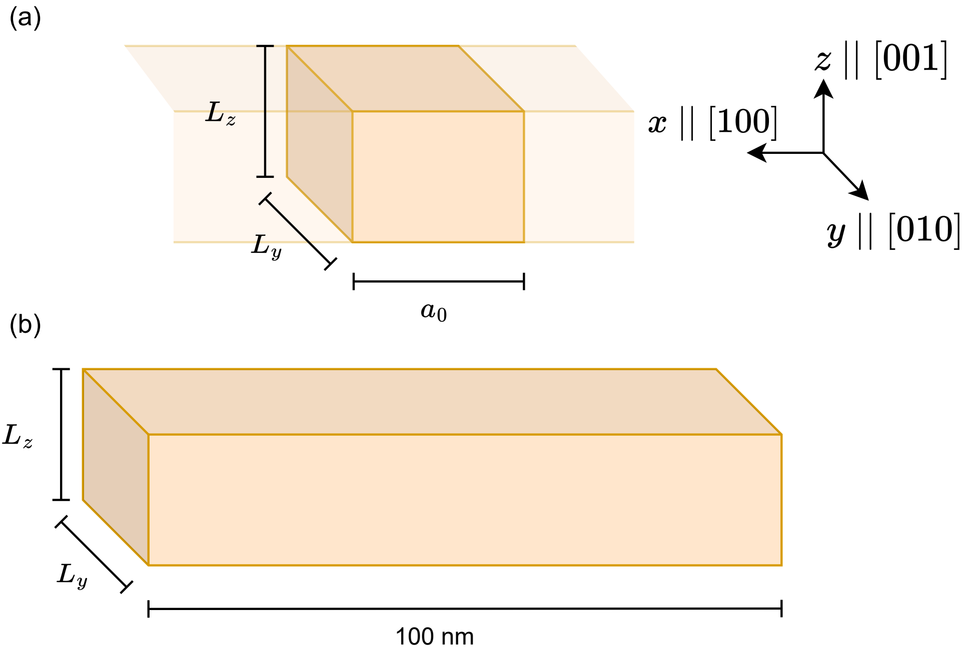


Figure 5.1. The geometry of the nanowire in NEMO3D simulations (not to scale) for (a) periodic boundary conditions in x and (b) closed boundary conditions x. The y and z boundary conditions are closed. L_y and L_z (in nm) are the dimensions in the y and z direction. a_0 is the lattice constant, the length of one unit cell. The nanowire unit cell is repeated along x in the periodic case.

Both the periodic and closed boundary condition nanowire geometries will be modelled in this chapter. The periodic nanowire is used to calculate the band structure of the nanowire and the closed nanowire is used for determining the effective g-factor anisotropy of the lowest conduction subband.

Surface Ion termination

By default, NEMO3D creates the nanowire from an eight atom unit cell of the zincblende lattice, four anions and four cations, as shown by the solid atoms in Figure 5.2. The opposing faces of this unit cell are made up of different ions, one side is the cation (In) and the other is the anion (As). For this reason, the default unit cell created by NEMO3D will be referred to as *mixed-ion terminating*. The standard zincblende crystal lattice with tetrahedral point group symmetry T_d is actually *same-ion terminating*; all surfaces are one element (either In or As) [106]. An example of same-ion termination is given by the addition of opaque atoms to the outer faces of the NEMO3D unit cell in Figure 5.2.

A mixed-ion terminating lattice has lower symmetry than the same-ion terminating lattice due to the differing potentials of the ions making up the opposing faces [106]. This is a spatial inversion asymmetry which can induce a spin-orbit interaction in the nanowire. Nanowire devices typically rest on substrates which induce spin-orbit interactions in a similar fashion to the mixed-ion termination. In attempts to gain a complete understanding of the spin-orbit interactions in the nanowire it is ideal to begin with the highest symmetry structure possible, the same-ion terminating nanowire with no substrate so that no spin-orbit interaction is induced from surface effects. Mixed-ion termination will be included later to simulate how a substrate influences the nanowire.

NEMO3D can create same-ion terminating nanowires by including an extra atomic layer of As atoms on the two outer faces of the nanowire in the confinement directions as depicted in Figure 5.2. Currently NEMO3D can only add atomic layers to two axes simultaneously, that is y and z can be same-ion terminating but not the x axis. Fortunately, this should not effect results because same-ion termination added along the axis of the nanowire x was found to have negligible influence on the effective g-factor anisotropy. For this chapter, it is assumed that same-ion termination is used in the confinement directions unless otherwise specified.

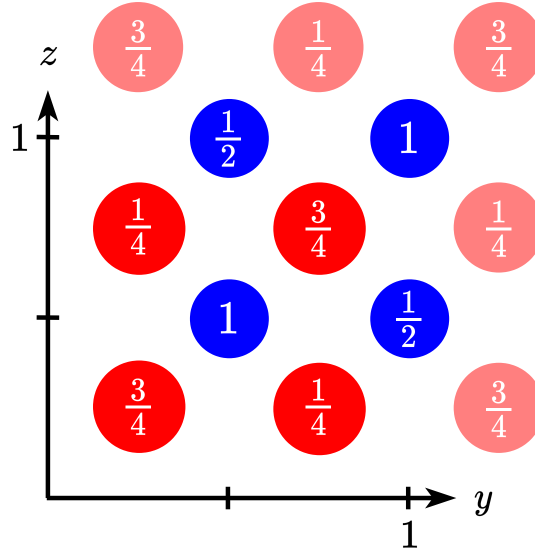


Figure 5.2. The structure of the NEMO3D unit cell looking down the x axis. The axes are in units of the lattice constant a_0. The white numbers on each atom denote the position in x. The cations (In) are in blue and the anions (As) in red. The solid atoms are the default mixed-ion terminating unit cell. The opaque atoms represent the process of adding anions to the outer surfaces to make it same-ion terminating.

5.1.3. Nanowire band structure

Figure 5.3 is the band structure resulting from the NEMO3D simulation of the periodic boundary condition InAs nanowire from Figure 5.2a with diameter, $L_y = L_z = L = 10\,\text{nm}$. The tight-binding Hamiltonian was solved for 20 eigenvalues (10 conduction and 10 valence band) at 40 k-space points along the periodic axis of the wire, [100], to give the nanowire band structure.

The confinement of an electron quantises the components of the wavevector $\mathbf{k}$ in the confinement directions and creates subbands in the band structure, as discussed in Section 2.1.2. Comparing the nanowire band structure to the bulk InAs band structure Figure 4.3, the nanowire confinement has lifted the conduction band minimum from 0.5942 eV in bulk InAs to 0.7264 eV and also lifted the valence band maximum in hole energy from 0.2243 eV to 0.1866 eV. The subbands present all appear to be from the first band of the conduction band and valence band given their similar shape. The other bands would be located higher in energy.

The inset of Figure 5.3 zooms in on the two lowest conduction subbands, E_1 and E_2 near $k_x = 0$. E_1 appears to be spin degenerate while the spin degeneracy of E_2 has been lifted to form two spin states $E_{2\uparrow,2\downarrow}$ where $\uparrow,\downarrow$ denotes the up or down spin state. Recall from Section 2.4.3 that the spin degeneracy can be lifted by the breaking of time-reversal symmetry or spatial inversion symmetry. Since there is no applied magnetic field this lifting of spin degeneracy must be the result of spin-orbit interactions breaking spatial inversion symmetry. From these results, one can conclude that the NEMO3D simulation is successfully capturing the physics of quantum confinement (subbands) and spin-orbit interaction (lifting of spin-degeneracy). However, it is hard to determine the magnitude of the spin-orbit interactions present from these band structure plots. A more effective method of determining the spin-orbit interactions present is by looking at the anisotropy in the nanowire's effective g-factor with magnetic field orientation. This method will be explored in the following sections.

The focus of this chapter will be modelling the effective g-factor anisotropy and spin-orbit interactions of the lowest conduction subband, E_1 (at $k_x = 0$), because it is the main subband of interest for nanowire devices used in Majorana zero mode research and spin-orbit qubits [22, 23, 15, 120]. The E_1 subband is typically studied because it is easy to access experimentally with doping or external gates. Additionally, for small confinement lengths L, the subband splitting is large and the higher subbands e.g. E_2 can be ignored. This is known as the lowest subband limit.

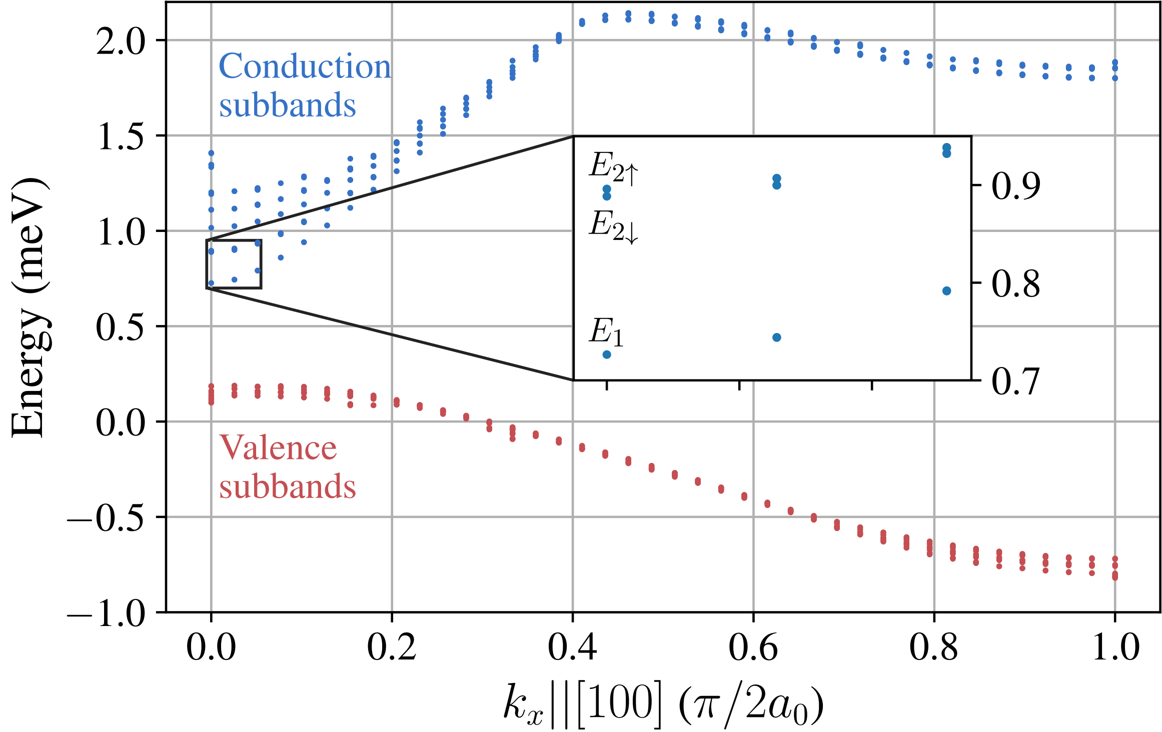


Figure 5.3. The NEMO3D band structure of the periodic InAs nanowire, Figure 5.2, with $L = 10\,\text{nm}$ when solving for 20 bands at 40 wavevector points, k_x, along [100]. The band structure shows quantised subbands due to the confinement of the nanowire. E_1 is the lowest conduction subband and $E_{2\uparrow}$ and $E_{2\downarrow}$ are the spin states of the second conduction subband.

5.2. Effective g-factor dependence on nanowire confinement

In this section, the lowest conduction subband, E_1, effective g-factor is modelled for the closed [100] InAs nanowire (see Figure 5.1b) with various square cross-section side lengths or 'diameters', $L = L_y = L_z$. The results show that the confinement of the nanowire influences the magnitude and sign of the effective g-factor, in agreement with previous works [106, 121] and the bulk g-factor of InAs, -15.3 [113]. They also confirm that the nanowire length of 100 nm is a good estimate of bulk InAs.

A magnetic field, $B = 2\,\text{T}$, is applied along the nanowire and the two spin states of the lowest conduction subband, $E_{1\uparrow,1\downarrow}$, are determined from NEMO3D for different nanowire diameters, L. The splitting of these states $\Delta E_1 = E_{1\uparrow} - E_{1\downarrow}$ is used to calculate

the effective g-factor

$$g^* = \frac{\Delta E_1}{\mu_B B}. \tag{5.1}$$

Figure 5.4 is the lowest conduction subband g-factor, g^*, as a function of the nanowire diameter L. For unit cell sized diameters, the nanowire g-factor approaches the electron spin g-factor, $g_0 \approx 2.002$. As the nanowire diameter increases, g^* decreases and flips sign at around 4.5 nm before approaching the bulk InAs g-factor for large L. At L =100 nm the absolute effective g-factor is approximately -14.8 in good agreement with the magnitude of the bulk g-factor. This result confirms that the length of 100 nm is a good approximation for the bulk direction of the closed nanowire.

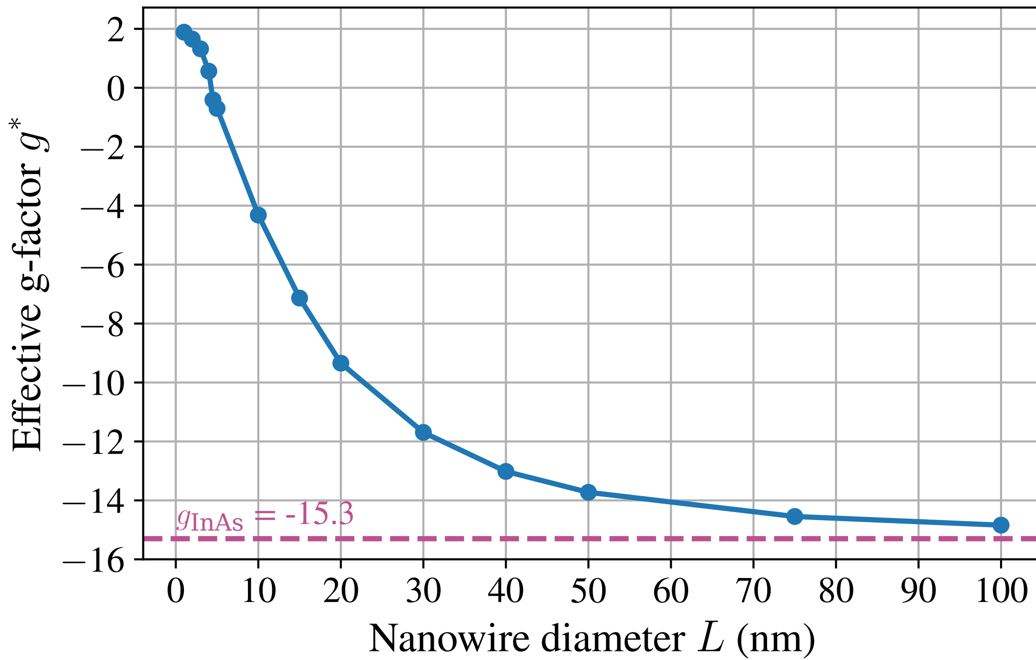


Figure 5.4. The effective g-factor for the lowest conduction subband E_1 in the closed [100] InAs nanowire as a function of nanowire diameter, L. The dashed line is the effective g-factor of bulk InAs, g_{InAs}.

NEMO3D outputs the eigenvalues in order of energy and not spin state. This means the difference of the two lowest conduction band eigenvalues is actually the absolute value of the spin splitting $|\Delta E_1|$ and hence the absolute value of the g-factor $|g^*|$. The sign of the effective g-factor was determined by checking the polarisations of the lowest conduction subband spin states. The expectation values of the components of the spin

operator, $\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}$, averaged over all atoms in the nanowire are given by

$$\begin{aligned} \langle S_i \rangle = \langle \Psi_n | S_i | \Psi_n \rangle &= \sum_{j,l}^{N} \langle \Psi_n | \mathbf{r}_j \rangle \langle \mathbf{r}_j | S_i | \mathbf{r}_l \rangle \langle \mathbf{r}_l | \Psi_n \rangle \\ &= \frac{\hbar}{2} \sum_{j,l}^{N} \sum_{\mu,\upsilon}^{20} \Psi^*_{n\mu}(\mathbf{r}_j)(\sigma_i)_{\mu\upsilon} \Psi_{n\upsilon}(\mathbf{r}_l) \end{aligned} \tag{5.2}$$

where $S_i = \frac{\hbar}{2}\sigma_i$ and σ_i are the Pauli matrices for $i = x, y, z$. $|\Psi_n\rangle$ is the nth eigenvector where n runs over the spin states of the subbands, j and l denote each of the N atomic sites and μ and υ index the atomic orbitals of which there are 20 for the $sp^3d^5s^*$ tight binding model with spin.

The average spin expectation values for the lowest subband spin states as functions of nanowire diameter are shown in Figure 5.5. The spins are mainly polarised in the x direction for small and large diameters as expected for a magnetic field applied along x. The spin polarisations in the x direction flip at around $L = 4.5$ nm and at this same point the expectation values $\langle S_y \rangle$ and $\langle S_z \rangle$ reach a peak, although the y polarisation is much smaller. The magnitude of $\langle S_z \rangle$ is close to $\hbar/2$ suggesting the spins flip through the z axis with a small y component. The energy splitting of the lowest conduction subband is defined as $\Delta E_1 = E_{1\uparrow} - E_{1\downarrow}$. From equation (5.1), g^* is positive for nanowire diameters less than 4.5 nm and negative for diameters larger than this. Hence, the effective g-factor from NEMO3D matches the electron g-factor g_0 for small nanowire diameters and agrees with the bulk InAs value -15.3 at large nanowire diameters, $L = 100$ nm.

These results agree with the works of [106, 121, 122]. Pryor et al. attribute the suppression of the g-factor magnitude with increased confinement to quenching of the orbital angular momentum of the electron [121]. This can be explained as follows. The spin-orbit interaction in the semiconductor couples the electron orbitals (or subbands) which creates an enhanced g-factor. The closer the orbitals are together the stronger the coupling and the larger the g-factor. For small confinement diameters, the energy splitting between the subbands (orbitals) becomes much larger than the Zeeman splitting of the spins resulting in less orbital coupling and a smaller g^*. In the limit of no orbital coupling, the g-factor is that of the electron's spin g_0. The effective g-factor of InAs nanowires has been observed experimentally to be larger then -15.3. However, this is mainly due to the nanowire chemical potential being large enough for electrons to occupy higher subbands which have stronger spin-orbit interactions and thus a larger g-factor

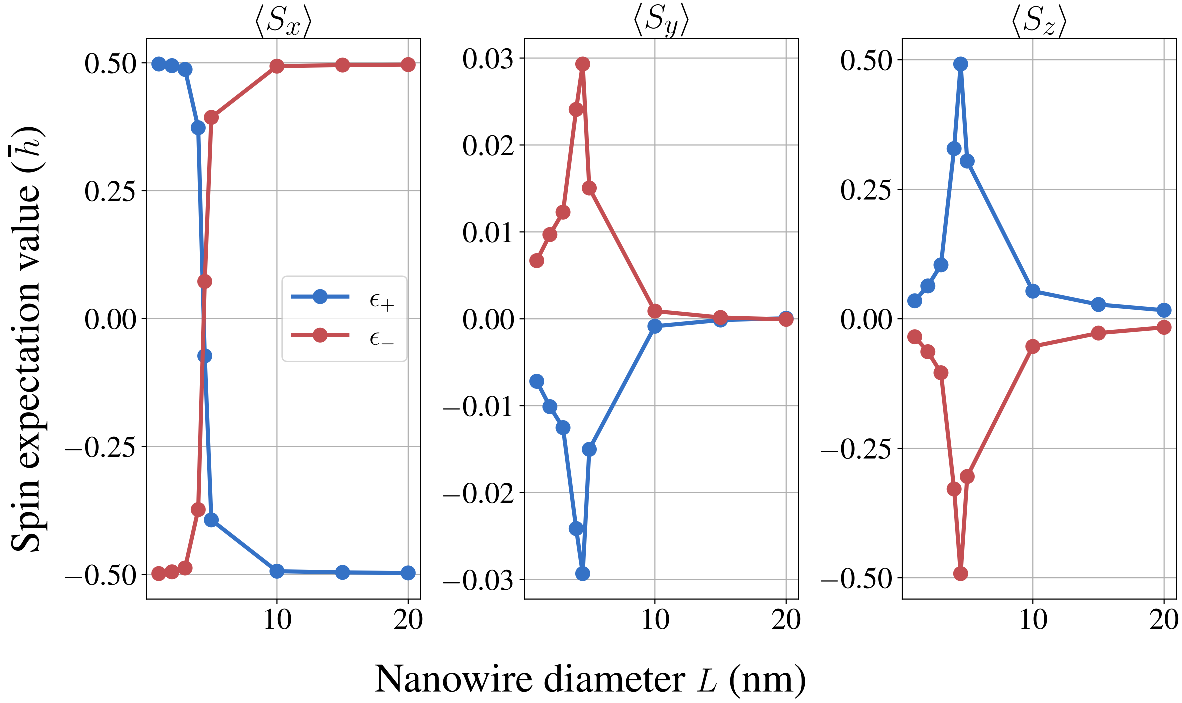


Figure 5.5. The spin expectation values (in units of $\hbar$) of the lowest conduction subband against the nanowire diameter L for a nanowire of length 100 nm with $B_x = 2$ T. The spin states flip at around $L = 4.5$ nm through the z axis shown by a sharp peak in $\langle S_z \rangle$.

[122]. For this work, the focus will be on the lowest conduction subband of the nanowire which should not have a g-factor larger than −15.3.

5.3. Effective model of the nanowire g-factor anisotropy

The effective g-factor anisotropy of the nanowire is a result of the spin-orbit interactions present in the InAs semiconductor. The g-factor anisotropy of the nanowire's lowest conduction subband, E_1, can be determined from NEMO3D. However, in order to understand the anisotropy and quantify the spin-orbit interactions present in the nanowire it is useful to have a model of the g-factor in terms of the well established semiconductor spin-orbit interactions, namely Rashba and Dresselhaus. This effective model can be fit to the NEMO3D results with the strength of spin-orbit interactions as free parameters. In this section, an effective model is derived for the lowest conduction subband g-factor of a [100] nanowire, based on the work of Sakr [123]. The model includes the Rashba

spin-orbit interaction due to applied electric fields and/or surface-ion termination and an intrinsic Dresselhaus term due to the zincblende crystal lattice. The same principle can be used to derive effective models for [110] and [111] nanowires. The only difference is the form of the Dresselhaus term due to the different orientation of the crystal lattice.

5.3.1. The nanowire geometry for g-factor anisotropy

The InAs nanowire geometry considered for the g-factor anisotropy is shown in Figure 5.6. The magnetic field orientation is varied in the plane of the nanowire (xy-plane) with the magnetic field angle, θ, measured from the nanowire axis. The shape of the g-factor anisotropy for the nanowire in this magnetic field plane is only influenced by the spin-orbit fields in said plane. Spin-orbit fields perpendicular to the magnetic field plane should only manifest as an offset to the anisotropy because they are perpendicular to the magnetic field for all θ. According to equation (2.39), an electric field induces a Rashba spin-orbit field in the $\hat{\mathbf{E}} \times \hat{\mathbf{k}}$ direction. The nanowire has $\mathbf{k} = k_x\hat{x}$ as the dominant wavevector, so any electric fields E (or mixed-ion termination) are applied perpendicular to the plane of the magnetic field to induce spin-orbit fields in said plane.

5.3.2. The effective Hamiltonian

The effective Hamiltonian of the nanowire is constructed by first considering a two dimensional Hamiltonian for a quantum well confined in the $z||[001]$ direction with Zeeman and spin-orbit interactions. The one-dimensional nanowire is then created by confining in the y direction with an infinite potential well to create the closed (or hard wall) boundary conditions. The resulting Hamiltonian is

$$H = \frac{\hbar^2(k_x^2 + k_y^2)}{2m^*} + H_Z + H_D^{\mathrm{2D}} + H_R^{\mathrm{2D}} + V(y), \tag{5.3}$$

with

$$V(y) = \begin{cases} 0 & 0 < y < L_y, \\ \infty & \text{otherwise.} \end{cases} \tag{5.4}$$

Here $(k_x, k_y) = -i(\partial_x, \partial_y)$. The vector potential is assumed to be negligible in the weak field limit where the confinement of the nanowire is smaller than the magnetic field length

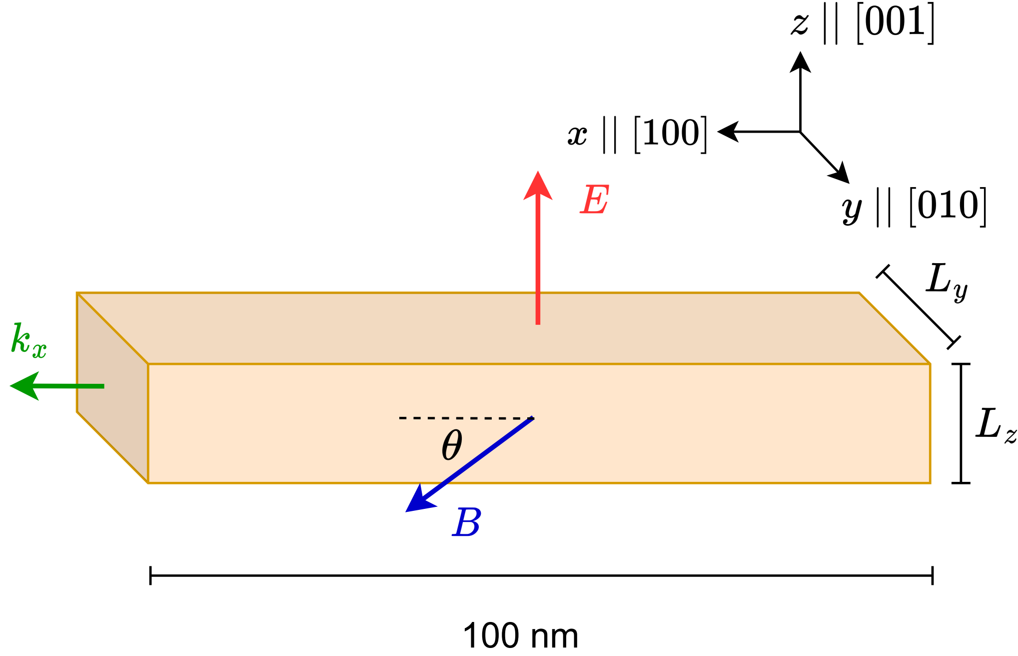


Figure 5.6. The [100] nanowire geometry for determining the g-factor anisotropy . A closed boundary condition wire with a square/rectangular cross-section of side length L_y and L_z. The magnetic field is applied in the xy-plane with angle θ measured from the nanowire axis. The electric field is applied perpendicular to the magnetic field plane and the unconfined wavevector of the nanowire, k_x.

($l_B = \sqrt{\frac{\hbar}{eB}}$) and the phase gained due to the vector potential is small [63]. Instead the magnetic field is introduced through the Zeeman term H_Z, (2.36), with g the quantum well g-factor, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$ and the magnetic field $\mathbf{B} = B\hat{n}$ with $\hat{n} = \cos\theta\hat{x} + \sin\theta\hat{y}$. The Dresselhaus and Rashba spin-orbit terms, $H_D^{2\mathrm{D}}$ and $H_R^{2\mathrm{D}}$, are that of the quantum well derived earlier, equations (2.40) and (2.41) respectively. For simplicity, k_z and the confinement potential in z are ignored by assuming the system is in the lowest quantum well subband which is valid if L_z is small.

The model starts from the quantum well system because: (1) the magnetic field is varying in the quantum well plane and only in-plane spin-orbit fields influence the shape of the g-factor anisotropy. (2) The linear in k quantum well spin-orbit terms are simpler than the bulk spin-orbit interactions, equations (2.38) and (2.39). Out-of-plane spin-orbit fields may influence the magnitude of the quantum well g-factor g but not the shape of the anisotropy. Additionally, for the $k_x = 0$ case, as in the NEMO3D results of the closed nanowire, the transverse k_y terms of spin-orbit interactions actually dictate the anisotropy.

5.3.3. Deriving the effective g-factor expression

An expression for the effective g-factor can be determined by first solving for the dispersion relation of the Hamiltonian, determining the lowest subband splitting and then dividing out by the Zeeman energy, as in equation (5.1). The Hamiltonian (5.3), can be solved by assuming that the spin-orbit terms are weak, so that they can be treated as a perturbation to the rest of the Hamiltonian. Following time-independent perturbation theory (Appendix A.5), the Hamiltonian is rewritten as an unperturbed Hamiltonian plus a spin-orbit perturbation

$$H = H_0 + H' \tag{5.5}$$

with

$$H_0 = \frac{\hbar^2(k_x^2 + k_y^2)}{2m^*} + \frac{g\mu_B B}{2}(\cos\theta\sigma_x + \sin\theta\sigma_y) + V(y) \tag{5.6}$$

and the perturbation $H' = H_R^{2D} + H_D^{2D}$. The eigenfunctions of H_0 take the separable form

$$\langle \mathbf{r} | n, \eta, k^{(0)} \rangle = e^{ikx}\phi_n(y)\,|\eta\rangle \tag{5.7}$$

with k the wavevector in the x direction, $\phi_n(y) = \langle y|n\rangle$ the nth transverse subband envelop function (or eigenfunction of the infinite potential well) and $|\eta\rangle$ the spin eigenfunctions along the direction of the magnetic field with spin index $\eta = \pm$. The transverse and spin eigenfunctions are complete sets and satisfy the orthogonality conditions: $\langle m|n\rangle = \delta_{mn}$ and $\langle \eta'|\eta\rangle = \delta_{\eta'\eta}$. The corresponding unperturbed eigenvalues take the form

$$E_{n\eta}^{(0)}(k) = \frac{\hbar^2 k^2}{2m^*} + \epsilon_n + \frac{g\mu_B B}{2}\eta, \tag{5.8}$$

where $\epsilon_n = \hbar^2 n^2 \pi^2 / 2m^* L_y^2$ are the energies of the transverse infinite potential well subbands in y.

When determining the eigenvalue expression (5.8) and the eigenvalue corrections below it is useful to apply a unitary transformation, $U = e^{-i\theta/2\sigma_z}$, to H as shown in Appendix A.4.2. This transformation corresponds to a counter-clockwise rotation of the spin basis by θ about the z axis such that $\cos\theta\sigma_x + \sin\theta\sigma_y \rightarrow \sigma_x$. In this new basis, σ_x polarises spins along the direction of the magnetic field $\hat{n}$, with eigenvalues η and eigenfunctions $|\eta\rangle$. The unitary transform, introduces a θ dependence to the spin-orbit

term H' which can be written as

$$\tilde{H}' = U^{\dagger} H' U = -\alpha_R(k_x(s\sigma_x + c\sigma_y) - k_y(c\sigma_x - s\sigma_y)) + \beta_D(k_x(c\sigma_x - s\sigma_y) - k_y(s\sigma_x + c\sigma_y)) \tag{5.9}$$

where the notation $c = \cos\theta$ and $s = \sin\theta$ has been introduced and the following transformations $e^{i\theta/2\sigma_z}\sigma_x e^{-i\theta/2\sigma_z} = c\sigma_x - s\sigma_y$ and $e^{i\theta/2\sigma_z}\sigma_y e^{-i\theta/2\sigma_z} = s\sigma_x + c\sigma_y$ from Appendix A.4.2 were used. In this new spin basis, the spin states are polarised along σ_x and the Pauli matrices act on the spin states as follows

$$\begin{aligned} \sigma_x \left|\eta\right\rangle &= \eta \left|\eta\right\rangle, \\ \sigma_y \left|\eta\right\rangle &= -i\eta \left|-\eta\right\rangle. \end{aligned} \tag{5.10}$$

Now the first order correction of the eigenvalues in time-independent perturbation theory is given by

$$E^{(1)}_{n\eta}(k) = \left\langle n, \eta, k^{(0)} \middle| \tilde{H}' \middle| n, \eta, k^{(0)} \right\rangle. \tag{5.11}$$

From the operation of Pauli matrices on the spin states $|\eta\rangle$, equation (5.10), only the σ_x terms of equation (5.9) survive in this expectation value. The transverse subband eigenstates are acted on by the k_y operator which results in $\langle n|k_y|n\rangle = \langle k_y\rangle_{nn} = 0$, from the definition of the infinite potential well eigenstates, equation (A.5). The remaining terms are linear in k, resulting in the first order correction

$$E^{(1)}_{n\eta}(k) = \left(-\alpha_R \sin\theta + \beta_D \cos\theta\right)\eta k. \tag{5.12}$$

This linear in k correction could also be derived from the typical nanowire spin-orbit terms that are used in literature, equations (2.42) and (2.43). However, since NEMO3D solves for eigenvalues at $k = 0$ this correction vanishes. In the $k = 0$ case, the anisotropy of the spin splitting and hence the effective g-factor comes from the second order correction to the eigenvalues.

The second order correction of the eigenvalues in time-independent perturbation theory is given by

$$E^{(2)}_{n\eta}(k) = \sum_{m\neq n}\sum_{\eta'} \frac{\left|\left\langle m, \eta', k^{(0)} \right| \tilde{H}' \left|n, \eta, k^{(0)}\right\rangle\right|^2}{E^{(0)}_{n\eta} - E^{(0)}_{m\eta'}}. \tag{5.13}$$

For the $k = 0$ case, only the k_y terms of equation (5.9) contribute to the second order correction. There are two types of matrix elements of the perturbation H' which contribute to the second order correction. These matrix elements depend on whether the spin indexes η and η' are equivalent or not. If they are equivalent, $\eta' = \eta$ only the σ_x terms of $\tilde{H}'$ survive

$$\langle m, \eta | \tilde{H}' | n, \eta \rangle = \langle k_y \rangle_{mn} \eta (\alpha_R c - \beta_D s). \tag{5.14}$$

If the spins are opposite, $\eta' = -\eta$, only the σ_y terms survive because they flip the spin of $|\eta\rangle$

$$\langle m, -\eta | \tilde{H}' | n, \eta \rangle = \langle k_y \rangle_{mn} (i\eta) (\alpha_R s + \beta_D c). \tag{5.15}$$

The k terms have been neglected because they vanish for $k = 0$. The $\langle m | k_y | n \rangle = \langle k_y \rangle_{mn}$ term is non-zero for certain subbands m and n which will be discussed below. The second order correction is then given by

$$E^{(2)}_{n\eta}(k = 0) = \sum_{m \neq n} \left[\frac{\left| \langle k_y \rangle_{mn} \right|^2}{\epsilon_n - \epsilon_m} (\alpha_R c - \beta_D s)^2 + \frac{\left| \langle k_y \rangle_{mn} \right|^2}{\epsilon_n - \epsilon_m + \eta E_Z} (\alpha_R s + \beta_D c)^2 \right] \tag{5.16}$$

where $E_Z = g\mu_B B$ and $\eta^2 = 1$. Only the second term of this correction has η dependence and hence only this type of term contributes to the spin splitting. The expectation value $\langle k_y \rangle_{mn}$ couples the current subband n to another subband m. For the infinite potential well eigenstates (A.5) this expectation value is given by

$$\langle k_y \rangle_{mn} = \begin{cases} 0, & \text{if } m \text{ and } n \text{ have same parity,} \\ -\frac{4imn}{L_y(m^2 - n^2)}, & \text{if } m \text{ and } n \text{ have opposite parity.} \end{cases} \tag{5.17}$$

The largest coupling for the lowest subband $n = 1$ is $\langle k_y \rangle_{21}$ since the higher subbands $n = 4, 6, ...$ are further away in energy. The dispersion relations can be determined to good approximation by considering this coupling, $\langle k_y \rangle_{21}$, only because the coupling to higher subbands is negligible in comparison.

From (5.8) and (5.16), the effective g-factor of the lowest conduction subband up to second order in the spin-orbit coupling perturbation is

$$g^* = g\left\{1 + \frac{2\left|\langle k_y\rangle_{21}\right|^2}{E_Z^2 - (\epsilon_1 - \epsilon_2)^2}(\alpha_R^2 + \beta_D^2)\sin^2(\theta + \phi)\right\} \tag{5.18}$$

where the phase offset ϕ has been introduced such that $\tan\phi = \frac{\beta_D}{\alpha_R}$. To second order, the ratio of spin-orbit interaction strengths affects the anisotropy by introducing a phase shift and the strength of the spin-orbit interactions determines the anisotropy amplitude.

Consider the plot of the effective g-factor expression (5.18) for three different ϕ in Figure 5.7. If $\phi = 0$, only Rashba present, the second term of the g-factor takes a $\sin^2\theta$ shape and if $\phi = \pi/2$, only Dresselhaus present, the second term has a $\cos^2\theta$ shape. If both Rashba and Dresselhaus are present at equal strength the amplitude is doubled and the phase shift is $\phi = \pi/4$. It is also worth noting that if the subband splitting is larger than the Zeeman splitting, $|\epsilon_1 - \epsilon_2| > E_Z$, as it is for all simulations in this work, the second term is negative and g is the maximum value of the anisotropy. The resulting anisotropy for the Rashba only case is effectively a '$1 - \sin^2\theta = \cos^2\theta$' shape and for the Dresselhaus only case is actually a '$1 - \cos^2\theta = \sin^2\theta$' shape.

The derivation of this effective model demonstrates that the g-factor anisotropy at $k = 0$ is created by the coupling between the transverse subbands. This coupling is due to the confined wavevector k_y terms in the spin-orbit interactions. If $k \neq 0$, the standard first order k_x terms of Rashba and Dresselhaus would dominate the anisotropy.

Comparison to experiment

A common experimental method for determining the orientation of spin-orbit fields in a nanowire, used by Wang et al. [24] and Nadj-Perge et al. [23], is to find the minimum of the magnetic field anisotropy of an anti-crossing between singlet and triplet states of an InAs or InSb nanowire double quantum dot. The shape of these anisotropies is analogous to that of the g-factor anisotropy. The groups made their measurements on [211] and [111] grown nanowires, respectively. Both papers make use of the result from two band models of the two electron system, which state that the orientation of the spin-orbit field in the nanowire – whether that be Rashba, Dresselhaus or a combination – corresponds to the magnetic field direction when the anti-crossing between singlet and triplet states is a minimum [120, 124]. This is similar to the theory of the helical states in Section 2.4.4.

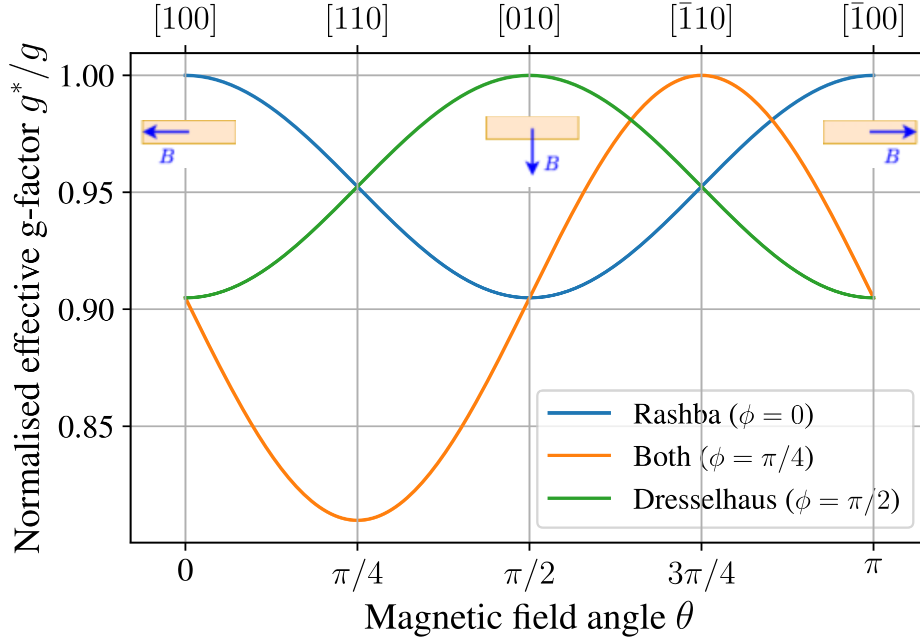


Figure 5.7. The effective model g-factor anisotropy normalised by g with $L_y = 10\,\text{nm}$ and $B = 1\,\text{T}$ for three different cases: $\phi = 0$, Rashba only with $\alpha_R = 0.1\,\text{eVnm}$; $\phi = \pi/4$, equal Rashba and Dresselhaus $\alpha_R = \beta_D = 0.1\,\text{eVnm}$; and $\phi = \pi/2$, Dresselhaus only with $\beta_D = 0.1\,\text{eVnm}$. The cartoons indicate the direction of the magnetic field with respect to the nanowire from Figure 5.6 at different θ.

When the magnetic field and spin-orbit field are perpendicular, the anti-crossing between the k split spin-states is the largest and when they are parallel, it is a minimum. Wang et al. attribute the minima to a Rashba SO field due to the substrate beneath the nanowire and a possible out-of-plane Dresselhaus component due to the [211] nanowire crystal structure [24]. Nadj-Perge et al. finds a minimum in the g-factor anisotropy when the magnetic field is oriented almost perpendicular to the nanowire and they conclude this is due to a spin-orbit field. Given the nanowire is zincblende and grown along the [111] direction they conclude that there is no presence of Dresselhaus and that the spin-orbit is due purely to a Rashba interaction induced by the substrate [23].

The effective model derived here agrees with this intuition that the g-factor is at a minimum when the magnetic field is parallel with the spin-orbit field. The Rashba dominated shape in Figure 5.7 has a minimum at $\theta = \pi/2$ perpendicular to the nanowire as expected for the Rashba spin-orbit interaction, equation (2.43). The Dresselhaus dominated shape on the other hand has a minimum along the nanowire, $\theta = 0$ or π, which agrees with the orientation of the Dresselhaus spin-orbit interaction for a nanowire along [100], equation (2.42).

5.4. Effective g-factor anisotropy of square cross-section nanowires

In this section, the effective g-factor anisotropy of the square cross-section [100] InAs nanowire's lowest conduction subband, E_1, is calculated with NEMO3D for various applied electric fields, surface ion terminations and confinement lengths. The effective model expression for the g-factor, equation (5.18), is fit to the NEMO3D g-factor anisotropies with g, α_R and β_D as free parameters. The effective model fit will determine the spin-orbit composition of the nanowire and assist in the interpretation of the NEMO3D results.

5.4.1. Electric field dependence

A same-ion terminating [100] InAs nanowire with no substrate is expected to have Dresselhaus spin-orbit interaction due to the bulk semiconductor crystal along the nanowire but no Rashba spin-orbit interaction. One way to induce Rashba in such a nanowire is to apply an external electric field. If the electric field is applied perpendicular to the magnetic field plane (in the z direction), from equation (2.39) the induced Rashba should be in the plane of the applied magnetic field and influence the g-factor anisotropy. From equation (2.41), the strength of Rashba should have a linear dependence on the applied electric field, $\alpha_R = \alpha_0 \langle E \rangle$. According to equation (2.40), Dresselhaus should be independent of the applied electric field.

These hypotheses were tested by performing NEMO3D simulations of the g-factor anisotropy for a square cross-section, $L = L_y = L_z = 10\,\mathrm{nm}$, same-ion terminating InAs nanowire with an applied magnetic field of magnitude, $B = 5\,\mathrm{T}$, and various applied electric fields E in the configuration of Figure 5.6. The nanowire cross-section side length of 10 nm was chosen over more realistic lengths of 50 nm to 80 nm because the smaller number of atoms allows for efficient computation times, typically less than an hour. The large magnetic field of 5 T was chosen so that spin states of the lowest subband were easily resolved by NEMO3Ds Lanczos algorithm. For large subband splitting, the g-factor anisotropy should be independent of the magnitude of the magnetic field as in equation (5.18) which was confirmed for the 10 nm nanowire in Appendix E.1.

Figure 5.8 is the effective g-factor anisotropy of the nanowire's lowest conduction subband, E_1, for various applied electric fields, E. The filled circles are data from the tight-binding model and the solid lines are the effective model fit to the data with the fit

parameters g, α_R and β_D specified in the plots. The inset of each plot is a y slice of the probability density of the lowest subband spin state, $|\psi_{1\downarrow}|^2$, at the centre of the nanowire. The shape of the anisotropy for the $E = 0$ case has a $\phi = \pi/2$ phase shift, resembling the Dresselhaus shape of the effective model in Figure 5.7. The fit parameters indicate that Rashba is negligible, $\alpha_R = 0\,\text{eVnm}$, and Dresselhaus is present, $\beta_D = 0.2\,\text{eVnm}$. This is as expected for the same-ion terminating nanowire with no structural inversion asymmetry from a substrate, mixed-ion termination or an applied electric field.

Figure 5.9 is the spin-orbit constants, α_R and β_D, from the effective model fits in Figure 5.8 as functions of electric field E. For non-zero electric fields Rashba is induced, however, it is only approximately linear with electric field for small applied electric fields ($E < 10\,\text{mV/nm}$). After about $20\,\text{mV/nm}$, α_R begins to decrease with electric field. This is observed in the g-factor anisotropy plots of Figure 5.8 as the phase shift ϕ decreasing for $E < 10\,\text{mV/nm}$ and then increasing towards $\phi = \pi/2$ again after $E = 20\,\text{mV/nm}$. A linear regression fit (purple dot-dashed line) to the values of α_R for the first three electric fields gives an estimate of the value of α_0, $12.08\,\text{enm}^2$. This value is almost exactly an order of magnitude larger than the value predicted by $\mathbf{k}\cdot\mathbf{p}$ theory for an InAs quantum well, $\alpha_0 = 1.171\,\text{enm}^2$ [51] and a few orders of magnitude larger than that of experimental works [125]. The relation $\alpha = \alpha_0 \langle E \rangle$, is an approximation for a quantum well system, so the disagreement in the values of α_0 is likely due to the contribution of the extra confining potential present for a nanowire.

Counter to the original hypothesis, the Dresselhaus constant increases almost linearly with electric field and experiences more than a three fold increase over $50\,\text{mV/nm}$. This can be seen in Figure 5.8, by the $\phi = \pi/2$ g-factor anisotropy increasing in amplitude for large electric fields. A linear regression fit (black dotted line) to the Dresselhaus constant in Figure 5.9 found a linear Stark coefficient (slope) of $10.6\,\text{enm}^2$ similar to that of the Rashba constant in the small electric field region.

These results can be understood by observing the behaviour of the probability densities of $E_{1\downarrow}$ with electric field in Figure 5.8. As the applied electric field increases, the electron wavefunction is confined toward the bottom surface of the nanowire. Recall from equation (2.40), that the Dresselhaus constant is $\beta_D \propto \langle k_z^2 \rangle$ and for infinite potential well envelop functions $\langle k_z^2 \rangle \sim (\pi/L_z)^2$. Hence, increasing the confinement of the wavefunction in the z direction, which corresponds to reducing L_z, results in an increase in the Dresselhaus constant, β_D. The decrease in the Rashba constant, α_R, with electric field after around $20\,\text{mV/nm}$ can be explained by the electron wavefunction being pushed away from the top surface of the nanowire. It seems that if the wavefunction probability

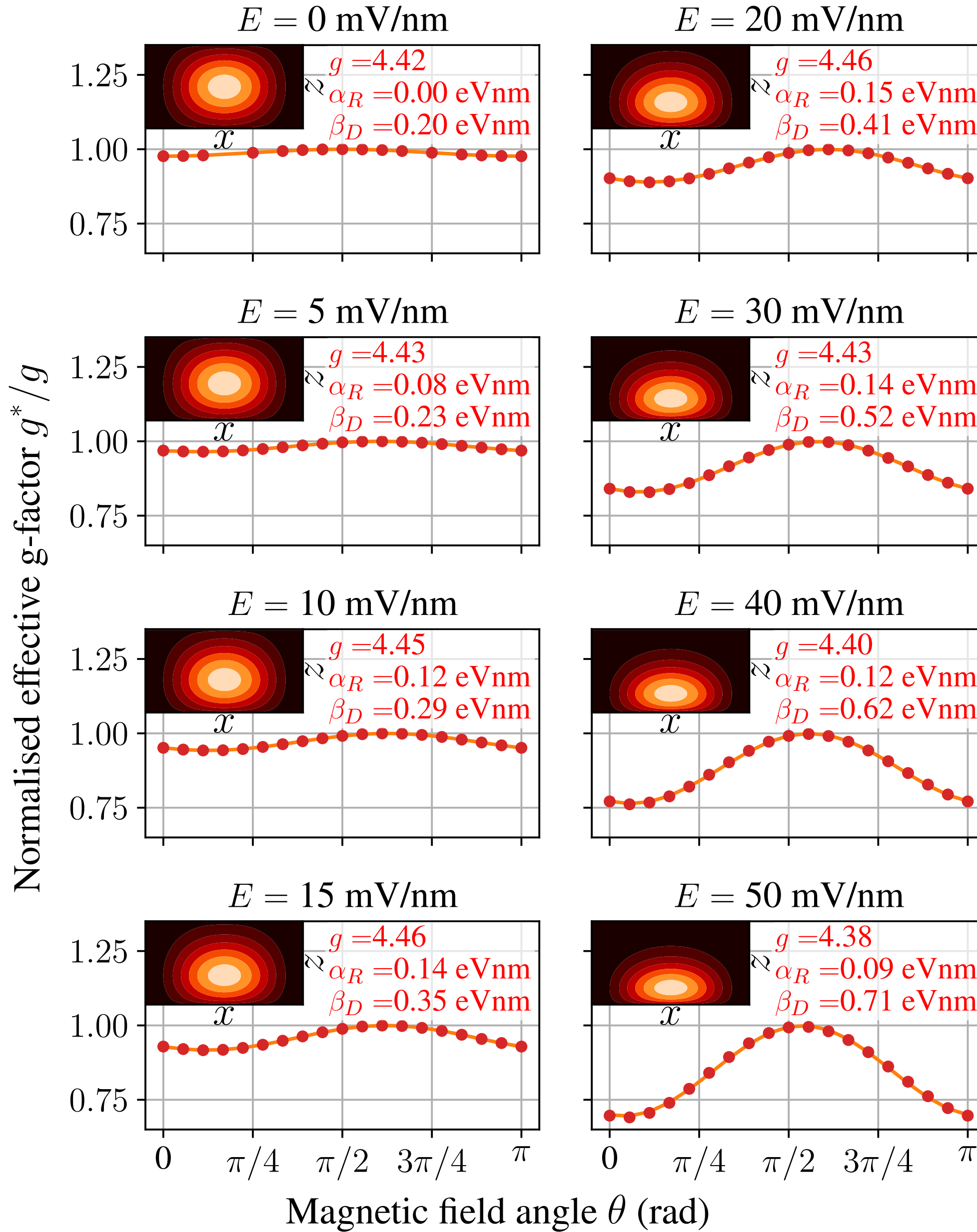


Figure 5.8. The normalised g-factor anisotropy of E_1 for the same-ion terminating, square cross-section $L = L_y = L_z = 10\,\text{nm}$ InAs nanowire for different electric field strengths E. The red filled circles are the data points from the tight-binding $sp^3d^5s^*$ model and the orange solid lines are the effective model fits to the data. The fit parameters of the effective model, g, α_R and β_D are displayed in red. The top left inset of each plot is a y-slice of the probability density of the electron's $E_{1\downarrow}$ state in the nanowire with dimensions x and z from Figure 5.6.

is not centred between the nanowire surfaces it no longer feels the asymmetry due to the external electric field. Additionally, Moroz and Barnes [126] showed that a confinement

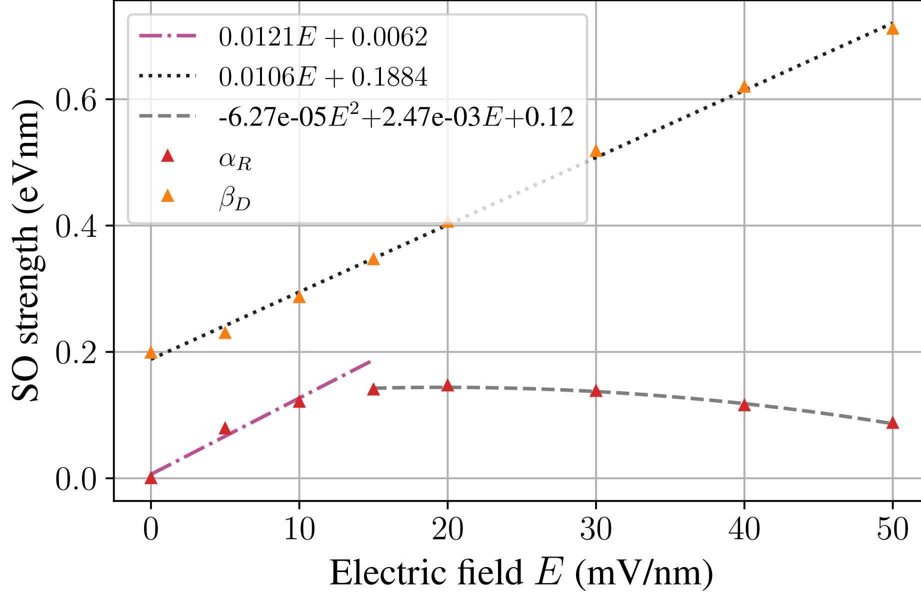


Figure 5.9. The Rashba and Dresselhaus spin-orbit strength for the same-ion terminating nanowire as a function of electric field E. The fit parameters, α_R and β_D, from Figure 5.8 are plotted against electric field. The linear and quadratic regressions for the data are displayed as dashed and dotted lines.

potential can produce an electric field, $E_{\mathrm{con}} = -\boldsymbol{\nabla} V_{\mathrm{con}}(z) \propto z$, which can induce a spin-orbit interaction in addition to that produced by the external electric field. Since the position of the electron wavefunction changes in z with external electric field the strength of this confinement electric field will also change which could be a reason for the lack of linearity in α_R's external electric field dependence. A quadratic fit (grey dashed line) to α_R as a function of electric field after the wavefunction begins to shift, gives the quadratic and linear Stark coefficients $-62.7\,\mathrm{enm}^3/\mathrm{V}$ and $2.47\,\mathrm{enm}^2$.

The results presented here seem to agree with the experimental works of Liang et al. [79, 127]. They observed that applying an electric field to a gated nanowire did not increase Rashba but rather decreased it. The solution to this was to employ a double gate setup, a potential applied between a top and bottom gate, which ensured that an electric field could be applied without altering the electron density. The Rashba was tuned from $0.5 \times 10^{-11}\,\mathrm{eVm}$ to $3 \times 10^{-11}\,\mathrm{eVm}$ with an applied electric field on the order of $10^6\,\mathrm{V/cm}$ or $10^2\,\mathrm{mV/nm}$ [127]. One could implement a double gate potential in NEMO3D by first solving the applied potential across a double gate setup self consistently for a given nanowire geometry and then applying this potential to the NEMO3D tight-binding Hamiltonian. This modelling could validate the experimental results of Liang et al. and allow for further electrical tuning of Rashba.

5.4.2. Mixed-ion termination

The previous section revealed that an external electric field can induce a Rashba spin-orbit interaction provided the electron wavefunction is not shifted too far from the nanowire surface. Rashba can also be induced by an electric field due to asymmetric confinement i.e. a difference between the intrinsic potentials of the surface/interface at opposite faces of the nanowire. One way to create this asymmetric confinement is to construct the nanowire with mixed-ion termination (see Section 5.1.2) on at least one set of opposite surfaces. The ions on opposite faces of the nanowire have different surface potentials creating a structural inversion asymmetry which can induce Rashba . In this section, the square cross-section $L = 10\,\text{nm}$ nanowire is modelled as in Section 5.4.1 but with mixed-ion termination on opposing nanowire faces normal to the z axis. This will be referred to as mixed-ion termination along z. The atom positions in the cross-section of the nanowire for the same-ion termination and mixed-ion termination in z cases are shown in Figure 5.10.

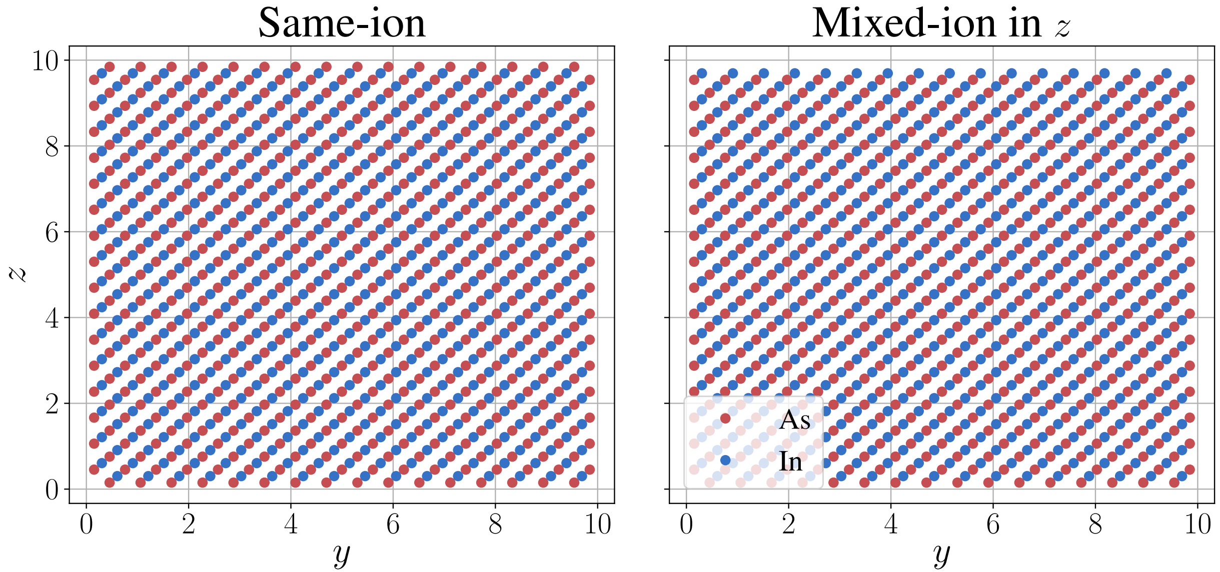


Figure 5.10. The atom positions of two neighbouring x slices (333rd and 334th atom) of the square cross-section $L = 10\,\text{nm}$ nanowire plotted on the same axes for: (left) the same-ion termination case and (right) the mixed-ion termination in z case.

The effective g-factor anisotropy for E_1 of the mixed-ion terminating $L = 10\,\text{nm}$ nanowire is simulated in NEMO3D for various applied electric fields E and the effective model is fit to the results. Figure 5.11 shows the results in the same format as in Figure 5.8 with the NEMO3D data as filled circles, the effective model fit as solid lines

and $|\psi_{1\downarrow}|^2$ in the insets. For the $E = 0$, case the anisotropy has a phase shift of $\phi \approx \pi/4$ representing the presence of both Rashba and Dresselhaus. The Rashba constant, α_R, is more than double the maximum α_R achieved with an external electric field in the same-ion terminating nanowire. The Dresselhaus strength, β_D, is almost double that of the same-ion case. From this result one can conclude that the mixed-ion termination induces a Rashba perpendicular to the nanowire (x) and the mixed-ion termination (z), along the y direction. The fact that the mixed-ion termination increases the Dresselhaus spin-orbit interaction present in the nanowire suggests that the mixed-ion termination could be an interface inversion asymmetry as discussed in [128].

Figure 5.12 is the spin-orbit constants from the effective model fits of Figure 5.11 as functions of the applied electric field. The Dresselhaus constant increases with electric field, as it did in the same-ion terminating case, due to the confinement of the electron wavefunction towards one surface. The Rashba constant decreases with electric field as before. Unlike the same-ion termination case, there is no increase in Rashba for small electric fields. This is because the mixed-ion termination Rashba constant at $E = 0$ is larger than the electric field Rashba constant and likely more sensitive to the electron wavefunction position between mixed-ion terminating surfaces. Higher electric fields were applied for the mixed-ion terminating nanowire. As E approaches 100 mV/nm, Rashba approaches zero while Dresselhaus reaches 1 eVnm. The linear Stark coefficient for the Dresselhaus constant, β_D, is 7.53 enm^2 which is smaller than in the same-ion nanowire. For the Rashba constant a quadratic regression revealed a quadratic Stark coefficient of 24.1 enm^3/V and a linear Stark coefficient of -6.40 enm^2. The sign of the stark coefficients have changed from the same-ion terminating nanowire to the mixed-ion termination nanowire. The Rashba constant in the mixed-ion termination case experiences more of a linear decrease. At higher electric fields the effective g-factor anisotropies in Figure 5.11 begin to deviate from the effective model. This is likely due to higher order effects and/or the violation of the weak spin-orbit interaction assumption of the effective model, since $\beta_D = 1$ eVnm is large. Hence, the effective model may be limited to modelling spin-orbit interaction fields less than 1 eVnm.

Adding a substrate

The introduction of mixed-ion termination is one way of inducing an asymmetric confinement potential in the z direction. Another option is to add a substrate. The effective g-factor anisotropy of the nanowire's E_1 was also modelled with a SiO2 substrate on one

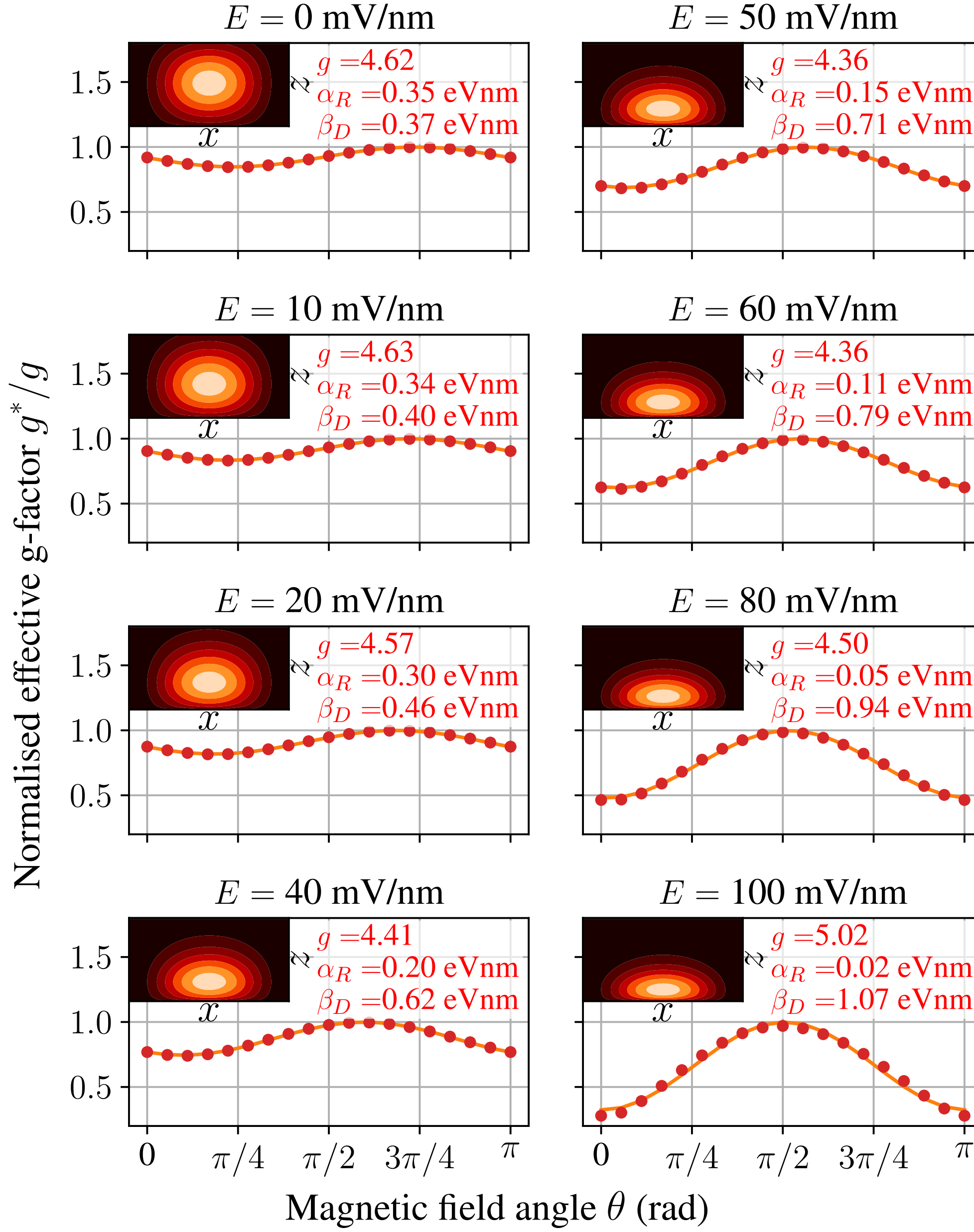


Figure 5.11. The normalised effective g-factor anisotropy of E_1 for the mixed-ion termination along z, $L = 10\,\mathrm{nm}$ nanowire with different external electric field strengths E. The red filled circles are the data points from the tight-binding $sp^3d^5s^*$ model and the orange solid lines are the effective model fit to the data. The fit parameters of the effective model, g, α_R and β_D are displayed in red. The top left inset is a y-slice of the probability density of the electron's $E_{1\downarrow}$ state in the nanowire.

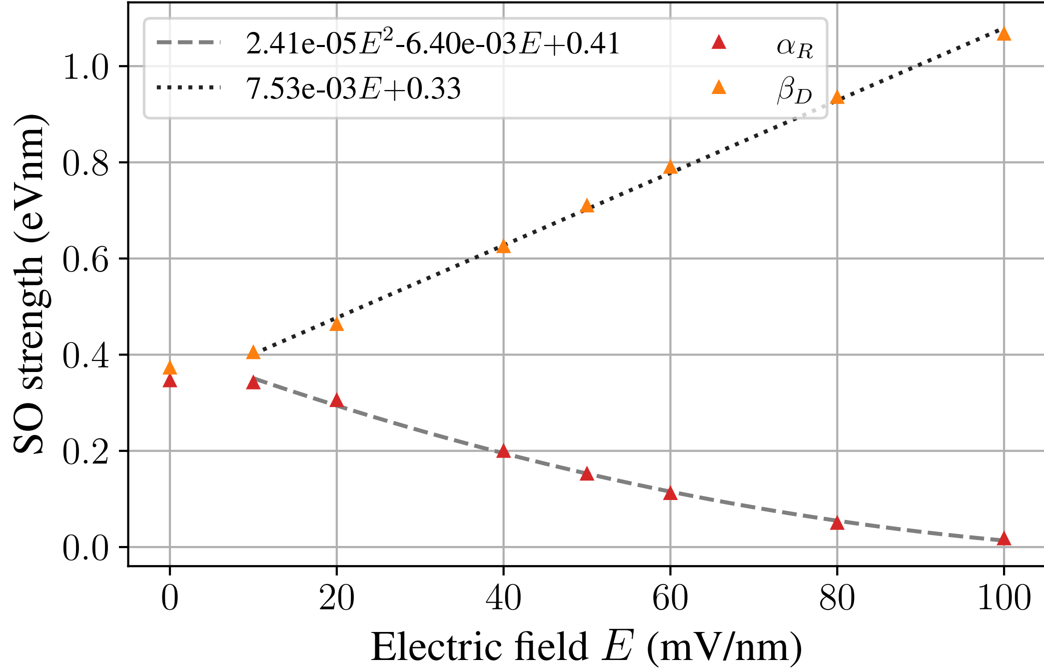


Figure 5.12. The Rashba and Dresselhaus spin-orbit strength as a function of electric field E for the mixed-ion terminating $L = 10\,\text{nm}$ nanowire. The spin-orbit fit parameters, α_R and β_D, from Figure 5.11 are plotted against electric field.

of the faces normal to the z direction. This was achieved in NEMO3D by printing out a geometry of atom positions which is longer in the z direction and redefining the excess atoms in the z direction as SiO2 atoms. For this simulation strain was not considered. A 5 nm substrate of SiO2 was added on top of both a same-ion terminating nanowire and a nanowire with mixed-ion termination in z. The atom positions of two neighbouring x slices through the centre of each of these nanowires is shown in Figure 5.13. SiO2 is made up of Si and O atoms but for simplicity NEMO3D approximates the species as a single atom lattice with empirical parameters that fit to the bulk material [111].

The resulting g-factor anisotropies of the two nanowires are compared with that of the same-ion terminating nanowire with no substrate in Figure 5.14. Similar to introducing mixed-ion termination to the nanowire with no substrate, adding a substrate introduces Rashba and increases Dresselhaus. The addition of the substrate to the mixed-ion nanowire introduces a larger Rashba constant and smaller Dresselhaus constant than the substrate added to the same-ion nanowire. Interestingly, the amplitude of the anisotropy is much larger in the substrate and same-ion terminating case. These results confirm that the substrate has a similar influence on the g-factor anisotropy to mixed-ion termination.

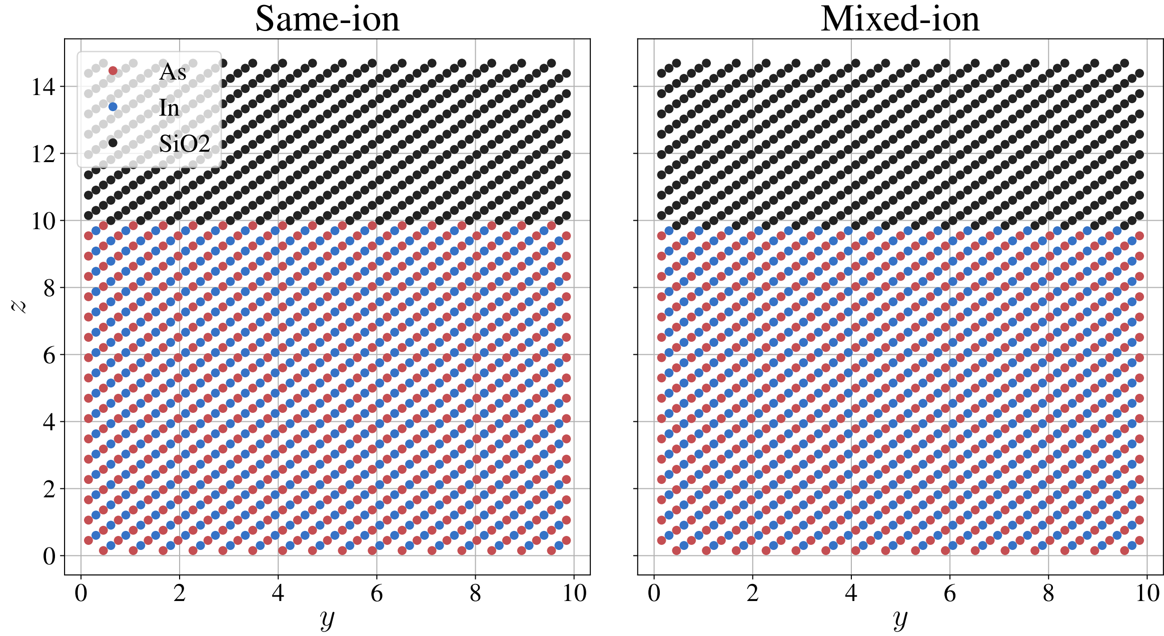


Figure 5.13. The atom positions of two neighbouring x slices of the $L = 10\,\text{nm}$ nanowire with a 5 nm substrate of SiO2 on top, plotted on the same axes for: (left) same-ion termination and the (right) mixed-ion termination.

This is also a good example of how NEMO3D could be employed to model real nanowire devices on substrates.

5.4.3. Larger square cross-section nanowires

The magnitude of the spin-orbit constants for the 10 nm square cross-section nanowire is on the order of 10^{-10} eVm. This is about an order of magnitude larger than that typically observed in experiments on gated InAs nanowires: 10^{-11} eVm [125, 127, 129, 130, 131, 132]. However, the nanowires tested in these experiments have larger diameters (50 nm to 100 nm) than the effective 10 nm diameter nanowire modelled here. The tighter confinement of the 10 nm nanowire is expected to increase the spin-orbit interaction compared to less confined, larger nanowires.

This section uses NEMO3D to model the effective g-factor anisotropy of a nanowire's E_1 with different diameters (or square cross-section side lengths L). The results demonstrate that a smaller nanowire diameter yields a smaller g-factor and larger spin-orbit constants. The larger nanowires have spin-orbit strengths comparable to experimental results which verifies NEMO3D and the effective model results. It also demonstrates that NEMO3D

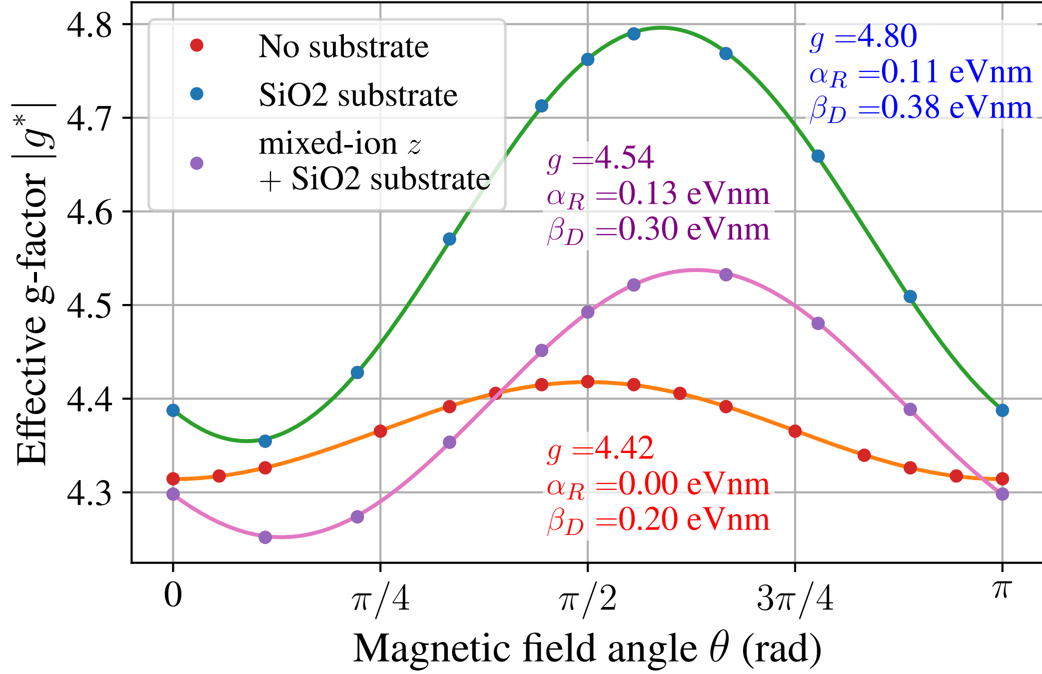


Figure 5.14. The effective g-factor anisotropy for square cross-section $L = 10\,\mathrm{nm}$ nanowires with a SiO2 substrate on the top z face compared with the no substrate same-ion terminating case. The magnetic field applied is $B = 5\,\mathrm{T}$. The filled circles are the NEMO3D data and the solid lines are the effective model fit to the data with the fit parameters indicated in the plot.

can model the g-factor anisotropies of larger diameter nanowires with longer, but still reasonable, computation times.

Effective g-factor anisotropy as a function of nanowire diameter

The effective g-factor anisotropy of E_1 was determined from NEMO3D for square cross-section, mixed-ion in z, InAs nanowires with different L. The effective model was fit to these g-factor anisotropies to determine the spin-orbit constants, α_R and β_D. The anisotropies and the effective model fits are shown in Figure 5.15. The maximum value of the effective nanowire g-factor (or the quantum well g-factor), g, increases with increasing nanowire diameter while the amplitude of the g-factor anisotropy decreases. The value of g increases with larger diameters due to the larger confinement lengths, which agrees with the results of Figure 5.4. The reduction in the anisotropy amplitude is due to the reduced strength of the spin-orbit interactions in the nanowire and weaker coupling between the subbands $\langle k_y \rangle_{21}$. The Dresselhaus interaction is reduced because there is

less confinement, L gets larger, and hence $\beta_D = \gamma \langle k_z \rangle \sim \left(\frac{\pi}{L}\right)^2$ decreases. The Rashba interaction decreases because the majority of the electron wavefunction gets further from the mixed-ion surfaces as the nanowire's size increases.

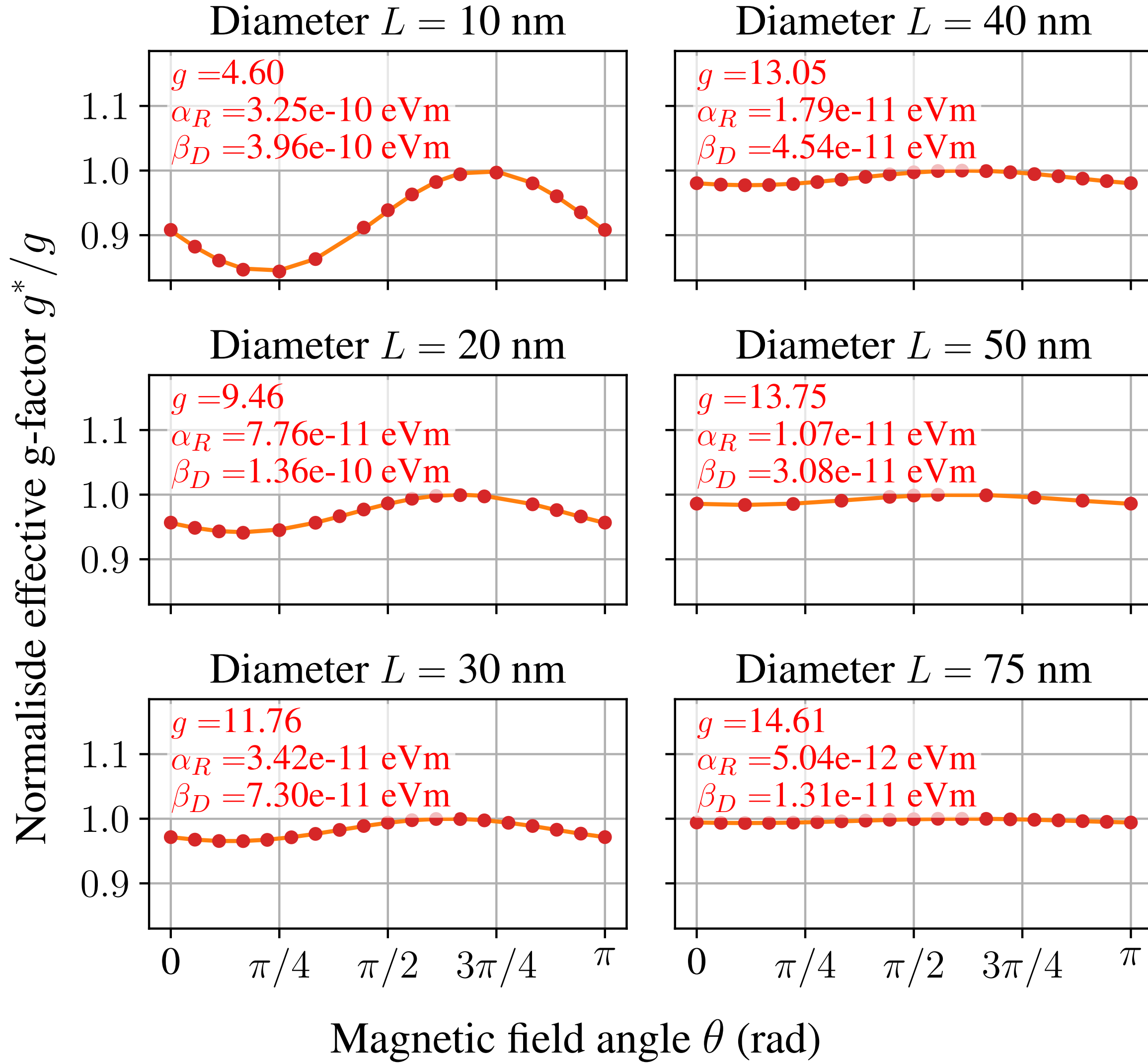


Figure 5.15. The normalised effective g-factor anisotropy of the lowest conduction subband E_1 for various diameters L of square cross-section, mixed-ion in z, InAs nanowires. The red filled circles are the NEMO3D data and the solid orange lines are the effective model. The fit parameters of the model are displayed in red.

Figure 5.16 is the spin-orbit constants as a function of L along with various experimentally observed values of the Rashba constant in InAs nanowires of different sizes. The magnitude of the spin-orbit interaction constants at larger diameters are comparable to the experimental results of Rashba in 50 nm to 100 nm diameter nanowires [125, 127, 129, 130, 131, 132]. The experimental results have larger Rashba than the

NEMO3D nanowires but this should not be too much of a concern considering the measured nanowires would also experience spin-orbit interaction from strain and structural inversion asymmetry of the substrates. In addition, the nanowires measured in experiments were not grown along the [100] direction which was modelled in NEMO3D. Other nanowire growth directions may experience larger Rashba. Overall, one can say that the NEMO3D results have qualitative agreement with the experimentally observed Rashba constants in InAs nanowires.

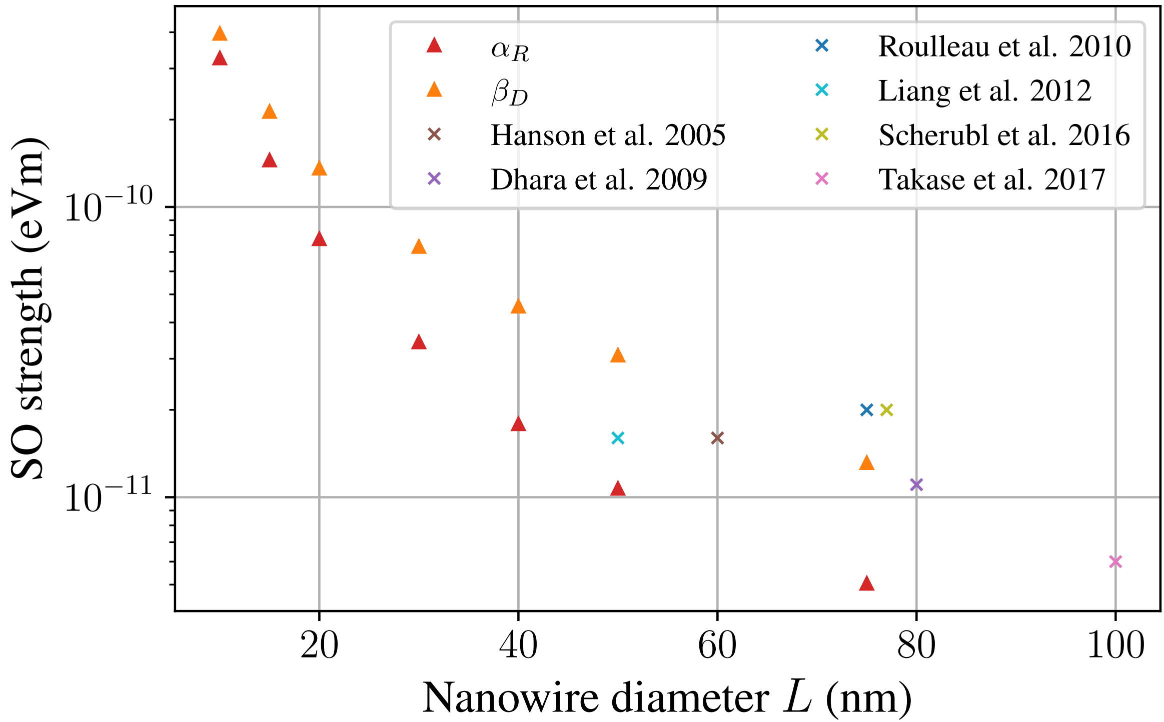


Figure 5.16. Spin-orbit constants as a function of square cross-section nanowire diameter (cross-section side-length). The values of α_R and β_D are from the effective model fits in Figure 5.15. Observed Rashba constants from various experiments on InAs nanowires [125, 127, 129, 130, 131, 132] are shown by '$\times$'s.

One may wonder whether it is worth considering the anisotropy given the minimal change in the g-factor at the larger diameters typically used in experiments. However, note that here the $k = 0$ case is considered which has the smallest anisotropy. The anisotropy is expected to be larger for $k \neq 0$ due to the first order and second order terms in k. When performing conductance measurements the non-zero wavevectors can contribute to the transport of electrons depending on the chemical potential of the system and hence these non-zero k contributions could result in larger g-factor anisotropies than seen here.

Computational resources for larger nanowires

The nanowires used in real devices are typically larger than the $L = 10$ diameter nanowire used throughout this chapter. For comparison to experiment and informing device design it is important to be able to model larger nanowires and do so efficiently. A summary of the computational resources required for the NEMO3D simulations of different diameter nanowires is given in Table 5.1. NEMO3D searches a user specified energy range for a given number of eigenvalues. The less eigenvalues being solved for the faster the simulation and the less resources required. The eigenvalue's corresponding eigenvectors and probability densities can also be solved for if desired. Typically, solving for either the eigenvectors or probability densities along with the eigenvalues doubles the simulation time compared to only solving for the eigenvalues. For this work the eigenvectors were generally not needed however some probability densities were required. The table lists the computational requirements for different nanowire diameters, number of eigenvalues solved for and whether the probability densities were calculated.

Table 5.1. A table of the computer resources required for the NEMO3D simulations of different sized nanowires.

L (nm)	No. eigenvalues	Probability density	No. CPUs	Memory (GB)	Time (min)
10	2	No	16	6	10-20
10	2	Yes	16	6	30-40
10	6	Yes	16	7	30-40
15	2	No	16	8	30-40
20	2	No	32	16	40
30	2	No	32	25	80-90
40	2	No	48	44	120
50	2	No	48	60	170
75	2	No	96	150	300
75	2	Yes	96	200	500
100	2	Yes	144	380	780

For the 10 nm nanowires the simulations are very efficient, typically completed within the hour, but for larger nanowires the simulation time approaches the order of hours. It is possible to simulate a $L = 100$ nm nanowire (at which point the geometry is a cube rather than a wire) in less than a day for only two eigenvalues. These are rough

guides on computation times, the block Lanczos solver can take a longer amount of time depending on the requested energy range within which the algorithm searches for eigenvalues. One must also keep in mind that one NEMO3D simulation of two eigenvalues only corresponds to a single data point of the g-factor anisotropy. In order to create an anisotropy, 10 to 20 of these NEMO3D simulations are required. This is not a problem for a supercomputer with many nodes where multiple NEMO3D simulations of this size can be run simultaneously. However, it could be a time consuming process if the computing resources only allowed for one NEMO3D simulation to run at a time.

In summary, if enough supercomputing resources are available to the user it is possible to determine the g-factor anisotropy of nanowires with diameters up to $L = 100\,\text{nm}$ in NEMO3D within a day. It is likely that nanowires of length greater than 100 nm (with small L) could also be simulated within the time frame of a few days. In comparison to other first principles calculations performed on supercomputing clusters this is a considerably low computational time. As such, employing NEMO3D to motivate and check nanowire device design is a feasible option if the computational resources are readily available.

5.5. Rectangular cross-section nanowires

The nanowires modelled in this chapter so far have a square cross-section with equal confinement in both y and z directions, $L = L_y = L_z$. The [100] nanowires do not always have a square cross-section but can be grown with rectangular cross-sections [117]. In order to understand the influence of the nanowire cross-section shape on the g-factor and spin-orbit interactions it is worth modelling the effective g-factor anisotropy of a rectangular cross-section [100] nanowire. The rectangular cross-section geometry was constructed by starting with the geometry in Figure 5.6 and increasing the length of confinement in the y direction, L_y, while maintaining z confinement at $L_z = 10\,\text{nm}$. The effective g-factor anisotropy of E_1 was then determined from NEMO3D as before. Originally, the only expected difference between the square and rectangular cross-section nanowires was a difference in Dresselhaus strength β_D due to the reduced confinement in y. However, this expectation was incorrect, and rather the rectangular cross-section nanowire was found to have a Rashba dominant anisotropy with an amplitude dependent on the confinement in y.

5.5.1. Effective g-factor anisotropy

Figure 5.17 is the effective g-factor anisotropy of E_1 for a same-ion terminating nanowire with a rectangular cross-section of dimensions $L_y \times L_z = 20$ nm $\times$ 10 nm and that of the previously modelled same-ion terminating square cross-section nanowire with $L_y = L_z = 10\,\text{nm}$. The anisotropy is shown for the magnetic field varying in the xy-plane with a magnitude of $B = 5\,\text{T}$. The effective model expression (5.18) is fit to the g-factor anisotropies and the fit parameters displayed in the figure.

The anisotropy of the rectangular cross-section nanowire has a larger g than the square cross-section nanowire due to the larger confinement length in the y direction. The larger confinement length in y has also led to a Rashba dominant anisotropy with no phase shift $\phi = 0$. There is no presence of Dresselhaus compared to the square cross-section nanowire which had $\beta_D = 0.20\,\text{eVnm}$ and no Rashba interaction. This is odd considering there is no structural inversion asymmetry from mixed-ion termination or an applied electric field to induce the Rashba contribution.

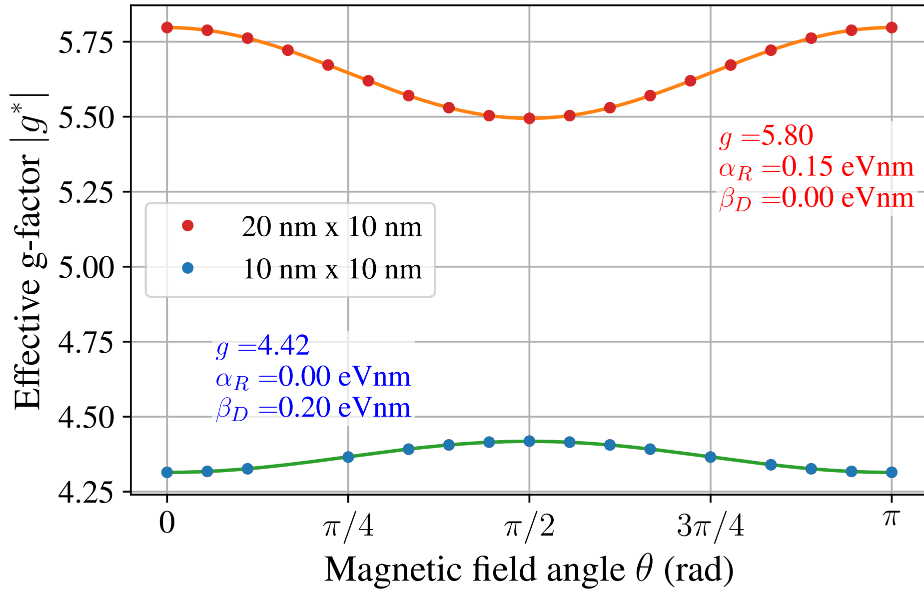


Figure 5.17. Effective g-factor anisotropy for E_1 of a rectangular cross-section $(20\,\text{nm} \times 10\,\text{nm})$ nanowire and the square cross-section $(10\,\text{nm} \times 10\,\text{nm})$ nanowire both same-ion terminating. The filled circles are the NEMO3D data and the solid lines are the effective model fit to the data with the fit parameters indicated in the plot.

In attempts to understand the origin of the Rashba spin-orbit interaction, the lowest conduction subband probability density $|\psi_{1\downarrow}|^2$ is plotted for the different nanowire cross-

sections lengths L_y. Figure 5.18 is an x slice of $|\psi_{1\downarrow}|^2$ in the centre of the nanowire for different nanowire cross-sections. All nanowires are same-ion terminating in y and z. The square cross-section nanowire's wavefunction is not spherical but deformed at the edges of the nanowire. For the rectangular cross-section nanowires, the wavefunction is less deformed in the y direction and there is a lower probability density at the surfaces normal to the y axis than in the square cross-section nanowire. The fact that there is a higher density of the electron wavefunction at the z surfaces than the y surfaces may be the cause for this Rashba spin-orbit interaction in the rectangular cross-section nanowires. Additionally, Moroz and Barnes [126] state that a confinement potential, V_z, creates an electric field, $E = -\nabla V_z$, that can induce a Rashba spin-orbit interaction perpendicular to the nanowire. Although this is a smaller spin-orbit interaction than that from external electric fields it can be dominant in the absence of electric fields. Since V_z is a smaller confinement region than V_y it should result in a larger confining electric field which will produce a larger Rashba spin-orbit interaction. However, this does not explain how α_R is larger than $\beta_D \propto (\pi/L_z)^2$. One possible explanation is that the confinement in the y direction contributes to the Dresselhaus interaction along the nanowire. By increasing L_y a component of the Dresselhaus constant $\beta_{Dy} \propto \frac{1}{L_y^2}$ decreases and as a result the total $\beta_D = \beta_{Dy} + \beta_{Dz}$ decreases and α_R becomes dominant.

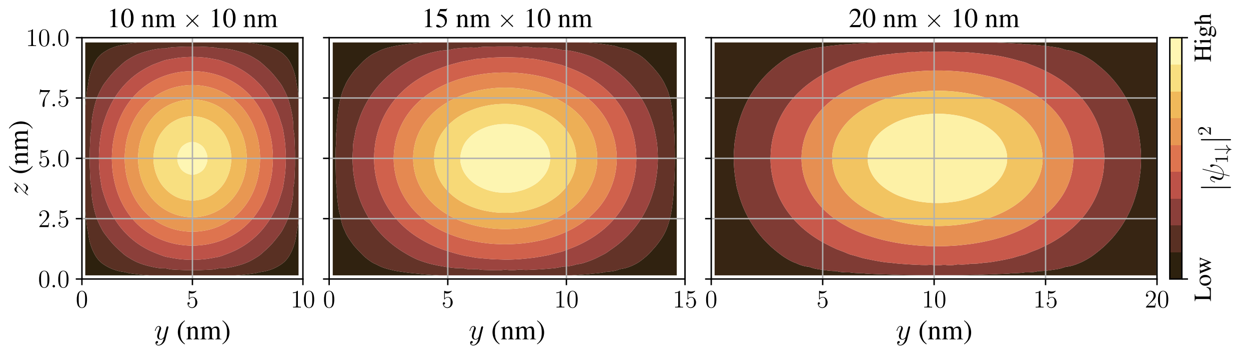


Figure 5.18. The lowest spin state of the first conduction subband electron probability density in the cross-section of the various nanowires.

An alternative explanation was that the wavefunction in the rectangular cross-section nanowire may be slightly off-centre which could give the wavefunction different values at opposite interfaces and create a strong structural inversion asymmetry. However, examination of y slices taken at the centre of each probability density in Figure 5.18 and plotted in Figure 5.19, reveal that the centre of the wavefunction does not move significantly in the z direction when going from square to rectangular cross-section. The shape of the wavefunction at the edges of the nanowire does bend more significantly for

larger L_y but the total wavefunction is symmetric about the centre of the nanowire. This should not produce inversion asymmetry and is unlikely to be the cause of the Rashba spin-orbit coupling.

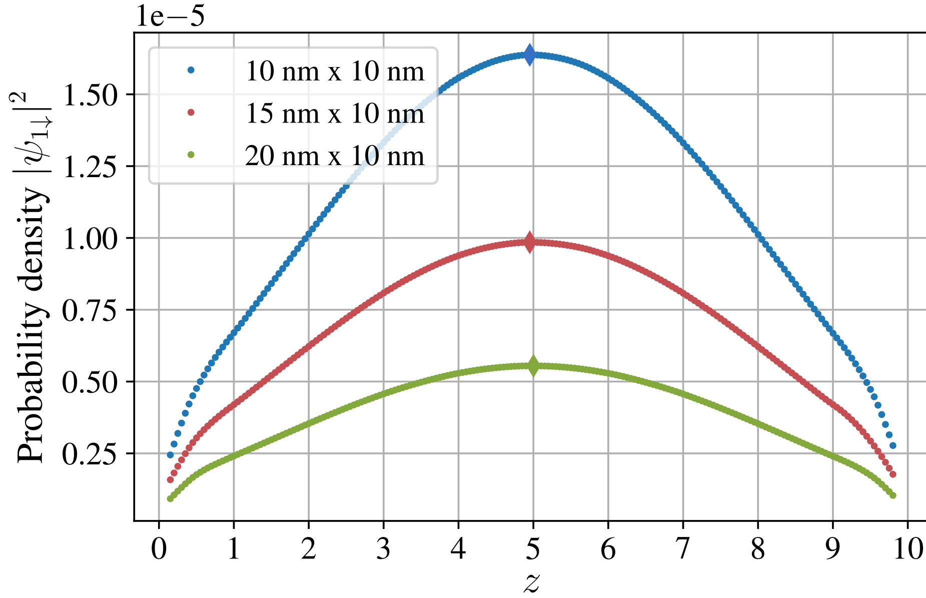


Figure 5.19. A y slice of the lowest spin state of the first conduction subband electron probability density $|\psi_{1\downarrow}|^2$ taken at the centre of the different nanowire cross-sections Figure 5.18. The diamond identifies the maximum of the probability density.

To further understand how the confinement in the y direction influences the anisotropy, the y dimension of the rectangular cross-section nanowire is varied from 10 nm to 100 nm. Figure 5.20 is the effective g-factor anisotropies of same-ion terminating rectangular cross-section nanowires for the different L_y and $L_z = 10$ nm. Again the magnitude of the offset of the effective g-factor increases as L_y increases, as expected from previous results. The square cross-section nanowire has the Dresselhaus dominant ($\phi = \pi/2$) g-factor anisotropy but as L_y increases and reaches 11 nm the g-factor anisotropy becomes Rashba dominant ($\phi = 0$). As L_y increases, α_R increases until some point after $L_y = 15$ nm where α_R begins to decrease with increasing L_y until there is no spin-orbit interaction present at $L_y = 100$ nm resulting in no g-factor anisotropy. The $L_y = 100$ nm case is essentially a quantum well given that the dimensions in x and y are equivalent at the approximated bulk length of 100 nm. From the effective model, an isotropic g-factor is expected at $\mathbf{k} = \mathbf{0}$ for a quantum well with no confinement in y direction. Without confinement in y, there is no quantisation of the k_y terms in equation (5.3) so the k_y terms in the spin-orbit interaction are zero and there is no effective g-factor anisotropy.

Note that the $L_y = 100\,\text{nm}$ case had to be calculated with an applied magnetic field of $B = 2\,\text{T}$ because larger magnetic fields would cause a spin splitting larger than the subband splitting and make it difficult to identify the lowest subband spin states. This weaker magnetic field should not change the results as discussed in Appendix E.1.

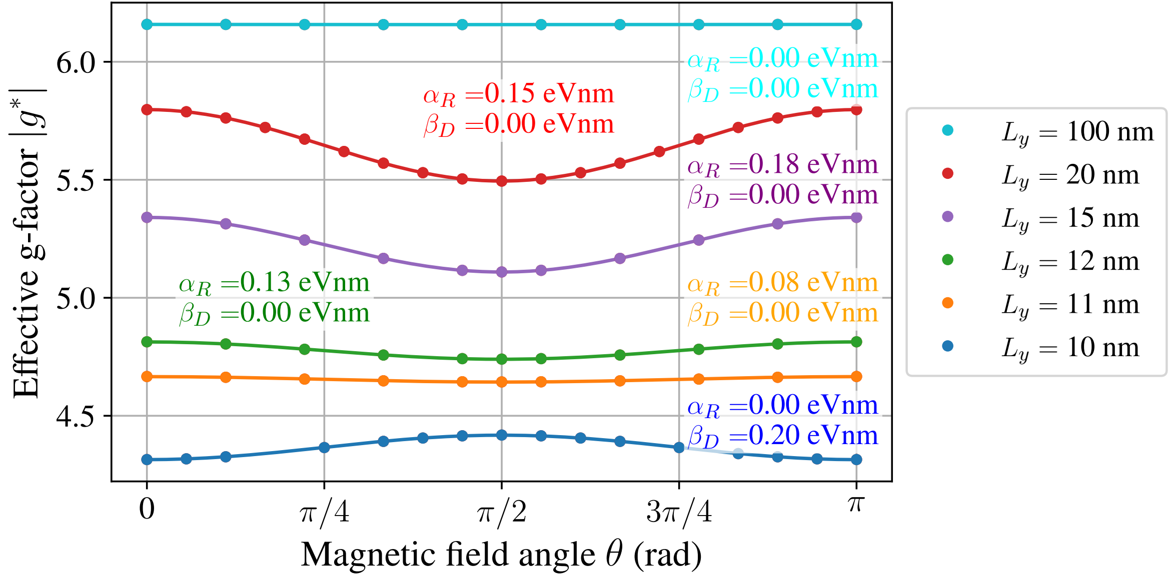


Figure 5.20. The effective g-factor anisotropy of the rectangular cross-section nanowire for various L_y and $L_z = 10\,\text{nm}$. All the anisotropies were calculated with a magnetic field of $B = 5\,\text{T}$ except for the $L_y = 100\,\text{nm}$ case which was calculated with $B = 2\,\text{T}$.

The significance of the point where the Rashba spin-orbit interaction becomes dominant over the Dresselhaus is not fully understood. However, it seems that this is just the point where the confinement in the z direction is stronger than that in the y direction. As the confinement in the y direction increases further the coupling between the transverse subbands decreases until the g-factor anisotropy disappears as expected from equation (5.18).

5.5.2. Mixed-ion termination and applied electric fields

The rectangular cross-section nanowire can induce a Rashba spin-orbit interaction when the y confinement is larger than the z confinement. Since this Rashba dominant anisotropy occurs in a same-ion terminating nanowire it is worth testing how the introduction of mixed-ion termination and electric fields influences this g-factor anisotropy. Figure 5.21 is the effective g-factor anisotropies of E_1 for the $20\,\text{nm} \times 10\,\text{nm}$ cross-section nanowire

with same-ion termination, mixed-ion termination and an applied electric field of $E_z = 10\,\mathrm{mV/nm}$ for the same-ion termination nanowire.

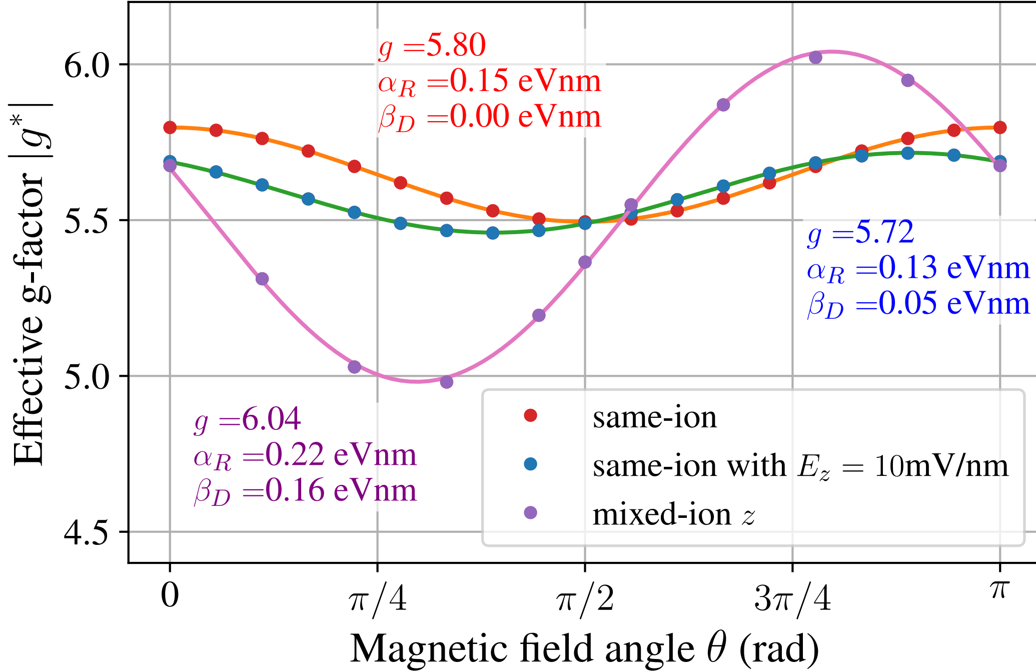


Figure 5.21. The effective g-factor anisotropy for a 20 nm × 10 nm rectangular cross-section nanowire with different ion termination and applied electric fields in z direction. The filled circles are the NEMO3D data and the solid lines are the effective model fit to the data with the fit parameters indicated in the plot.

Introducing an electric field in the z direction to the rectangular cross-section nanowire alters the anisotropy as expected from Section 5.4.1. The confinement from the electric field increases the Dresselhaus constant and decreases the Rashba constant because the wavefunction is confined towards one surface. Introducing mixed-ion termination in the z direction increases the anisotropy dramatically. It increases the Rashba strength and introduces Dresselhaus both stronger than the electric field case as one would expect from the results of Section 5.4.2.

The dependence of the spin-orbit constants of the rectangular cross-section nanowire on electric field is determined from g-factor anisotropy plots as was done for the square cross-section nanowire. Figure 5.22 is the g-factor anisotropies with the effective model fits and the y slice of the probability density $|\psi_{1\downarrow}|^2$ as insets. Figure 5.23 is the electric field dependence of the spin-orbit constants, α_R and β_D. Similar to the mixed-ion square cross-section nanowire case, the Dresselhaus strength increases almost linearly with electric field due to the increasing confinement of the wavefunction. The Rashba strength decreases

with electric field in a linear fashion as the electron wavefunction is pushed towards one surface and unlike the square cross-section nanowires the Rashba actually passes through zero and changes sign. The linear Stark coefficient for Dresselhaus is $6.35\,\mathrm{enm}^2$ which is less than the square cross-section nanowire likely due to less confinement in y. The linear Stark coefficient for the Rashba is $-2.31\,\mathrm{enm}^2$. This coefficient is smaller than that of the square cross-section nanowires. Notable features of Figure 5.23 are that there are applied electric fields for which there is only one type of spin-orbit interaction present, as well as a point where α_R and β_D have the same magnitude. If the spin-orbit interactions could somehow be anti-aligned as is predicted to occur in [110] InAs nanowires one could created a cancellation of the spin-orbit interactions at the electric field where $\alpha_R = \beta_D$ [54]. From these results, one can conclude that the mixed-ion termination and applied electric fields influence the g-factor anisotropy and spin-orbit interaction of the rectangular cross-section nanowire in a similar manner to that of the square cross-section nanowire.

5.6. Discussion of results and applications

The results of this chapter have demonstrated NEMO3D's ability to determine the influence of atomic level alterations to the nanowire geometry on the lowest conduction subband effective g-factor anisotropy and spin-orbit interactions. The key features discovered were that the [100] grown nanowire's Dresselhaus spin-orbit interaction can be strengthened by applying electric fields and altering the surface structure with mixed-ion termination or substrates. The Rashba spin-orbit interaction can be tuned by weak electric fields but is influenced mainly by surface ion termination and the addition of substrates. The g-factor anisotropy and spin-orbit composition can also be controlled by varying the in-plane width of the rectangular cross-section of a nanowire. The [100] nanowire is the only growth direction of the nanowire modelled here. In the future, it will be worth modelling other nanowire growth directions i.e. [110] and [111] as they are predicted to have interesting properties due to the Dresselhaus interactions induced by the different crystal lattice orientations [122, 133, 134].

The results of the NEMO3D modelling have been verified by comparing to g-factor anisotropies and spin-orbit constants observed in experimental works. The simulations could be further verified by performing low temperature conductance measurements on InAs nanowires and varying the direction of an applied magnetic field. The experiment

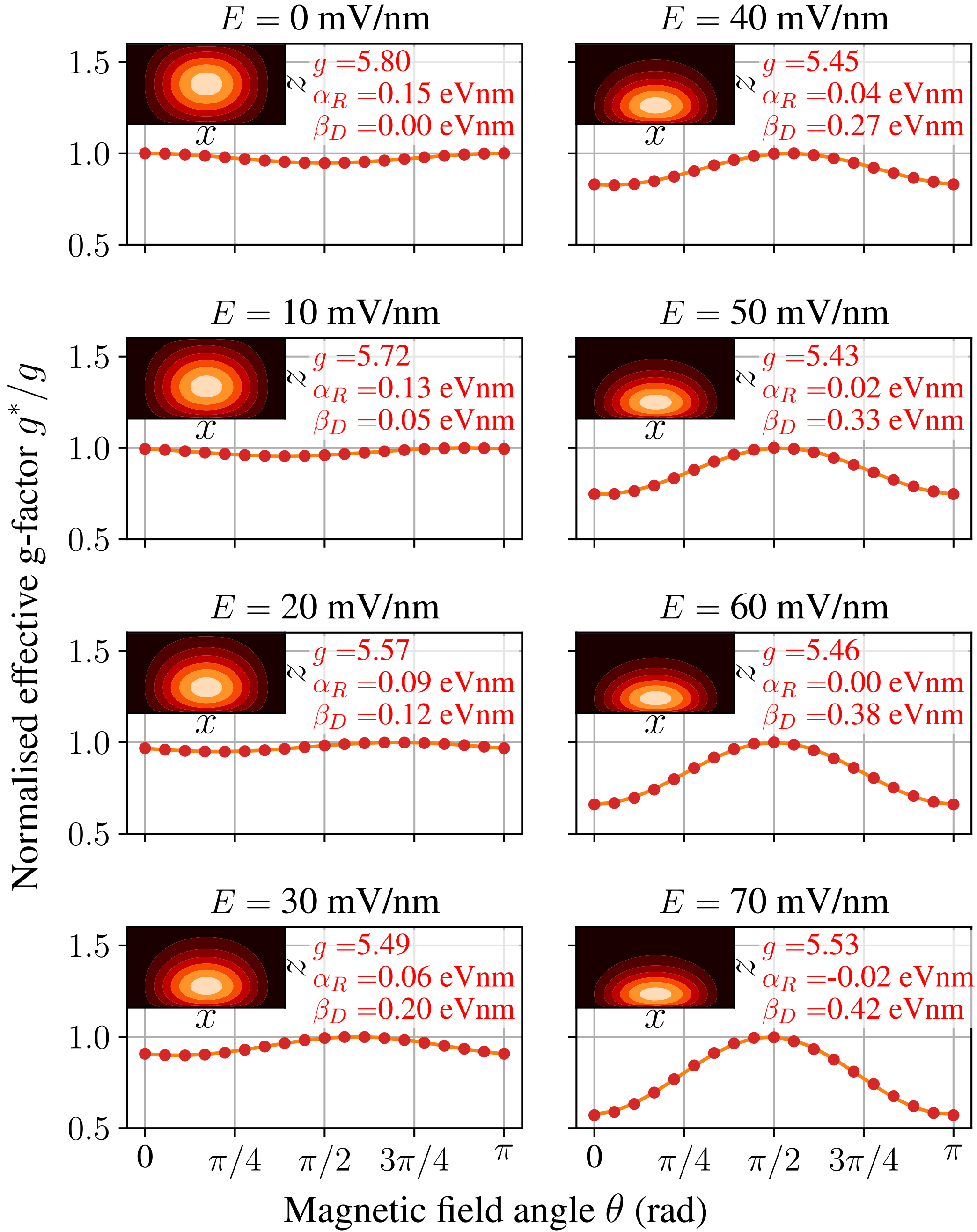


Figure 5.22. The normalised g-factor anisotropy of E_1 for the same-ion terminating 20 nm × 10 nm cross-section nanowire for different electric field strengths E. The fit parameters of the effective model, g, α_R and β_D are displayed in red. The top left inset of each plot is a y-slice of the probability density of the electron's $E_{1\downarrow}$ state in the nanowire.

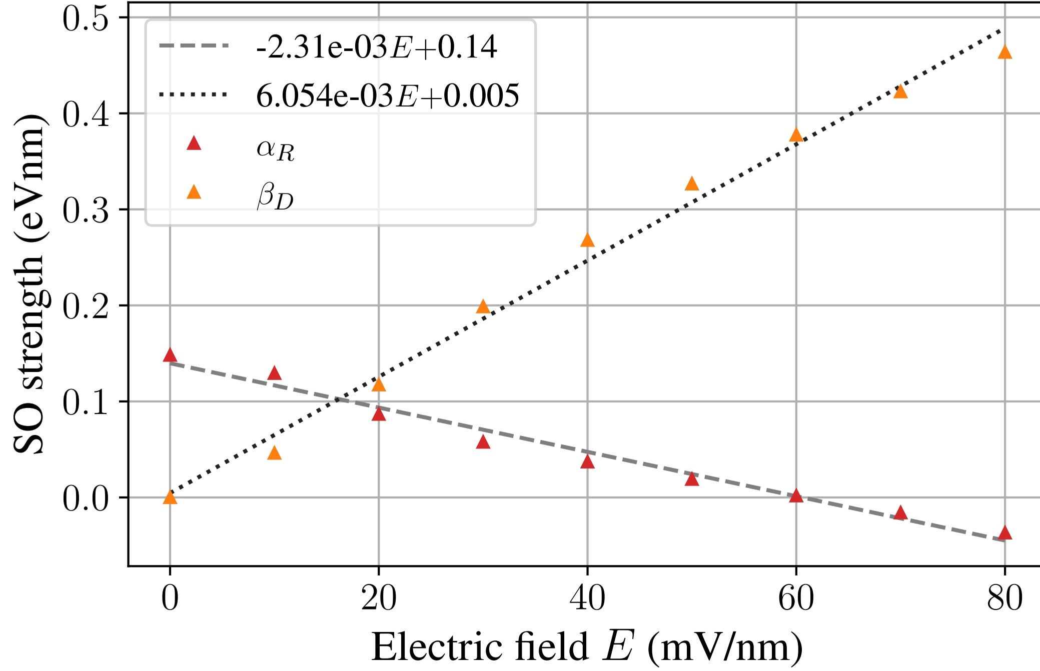


Figure 5.23. The Rashba and Dresselhaus spin-orbit strength as a function of electric field E for the same-ion terminating 20 nm × 10 nm cross-section nanowire. The spin-orbit fit parameters, α_R and β_D, from Figure 5.22 are plotted against electric field.

would produce conductance versus gate voltage plots similar to that of [62, 68, 126] or Figure 2.6. The width of the dip in the conductance due to the spin gap should decrease and increase depending on the angle of the applied magnetic field which corresponds to the effective g-factor anisotropy modelled by NEMO3D. This would mean that for nanowires with different surface-ion terminations the conductance dip would be maximised at different magnetic field angles. However, if the nanowire had varying surface-ion terminations along its length the conductance measurements of the nanowire would only reveal the smallest spin gap of the different surface-ion termination regions and not the true spin gap in each region. Hence, another experimental method may be required to distinguish between regions of differing surface-ion termination and determine the true spin gap. For example, it may be possible to measure the spin splitting of gate defined quantum dots in each region.

The knowledge gained from the NEMO3D simulations of the effective g-factor anisotropy could inform the design of a wide array of nanoscale devices. For instance, spin-orbit qubits that are typically constructed from gate defined double quantum dots in III-V nanowires [22, 23]. The spin-orbit dependent anti-crossing occurring between

singlet and triplet states defines the Pauli blockade of the system which is the basis for qubit operation. The [100] nanowire's g-factor anisotropy with Rashba and Dresselhaus spin-orbit interactions could allow for more control or at least another option for control of the qubit along with the standard Rashba spin-orbit interaction anisotropy seen in the commonly used [111] nanowires.

Another area of interest are spintronic devices, such as the spin transistor proposed by Datta to improve on current electronic integrated circuits [25]. This device operates by electrically controlling the spin current travelling through a nanowire between two ferromagnetic electrodes. If there is a spin-orbit interaction in the nanowire it can cause spin precession. With the correct spin-orbit interaction in the nanowire the spins ejected from the source electrode can experience an exact π-flip by the time they reach the drain electrode creating an on-state. If the spin is not completely flipped by π, the transistor would be in the off-state. From the simulations above it may be possible for one to construct a spin transistor from the [100] nanowire. For example, a rectangular cross-section nanowire with a Rashba dominant g-factor anisotropy, corresponding to a spin-orbit field perpendicular to the nanowire, could be electrically tuned between Rashba and Dresselhaus dominant states of the nanowire to switch the transistor between the on- and off-state. It would also be of interest to model the spin-orbit interactions in the [110] nanowire. The [110] growth direction is predicted to allow for a cancellation of Rashba and Dresselhaus at the correct electrical tuning which could be another method of setting the transistor to its off-state.

A well-established application of these nanowires is the semiconductor-superconductor heterostructure which can form a topological superconducting state and host Majorana zero modes. The main requirements for the nanowire to form this state is: (1) a large g-factor in the direction of the applied magnetic field to ensure minimal suppression of the superconductivity in the bulk superconductor and (2) a large spin-orbit interaction to create a stable topological superconducting state. The topological state's sensitivity to disorder is said to be minimised when $E_Z/E_{\text{SO}} < 1$ is satisfied [72], as discussed in Section 2.4.4.

The typical Zeeman splitting required to satisfy the topological criterion (2.46) and achieve a topological phase is $E_Z = 1\,\text{K}$ ($\approx 0.086\,\text{meV}$) according to [17, 72]. For the 50 nm nanowire with a g-factor of 14 this would require applying a magnetic field of 0.1 T. The largest Rashba constant α_R observed in this modelling is of the order of 0.1 eVnm corresponding to an $E_{\text{SO}} = \frac{m^*\alpha_R^2}{2\hbar^2} \approx 1.5\,\text{meV}$ which satisfies the stability condition, $E_Z/E_{\text{SO}} < 1$. However, these values are for small nanowire diameters less than

$L = 20\,\text{nm}$ which are not as feasible to grow and reported to have poor conductivity [132]. From the results of a larger diameter nanowire typically used in experiments, the 50 nm nanowire in Figure 5.15, the Rashba constant is $\alpha_R = 0.01\,\text{eVnm}$ corresponding to $E_{\text{SO}} \approx 0.015\,\text{meV}$ which does not satisfy the condition $E_Z/E_{\text{SO}} < 1$. However, the combined magnitude of α_R and β_D of 0.04 eVnm from Figure 5.15 gives $E_{\text{SO}} \approx 0.24\,\text{meV}$ which would satisfy the stability condition. The problem is that Dresselhaus is not acting perpendicular to the nanowire which is a requirement to create the topological gap.

One way to potentially induce a larger Rashba would be to tune with a top and bottom gate structure as done by [127]. An alternative solution would be to use [110] grown nanowires. In this nanowire orientation the Rashba and Dresselhaus are predicted to align along the same direction, perpendicular to the nanowire [54, 128]. Based on the calculation above the combined spin-orbit interaction perpendicular to the nanowire would meet the criteria for protection against disorder. Many researchers have also recently begun researching the g-factor anisotropy of holes in III-V semiconductors because the larger spin-orbit interactions of holes may create more stable topological states [135, 136, 137].

According to [23], the superconductivity of the heterostructure is suppressed least when the magnetic field is in the plane of the substrate. Typically, the magnetic field is applied along the nanowire [17, 15, 16]. From the results of this chapter, the best [100] nanowire to use would be the rectangular cross-section nanowire because the g-factor is maximised along the nanowire. With a substrate below the nanowire the g-factor anisotropy and the Rashba would be increased, however, this would also induce a Dresselhaus spin-orbit coupling along the nanowire which shifts the g-factor maximum away from the [100] direction. It is for this reason that [111] nanowires are typically employed because they possess inversion symmetry and have no Dresselhaus [23, 62, 67, 122]. Hence, it would be useful to model the [111] nanowire and the [110] nanowire in NEMO3D to see if a stronger spin-orbit interaction can be achieved perpendicular to the nanowire which satisfies the condition $E_Z/E_{\text{SO}} < 1$ and maximises the g-factor along the nanowire.

5.7. Summary

The effective g-factor anisotropy of the lowest conduction subband for a [100] nanowire has been successfully modelled using NEMO3D. The spin-orbit interactions, Rashba and Dresselhaus, were extracted from the g-factor anisotropies using an effective model constructed from the quantum well Rashba and Dresselhaus interaction terms. The influence of applied electric fields, surface-ion termination and the nanowire cross-section on the effective g-factor anisotropy and spin-orbit interactions were studied. The applied electric field induced a Rashba spin-orbit interaction for small electric fields but for larger electric fields it increased the Dresselhaus interaction and maximised the g-factor perpendicular to the nanowire. The introduction of mixed-ion termination perpendicular to the magnetic field induces a Rashba spin-orbit interaction and increases the Dresselhaus interaction present in the nanowire. The mixed-ion termination induces a stronger Rashba spin-orbit interaction than the applied electric fields and is analogous to adding a substrate to the nanowire. The square cross-section nanowire has a Dresselhaus dominant spin-orbit interaction with the g-factor maximised perpendicular to the nanowire but if the nanowire has a rectangular cross-section the g-factor anisotropy can change to maximised along the nanowire with a Rashba dominant spin-orbit interaction.

The atomistic detail of the surface ion termination, substrate interface and the shape of the nanowire cross-section can only be modelled in an atomistic model like NEMO3D. This atomistic method of modelling can provide device engineers with the information required to design various nanowire devices with specific g-factor anisotropies and/or spin-orbit interactions. This includes but is not limited to, spin-orbit qubits, spin transistors or topologically protected qubits which can host Majorana zero modes.

The possible directions for future research is vast. There are a plethora of variables that could be considered in NEMO3D such as strain, different material interfaces, the other nanowire growth directions ([110] and [111]) and the g-factor anisotropy of holes. The most beneficial area to research first would be the g-factor anisotropy for the [110] and [111] nanowires because this could provide insight into the optimal nanowire growth direction for devices. There are also multiple improvements that could be made to current modelling methods. In order to increase the electric control of the Rashba spin-orbit interaction, it would be useful to employ a top and bottom gate model so that the electron wavefunction density is not confined to one surface. The modelling of NEMO3D could also be improved by adding functionality for three dimensional same-ion termination and constructing hexagonal cross-section [111] nanowires.

Chapter 6.

Conclusion

This thesis has explored features of various semiconductor nanostructures with applications in spintronics and quantum information technology. A form of quantum transport, single electron charging, was demonstrated through dopants in the silicon channel of a FinFET using a low temperature measurement setup. The Coulomb blockade spectrum identified two bound states, D^0 and D^-, for a handful of dopants in the corners of the silicon channel. This Coulomb blockade measurement validated the low temperature measurement setup to temperatures as low as 11 K. In the future, temperatures of 4 K and below should be achievable by making some improvements to the measurement setup and cooling procedure. Lower temperatures will allow for better resolution of Coulomb spectrums and quantum transport through the quantised subbands of semiconducting nanostructures.

The field of nano-electronic modelling was then introduced. In particular, the semi-empirical tight-binding modelling tool NEMO3D was described, including the tight-binding formalism, the inclusion of electromagnetic and spin-orbit coupling and the advantages of NEMO3D's atomistic approach over ab initio and continuum models. An example of NEMO3D's ability to model the electronic states of nanostructures was presented, showcasing the results from a recently published work that modelled the electrical dependence of valley state coupling in a gate-defined quantum dot within a silicon quantum well.

The core work of this thesis was modelling the magnetic field anisotropy of the lowest conduction subband effective g-factor in a [100] InAs nanowire. A modelling method was presented for determining the g-factor anisotropy of the nanowire under different external electric field strengths, surface-ion terminations and confinement conditions, and then for extracting the corresponding spin-orbit interactions of the nanowire from this anisotropy.

The g-factor anisotropy was modelled using the NEMO3D $sp^3d^5s^*$ 20 band tight-binding model. The spin-orbit constants for the common Rashba and Dresselhaus spin-orbit interaction, α_R and β_D, were determined by fitting the g-factor expression of a one band effective model to the NEMO3D results. NEMO3D captures the atomic resolution of the system and crystal asymmetries caused by different surface-ion terminations. The effective model interprets how these atomic details and crystal symmetries influence the g-factor anisotropy and the spin-orbit interactions present in the nanowire.

The key results are as follows. An external electric field can induce and tune the Rashba spin-orbit interaction but it also increases the Dresselhaus spin-orbit interaction by enforcing additional confinement on the electron wavefunction. A nanowire with mixed-ion termination, or a substrate, has reduced crystal symmetry which induces a strong Rashba (larger than that of external electric fields) and also increases the strength of Dresselhaus. The g-factor anisotropy and spin-orbit interaction are influenced by the extent of confinement and similarly the geometry of the cross-section. The nanowires with a smaller effective diameter have larger spin-orbit interaction constants but a smaller effective g-factor than those with a large effective diameter. The spin-orbit interaction constants and effective g-factors of the large diameter nanowires were similar to that reported in experiments. Elongating the nanowire cross-section from square to rectangular was found to change the maximum of the g-factor anisotropy from perpendicular to the nanowire axis to along the nanowire axis and resulted in the spin-orbit interaction transitioning from Dresselhaus to Rashba dominant.

The insights gained from this work and the modelling method can inform the design and fabrication of nanowires for a variety of spintronic and quantum computing applications, such as spin-orbit qubits, spin transistors and Majorana zero modes in topological qubits. However, this work has far from considered all the parameters that influence a III-V nanowire's effective g-factor anisotropy and as such there are many avenues for further research. The most prominent is applying the modelling method to [110] and [111] grown nanowires due to their interesting Dresselhaus spin-orbit interactions which are distinct from the [100] nanowire and may be advantageous for certain applications. It is also of interest to model the applied electric fields in a top-bottom gate fashion so that Rashba can be tuned over a larger range of electric fields without the electron wavefunction being confined to a surface. Finally, one crucial task for validating these g-factor anisotropy results is to perform conductance measurements on InAs nanowires in the presence of different magnetic field orientations using the low temperature measurement setup and a vectorial magnet.

Appendix A.

Useful Derivations

A.1. A review: The 1D infinite potential well

This derivation can be found in most undergraduate quantum mechanics textbooks. Here the derivation follows closely to that of [29]. Consider an electron free to move in one dimension, the x axis, and confined by an infinite potential well of length L

$$V(x) = \begin{cases} 0 & 0 < x < L \\ \infty & x < 0 \text{ and } x > L. \end{cases} \tag{A.1}$$

The 1D Schrödinger equation describing this electron is

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + V(x)\psi(x,t), \tag{A.2}$$

where m is the mass of the electron, $\hbar$ is Planck's reduced constant and $\psi(x,t)$ is the electron wavefunction.

Since $V(x)$ is time independent the Schrödinger equation can be separated into two equations, one for the spatial dependent part of the electron wavefunction $\phi(x)$ and the other for the time dependent part of the wavefunction $\chi(t)$, where $\psi(x,t) = \phi(x)\chi(t)$. Assuming the equations have energies eigenvalues $E = \hbar\omega$, these equations are written

$$i\hbar\frac{\partial}{\partial t}\chi(t) = E\chi(t), \tag{A.3}$$

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\phi(x) + V(x)\phi(x) = E\phi(x). \tag{A.4}$$

The time dependent equation (A.3) has eigenfunctions of the form $\chi(t) = e^{-iE/\hbar t} = e^{-i\omega t}$ with $E = \hbar\omega$ the corresponding energy eigenvalue of $i\hbar\frac{\partial}{\partial t}$. For the spatial dependent equation (A.4), also known as the time independent Schrödinger equation, $\phi(x) = 0$ unless $0 < x < L$ because of the infinite potential well $V(x)$. Inside the well the eigenfunctions of the second order spatial derivative are determined by enforcing continuity of the wavefunction and its first derivative at the boundaries, $\phi(0) = \phi(L) = 0$. The second derivative can be discontinuous as the potential is discontinuous. The resulting eigenfunction of $\frac{\partial^2}{\partial x^2}$ is $\phi_n(x) = A\sin(kx)$ with the wave vector $k = \frac{\sqrt{2mE}}{\hbar} = \frac{n\pi}{L}$ for $n = 1, 2, 3, ...$

Combining theses two solutions, the full solution to the electron in a 1D infinite potential well problem is

$$\psi_n(x,t) = \begin{cases} \sqrt{\frac{2}{L}}\sin{(k_n x)}e^{i\omega_n t} & 0 < x < L \\ 0 & \text{otherwise} \end{cases} \tag{A.5}$$

with $k_n = n\pi/L$ and $E_n = \hbar\omega_n = n^2\hbar^2\pi^2/2mL^2$. The value of the constant A was determined using the normalisation condition of the electron wavefunction $\int_{-\infty}^{\infty} |\psi_n(x)|^2 \, dx = 1$. This is the requirement that the norm square of the wavefunction can be interpreted as a probability density distribution. The single electron must be somewhere in space.

To summarise, confining an electron in one dimension, x, quantises the wavevector k and the resonant frequency $\omega = E/\hbar$ in the x direction. The electron has discrete energy levels or eigenenergies that it can reside in.

A.2. An infinite box: Periodic boundary conditions

The infinite potential well above is an example of 'closed' boundary conditions, where the electron wavefunction must be zero at the boundaries. Another common choice of boundary conditions is periodic boundary conditions where $x = L$ is equivalent to $x = 0$. This one dimensional system could be thought of as a ring or a box of infinite length. This boundary condition is appropriate when one is only interested in the properties of the interior of the box and the ends of the box do not matter e.g. a bulk semiconductor.

For periodic boundary conditions the eigenfunctions take the form [29]

$$\begin{aligned}
&\psi_n(x,t) = \psi_0 e^{i(k_n x - \omega_n t)},\\
&\psi_n(x,t) = \psi_0 e^{-i(k_n x + \omega_n t)},\\
&\text{with} \;\; k_n = \frac{2n\pi}{L},\;\; E_n = \hbar\omega_n = \frac{2\hbar^2\pi^2 n^2}{mL^2}
\end{aligned} \tag{A.6}$$

and ψ_0 is the normalisation constant. There are now half as many allowed k values but two eigenfunctions for each eigenvalue. The first $+k_n$ state is considered as the state travelling in the positive x direction and the second $-k_n$ state is considered as the state travelling in the negative x direction. In nanostructures, such as quantum wells and nanowires the electrons can often be considered as plane waves modulated by a periodic (Bloch) function travelling in positive or negative k directions of the nanostructure.

A.3. The Pauli matrices and identities

The Pauli matrices are defined as [138]

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A.7}$$

and the corresponding Pauli vector is defined as

$$\boldsymbol{\sigma} = \sigma_x \hat{x} + \sigma_y \hat{y} + \sigma_z \hat{z}. \tag{A.8}$$

The Pauli matrices obey some basic identities as defined in [138]. The multiplication of two Pauli matrices is given by

$$\sigma_j \sigma_k = \delta_{jk} I + i\epsilon_{jkl}\sigma_l \tag{A.9}$$

where $j, k = x, y, z$, I is the identity of order 2, δ_{jk} is the Kronecker delta and ϵ_{ijk} is the Levi-Cevita symbol.

This leads to the following commutator and anti-commutator relations

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l, \tag{A.10}$$

$$\{\sigma_j, \sigma_k\} = 2\delta_{jk}I, \tag{A.11}$$

and the relation of the Pauli vector to the dot product and cross product

$$(\boldsymbol{a}\cdot\boldsymbol{\sigma})(\boldsymbol{b}\cdot\boldsymbol{\sigma}) = (\boldsymbol{a}\cdot\boldsymbol{b})I + i(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{\sigma}. \tag{A.12}$$

where **a** and **b** are three vectors.

The general eigenspinors for a spin orientation defined in spherical coordinates by θ and ϕ are [138]

$$\chi_+ = \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\sin\theta/2 \end{pmatrix} \tag{A.13}$$

$$\chi_- = \begin{pmatrix} \cos(\theta/2) \\ -e^{i\phi}\sin\theta/2 \end{pmatrix}. \tag{A.14}$$

A.4. Unitary transformations

This section describes the steps and identities used to apply the unitary transformation in Section 5.3.3 when deriving the one-band effective model. The process requires the Pauli matrix identities from Appendix A.3 and the Baker-Campbell-Hausdorff formula.

A.4.1. The Baker-Campbell-Hausdorff formula

The Baker-Campbell-Hausdorff formula [139] states that a unitary transformation of e^{-X} on Y is given by

$$e^X Y e^{-X} = \sum_{n=0}^{\infty} \frac{[(X)^n, Y]}{n!} \tag{A.15}$$

where the iterative commutator is defined as $[(X)^n, Y] = [X, \cdots [X, [X, Y]] \cdots]$ with $[(X)^0, Y] \equiv Y$. This can alternatively be written as

$$e^X Y e^{-X} = Y + [X, Y] + \frac{1}{2!}[X, [X, Y]] + \frac{1}{3!}[X, [X, [X, Y]]] + \cdots . \tag{A.16}$$

A.4.2. Rotation of the spin basis

When performing perturbation theory on equation (5.3) it is useful to align the spin axis x along the magnetic field direction $\hat{n} = \cos\theta \hat{x} + \sin\theta \hat{y}$. This can be achieved by applying the unitary transform $U = e^{-i/2\theta\sigma_z}$, a rotation of θ about the z axis. The unitary transformation only influences the Pauli matrices. Applying this unitary transformation to σ_x and using the Baker-Campbell-Hausdorff formula (A.16) one obtains

$$\begin{aligned} U^\dagger \sigma_x U = e^{i/2\theta\sigma_z} \sigma_x e^{-i/2\theta\sigma_z} &= \sigma_x + \frac{i\theta}{2}[\sigma_z, \sigma_x] + \frac{1}{2!}\left(\frac{i\theta}{2}\right)^2 [\sigma_z, [\sigma_z, \sigma_x]] \\ &+ \frac{1}{3!}\left(\frac{i\theta}{2}\right)^3 [\sigma_z, [\sigma_z, [\sigma_z, \sigma_x]]] + \cdots \end{aligned} \tag{A.17}$$

From the Pauli matrix commutator relation (A.10), the commutators are

$$\begin{aligned} [\sigma_z, \sigma_x] &= 2i\sigma_y, \\ [\sigma_z, [\sigma_z, \sigma_x]] &= 2i[\sigma_z, \sigma_y] = -(2i)^2 \sigma_x = 4\sigma_x, \\ [\sigma_z, [\sigma_z, [\sigma_z, \sigma_x]]] &= 4[\sigma_z, \sigma_x] = 8i\sigma_y. \end{aligned}$$

Substituting these commutator expressions back into (A.17) and rearranging gives

$$\begin{aligned} U^\dagger \sigma_x U &= \sigma_x \left[1 - \frac{\theta^2}{2!} + \cdots\right] - \sigma_y \left[\theta - \frac{\theta^3}{3!} + \cdots\right] \\ &= \sigma_x \cos\theta - \sigma_y \sin\theta \end{aligned} \tag{A.18}$$

where the last equality is true from the Taylor expansion of $\cos\theta$ and $\sin\theta$ about $\theta = 0$. Similarly, it can be shown that the unitary transform U applied to σ_y gives

$$U^\dagger \sigma_y U = \sigma_x \sin\theta + \sigma_y \cos\theta. \tag{A.19}$$

Applying the transformation to the Zeeman term of equation (5.3) with $\mathbf{B} = B\hat{n}$ gives

$$\begin{aligned} U^{\dagger}H_{Z}U &= \frac{g\mu_{B}B}{2}\left\{\cos\theta U^{\dagger}\sigma_{x}U + \sin\theta U^{\dagger}\sigma_{y}U\right\} \\ &= \frac{g\mu_{B}B}{2}\left\{\cos\theta(\sigma_{x}\cos\theta - \sigma_{y}\sin\theta) + \sin\theta(\sigma_{x}\sin\theta + \sigma_{y}\cos\theta)\right\} \\ &= \frac{g\mu_{B}B}{2}\left\{\sigma_{x}(\cos^{2}\theta + \sin^{2}\theta) + \sigma_{y}(-\cos\theta\sin\theta + \sin\theta\cos\theta)\right\} \\ &= \frac{g\mu_{B}B}{2}\sigma_{x} \end{aligned} \tag{A.20}$$

where the identity $\cos^2\theta + \sin^2\theta = 1$ was used to produce the last equality. This proves that the unitary transformation $U = e^{-i/2\theta\sigma_z}$ can be used to align the x spin matrix with the magnetic field direction $\hat{n}$. The rest of the Hamiltonian can be transformed by applying the unitary transformation, $U^{\dagger}HU$, and using equations (A.18) and (A.19).

A.5. Time-independent perturbation theory

The time-independent perturbation theory [138] is used to derive the effective model expression of the effective g-factor for the lowest subband of a nanowire. Time-independent perturbation theory starts with an unperturbed time-independent Hamiltonian, H_0, with known eigenvalues $E_n^{(0)}$ and eigenfunctions $\left|n^{(0)}\right\rangle$ with $n = 1, 2, 3, \cdots$. Additionally the energy spectrum is assumed to be non-degenerate. A weak perturbation H' is then introduced to the Hamiltonian with a dimensionless parameter λ ranging continuously between 0 to 1 with $\lambda = 0$ corresponding to no perturbation and $\lambda = 1$ the full perturbation. The perturbed Hamiltonian can be written

$$H = H_0 + \lambda H'. \tag{A.21}$$

The eigenvalues and eigenfunctions of this perturbed Hamiltonian satisfy the time-independent Schrödinger equation

$$H\left|n\right\rangle = E_n\left|n\right\rangle. \tag{A.22}$$

The goal of perturbation theory is to determine the eigenfunctions $|n\rangle$ and eigenvalues E_n in terms of the unperturbed eigenfunctions $\left|n^{(0)}\right\rangle$ and eigenvalues $E_n^{(0)}$. This can be

achieved for a weak perturbation by writing them as a power series expansion in λ

$$\begin{aligned} |n\rangle &= \left|n^{(0)}\right\rangle + \lambda\left|n^{(1)}\right\rangle + \lambda^2\left|n^{(2)}\right\rangle + \cdots \\ E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \end{aligned} \tag{A.23}$$

These expansions can be substituted back into the time independent Schrödinger equation (A.22). By matching the coefficients of each power of λ an infinite series of simultaneous equations is created.

The zeroth order equation is simply the unperturbed time independent Schrödinger equation

$$H_0\left|n^{(0)}\right\rangle = E_n^{(0)}\left|n^{(0)}\right\rangle. \tag{A.24}$$

The first order equation can be used to derive the first order correction to energy

$$E_n^{(1)} = \left\langle n^{(0)}\middle|H'\middle|n^{(0)}\right\rangle, \tag{A.25}$$

and the first order correction to the eigenfunctions

$$\left|n^{(1)}\right\rangle = \sum_{k\neq n}\frac{\left\langle k^{(0)}\right| H'\left|n^{(0)}\right\rangle}{E_n^{(0)} - E_k^{(0)}}\left|k^{(0)}\right\rangle. \tag{A.26}$$

The second order equation can be used to derive the second order correction to the energy

$$E_n^{(2)} = \sum_{k\neq n}\frac{\left|\left\langle k^{(0)}\right| H'\left|n^{(0)}\right\rangle\right|^2}{E_n^{(0)} - E_k^{(0)}}. \tag{A.27}$$

Higher order corrections can be derived but the calculations become tedious. Typically, only the second order energy corrections is needed which is the case for the effective model in Section 5.3.3.

Appendix B.

Measurement techniques

The original intention of this thesis was to perform quantum transport measurements on InAs nanowires like those modelled in Chapter 5. However, due to delays experienced by our collaborators in preparing the InAs nanowires these measurements could not be performed within the MPhil candidature. The results of the InAs nanowire modelling became the focus of the thesis. This section and Appendix C are included to describe the low temperature measurement setup that was prepared during candidature to perform single electron charging through dopants in silicon Chapter 3, and was intended for quantum transport measurements of the InAs nanowires.

The low temperature experimental setup can be split into two main parts: the measurement setup and the cooling system. The measurement setup consists of:

- a cryostat sample area where a chip of devices can be placed,
- a home-made control box connecting the sample area to the outputs/inputs of instruments,
- instruments interfaced to a computer for remote control,
- code that instructs the instruments to output and/or measure.

The cooling system consists of:

- a cryostat with a sample area inside a vacuum chamber,
- a pump to create the vacuum and draw liquid helium through the cryostat flow line,
- a dewar full of liquid helium that the cryostat can be inserted into,

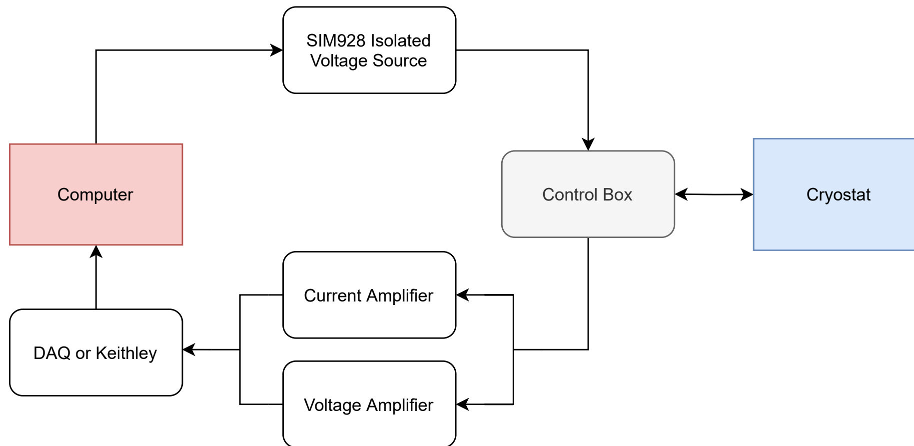


Figure B.1. The measurement setup summarised as a diagram.

- a temperature controller which senses the cryostat's temperature and can control the cryostat's internal resistor to heat the sample.

This appendix will focus on the preparation and room temperature testing of the measurement setup for a single gate transistor used in Chapter 3. Appendix C will outline the preparation and operation of the cooling setup.

B.1. Measurement setup

The measurement setup can be summarised by Figure B.1. This section will describe the different components of the diagram and discuss the measurement philosophy of the setup.

B.1.1. Computer

Beginning from the left side of Figure B.1, the lab computer is a desktop set up with CentOS Linux 8. The computer is connected to the instruments e.g. voltage sources and Keithley via a GPIB card installed in the back. An open-source Python 3 package Qcodes is used to interfacing with the drivers that control the instruments easily. Qcodes is developed by the Copenhagen/Delft/Sydney/Microsoft quantum computing consortium and is widely used in the field of quantum materials [80]. This package was chosen because of its simple measurement functions, neat database saving and the fact that

most of the instruments already have virtual instrument wrappers written by the Qcodes community.

B.1.2. Instruments

The different instruments are connected to the computer via GPIB or inserted directly into the computer. Above the computer in the diagram is the Standford Research System Small Instrument Module 928 (SRS SIM928) isolated voltage source used to output DC voltage to a device in the sample area. Below the computer is an instrument used to measure voltages from the amplifiers; either a National Instruments PCI 6221 data acquisition card (DAQ) with a BNC 2090A accessory or the Keithley 2000 Multimeter. The Femto DLPCA-200 current amplifier before the DAQ/Keithley measures current from the circuit in the sample area and outputs it as an amplified voltage. In a similar way a voltage amplifier could be used to amplify voltage from the device to the Keithley or the DAQ. The voltage sources, DAQ and Keithley 2000 are all controlled by the computer while the current amplifier is set manually in the lab. Each instrument is described briefly below.

SRS SIM928 Isolated Voltage Source

The SIM928 voltage sources are modules that slot into a SRS SIM900 Mainframe. The SIM900 provides power, computer interfacing via GPIB and clock synchronisation to up to eight modules. The SIM900 is powered from the grid and provides power to the control circuitry of the SIM928 modules. However, the output stage of each module is powered by its own internal battery so that voltage output is isolated from grid noise. The SIM928 has a resolution of 1 mV, maximum output voltage of ±20 V, typical noise of 10 μV_{rms} (1 kHz bandwidth) and maximum output current of ±10 mA. The SIM900 is shown in Figure B.2 with four SIM928 modules and a SIM911 Preamp module that was not used here.

Keithley 2000 Multimeter

The Keithley 2000 is a multimeter that can be used to measure DC/AC voltages, DC/AC current and resistance. In this measurement setup it is only used to measure output DC voltages from the current or voltage amplifier. The Keithley 2000 can make DC

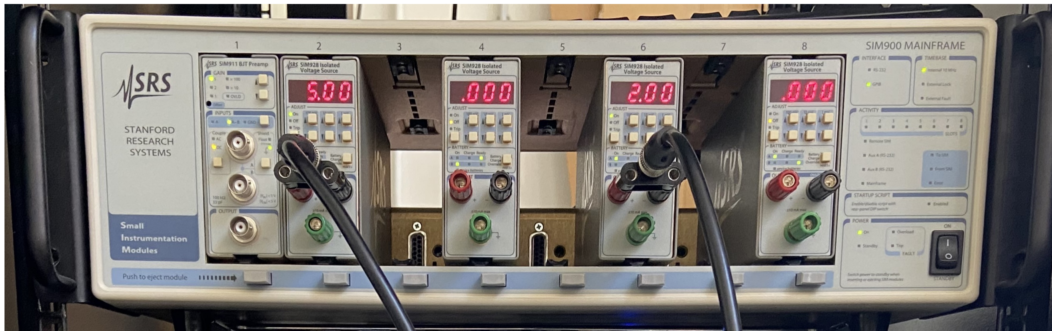


Figure B.2. The SIM900 with four SIM928 modules and a SIM911 module installed. Two of the SIM928 modules are outputting voltage via coaxial cables.

voltage measurements from 0.1 µV to 1 kV with the minimum resolution of 0.1 µV. The multimeter is connected to the computer via GPIB and measures DC voltage via banana plugs as seen in Figure B.3.

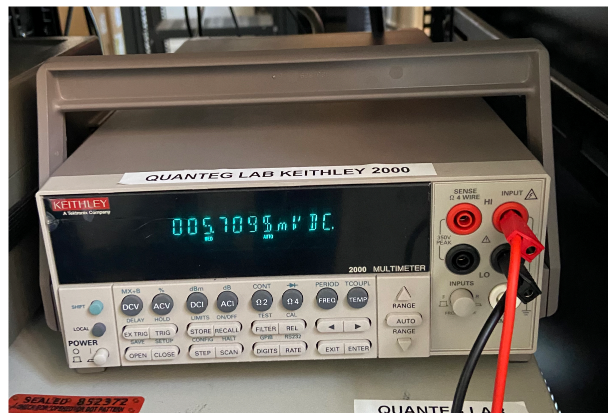


Figure B.3. Keithley 2000 multimeter setup to measure DC voltage across two banana plugs.

PCI 6221 DAQ card with BNC2090A accessory

The PCI 6221 DAQ card is installed into the computer and connected to the external rack accessory BNC2090A via a 68 pin connector. The BNC2090A has 2 BNC analog outputs (AO) and 16 BNC analog inputs (AI). For this project we use only the AIs for measuring voltages. The AOs will not be used because the SIM928 will see less grid noise than the DAQ which is powered by the grid.

Femto DLPCA-200 Variable Gain Low Noise Current Amplifier

The Femto DLPCA-200 is a current amplifier with a variable gain from 10^3 V/A to 10^{11} V/A and bandwidth from 1 Hz to 500 kHz. The voltage output can be AC or DC coupled and the bandwidth can be switched to 10 Hz for low noise DC measurements. The current amplifier also has an adjustable bias voltage. The connected current amplifier is shown in Figure B.4.

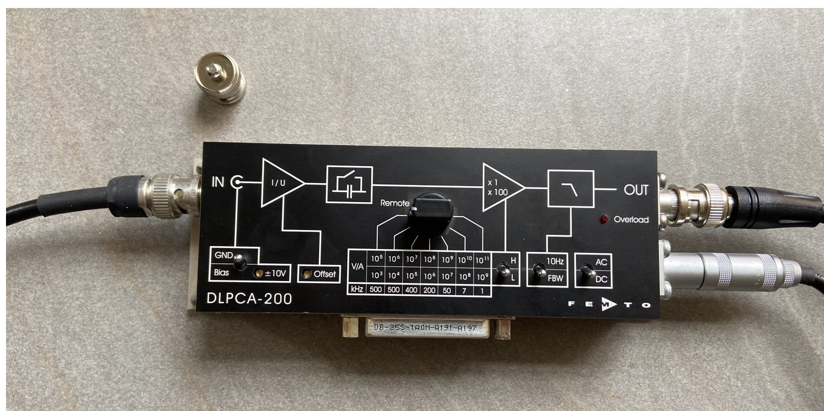


Figure B.4. The Femto DLPCA-200 variable gain low noise current amplifier receiving current input via a coaxial cable (left) and outputting voltage via a coaxial cable (right). The input has no voltage bias, the gain is set to 10^3 V/A and the output is full bandwidth with DC coupling. The connection on the bottom right is the power cable.

B.1.3. Control Box

The home-made control box has 16 male BNC connector inputs for instruments to connect to and a 24 pin cable that runs out of the box to the sample area of the fridge. Each of the 16 BNC connectors are wired to a switch immediately above it that can be grounded, floating or connected through the 24 pin cable to a pin in the cryostat sample area. Each of the wires from the BNC connector to the pin on the cryostat sample area have a resistance of 105 Ω. The control box is also used as the place where the experimentalist may ground themselves with an electrostatic discharge (ESD) grounding bracelet when transferring devices in and out of the cryostat sampling area or handling the cryostat. The control box setup is shown in Figure B.5.

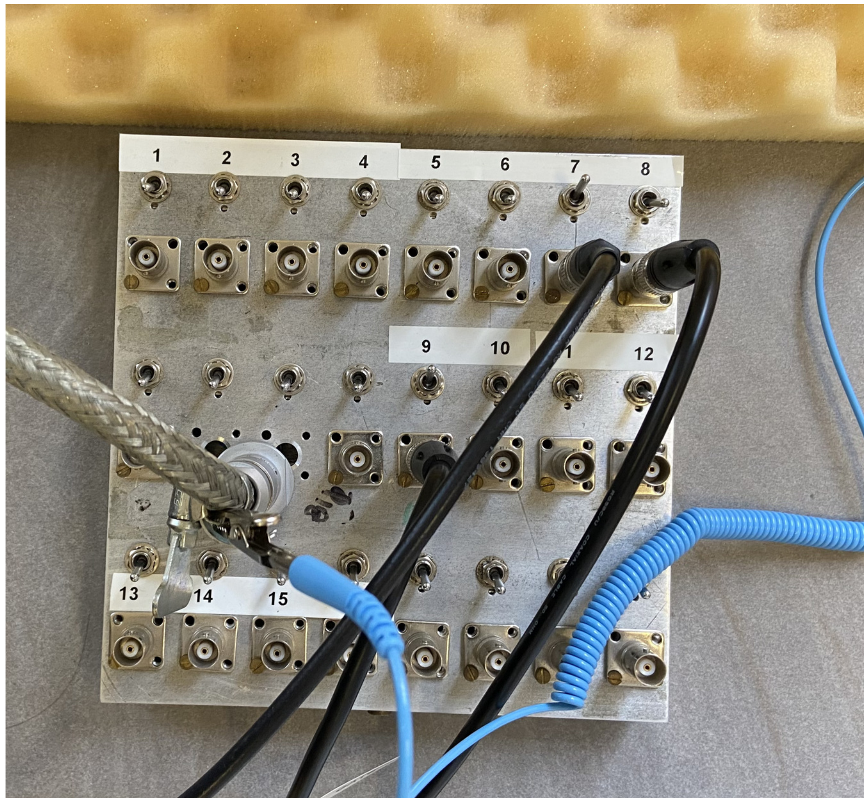


Figure B.5. The home-made control box with the 16 BNC connectors that run to the cryostat sampling area labelled. On the left is the 24 pin cable that runs to the cryostat and the blue cable connected to the casing of the cryostat cable runs to an electrostatic grounding bracelet. Note that connectors 7, 8 and 9 are connected to instruments via coaxial cables. Switch 7 and 9 are in the connected position, 8 is in the floating position and all other switches are off.

B.1.4. Cryostat

The last part of the measurement setup, is the cryostat. In this thesis an Oxford Instruments Compact Variable Temperature Insert (VTI) - Sample in vacuum cryostat is inserted into a large storage dewar of liquid helium to cool it down to 4 K. The sample area of the VTI is inside an Inner Vacuum Chamber (IVC) that must be pumped to vacuum prior to cooling. The sample area has a 16 pin female dual inline holder where male chip holders carrying device chips may be placed. The 16 pins connect to the inner pin of the BNC connectors on the control box so that one can apply and measure potentials of the device in the sample area. The VTI has a heat exchanger for heating the sample once cryogenic temperatures are reached and a Cernox 1050 calibrated thermometer for sensing the temperature of the heat exchanger. The heat exchanger and thermometer can be controlled by a temperature controller via the connection of a 6 pin

cable that was built in the lab. The VTI has Radio Frequency (RF) lines and a fibre optic that run to the sample area. The VTI also has a heat exchanger flow line that can be activated using an external pump to draw liquid helium through a coil above the IVC to cool the sample close to 1 K temperatures. The pressure of the helium flow can be controlled by releasing the needle valve. Figure B.6 is an annotated picture of the VTI.

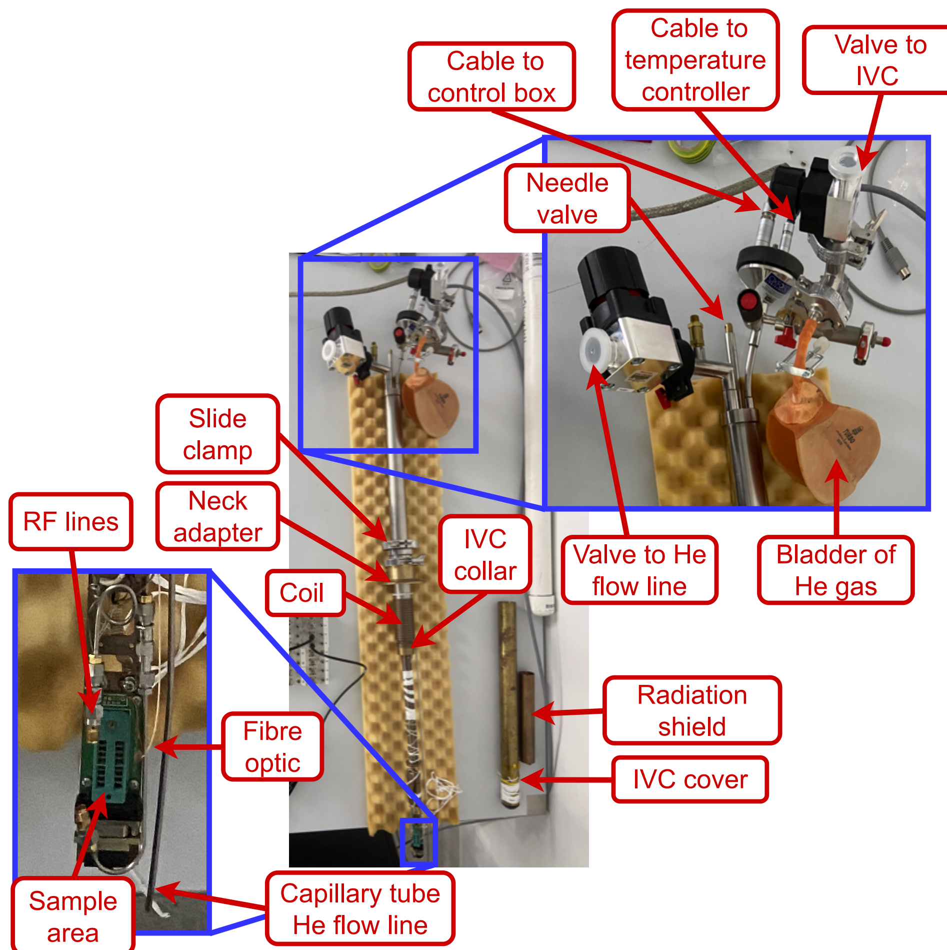


Figure B.6. The Oxford Instruments Compact Variable Temperature Insert (VTI) with the key components labelled.

B.1.5. Measurement philosophy

First, one must have a chip of devices in a chip holder with devices bonded to certain pins of the chip holder. This chip holder is placed in the sample area of the VTI. The chip holder pins are now connected (through the grounding switches) to the control box where the voltage sources and amplifiers can be connected to supply voltage or measure currents and voltages. The current or voltage amplifier output is connected to either the

DAQ card or the Keithley which will measure the output voltage of the amplifier. The voltage sources, DAQ card and Keithley are all connected to the computer and controlled by Python code.

To make a measurement of one of the devices on the chip, one must first ensure that the voltage sources and the current amplifier are connected to the correct pins, turn the switches to these pins from ground to on and run the prepared Python code for the measurement. The voltage sources will supply voltage to the devices and the current amplifier will send the current output as a voltage to be measured by the DAQ or Keithley. The results will then be saved on the computer using the database features of Qcodes for post experiment analysis.

B.2. Testing the measurement setup

This section outlines a room temperature test performed on the measurement setup to prepare it for measurements at cryogenic temperatures. The main setup used in this thesis is a simple single gate transistor setup for measuring IV curves.

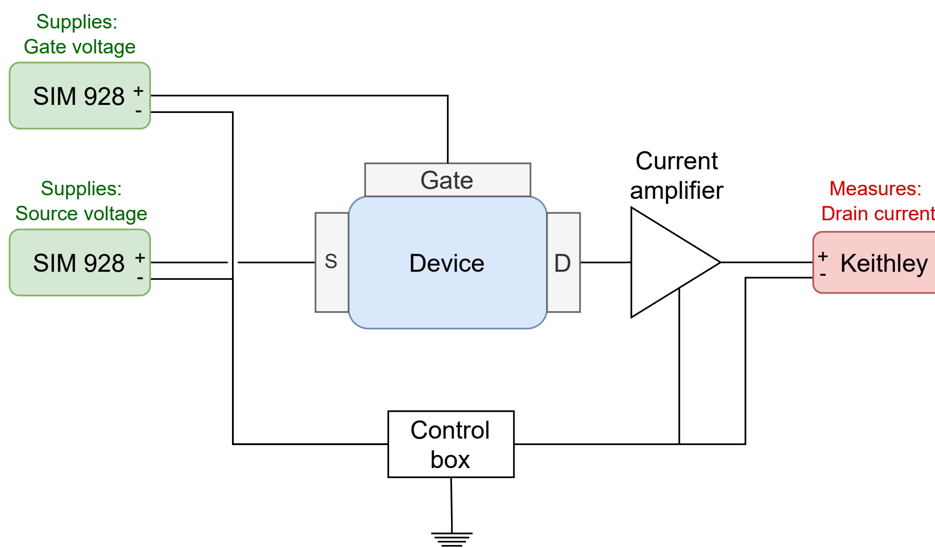


Figure B.7. The single gate transistor setup.

B.2.1. Transistor IV curve setup

The setup used for the measurement of single gate transistors is shown in the circuit diagram of Figure B.7. The SIM 928 isolated voltage sources supply the gate voltage and source voltage from their positive terminals. The current amplifier takes the drain current as an input and outputs a gained voltage measured by the positive terminal of the Keithley. The negative terminal of the SIM 928 voltage sources, the Keithley and the current amplifier are all connected to the control box as a common ground. The measurement setup was tested for a number of devices including resistors, Zener diodes and bipolar junction transistors. For brevity the resistor is discussed and not the transistor and Zener diode test.

The first basic test performed was producing the characteristic I-V curve of a 100 kΩ resistor. The resistor was placed directly in the sample area and the instruments connected to the corresponding control box pins to form the single gate transistor circuit without the gate. The current was measured for a voltage sweep from 100 mV to 120 mV with a current amplifier gain of 10^6 V/A. The result of the sweep is the IV curve shown in Figure B.8. The slope of the line corresponds to 102.3 kΩ which is within the 5% tolerance of the resistor. The wires from the control box to the cryostat each have a resistance of about 100 Ω, but this is less than 1% of the resistors value. The offset of the 0.01 µA was later discovered to be from the offset of the current amplifier.

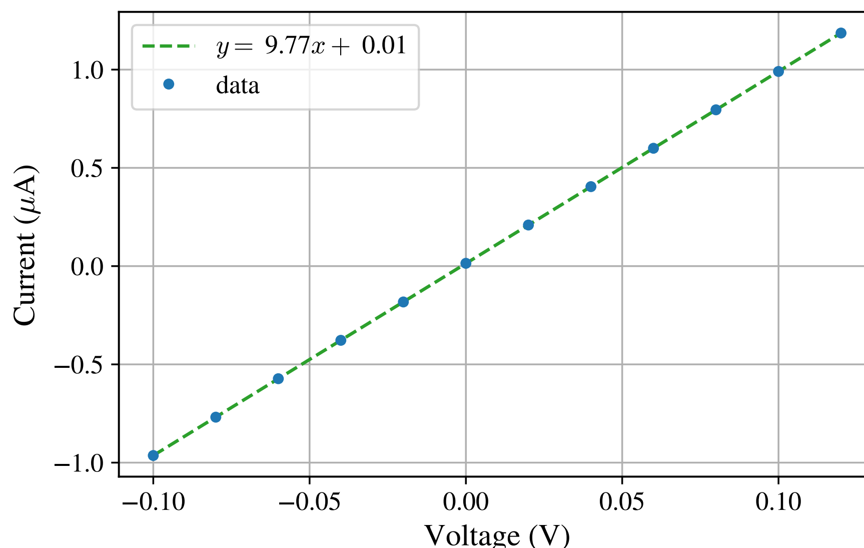


Figure B.8. The IV curve of the resistor in the VTI sample area measured by the current amplifier to the Keithley for different SIM928 voltages. The blue filled circles are the data and the green dashed line is the linear fit.

The single gate transistor setup was successful in measuring resistors, Zener diodes and transistors according to their manufacturers specifications. Hence, the setup was ready for the cryogenic measurements performed in Chapter 3.

Appendix C.

Cryogenic cooling

Quantum mechanical phenomena are more easily observed in a low temperature system where only the lowest energy states of a material are occupied and any changes to the system cannot be due to thermal excitations. The types of phenomena that are observable depend on the amount of thermal energy available to the system. For the measurements in this thesis, temperatures close to absolute zero are required. The conventional method for achieving these temperatures is by thermal contact with cryogenic liquids such as nitrogen and helium which have boiling points much lower than room temperature, 77 K and 4 K respectively. By bringing a sample in contact with these cryogenic liquids the sample can reach temperatures close to the boiling point of the liquid. By pumping helium through a flow line adjacent to the sample, temperatures of 1 K can be reached and by using more complex systems such as a dilution fridge temperatures on the order of 1 mK can be achieved. This chapter will discuss cooling with liquid helium, how one would cool a sample to 1 K, thermometry and an in depth description of the cooldown procedure used during the experiments of this thesis.

C.1. Cooling with liquid helium

From the second law of thermodynamics, when two isolated systems each at thermal equilibrium are brought into thermal contact, after a given period of time the systems will reach thermal equilibrium i.e. share the same temperature. This is the basis of cryogenic cooling with liquid helium. In the following experiments, the sample is cooled to 4 K by attaching it to the cold finger of the VTI and inserting it into an open neck dewar of liquid helium. The sample on the cold finger is inside a rough vacuum chamber made of

copper that is flushed with helium gas to provide thermal contact between the sample and the liquid cryogen. After a period of time the sample will cool to the temperature of liquid helium (4 K), that is thermal equilibrium will be reached.

C.1.1. Cooling to 1K

To reach temperatures below that of liquid helium's boiling point the cryostat has a flow line that can be pumped to create a continuous flow of liquid helium and cold gas through the cryostat's heat exchanger. The flow line takes in liquid helium from a small capillary tube at the bottom of the VTI and runs up the cryostat through a heater exchanger which wraps around the top of the finger in a coil and out to the pump. Refer to Figure B.6. The flow of cold gas through the heat exchanger cools the vacuum chamber and the sample space. The pressure of the cold gas running through the flow line can be controlled by the needle valve which in turn controls the temperature [140].

The heat exchanger is installed with a resistor thermometer, Cernox 1050, and a heating resistor that are connected to the temperature controller. Once the base temperature of approximately 1 K is reached the temperature controller can be used to increase the temperature of the heat exchanger and in turn warm the sample to desired temperatures. The sample is not in direct thermal contact with the helium reservoir so the heating of the sample should not create excess evaporation of the liquid helium which would occur otherwise [141].

C.2. Thermometry

Changes in a body's temperature can lead to changes in the physical properties of the body e.g. phase, dimensions, vapour pressure, electrical resistance. These changing properties can be utilised as a means for measuring temperature or temperature change.

For temperatures above $T \approx 50\,\mathrm{mK}$ one of the most useful measuring systems is the resistance thermometer. The electrical resistance in a resistor varies as the temperature changes. By determining the resistance of the resistor one can determine the temperature. The resistance thermometer used in this setup uses a direct determination measurement of the resistances but there are other methods such as Wheatstone bridge techniques [142] which can offer greater resolution with the risk of complicating the setup.

The direct determination measurement uses a potentiometric setup in which a current is passed through the resistor and the voltage across it is measured. The VTI resistor uses a four wire setup in which the two outer leads supply the current and the two inner leads measure voltage, see Figure C.1. The benefit of the four wire setup over a two wire setup is two-fold; the lead resistance is not included in the measurement and the high output impedance of the current source guarantees a fixed current independent of the lead resistance [142]. The high input impedance of the voltmeter ensures that little current will flow into the voltage leads thus avoiding any significant potential drop in the leads when a measurement is made.

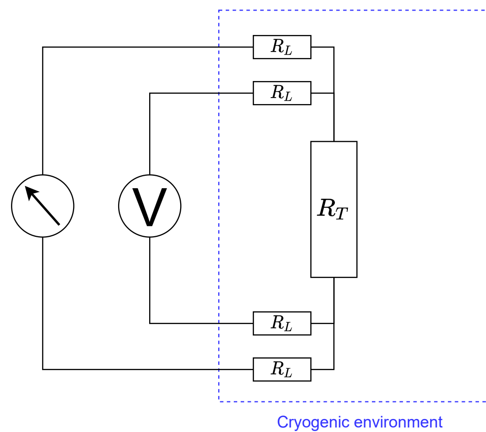


Figure C.1. The four wire setup used to measure the resistance thermometer R_T without being affected by the lead resistance R_L. Based on Figure 3.6b in [142].

The VTI resistor is a Cernox 1050 which is a thin-film 'ceramic' thermometer based on zirconium oxynitride. These are widely used because they are not as sensitive to magnetic fields as the metallic resistance thermometers. The thermometers have a negative $\frac{\mathrm{d}R}{\mathrm{d}T}$ being around $10^5\,\Omega$ at 1 K, $10^2\,\Omega$ at 100 K and 50 Ω at room temperature (300 K) [142]. The Cernox 1050 resistance thermometer is made for temperature sensing from 1 K to 400 K which is suitable for the cryogenic experiments performed in our laboratory. The temperature controller must be calibrated with a calibration curve of the thermometer's resistance against temperature in order to accurately sense the temperature.

C.3. The cooling procedure

The following cooling procedure was used in the cryogenic measurements of Chapter 3. The primary objective of the cooling procedure is to efficiently and safely cool devices

below 4 K for a long enough period of time that measurements can be performed. This section will outline and describe the required equipment, safety considerations and the step by step procedure of cooling the devices.

C.3.1. Required equipment

Similar to the measurement setup there are many essential components to a the cryogenic cooling setup. A list of required components is given below. A brief description of the main components; the cryostat, the pump setup and the temperature controller are also given.

1. Oxford Variable Temperature Insert (VTI) - Sample in vacuum
2. Vacuum pump setup (minimum pump speed of $8\,\mathrm{m^3h^{-1}}$)
3. Temperature Controller
4. 100 L storage dewar full of liquid helium with neck adapter to fit the VTI
5. Silicon Grease
6. Helium gas
7. Safety equipment
    a) Cryogenic Gloves
    b) Closed shoes
    c) Lab Coat
    d) Face shield or goggles, NOT safety glasses
    e) Oxygen Sensor (if the lab is too small to contain all the helium if it were to evaporate)

Cryostat

As introduced in the appendix B.1.4, the cryostat used for cooling to cryogenic temperatures in this work is the Oxford Instruments Compact Variable Temperature Insert (VTI) - Sample in vacuum. The devices are stored in a vacuum chamber and reach cryogenic temperatures by insertion of the VTI into a dewar of liquid helium.

By thermal contact with the liquid helium alone devices in the VTI can reach temperatures of 4 K. By pumping the flow line of the VTI which draws helium through the coil of the VTI, temperatures close to 1 K may be reached. There was not much preparation required for the VTI other than checking that the lines to the sample area pins are connected and the flow line is clear.

Vacuum Pump Setup

The pump station consists of a turbo molecular pump, the Agilent TwissTorr 74 FS, backed by a scroll pump, the Agilent IDP-7 dry scroll pump and two gauges; a Convectron gauge and a pirani gauge. The pump station is connected to the lines on the head of the VTI which lead to the helium flow line and IVC. A detailed schematic of the setup is given in Figure C.2. The scroll pump is used to back the turbo molecular pump and draw helium through the VTI flow line. The turbo molecular pump is used to pump the IVC to a rough vacuum < 0.5 mbar. It is possible to reach a rough vacuum of 10^{-2} mbar using the scroll pump, however, the turbo molecular pump can speed up the process and achieve a vacuum pressure as low as 10^{-10} mbar. The 275 Mini-Convectron gauge before the scroll pump can measure atmospheric pressure to 10^{-4} mbar while the Agilent FRG-700 full range pirani gauge can measure from just below atmospheric pressure to 10^{-9} mbar. The bladder of helium gas is used to flush the vacuum chamber and to release a small amount of helium gas into the IVC for thermal conductance. All the valves used are manually operated screw valves for maximum control of the pump lines. The operation of the vacuum pump setup will be described in appendix C.3.3.

The vacuum pump station was ordered in parts and built in the laboratory on a mobile trolley. Figure C.3 is an annotated image of the setup. All pumps and gauges were calibrated by the manufacturers. The offset of the Convectron gauge had to be tuned to atmospheric pressure using the offset screw.

Temperature Controller

The Lakeshore model 340 temperature controller is used to measure and control the temperature of the VTI. The temperature controller has a sensor that can measure the temperature of the VTI's resistor thermometer and a current source for sending current to the resistive heater in the VTI.

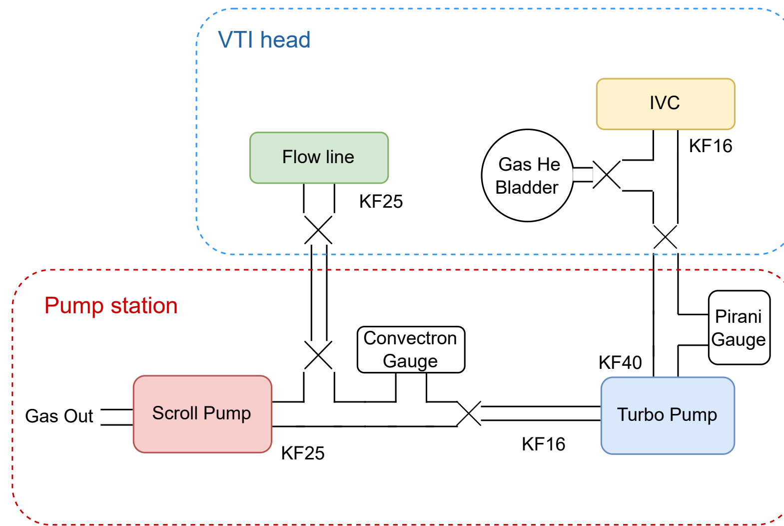


Figure C.2. A schematic of the vacuum pump setup. The bottom is the setup on the mobile pump station and the top is the connections on the head of the VTI leading to the helium flow line and inner vacuum chamber (IVC). The $\times$'s correspond to manual valves and the flange of inputs/outputs is labelled.

The 6 pin cable to connect the sensor and heater to the VTI was built in the lab. There are four pins for the sensor (I^+, I^-, V^+, V^-) and two for the heater (+,−). The manual recommends that the heater and the sensor connections be split into separate wires to reduce noise, however, for simplicity and neater cabling they were both put in a single cable 10 wire cable with twisted pairs and shielding. The temperature controller was calibrated for sensing the VTI's Cernox resistor using a temperature response curve from the Lakeshore disc. The response curve converts the sensor readings from sensor units to temperature units, Kelvin or Celsius.

C.3.2. Safety

Standard plastic gloves were worn whenever handling the VTI components or device chips so that no impurities or grease are present when being cooled in the liquid helium. This reduces the risk of damaging devices and dangerous explosions due to trapped oxygen freezing. Personal protective equipment was also worn when inserting and removing the VTI from the dewar of liquid helium.

One of the biggest dangers when working with liquid helium is asphyxiation, the lack of oxygen in the air due to its reaction with helium gas. The lab in which the experiment was performed is large enough that asphyxiation is not a danger but if it were performed

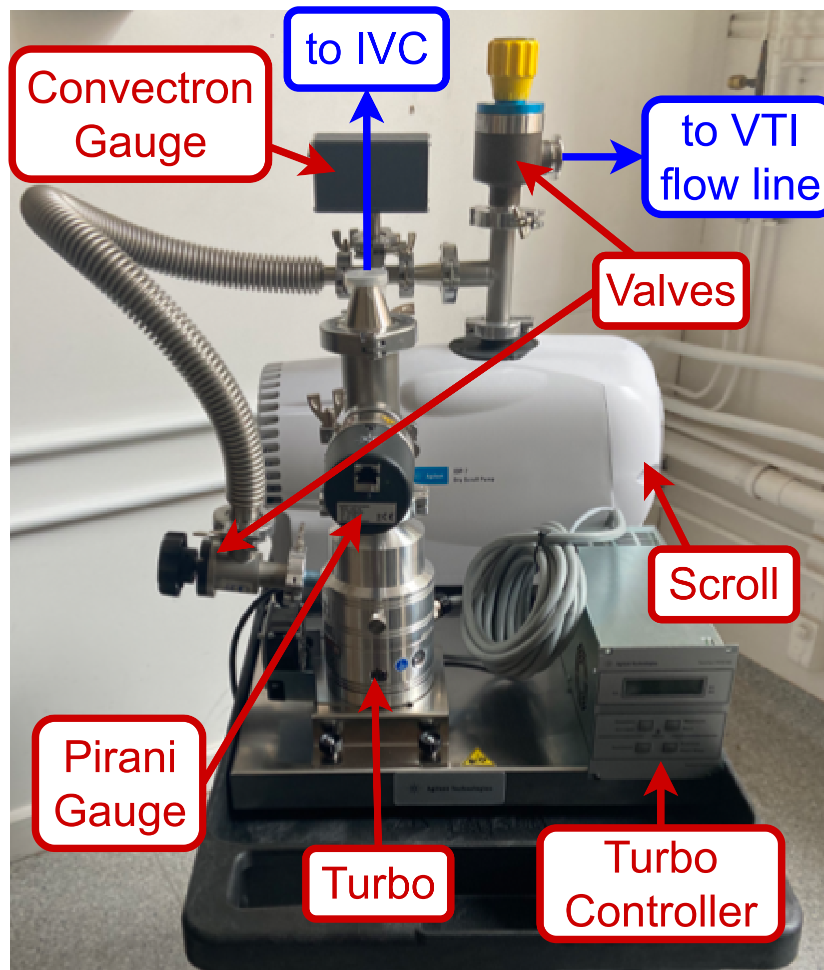


Figure C.3. An annotated image of the pump station. The red arrows identify the different components and the blue arrows indicate which parts of the VTI the pump lines connect to via bellows. Compare to the schematic Figure C.2.

in a smaller lab where the volume of evaporated helium from a full dewar is larger than the volume of oxygen in the room an oxygen sensor would be required.

C.3.3. The cool down procedure

Prior to beginning the cooling procedure the following checks were made:

- All instruments and the computer were on
- The instruments and computer were connected properly
- The control box was connected to the VTI

- The temperature cable was connected to VTI (and temperature controller read room temperature)
- All control box switches were grounded

After these checks the device chip was inserted into the sample area. The instruments were then connected to the control box pins based on the desired measurement circuit and the configuration of the devices on chip. It is good practice to test the devices at room temperature before cooling. For troubleshooting reasons one should also test the devices at each step of the cooling procedure e.g. after attaching the IVC cover, after pumping and after inserting VTI into the dewar. Then if the device breaks the experimentalist knows at which step the device stopped working and can easily determine the reason for device failure.

Once the measurement setup was ready and the device chip to be measured was secured in the VTI's sample area, the following cooling procedure was performed:

1. Prepare the IVC of the VTI
2. Insert the VTI into the dewar of liquid helium
3. Cool down measurement sweeps
4. Pump liquid helium through the VTI flow line
5. Take measurements above the base temperature

The following sections describe in detail these steps.

Preparing the IVC

In this step, the IVC cover was sealed in place over the VTI sample area by creating a rough vacuum. With the device chip in the sample area of the VTI and all its pins grounded the radiation shield was carefully slid on and screwed into place. The slide clamp for lowering the VTI into the dewar and the neck adapter for the dewar were carefully slid over the sample area to just above the coil of the VTI. A small amount of silicon vacuum grease was applied to the IVC collar of the VTI and the inside of the IVC cover. The IVC cover was then carefully slid over the end of the VTI sample area ensuring that the capillary tube of the flow line remains outside of the cover. White

plumbing tape was wrapped around the flow line and the IVC to keep the capillary tube in place. The IVC cover was pushed firmly in place.

Before connecting the pump to the IVC pump line, the bladder depicted in Figure B.6 was completely filled with helium gas, clamped off and attached to the nozzle on the IVC line. This helium must be used later to flush the IVC, hence, it is important the bladder is clamped so that the helium gas is not lost. With the pump station disconnected from the VTI lines and both valves before the scroll pump closed; the scroll pump, the turbo pump controller and gauges were turned on. Refer to Figure C.2 and Figure C.3. The setup was disconnected from the VTI so that the devices would not experience any sudden potential changes produced from turning on the pump setup. With the turbo pump off and the valve of the IVC closed the input line of the turbo pump was connected to the IVC. With the valve to the bladder of helium gas and the valve before the turbo closed, the valve after the turbo pump was opened. The valve before the turbo pump was then opened and the scroll pump used to pump the IVC.

Once the Convectron gauge read approximately 100 mTorr or 0.1 mbar on the Pirani gauge, the IVC valve was closed and the system flushed with some helium gas by slightly opening the valve to the helium gas bladder and closing it immediately after. The turbo pump was then turned on to achieve a vacuum of about 10^{-2} mbar as read on the Pirani gauge. Once the rough vacuum was achieved the valve before the turbo was closed and a small amount of gas helium was let into the IVC by slightly opening the valve to the bladder. The gas helium provides thermal contact with the liquid helium outside the IVC when the VTI is placed in liquid helium.

The rough vacuum firmly sealed the IVC cover onto the VTI. The turbo pump was spun down and then the valve after the turbo pump closed so that the pump setup could be disconnected at the valve of the IVC line. Once disconnected, the other components of the pump setup were turned off.

Insert the VTI into dewar

All the valves leading to the head of the VTI were sealed. The slide clamp and the neck adapter were placed just above the coil. The 100 L dewar of liquid helium was brought near the VTI so that the VTI may be inserted. The clamp holding the dewar lid on was opened and the dewar lid carefully removed and placed on an empty clean bench. The VTI was then slowly and carefully lowered into the dewar. Once the neck adaptor

rested on the dewar's neck it was clamped in place. The VTI was sticking out of the dewar as pictured in Figure C.4. The VTI was left in this position until the temperature controller reached a steady temperature.

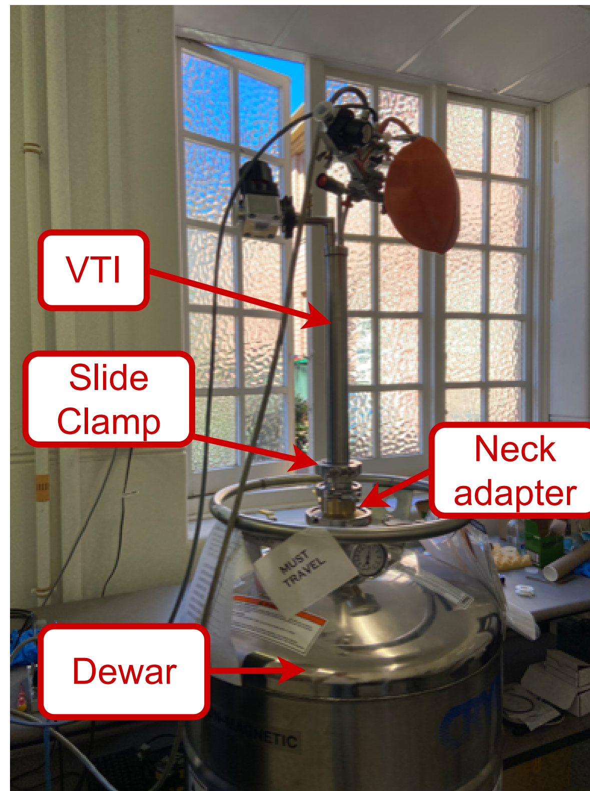


Figure C.4. The VTI after initial insertion into the dewar of liquid helium.

The liquid helium will evaporate on contact with anything warmer than 4.2 K, so the VTI had to be inserted slowly to reduce the amount of evaporation. Once, the VTI and helium reached equilibrium the VTI could be lowered further into the dewar in small increments. Ideally, the gas out of the liquid helium dewar goes to a helium liquefier to be recovered but the laboratory was not yet setup for this.

Initial cool down measurement

Once the VTI was inserted the device immediately started cooling. At this point an infinite sweep measurement of the device was initiated to ensure the device was still functioning and to observe the evolution of its operation as it approached lower temperatures. If the device did fail at some point during cool down, this would provide a good record of what happened in the lead up to the breaking of the device.

As mentioned above, to fully immerse the VTI in liquid helium the VTI was gradually lowered in increments to reduce excessive evaporation of liquid helium. One must follow a process of opening the clamp, further inserting the VTI by a small amount, closing the clamp, waiting for the temperature of the VTI and gas helium pressure from evaporation to settle before repeating the same steps. This was repeated until the VTI was fully immersed.

Pump the flow line

Once the VTI is fully inserted into the dewar of liquid helium the device should be 4 K because of thermal contact with the liquid helium. In order to cool the device further to 1 K the VTI flow line must be utilised. The scroll pump and the Convectron gauge were turned on while disconnected from the VTI. The scroll pump line was connected to the VTI flow line and the valves to the flow line opened. The needle valve was then opened slightly to draw helium through the flow line and the coil above the IVC. The needle valve was adjusted until the desired temperature was reached. The further the needle valve is opened the more helium is released.

Measurements at higher temperatures

From the base temperature reached, measurements can be made in successive, positive increments of temperature. This could be done either by heating the VTI's resistor with the temperature controller using a feedback loop or by gradually lifting the VTI out of the helium dewar and letting it equilibrate at higher temperatures. The latter method is much simpler to perform but causes a larger quantity of liquid helium to evaporate and cannot accurately select a temperature for the system. The temperature controller method requires more setup and calibration of the feedback loop but it gives more accurate results. For these experiments the latter method was employed for simplicity.

C.3.4. The warm-up procedure

The VTI was taken back to room temperature by performing the cool down procedure in reverse. The VTI was incrementally lifted from the lowest clamp position to the top ensuring minimal liquid helium evaporation. Once the VTI reached thermal equilibrium at the top of the dewar it was carefully removed. The VTI was left to warm up for a few

hours before the IVC vaccum was released. If the vacuum is released when the IVC is too cold, the air entering the vacuum would expand rapidly in the IVC and could break the wires of the device. After letting the VTI warm to room temperature the IVC valve was opened slowly to release the vacuum. The IVC cover was then carefully removed as to not damage devices. The devices were then removed from the sample area and stored in the lab.

Appendix D.

Quantum transport through dopants in silicon supplementary

D.1. Determining the energy barrier and conduction band minimum

The Coulomb peaks occurring in the current-voltage trace of Figure 3.7b at sub-threshold voltages can be due to the bound states of donors or the quantum dot formed by the energy barriers and the corners of the silicon channel. In this section, the location of the conduction band minimum is determined to differentiate between the donor and quantum dot states. At higher temperatures the current through the corner regions of the FinFET is dominated by thermionic emission of the carriers over the channel and the energy barriers E_B in Figure 3.8b. The thermionic emission model can be used to determine the height of the energy barrier and the cross-sectional area of the current flow. From the energy barrier height as a function of gate voltage, the conduction band minimum can be determined.

Thermionic emission is a conduction mechanism associated with a potential barrier. The electrons are transported over the barrier by thermal excitation. The thermionic-emission current density produced from thermally excited electrons can be determined by integrating the number of carriers above the barrier with energy E_B. The resulting current density is [78]

$$J = A^{*}T^{2}\exp\left(-\frac{E_B}{k_B T}\right) \tag{D.1}$$

where k_B is the Boltzmann constant and A^* is the effective Richardson's constant. For electrons in silicon the effective Richardson constant is $A^* = 2.1A$ with $A = 120\,\mathrm{Acm^{-2}K^{-2}}$ the Richardson constant [78]. The thermionic-emission current is then the current density equation (D.1) multiplied by the cross-sectional area of current flow S

$$I = SJ = SA^*T^2 \exp\left(-\frac{E_B}{k_B T}\right). \tag{D.2}$$

The natural log of equation (D.2), after rearranging, gives $\log\left(I/T^2\right)$ as a linear function of $1/T$

$$\log\left(I/T^2\right) = -\frac{E_B}{k_B}\frac{1}{T} + \log\left(SA^*\right). \tag{D.3}$$

Linear regression fits were performed for FinFET2's drain current data, $\log(I_D/T^2)$ against $1/T$. Figure D.1 is the drain current of FinFET2 as a function of $1/T$ for different gate voltages. The solid lines are created using the results of the linear regression fits and rearranging equation (D.3). The slope and intercept of these fits was used to determine the barrier energy E_B and the cross-sectional area of current flow S, respectively. The FinFET current is dominated by thermal emission above approximately 220 K.

The barrier height as a function of gate voltage is plotted in Figure D.2. The barrier energy decreases linearly with increased gate voltage. Oddly, at higher gate voltages the energy barrier is negative. From the works of Sellier et al. [44, 75], one would expect the energy barrier to decrease linearly with voltage while the conduction band of the channel is above the energy barriers. However, once the conduction band edge was below the energy barriers expected the barrier height to approach a plateau so that a small barrier height remains. The spacer regions in this FinFET may have a higher number of dopants and as a result have lower energy barriers that decrease with the conduction band below the Fermi level. However, given the small number of data points and smaller R^2 values in some of the thermionic emission fits, it is likely that more drain current data at higher temperatures is required.

The linear fit to the barrier height as a function of gate voltage in Figure D.2 has a $R^2 = 0.999$. The slope of the fit determines the coupling between the barrier and the gate, $\alpha = (1/e)(dE_B/dV_G) = 0.42 \pm 0.04$. The work of Sellier et al. [44, 75] with similar devices reports a coupling $\alpha = 0.68$. The weaker coupling in our devices could be due to a gate of different dimensions or old wire bonds. The y-intercept of the fit gives a barrier

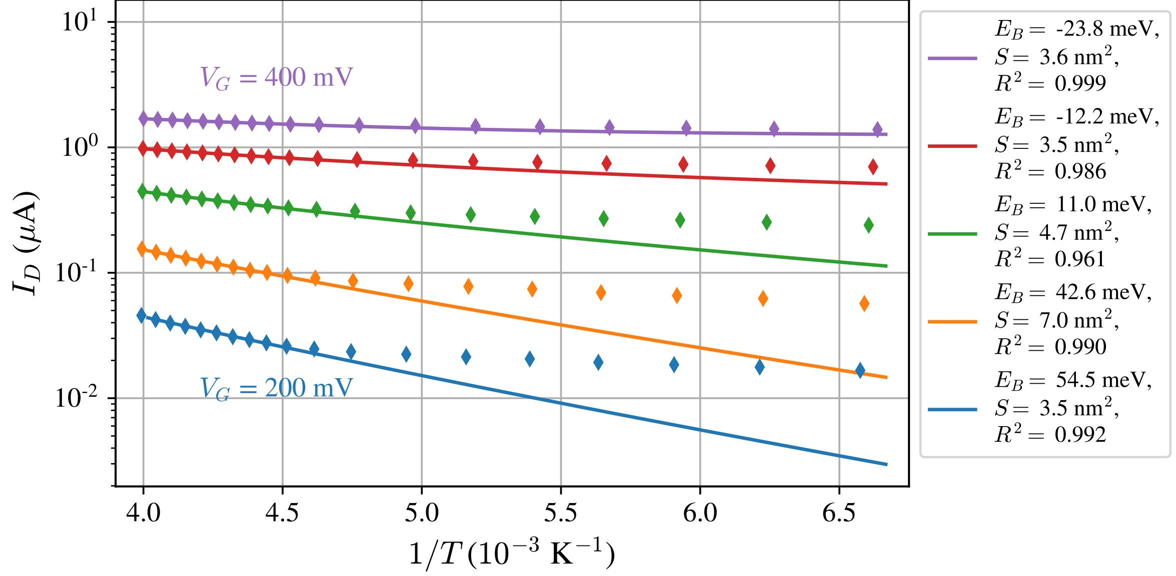


Figure D.1. Thermionic emission of the FinFETs. The diamonds are the FinFET2 drain current data for gate voltages steps of 50 mV. The solid lines represent the thermionic emission fit, equation (D.2), with the corresponding barrier energy E_B, cross-section S and R^2 value.

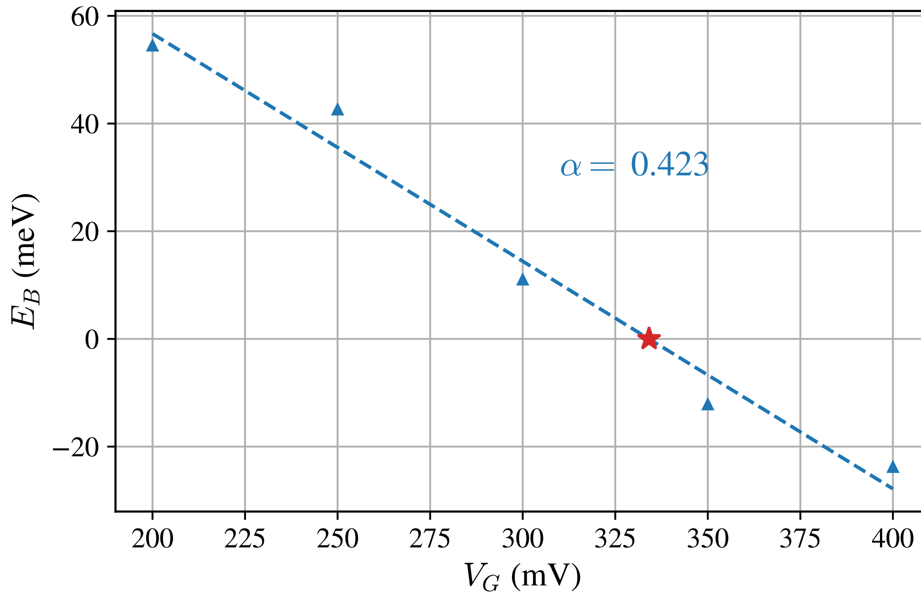


Figure D.2. The energy of the barrier as a function of gate voltage. The triangles are the values extracted from the thermionic emission fits and the dashed line is a fit to these values. The red star is the x intercept of the fit $V_G = 330 \pm 10$ mV. The slope of line is $-\alpha = dE_B/edV_G$.

height of 140 ± 10 mV at zero gate voltage, less than the 220 mV reported for similar devices [75]. The x-intercept of this fit, 330 ± 10 mV, corresponds to the point where the conduction band minimum matches the Fermi level $E_C = E_F$. Referring to Figure 3.8 and assuming the conduction band edge and barriers decrease together, the Fermi level and conduction band edge would align when the energy barriers are zero. Above gate voltages of 330 mV the low temperature current will rise due to the conduction band states and below 330 mV is the region where dopants states should occur.

The cross-sectional area of current flow S has an average value of $4.5 \pm 0.5\,\text{nm}^2$ across the five fits. This agrees with the work of Sellier [75] which found similar devices to have a cross section of $4\,\text{nm}^2$. This also confirms the corner effect since the cross-section of current flow is much smaller than the silicon fins cross-sectional area of at least $2000\,\text{nm}^2$.

D.2. Limits of the measurement setup

The results of this experiment are limited by two main factors. Firstly, the current measurements are limited by the resolution and range of the current amplifier. Secondly, the devices were only cooled to a temperature of 11 K and ideally temperatures of 4 K and below should be reached to measure clear Coulomb peaks.

D.2.1. Current amplifier gain and offset

The current amplifier measures the drain current and outputs it as a gained voltage which in this setup is measured by the Keithley 2000 multimeter. The current amplifier can output a maximum DC voltage of ± 10 V and has a switchable gain from 10^3 V/A to 10^{11} V/A. At the 10 V range, the Keithley multimeter has a resolution of 10 µV. This should give six orders of magnitude for current measurements at any current amplifier gain (10 µV to 10 V). However, due to the current amplifier's offset and the fixed output voltage range, the range of currents that can be read in a single stability diagram is limited.

During initial gate voltage sweeps in this low temperature run the output voltage of the current amplifier approached ~ 10 mV at zero gate voltage. This occurred for all sweeps irrespective of the current amplifier gain, see Figure D.3a. Upon investigation, this offset was due to the imbalance of the transistors in the current amplifier and could

be reduced by turning a small 'offset' screw on the current amplifier itself. Since this discovery was mid-cool down, the offset screw of the current amplifier was adjusted to minimise the output voltage without a reference current. The minimum offset that could be achieved mid-cool down was 1 mV improving the resolution of the current by one order of magnitude as in Figure D.3b. Ideally, this offset would be calibrated to a known constant DC current source with an accurate multimeter prior to the experiment.

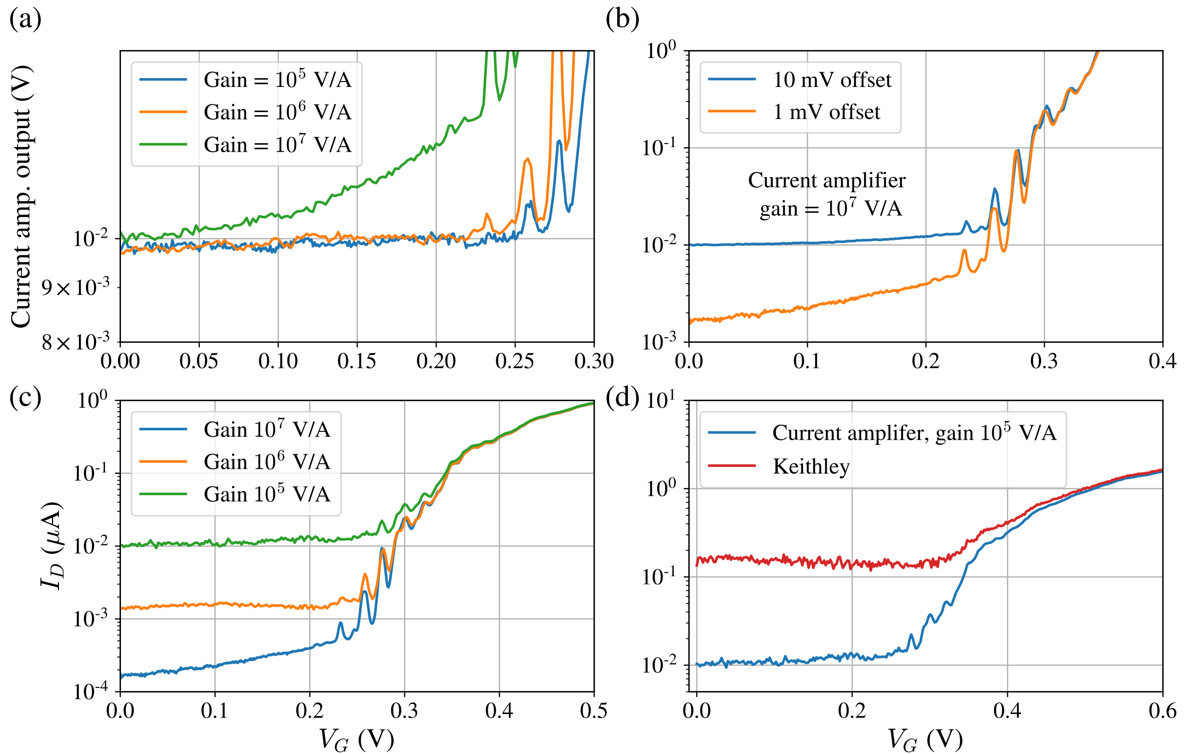


Figure D.3. Current amplifier troubleshooting. Gate voltage sweeps of FinFET2 with 1 mV source drain bias at $\sim$ 12 K. (a) The output of the current amplifier for different gains all approaching 10 mV at zero gate voltage. (b) The current amplifier output for the different offsets. (c) Drain current traces for various current amplifier gains. (d) Drain current traces of current measured with the current amplifier versus directly with the Keithley.

The range of the current amplifier voltage output limits the range of the currents that can be measured. The current resolution increases with the gain of the amplifier as seen in Figure D.3c. However, increasing the gain also reduces the maximum current output because the current amplifier can only output a maximum of 10 V. At 10^5 V/A gain the resolution is 10 nA and the maximum output is 100 µA while at 10^7 V/A gain the resolution is 100 pA and the maximum output is 1 µA. The Coulomb diamonds for these devices at 4 K and below typically span ±30 mV source drain voltage [44, 45, 46, 75]. The current at source drain voltages of 30 mV and gate voltages of 300 mV–400 mV is

the order of 10 µA. A stability with full Coulomb diamonds would require a current amplifier gain of 10^5 V/A so that these voltages do not overload the current amplifier. Unfortunately, this amplifier gain would not resolve Coulomb peaks of interest on the order of 100 pA (or 100 nS for conductance), see Figure 3.9 and [44, 45, 46]. The only way to resolve these peaks is to set the current amplifier gain to 10^7 V/A (or lower) which gives a resolution of 100 pA. However, then the current amplifier overloads at 1 µA and one has to settle for a small stability diagram like Figure 3.11 with ±10 mV source drain range. For a full stability diagram a current output from at least 10 pA to 10 µA, a range of 10^6, would be required. Two orders of magnitude larger than the current amplifier's output range in this experiment.

In attempts to solve this output range problem, the current amplifier was bypassed and the drain current was measured by the Keithley multimeter directly. The Keithley is sensitive to a larger range of current and is documented with a resolution of 1 nA. However, the minimum resolution of the Keithley's DC current measurement was found to be closer to 100 nA, see Figure D.3d. The current amplifier outputting voltage into the Keithley has a lower noise floor than measuring current with the Keithley directly.

Only the FinFET2 results were measured with the reduced current amplifier offset. The FinFET2 and FinFET1 stability diagram were taken at gains of 10^6 V/A and 10^7 V/A, respectively, and the Coulomb peak traces at 10^7 V/A. Ideally, a 10^7 V/A gain stability diagram would have been taken for FinFET2 to resolve the 10 pA peaks, however, the liquid helium in the storage dewar evaporated before this measurement and FinFET1 measurements could be reproduced with the reduced offset. Before future experiments it would be important to improve the range of drain current that can be measured in a single sweep. It may be possible to achieve this by interfacing the current amplifier with the computer and calibrating the offset with a reference current. Otherwise, a more accurate current amplifier or ammeter should be employed.

D.2.2. Problems encountered with cooling

The lowest temperature reached in these measurements was 11 K, with the VTI completely lowered into the dewar of liquid helium, which was less than half full ($\sim$ 40 L), and allowed to equilibrate for a day. The cryostat system should be able to reach temperatures below 4 K. There are three factors that may have contributed to this high temperature:

1. The VTI flow line valve was mistakenly not replaced after the previous cool down. As a result the flow line was open to warm air while the VTI was in the dewar which could raise the temperature of the devices and increased helium evaporation.

2. The remaining liquid helium in the dewar was not enough to cover the VTI and cool the devices to the boiling point of helium.

3. The Lakeshore was reading the incorrect temperature, whether that be due to poor calibration or problems with cabling, and the devices did actually reach 4 K.

Point 1 is likely the largest contribution to this higher temperature as it was determined prior to experiment that the valve was in fact left open during measurement. Considering that there is evidence of thermal fluctuations in the measurement results, namely a baseline current, broad Coulomb peaks and small Coulomb diamonds, it is unlikely to be the Lakeshore's calibration. The half full dewar is also likely to contribute to the high temperature but this is much harder to determine. Ideally, all these problems should be addressed before future experiments. The experimentalist should have a storage dewar of liquid helium close to full, the Lakeshore temperature controller should be recalibrated and cabling tested and all valves on the VTI should be fitted and sealed prior to cool down.

Appendix E.

Effective g-factor anisotropy of InAs nanowires supplementary

E.1. Magnetic field dependence of the effective g-factor anisotropy

The magnetic field applied in the simulations of Chapter 5 is quite large (5 T). The reason for this large applied magnetic field is so that the lowest subband spin states experience a large spin splitting and can be easily resolved by NEMO3D's Lanczos algorithm. Spin states that are too close together typically take longer to find or require a more intensive version of the Lanczos algorithm. Although this is a large magnetic field which is difficult to apply in practice, it should capture the same g-factor anisotropy as for smaller magnetic fields. This is because the subband splitting is the dominant energy scale in the confined systems explored in this work and therefore the influence of the magnetic field is negligible according to equation (5.18). To verify that applying a large magnetic field does not change the g-factor anisotropy, the effective g-factor of the lowest conduction subband was calculated for the $L = 10\,\mathrm{nm}$ mixed-ion terminating nanowire with different magnetic field magnitudes. The results are shown in Figure E.1. These results clearly illustrate that the magnitude of the magnetic field applied has negligible effect on the g-factor anisotropy for three different magnetic fields ranging from 1 T to 9 T. Hence, applying a magnetic field of 5 T for the sake of simplifying computations should not produce unphysical results.

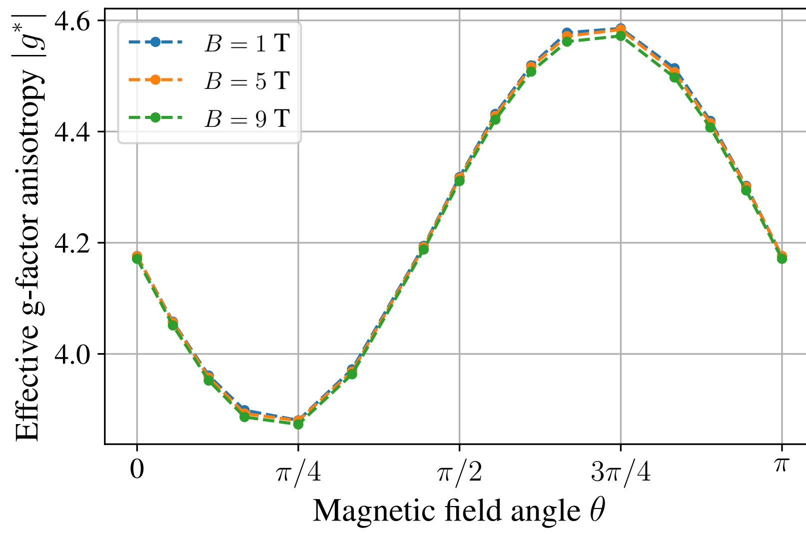


Figure E.1. The effective g-factor anisotropy of the lowest conduction subband of the $L = 10\,\text{nm}$ mixed-ion termination nanowire for different applied magnetic fields B.

Bibliography

[1] A. Chodos, *November 17 - december 23, 1947: Invention of the first transistor*, *APS News* **9** (2000) .

[2] W. Shockley, *Circuit element utilizing semiconductive material*, U.S. Patent 2569347A, 1948.

[3] F. M. Wanlass, *Low stand-by power complementary field effect circuitry*, U.S. Patent 3356858A, 1963.

[4] G. E. Moore, *Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp.114 ff.*, *IEEE Solid-State Circuits Society Newsletter* **11** (Sept., 2006) 33–35.

[5] S. K. Moore, *Another step toward the end of moore's law: Samsung and TSMC move to 5-nanometer manufacturing - [news]*, *IEEE Spectrum* **56** (June, 2019) 9–10.

[6] J. Liu, S. Mukhopadhyay, A. Kundu, S. Chen, H. Wang, D. Huang et al., *A reliability enhanced 5nm CMOS technology featuring 5th generation FinFET with fully-developed EUV and high mobility channel for mobile SoC and high performance computing application*, *IEEE International Electron Devices Meeting (IEDM)* (2020) .

[7] G. Yeap, S. S. Lin, Y. M. Chen, H. L. Shang, P. W. Wang, H. C. Lin et al., *5nm CMOS production technology platform featuring full-fledged EUV, and high mobility channel FinFETs with densest 0.021μm^2 SRAM cells for mobile SoC and high performance computing applications*, *2019 IEEE International Electron Devices Meeting (IEDM)* (2019) .

[8] T. S. M. Company, "3 nm technology." `https://www.tsmc.com/english/dedicatedFoundry/technology/logic/l_3nm`.

[9] S. M. Yakout, *Spintronics: Future technology for new data storage and communication devices*, *Journal of Superconductivity and Novel Magnetism* **33** (May, 2020) 2557–2580.

[10] A. Steane, *Quantum computing*, *Reports on Progress in Physics* **61** (Feb., 1998) 117–173.

[11] L. Gyongyosi and S. Imre, *A survey on quantum computing technology*, *Computer Science Review* **31** (Feb., 2019) 51–71.

[12] J. C. McCallum, B. C. Johnson and T. Botzem, *Donor-based qubits for quantum computing in silicon*, *Applied Physics Reviews* **8** (Sept., 2021) 031314.

[13] M. F. Gonzalez-Zalba, S. de Franceschi, E. Charbon, T. Meunier, M. Vinet and A. S. Dzurak, *Scaling silicon-based quantum computing using CMOS technology*, *Nature Electronics* **4** (Dec., 2021) 872–884.

[14] A. Y. Kitaev, *Unpaired majorana fermions in quantum wires*, *Physics-Uspekhi* **44** (2001) 131–136.

[15] J. D. Sau, S. Tewari, R. M. Lutchyn, T. D. Stanescu and S. D. Sarma, *Non-abelian quantum order in spin-orbit-coupled semiconductors: Search for topological majorana particles in solid-state systems*, *Physical Review B* **82** (Dec., 2010) 214509.

[16] L. Fu and C. L. Kane, *Superconducting proximity effect and majorana fermions at the surface of a topological insulator*, *Physical Review Letters* **100** (2008) .

[17] J. Alicea, *New directions in the pursuit of majorana fermions in solid state systems*, *Reports on Progress in Physics* **75** (2012) 36.

[18] D. Kim, Z. Shi, C. B. Simmons, D. R. Ward, J. R. Prance, T. S. Koh et al., *Quantum control and process tomography of a semiconductor quantum dot hybrid qubit*, *Nature* **511** (July, 2014) 70–74.

[19] M. Koch, J. G. Keizer, P. Pakkiam, D. Keith, M. G. House, E. Peretz et al., *Spin read-out in atomic qubits in an all-epitaxial three-dimensional transistor*, *Nature Nanotechnology* **14** (Jan., 2019) 137–140.

[20] L. Fricke, S. J. Hile, L. Kranz, Y. Chung, Y. He, P. Pakkiam et al., *Coherent control of a donor-molecule electron spin qubit in silicon*, *Nature Communications* **12** (June, 2021) .

[21] J. J. Pla, K. Y. Tan, J. P. Dehollain, W. H. Lim, J. J. L. Morton, D. N. Jamieson et al., *A single-atom electron spin qubit in silicon*, *Nature* **489** (Sept., 2012) 541–545.

[22] S. Nadj-Perge, S. M. Frolov, E. P. A. M. Bakkers and L. P. Kouwenhoven, *Spin-orbit qubit in a semiconductor nanowire*, *Nature* **468** (2010) 1084–1087.

[23] S. Nadj-Perge, V. S. Pribiag, J. G. van den Berg, K. Zuo, S. R. Plissard, E. P. A. M. Bakkers et al., *Spectroscopy of spin-orbit quantum bits in indium antimonide nanowires*, *Physical Review Letters* **108** (2012) .

[24] J.-Y. Wang, G.-Y. Huang, S. Huang, J. Xue, D. Pan, J. Zhao et al., *Anisotropic pauli spin-blockade effect and spin–orbit interaction field in an InAs nanowire double quantum dot*, *Nano Letters* **18** (July, 2018) 4741–4747.

[25] S. Datta and B. Das, *Electronic analog of the electro-optic modulator*, *Applied Physics Letters* **56** (Feb., 1990) 665–667.

[26] P. Senellart, G. Solomon and A. White, *High-performance semiconductor quantum-dot single-photon sources*, *Nature Nanotechnology* **12** (Nov., 2017) 1026–1039.

[27] T. Xu, H. Wang, X. Chen, M. Luo, L. Zhang, Y. Wang et al., *Recent progress on infrared photodetectors based on inas and inassb nanowires*, *Nanotechnology* **31** (2020) 19.

[28] Y. V. Nazarov and Y. M. Blanter, *Quantum Transport: Introduction to Nanoscience*. Cambridge University Press, 2009.

[29] S. Datta, *Quantum Transport: Atom to Transistor*. Cambridge University Press, New York, 2005.

[30] B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel et al., *Quantized conductance of point contacts in a two-dimensional electron gas*, *Physical Review Letters* **60** (Feb., 1988) 848–850.

[31] D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J. E. F. Frost et al., *One-dimensional transport and the quantisation of the ballistic resistance*, *Journal of Physics C: Solid State Physics* **21** (Mar., 1988) L209–L214.

[32] L. P. Kouwenhoven, *Quantum adiabatic electron transport in ballistic conductors*, in *Physics of Low-Dimensional Semiconductor Structures*, pp. 463–498. Springer

US, 1993. DOI.

[33] H. van Houten and C. W. J. Beenakker, *Comment on "conductance oscillations periodic in the density of a one-dimensional electron gas"*, *Physical Review Letters* **63** (Oct., 1989) 1893–1893.

[34] C. W. J. Beenakker, *Theory of coulomb-blockade oscillations in the conductance of a quantum dot*, *Physical Review B* **44** (July, 1991) 1646–1656.

[35] J. M. Thijssen and H. S. J. V. der Zant, *Charge transport and single-electron effects in nanoscale systems*, *physica status solidi (b)* **245** (Aug., 2008) 1455–1470.

[36] M. Born and W. Heisenberg, *Zur quantentheorie der molekeln*, in *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pp. 216–246. Springer Berlin Heidelberg, 1985. DOI.

[37] K. Likharev, *Single-electron devices and their applications*, *Proceedings of the IEEE* **87** (Apr., 1999) 606–632.

[38] M. A. Kastner, *The single-electron transistor*, *Reviews of Modern Physics* **64** (1992) .

[39] S. B. Field, M. A. Kastner, U. Meirav and J. H. F. Scott-Thomas, *Conductance oscillations periodic in the density of one-dimensional electron gases*, *Physical Review B* **42** (1990) .

[40] U. Meirav and E. B. Foxman, *Single-electron phenomena in semiconductors*, *Semiconductor Science and Technology* **11** (Mar., 1996) 255–284.

[41] A. Tilke, F. Simmel, R. Blick, H. Lorenz and J. Kotthaus, *Coulomb blockade in silicon nanostructures*, *Progress in Quantum Electronics* **25** (2001) 97–138.

[42] H. Ahmed, *Single electron electronics: Challenge for nanofabrication*, *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures* **15** (Nov., 1997) 2101.

[43] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha and L. M. K. Vandersypen, *Spins in few-electron quantum dots*, *Reviews of Modern Physics* **79** (2007) .

[44] H. Sellier, G. P. Lansbergen, J. Caro, S. Rogge, N. Collaert, I. Ferain et al., *Transport spectroscopy of a single dopant in a gated silicon nanowire*, *Physical Review Letters* **97** (Nov., 2006) 206805.

[45] G. P. Lansbergen, R. Rahman, C. J. Wellard, I. Woo, J. Caro, N. Collaert et al., *Gate-induced quantum-confinement transition of a single dopant atom in a silicon FinFET*, *Nature Physics* **4** (June, 2008) 656–661.

[46] G. P. Lansbergen, R. Rahman, J. Verduijn, G. C. Tettamanzi, N. Collaert, S. Biesemans et al., *Lifetime-enhanced transport in silicon due to spin and valley blockade*, *Physical Review Letters* **107** (Sept., 2011) 136602.

[47] M. Fuechsle, J. A. Miwa, S. Mahapatra, H. Ryu, S. Lee, O. Warschkow et al., *A single-atom transistor*, *Nature Nanotechnology* **7** (Feb., 2012) 242–246.

[48] M. Pierre, R. Wacquez, X. Jehl, M. Sanquer, M. Vinet and O. Cueto, *Single-donor ionization energies in a nanoscale CMOS channel*, *Nature Nanotechnology* **5** (Dec., 2009) 133–137.

[49] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck et al., *Silicon quantum electronics*, *Reviews of Modern Physics* **85** (2013) 961–1019.

[50] V. V. Shorokhov, D. E. Presnov, S. V. Amitonov, Y. A. Pashkin and V. A. Krupenin, *Single-electron tunneling through an individual arsenic dopant in silicon*, *Nanoscale* **9** (2017) 613–620.

[51] R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems*. Springer-Verlag Berlin Heidelberg, 2003, 10.1007/b13586.

[52] J. J. Sakurai, *Advanced Quantum Mechanics*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1967.

[53] L. Thomas, *The kinematics of an electron with an axis*, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **3** (Jan., 1927) 1–22.

[54] M. Wu, J. Jiang and M. Weng, *Spin dynamics in semiconductors*, *Physics Reports* **493** (2010) 61–236.

[55] G. Dresselhaus, *Spin-orbit coupling effects in zinc blende structures**, *Physical Review* **100** (1955) .

[56] E. I. Rashba, *Properties of semiconductors with an extremum loop. 1. cyclotron and combinational resonance in a magnetic field perpendicular to the plane of the loop*, *Sov. Phys. Solid State* **2** (1960) 1109–1122.

[57] E. Majorana, *Teoria simmetrica dell'elettrone e del positrone*, Il Nuovo Cimento **14** (Apr., 1937) 171–184.

[58] M. Z. Hasan and C. L. Kane, *Colloquium: Topological insulators*, Reviews of Modern Physics **82** (Nov., 2010) 3045–3067.

[59] R. M. Lutchyn, J. D. Sau and S. D. Sarma, *Majorana fermions and a topological phase transition in semiconductor-superconductor heterostructures*, Physical Review Letters **105** (Aug., 2010) 077001.

[60] Y. Oreg, G. Refael and F. von Oppen, *Helical liquids and majorana bound states in quantum wires*, Physical Review Letters **105** (Oct., 2010) 177002.

[61] A. Kitaev, *Anyons in an exactly solved model and beyond*, Annals of Physics **321** (2006) .

[62] J. Kammhuber, M. C. Cassidy, F. Pei, M. P. Nowak, A. Vuik, O. Gül et al., *Conductance through a helical state in an indium antimonide nanowire*, Nature Communications **8** (Sept., 2017) .

[63] P. Streda and P. Seba, *Antisymmetric spin filtering in one-dimensional electron systems with uniform spin-orbit coupling*, Physical Review Letters **90** (2003) .

[64] V. Mourik, K. Zuo, S. M. Frolov, S. R. Plissard, E. P. A. M. Bakkers and L. P. Kouwenhoven, *Signatures of majorana fermions in hybrid superconductor-semiconductor nanowire devices*, Science **336** (2012) 1003–1007.

[65] A. Das, Y. Ronen, Y. Most, Y. Oreg, M. Heiblum and H. Shtrikman, *Zero-bias peaks and splitting in an al–InAs nanowire topological superconductor as a signature of majorana fermions*, Nature Physics **8** (Nov., 2012) 887–895.

[66] H. Zhang, C.-X. Liu, S. Gazibegovic, D. Xu, J. A. Logan, G. Wang et al., *RETRACTED ARTICLE: Quantized majorana conductance*, Nature **556** (Mar., 2018) 74–79.

[67] P. Wojcik, A. Bertoni and G. Goldoni, *Tuning rashba spin-orbit coupling in homogeneous semiconductor nanowires*, Physical Review B **97** (2018) .

[68] Y. V. Pershin, J. A. Nesteroff and V. Privman, *Effect of spin-orbit interaction and in-plane magnetic field on the conductance of a quasi-one-dimensional system*, Physical Review B **69** (Mar., 2004) 121306.

[69] O. Gül, H. Zhang, J. D. S. Bommer, M. W. A. de Moor, D. Car, S. R. Plissard et al., *Ballistic majorana nanowire devices*, *Nature Nanotechnology* **13** (Jan., 2018) 192–197.

[70] S. M. Albrecht, A. P. Higginbotham, M. Madsen, F. Kuemmeth, T. S. Jespersen, J. Nygård et al., *Exponential protection of zero modes in majorana islands*, *Nature* **531** (Mar., 2016) 206–209.

[71] T. D. Stanescu, R. M. Lutchyn and S. D. Sarma, *Majorana fermions in semiconductor nanowires*, *Physical Review B* **84** (Oct., 2011) 144522.

[72] J. D. Sau, S. Tewari and S. D. Sarma, *Experimental and materials considerations for the topological superconducting state in electron- and hole-doped semiconductors: Searching for non-abelian majorana modes in 1d nanowires and 2d heterostructures*, *Physical Review B* **85** (Feb., 2012) 064512.

[73] P. Anderson, *Theory of dirty superconductors*, *Journal of Physics and Chemistry of Solids* **11** (Sept., 1959) 26–30.

[74] J. Bardeen, L. N. Cooper and J. R. Schrieffer, *Theory of superconductivity*, *Physical Review* **108** (Dec., 1957) 1175–1204.

[75] H. Sellier, G. P. Lansbergen, J. Caro, S. Rogge, N. Collaert, I. Ferain et al., *Subthreshold channels at the edges of nanoscale triple-gate silicon transistors*, *Applied Physics Letters* **90** (Feb., 2007) 073502.

[76] R. K. Maurya and B. Bhowmick, *Review of FinFET devices and perspective on circuit design challenges*, *Silicon* **14** (Sept., 2021) 5783–5791.

[77] C. Hu, "Modern semiconductor devices for integrated circuits." `https://www.chu.berkeley.edu/modern-semiconductor-devices-for-integrated-circuits-chenming-calvin-hu-2010/`, 2010.

[78] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices.* John Wiley and Sons, New Jersey, 3rd ed., 2007.

[79] D. Liang, J. Du and X. P. Gao, *InAs nanowire devices with strong gate tunability: Fundamental electron transport properties and application prospects: A review*, *Journal of Materials Science & Technology* **31** (6, 2015) .

[80] M. Corporation and K. Universitet, "Qcodes."

https://qcodes.github.io/Qcodes/, 2016.

[81] J. Fossum, J.-W. Yang and V. Trivedi, *Suppression of corner effects in triple-gate MOSFETs*, *IEEE Electron Device Letters* **24** (Dec., 2003) 745–747.

[82] P. H. C. Eilers, *Parametric time warping*, *Analytical Chemistry* **76** (Dec., 2003) 404–411.

[83] S. He, W. Zhang, L. Liu, Y. Huang, J. He, W. Xie et al., *Baseline correction for raman spectra using an improved asymmetric least squares method*, *Anal. Methods* **6** (2014) 4402–4407.

[84] A. K. Ramdas and S. Rodriguez, *Spectroscopy of the solid-state analogues of the hydrogen atom: donors and acceptors in semiconductors*, *Rep Prog Phys* **44** (1981) .

[85] M. Taniguchi and S. Narita, *d^- state in silicon*, *Solid State Commun* **20** (1976) 131–133.

[86] P. Norton, *Formation of the upper hubbard band from negative-donor-ion states in silicon*, *Physical Review Letters* **37** (July, 1976) 164–168.

[87] G. Klimeck, S. Ahmed, H. Bae, N. Kharche, S. Clark, B. Haley et al., *Atomistic simulation of realistically sized nanodevices using NEMO 3-d—part I: Models and benchmarks*, *IEEE Transactions on Electron Devices* **54** (Sept., 2007) 2079–2089.

[88] G. Klimeck, S. Ahmed, N. Kharche, M. Korkusinski, M. Usman, M. Prada et al., *Atomistic simulation of realistically sized nanodevices using NEMO 3-d—part II: Applications*, *IEEE Transactions on Electron Devices* **54** (Sept., 2007) 2090–2099.

[89] G. Klimeck, F. Oyafuso, T. B. Boykin, R. C. Bowen and P. von Allmen, *Development of a nanoelectronic 3-d (nemo 3-d) simulator for multimillion atom simulations and its application to alloyed quantum dots*, *Computer Modeling in Engineering & Sciences* **3** (2002) 601–642.

[90] S. Ahmed, S. Sundaresan, H. Ryu and M. Usman, *Multimillion-atom modeling of InAs/GaAs quantum dots: interplay of geometry, quantization, atomicity, strain, and linear and quadratic polarization fields*, *Journal of Computational Electronics* **14** (Mar., 2015) 543–556.

[91] R. Rahman, G. P. Lansbergen, J. Verduijn, G. C. Tettamanzi, S. H. Park, N. Collaert et al., *Electric field reduced charging energies and two-electron bound*

excited states of single donors in silicon, Physical Review B **84** (Sept., 2011) 115428.

[92] R. Rahman, G. P. Lansbergen, S. H. Park, J. Verduijn, G. Klimeck, S. Rogge et al., *Orbital stark effect and quantum confinement transition of donors in silicon*, Physical Review B **80** (Oct., 2009) 165314.

[93] N. Kharche, M. Prada, T. B. Boykin and G. Klimeck, *Valley splitting in strained silicon quantum wells modeled with 2° miscuts, step disorder, and alloy disorder*, Applied Physics Letters **90** (Feb., 2007) 092109.

[94] N. Kharche, M. Luisier, T. B. Boykin and G. Klimeck, *Electronic structure and transmission characteristics of SiGe nanowires*, Journal of Computational Electronics **7** (Jan., 2008) 350–354.

[95] A. Gardin, R. D. Monaghan, T. Whittaker, R. Rahman and G. C. Tettamanzi, *Nonadiabatic quantum control of valley states in silicon*, Physical Review B **105** (Feb., 2022) 075406.

[96] W. M. C. Foulkes, L. Mitas, R. J. Needs and G. Rajagopal, *Quantum monte carlo simulations of solids*, Reviews of Modern Physics **73** (Jan., 2001) 33–83.

[97] P. Yu and D. Cardona, *Fundamentals of semiconductors.* Springer-Verlag Berlin Heidelberg, 2010.

[98] D. R. Hartree, *The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods*, Mathematical Proceedings of the Cambridge Philosophical Society **24** (Jan., 1928) 89–110.

[99] V. Fock, *Naherungsmethode zur losung des quantenmechanischen mehrkorperproblems*, Zeitschrift fur Physik **61** (Jan., 1930) 126–148.

[100] P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Physical Review **136** (Nov., 1964) B864–B871.

[101] W. Kohn and L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Physical Review **140** (Nov., 1965) A1133–A1138.

[102] J. C. Slater and G. F. Koster, *Simplified LCAO method for the periodic potential problem*, Physical Review **94** (June, 1954) 1498–1524.

[103] J.-M. Jancu, R. Scholz, F. Beltram and F. Bassani, *Empirical spds* tight-binding*

calculation for cubic semiconductors: General method and material parameters, Physical Review B **57** (Mar., 1998) 6493–6507.

[104] P.-O. Löwdin, *On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals*, The Journal of Chemical Physics **18** (Mar., 1950) 365–375.

[105] M. Graf and P. Vogl, *Electromagnetic fields and dielectric response in empirical tight-binding theory*, Physical Review B **51** (Feb., 1995) 4940–4949.

[106] K. Gawarecki and M. Zielinski, *Electron g-factor in nanostructures:continuum media and atomistic approach*, Scientific Reports **10** (2020) .

[107] D. J. Chadi, *Spin-orbit splitting in crystalline and compositionally disordered semiconductors*, Physical Review B **16** (July, 1977) 790–796.

[108] T. B. Boykin, R. C. Bowen and G. Klimeck, *Electromagnetic coupling and gauge invariance in the empirical tight-binding method*, Physical Review B **63** (June, 2001) 245314.

[109] T. B. Boykin and P. Vogl, *Dielectric response of molecules in empirical tight-binding theory*, Physical Review B **65** (Dec., 2001) 035202.

[110] R. Peierls, *Zur theorie des diamagnetismus von leitungselektronen*, Zeitschrift fur Physik **80** (Nov., 1933) 763–791.

[111] T. B. Boykin, G. Klimeck and F. Oyafuso, *Valence band effective-mass expressions in the $sp^3d^5s^*$ empirical tight-binding model applied to a si and ge parametrization*, Physical Review B **69** (Mar., 2004) 115201.

[112] T. B. Boykin and G. Klimeck, *Practical application of zone-folding concepts in tight-binding calculations*, Physical Review B **71** (Mar., 2005) 115215.

[113] O. Madelung, *Semiconductors: Data Handbook*. Springer-Verlag Berlin Heidelberg, New York, 2004.

[114] P. Boross, G. Széchenyi, D. Culcer and A. Pályi, *Control of valley dynamics in silicon quantum dots in the presence of an interface step*, Physical Review B **94** (July, 2016) 035438.

[115] T. B. Boykin, G. Klimeck, M. A. Eriksson, M. Friesen, S. N. Coppersmith, P. von Allmen et al., *Valley splitting in strained silicon quantum wells*, Applied Physics

Letters **84** (Jan., 2004) 115–117.

[116] S. A. Fortuna and X. Li, *Metal-catalyzed semiconductor nanowires: a review on the control of growth directions*, *Semiconductor Science and Technology* **25** (Jan., 2010) 024005.

[117] W. Seifert, M. Borgström, K. Deppert, K. A. Dick, J. Johansson, M. W. Larsson et al., *Growth of one-dimensional nanostructures in MOVPE*, *Journal of Crystal Growth* **272** (Dec., 2004) 211–220.

[118] H. Xu, Y. Wang, Y. Guo, Z. Liao, Q. Gao, H. H. Tan et al., *Defect-free <110> zinc-blende structured InAs nanowires catalyzed by palladium*, *Nano Letters* **12** (Oct., 2012) 5744–5749.

[119] A. Mikkelsen, N. Sköld, L. Ouattara and E. Lundgren, *Nanowire growth and dopants studied by cross-sectional scanning tunnelling microscopy*, *Nanotechnology* **17** (May, 2006) S362–S368.

[120] M. P. Nowak, B. Szafran, F. M. Peeters, B. Partoens and W. J. Pasek, *Tuning of the spin-orbit interaction in a quantum dot by an in-plane magnetic field*, *Physical Review B* **83** (June, 2011) 245324.

[121] C. E. Pryor and M. E. Flatté, *Landé g factors and orbital momentum quenching in semiconductor quantum dots*, *Physical Review Letters* **96** (Jan., 2006) 026804.

[122] G. W. Winkler, D. Varjas, R. Skolasinski, A. A. Soluyanov, M. Troyer and M. Wimmer, *Orbital contributions to the electron g factor in semiconductor nanowires*, *Physical Review Letters* **119** (July, 2017) 037701.

[123] M. Sakr, *In-plane electron g-factor anisotropy in nanowires due to the spin–orbit interaction*, *Physica E: Low-dimensional Systems and Nanostructures* **64** (2014) 68–71.

[124] D. Stepanenko, M. Rudner, B. I. Halperin and D. Loss, *Singlet-triplet splitting in double quantum dots due to spin-orbit and hyperfine interactions*, *Physical Review B* **85** (Feb., 2012) 075416.

[125] K. Takase, Y. Ashikawa, G. Zhang, K. Tateno and S. Sasaki, *Highly gate-tuneable rashba spin-orbit interaction in a gate-all-around InAs nanowire metal-oxide-semiconductor field-effect transistor*, *Scientific Reports* **7** (Apr., 2017) .

[126] A. V. Moroz and C. H. W. Barnes, *Effect of the spin-orbit interaction on the band structure and conductance of quasi-one-dimensional systems*, Physical Review B **60** (Nov., 1999) 14272–14285.

[127] D. Liang and X. P. Gao, *Strong tuning of rashba spin-orbit interaction in single inas nanowires*, Nano Letters **12** (2012) 3263–3267.

[128] S. D. Ganichev and L. E. Golub, *Interplay of rashba/dresselhaus spin splittings probed by photogalvanic spectroscopy -a review*, physica status solidi (b) **251** (Mar., 2014) 1801–1823.

[129] Z. Scherubl, G. Fulop, M. H. Madsen, J. Nygard and S. Csonka, *Electrical tuning of rashba spin-orbit interaction in multigated inas nanowires*, Physical Review B **94** (2016) .

[130] A. E. Hansen, M. T. Björk, C. Fasth, C. Thelander and L. Samuelson, *Spin relaxation in InAs nanowires studied by tunable weak antilocalization*, Physical Review B **71** (May, 2005) 205328.

[131] S. Dhara, H. S. Solanki, V. Singh, A. Narayanan, P. Chaudhari, M. Gokhale et al., *Magnetotransport properties of individual InAs nanowires*, Physical Review B **79** (Mar., 2009) 121311.

[132] P. Roulleau, T. Choi, S. Riedi, T. Heinzel, I. Shorubalko, T. Ihn et al., *Suppression of weak antilocalization in InAs nanowires*, Physical Review B **81** (Apr., 2010) 155449.

[133] C. Lü, H. C. Schneider and M. W. Wu, *Electron spin relaxation in n-type InAs quantum wires*, Journal of Applied Physics **106** (Oct., 2009) 073703.

[134] A. Bringer, S. Heedt and T. Schäpers, *Dresselhaus spin-orbit coupling in [111]-oriented semiconductor nanowires*, Physical Review B **99** (Feb., 2019) 085437.

[135] C. Kloeffel, M. J. Rančić and D. Loss, *Direct rashba spin-orbit interaction in si and ge nanowires with different growth directions*, Physical Review B **97** (June, 2018) 235422.

[136] F. N. M. Froning, M. J. Rančić10.1103/PhysRevB.97.235422, B. Hetényi, S. Bosco, M. K. Rehmann, A. Li et al., *Strong spin-orbit interaction and g-factor renormalization of hole spins in ge/si nanowire quantum dots*, Physical Review

Research **3** (Jan., 2021) 013081.

[137] T. Zhang, H. Liu, F. Gao, G. Xu, K. Wang, X. Zhang et al., *Anisotropic g-factor and spin–orbit field in a germanium hut wire double quantum dot*, *Nano Letters* **21** (Apr., 2021) 3835–3842.

[138] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*. Cambridge University Press, 3rd ed., 2021.

[139] W. Rossmann, *Lie groups: An introduction through linear groups*. Oxford University Press Inc., 2002.

[140] N. H. Balshaw, *Practical cryogenics: an introduction to laboratory cryogenics*. Oxford Instruments Superconductivity Limited, England, 1996.

[141] M. N. Jirmanus, *Introduction to laboratory cryogenics*. Janis Research Company Inc., 1990.

[142] G. K. White and P. J. Meeson, *Experimental techniques in low-temperature physics fourth edition*. Oxford Univeristy Press Inc., New York, 2002.