QUANTUM ELECTRONICS IN SEMICONDUCTORS

C. H. W. Barnes
Cavendish Laboratory, University of Cambridge
Contents

1 The Free Electron Gas
  1.1 Reference material 3
  1.2 Introduction 3
  1.3 Si, GaAs and Al_{x}Ga_{1-x}As properties 4
    1.3.1 Real space lattice 4
    1.3.2 Reciprocal space lattice 4
  1.4 Effective mass theory 6
  1.5 Doping semiconductors 8
  1.6 Band engineering 8
    1.6.1 Modification of chemical potential and carrier densities 8
    1.6.2 Band bending 11
  1.7 The Si metal-oxide-semiconductor junction 13
  1.8 The GaAs-Al_{x}Ga_{1-x}As heterostructure 17
  1.9 Capacitor model 17
  1.10 Bi-layer heterostructures: Weak coupling 19
    1.10.1 Magnetotunneling spectroscopy [Eisenstein] 22
    1.11 Bi-layer heterostructures: Strong coupling [Davies] 25
  1.12 Exercises 27

2 Semi-classical electron transport.
  2.1 Reference material 28
  2.2 Introduction 28
  2.3 The non-linear Boltzmann equation 28
  2.4 The linear Boltzmann equation 31
  2.5 Semi-classical conductivity 33
  2.6 Drude conductivity 34
  2.7 Impurity scattering 36
  2.8 Exercises 38
Contents

3 Particle-like motion of electrons 39
  3.1 Sources 39
  3.2 Introduction 39
  3.3 The Ohmic contact 39
  3.4 The electron aperture 40
    3.4.1 The effective potential 41
  3.5 Cyclotron motion 41
  3.6 Experiments 44
  3.7 Detection of ballistic motion  L. W. Molenkamp [1] 44
  3.8 Collimation 45
  3.9 Skipping orbits  J. Spector [2] 45
  3.10 The cross junction  C. J. B. Ford [3] 46
  3.11 Chaotic Motion  C. M. Marcus [4] 47
  3.12 Classical refraction 49
  3.13 Refraction  J. Spector [5,6] 52
  3.14 Exercises 53
4 The quantum Hall and Shubnikov de Haas effects 56
  4.1 Sources 56
  4.2 Introduction 56
  4.3 Boltzmann prediction 56
  4.4 Conductivity 58
  4.5 Experiment 58
  4.6 Eigenstates in a magnetic Field 60
  4.7 Density of electrons in a Landau level 62
  4.8 Disorder broadening of Landau levels 63
  4.9 Oscillation of the Fermi energy 65
  4.10 Oscillation of the capacitance 68
  4.11 Conductivity and resistivity at high magnetic field 71
  4.12 A simple model 75
  4.13 Appendix I 78
  4.14 Exercises 79
5 Quantum transport in one dimension. 80
  5.1 Sources 80
  5.2 Introduction 80
  5.3 Experimental realization of the quasi-one-dimensional system [1] 81
  5.4 Eigenstates of an infinitely long quasi-one-dimensional system 81
  5.5 Kardynal et al [2] 83
  5.6 Density of States in a quasi-one-dimensional system 92
  5.7 Oscillation of the Fermi energy and Capacitance 94
    5.7.1 Macks et al [4] 94
### Contents

5.7.2 Drexler *et al* [5] 96  
5.8 The Conductance of a quasi-one-dimensional system 97  
5.9 Thomas *et al* [9] 98  
5.10 The Landauer Formalism 100  
5.11 The saddle point potential 103  
5.12 Determination of saddle point potential shape 106  
5.13 Exercises 111  
6 General quantum transport theory 113  
6.1 Sources 113  
6.2 Introduction 113  
6.3 Eigenstates of an infinite quasi-one-dimensional system 113  
6.4 Group Velocity 118  
6.5 Density of states 119  
6.6 Effective potential at high magnetic field [1] 119  
6.7 Two-terminal conductance in a magnetic field 120  
6.8 The saddle point in a finite magnetic field [2] 122  
6.9 Magnetic Depopulation [3,4] 126  
6.11 Edge states and obstacles 129  
6.12 The Quantum Hall Effect 130  
6.13 Exercises 137  
7 Quasi-particles in two dimensions 138  
7.1 Sources 138  
7.2 Introduction 138  
7.3 Electrical forces 139  
7.4 Landau’s theory of quasi-particles [2-4] 140  
7.5 Quasi-particles 143  
7.6 Quasi-particle decay 144  
7.7 Resistivity 145  
7.8 Temperature dependence of quasi-particle decay [5] 146  
7.9 Density of states dependence of quasi-particle decay 149  
7.10 Formation of new quasi-particles in a quantising magnetic field 149  
7.11 Willett, Goldman and Du [8,11-12] 151  
7.12 Composite fermions [16-17] 155  
7.13 Composite fermion effective magnetic field 158  
7.14 Composite fermion effective electric field 159  
7.15 Composite fermion Hamiltonian 160  
7.16 Fractional quantum Hall effect 160  
7.17 Fractional edge-state model [18] 161  
7.17.1 Fractional quantum Hall effect 161
Contents

7.17.2 Reflection of composite fermion edge states 162
7.18 Exercises 164
8 Quantum dots 165
8.1 Sources 165
8.2 Introduction 165
8.3 Quasi-zero dimensional systems [1,2] 165
8.4.1 Zero field limit 167
8.4.2 High field limit 168
8.5 Conductance of a quantum dot 169
8.6 McEuen et al [2] 171
8.6.1 Magnetic field dependence of resonant peaks [1,2] 171
8.7 Classical Coulomb blockade [4]. 172
8.8 Quantum Coulomb blockade 176
8.9 Artificial atoms 178
8.10 The Aharonov-Bohm effect [5] 179
8.12 Highfield Aharonov-Bohm effect. 183
8.13 Edge-state networks, dots and anti-dots. 184
8.15 Exercises 188
9 Quantum computation 190
9.1 Sources 190
9.2 Introduction 190
9.3 Classical vs Quantum computation 190
9.4 The DiVincenzo rules 190
9.5 The charge qubit 190
9.6 Fujisawa 190
9.7 g-factor in a semiconductor 190
9.8 Hanson 190
9.9 Single spin detection: ... 190
9.10 Interaction in the Hubbard model 190
9.11 Single-spin rotation 190
9.12 Surface acoustic wave current quantisation 190
9.13 The surface acoustic wave quantum processor 190
9.14 Exercises 190
Appendix 1 Occupation probabilities. 191
Appendix 2 Density of states in three dimensions. 193
Appendix 3 Band bending from a delta layer 194
Appendix 4  Density of states in two dimensions.  196
1
The Free Electron Gas

1.1 Reference material
[Kittel] Introduction to Solid State Physics, C. Kittel.

1.2 Introduction
It is a remarkable fact that a free-electron-like gas can be made to form in a semiconductor crystal. As an interacting Fermi gas, it has many complex properties and behaviours. In particular, it has been shown that by manipulating these gases with electric and magnetic fields, they can be made to exhibit all of the familiar quantum effects of undergraduate and postgraduate quantum courses. This having been said though, many of the experimental indicators of these quantum effects can only be fully understood once the basic electrostatic building blocks of the host semiconductor devices have been understood. In particular, since quantum effect are more easy to see in lower-dimensional systems, we concentrate here on the essential physics needed to understand semiconductor devices containing single, or many parallel two-dimensional electron or hole gases.

This section of the notes covers: the basic properties of Si, GaAs and AlₓGa₁₋ₓAs; effective mass theory; semiconductor doping; band engineering; the Si MOSFET; and the GaAs-AlGas heterostructure.
The Free Electron Gas

1.3 Si, GaAs and Al$_x$Ga$_{1-x}$As properties

1.3.1 Real space lattice

Intrinsic crystalline Silicon has a diamond lattice structure Fig. 1.1a. Its underlying space lattice is face-centred cubic (Fig. 1.1b) and its primitive basis has two identical atoms, one at co-ordinate (0, 0, 0) and the other at co-ordinate (1/4, 1/4, 1/4) measured relative to each lattice point. Each atom has four nearest neighbours that form a tetrahedron and the structure is bound by directional covalent bonds.

Intrinsic crystalline GaAs has a zincblende crystal structure (Fig. 1.2). This structure also has a space lattice that is face-centered cubic but the primitive basis has two different atoms, one at co-ordinate (0, 0, 0) and the other at co-ordinate (1/4, 1/4, 1/4) measured relative to each lattice point. Each atom has four nearest neighbours of the opposite type that form a tetrahedron.

Al$_x$Ga$_{1-x}$As has the same lattice structure and approximately the same lattice constant as GaAs but with occasional Al atoms substituting for Ga atoms.

1.3.2 Reciprocal space lattice

The reciprocal lattice of a face-centered cubic lattice is a body-centred cubic lattice (Fig. 1.3a).

The principal symmetry points in the body-centred cubic reciprocal-space lattice are: $\Gamma = (0, 0, 0)$, $X = (1, 0, 0) + 6$ equivalent points, $L = (1, 1, 1) + 8$ equivalent points (Fig. 1.3b).
1.3 Si, GaAs and Al$_x$Ga$_{1-x}$As properties

Fig. 1.2. Zincblende structure of GaAs $a = 5.6\,\text{Å}$.

Fig. 1.3. (a) Body-centred cubic lattice. (b) Brillouin zone boundaries for a face-centered cubic lattice.

Figs. 1.4a,b show the band structure of Si and GaAs along the directions defined by $X$ and $L$ starting at $\Gamma$. These directions are the most important for both band structures because their principal band minima, which define their band gaps, lie along them.

The band structure of a Al$_x$Ga$_{1-x}$As is similar to that of GaAs but it has a larger band gap, which increases with increasing Al content. It has a direct band gap for $x < 0.45$ - typically $x = 0.3$ is used.
1.4 Effective mass theory

At low temperature the valence-band states in intrinsic semiconductors are full and the conduction-band states are empty. As a result intrinsic semiconductors are insulating at low temperature because there are no free carriers. If by some means electrons are introduced to the conduction band then their Fermi surface will be defined by a constant energy surface. Figs 1.5a,b show constant energy surfaces close to the conduction band edges in Si and GaAs. The indirect band gap in Si gives rise to six degenerate constant energy surfaces. Once spin is taken into consideration this increases by a factor of two. The direct band gap in GaAs gives rise to a single spherical constant energy surface. Once spin is taken into consideration this also increases by a factor of two.

The Fermi surfaces for holes in both cases are approximately spherical for low carrier densities. They are centred around the Γ point, and formed from two subbands.

The energy dispersion for both electrons and holes is approximately parabolic for energies close to the band edges - just as is the case for electrons in free space. Conduction electrons (or holes) in semiconductors therefore behave
1.4 Effective mass theory

like free particles. The difference is that they respond to external fields as if they have a different mass from the free space mass $m_e$. This mass is referred to as the ‘effective mass’. In general, the free space dispersion for electrons has the form

$$
\Phi_{n,k}(r) = U_{n,k} e^{ik \cdot r} \quad (1.1)
$$

$$
E_k = \frac{\hbar^2}{2m_e} k^2 \quad (1.2)
$$

where $k$ is their wave vector. In general, the conduction electron dispersion has an elliptical form

$$
E_k = \sum_{i,j} \frac{\hbar^2}{2} (m^*)_{i,j}^{-1} k_i k_j \quad (1.3)
$$

where $k_i$ and $k_j$ are wave vectors measured from a band minimum or maximum in orthogonal directions and the $m^*_{i,j}$ define the effective mass tensor. These effective masses are typically different for different bands.

The dynamic properties of conduction electrons are therefore determined by a Schrödinger equation that takes into account: the effects of crystal band structure through an effective mass tensor $m_{i,j}$; the effects of unbalanced charge, external voltages, ohmic contacts and impurities through a potential energy term $V(x, y, z)$; and the effects of charged surface states through boundary conditions. The simplest case is for electrons in the GaAs conduc-
tion band. They have a single effective mass $m^* = 0.067m_e$, and therefore obey a Schrödinger equation of the form

$$-rac{\hbar^2}{2m^*} \nabla^2 \psi + V(x, y, z) \psi = E \psi \quad (1.4)$$

1.5 Doping semiconductors

A three-dimensional electron or hole gas can be created in a semiconductor by doping it with impurities. For example, a group IV semiconductor with group III dopants has occasional bonds with missing electrons. At finite temperatures valence electrons become excited and fill some of these localized holes leaving behind de-localized holes in the valence band. These carriers are then free to propagate through the semiconductor and give rise to free-electron-like hole conductivity. We consider two generic types of impurity: donor impurities; and acceptor impurities. Both contribute states in the band gap of a semiconductor. Donor impurity states typically lie close to the conduction band edge (Fig 1.6) and contribute conduction electrons. An intrinsic semiconductor doped with donor impurities is called an $n$-type semiconductor. Acceptor impurity states typically lie close to the valence band edge (Fig 1.6) and contribute holes to the conduction process. An intrinsic semiconductor doped with acceptor impurities is called a $p$-type semiconductor.

Three-dimensional electron or hole gases made by doping semiconductors are not ideal for studying quantum effects for two reasons: they are strongly disordered owing to the background of ionized impurities; and most quantum effects are more pronounced in lower-dimensional systems. Therefore, we will look at ways in which two-dimensional or lower dimensional electron or hole gases can be made in semiconductors. This is achieved through ‘band engineering’. There are two principal parts to band engineering: modification of the chemical potential and carrier densities; and bending bands with unbalanced charge.

1.6 Band engineering

1.6.1 Modification of chemical potential and carrier densities

Fig. 1.7 shows an idealised doped semiconductor with an acceptor concentration $N_A$ (m$^{-3}$) and donor concentration $N_D$ (m$^{-3}$). The donor states form a narrow band of energies around energy $E_D$ - just below the conduction band edge $E_C$. The acceptor states form a narrow band around energy
1.6 Band engineering

Fig. 1.6. Ionization energies for various impurities in Si and GaAs. D indicates a donor if it is below mid-gap and A an acceptor if it is above mid-gap. Data taken from [Sze].

$E_A$ - just above the valence band edge $E_V$. In general at some finite temperature a density $n_C$ (m$^{-3}$) of electrons will be excited into the conduction band and a density $n_D$ (m$^{-3}$) of electrons will still occupy the donor states and a density $p_V$ (m$^{-3}$) of holes will be excited into the valence band and a density $p_A$ (m$^{-3}$) of holes will still occupy the acceptor states.

When dopants are added to a semiconductor it remains charge neutral since both the semiconductor and the dopants are initially neutral. For charge neutrality without dopants the total number of electrons in the conduction band must equal the total number of holes in the valence band $n_C = p_V$. With dopants, the number of electrons, free plus bound $n_C + n_D$, must equal the effective number of donors $N_D - N_A$ plus the number of holes, free plus bound $p_V + p_A$.

$$n_D + n_C = N_D - N_A + p_A + p_V \quad (1.5)$$

where

$$n_D = N_D \langle n \rangle \quad (1.6)$$

$$p_A = N_A \langle p \rangle \quad (1.7)$$
Fig. 1.7. Definitions of electron and hole densities and energies.

\( \langle p \rangle \) is the average number of holes occupying each acceptor state and \( \langle n \rangle \) is the average number of electrons occupying each donor state. Their functional forms are derived in Appendix I. They predict a lower occupation than Fermi-Dirac statistics owing to the Coulomb repulsion entailed in multiple occupation of localized impurity states.

The free carrier densities \( n_C \) and \( p_V \) are determined by Fermi statistics through

\[
\begin{align*}
n_C &= \int_{E_C}^{\infty} g_C(\varepsilon)f(\varepsilon)d\varepsilon \\
p_V &= \int_{-\infty}^{E_V} g_V(\varepsilon)(1 - f(\varepsilon))d\varepsilon
\end{align*}
\]

where \( g_C(\varepsilon) \) and \( g_V(\varepsilon) \) are the conduction-band and valence-band densities of states, (derived for a spherical Fermi surface in Appendix II) and \( f(\varepsilon) \) is the Fermi-Dirac distribution function (Appendix I). The set of coupled equations \([1.5] \rightarrow [1.9]\) have as input \( N_A, N_D, E_C - E_V, E_D, E_A, m^* \), and \( T \) and can be solved for \( n_C, n_D, p_A, p_V \) and \( \mu \). Approximate solutions for these quantities are derived in many text books [Ashcroft and Mermin, Kittel, Sze, Kelly] and exact solutions can be found numerically. The most important results from these solutions are that: the larger \( N_A \) is, the closer \( \mu \) is to \( E_V \).
1.6 Band engineering

The chemical potential is the larger \( p_V \) is; and the larger \( N_D \) is, the closer \( \mu \) is to \( E_C \) and the greater \( n_C \) is.

1.6.2 Band bending

Equation [1.5] is correct for a homogeneous system in which charge neutrality is obeyed at each point. Local charge neutrality is not necessary to ensure the overall neutrality of a heterogeneous system though. Consider for example what happens when a piece of \( p \)-type semiconductor is joined to a piece of \( n \)-type semiconductor. At low temperatures the chemical potential is between the donor states and the conduction band edge in the \( n \)-type material (Fig 1.8a), but between the acceptor states and the valence band edge in the \( p \)-type material (Fig 1.8b). If the two systems are made from the same material, when they are joined together, the band edges must align. Since \( E_C \) and \( E_V \) are not affected by doping they will initially have different chemical potentials. This will cause electrons to flow from the higher chemical potential material (\( n \)-type) to the lower (\( p \)-type) to establish thermodynamic equilibrium (Fig 1.9). The \( n \)-type material then becomes positively charged and the \( p \)-type material becomes negatively charged in the vicinity of the join since they were initially charge neutral (Fig.1.10). This imbalance of charge results in the band edge profile bending increasingly upwards when passing through the positively charged region and increasingly downwards when passing through the negatively charged region to represent the attraction of electrons to positive charge and repulsion from negative charge (Fig 1.11). The magnitude of the band bending, due to a \( \delta \)-function layer of charge, is derived in Appendix III to give a more rigorous understanding of this. The conduction and valence band edges \( E_C(x, y, z) \) and \( E_V(x, y, z) \) are
The Free Electron Gas

Fig. 1.9. \( n - p \) junction before equilibration.

Fig. 1.10. Charge density distribution \( \rho \) in a \( n - p \) junction.

the position dependent effective electrostatic potential energies of free electrons and free holes respectively. An electron placed in the conduction band of Fig. 1.11 will accelerate downhill moving to the left \( (a_n = -\nabla E_C/m_n^*) \), and a hole placed in the valence band of Fig. 1.11 will accelerate uphill moving to the right \( (a_p = -\nabla E_V/m_p^*) \).

Once equilibrium is established in any electronic device the chemical potential \( \mu \) will be constant throughout. We will frequently choose \( \mu \) as the potential zero of the system in these notes. The general problem of calculating the equilibrium band edge profiles \( E_C(x, y, z) \) and \( E_V(x, y, z) \) and the charge density distribution \( \rho(x, y, z) \) is numerically intensive since one cannot assume local charge neutrality, Eqn. [1.5], only global charge neu-
One must solve Poissons equation to find $E_C$ and $E_V$ in terms of $\rho$ the unbalanced charge density:

$$-\nabla^2 \phi = \frac{\rho}{\varepsilon} \quad (1.10)$$

$$\phi = \frac{-E_C}{e} \quad (1.11)$$

$$E_V = E_c - E_g \quad (1.12)$$

$$\rho = e(N_D - N_A + p_V + p_A - n_C - n_D) \quad (1.13)$$

For boundary conditions to Eqn. [1.10], the values of $E_C$ and $E_V$ on the surface of the device can be chosen to be those obtained for bulk doped material. The potentials $E_C$ and $E_V$ are related to $\rho$ through equations [1.6 - 1.9] and therefore the set of coupled equations [1.6-1.13] can be solved iteratively until self-consistent band profiles and densities are found.

We are now in a position to understand how to use selective doping to make low-dimensional electron or hole gases in semiconductor materials through band engineering.

1.7 The Si metal-oxide-semiconductor junction

The Si metal-oxide-semiconductor (MOS) junction (Fig. 1.12) is a double plate capacitor consisting of a metal plate, a SiO$_2$ insulating spacer layer and a Si plate. The states of a metal and a semiconductor align at the
Fig. 1.12. Metal-Oxide-Semiconductor junction. $W_M$ is the metal work function and $W_S$ is the semiconductor work function.

Fig. 1.13. Metal-oxide-$p$-type semiconductor junction (a) Before equilibration (b) After equilibration.

vacuum level but, in general, have different work functions $W_M$ and $W_S$. This results in them having different chemical potentials and therefore some equilibration and band bending will occur when a metal and a semiconductor are brought together. In Fig. 1.12 we have neglected this.

If the semiconductor in a MOS junction is $p$-type we will have $W_M < W_S$ and on contact charge will flow from the metal into the semiconductor to achieve equilibrium as shown in Figs. 1.13a,b. This creates a region of unbalanced negative charge at the oxide-semiconductor interface giving rise to the band bending seen in Fig 1.13b.

If the semiconductor is $n$-type then $W_M > W_S$ and on contact charge will flow from the semiconductor into the metal to achieve equilibrium as shown
1.7 The Si metal-oxide-semiconductor junction

Fig. 1.14. Metal-oxide-n-type semiconductor junction. (a) Before equilibration (b) After equilibration.

in Figs. 1.14a,b. This creates a region of unbalanced positive charge at the oxide-semiconductor interface.

If a positive voltage is applied to the metal plate in Fig. 1.13b with respect to the substrate, the band edge profile in the vicinity of the surface gate will be increasingly bent downward as more as more negative charge is attracted to metal surface gate, filling acceptor states. At some point the conduction band edge \( E_C \) will dip below the chemical potential in the substrate and the semiclassical Eqn. [1.8] would predict that the triangular well that forms at the interface will fill with free conduction electrons Fig. 1.15.

Typically, at least for small surface-gate voltages, the width of triangular well is approximately equal to the Fermi wavelength of the conduction electrons and Eqn. [1.8] should not be used to calculate the well carrier density because it assumes zero Fermi wavelength. A quantum mechanical equivalent to Eqn. [1.8] must be used. In the effective-mass approximation, the Schrödinger equation for the well region has the form

\[
-\frac{\hbar^2}{2m^*} \nabla^2 \psi + E_C \psi = E \psi
\]

(1.14)

Which has solutions of the form

\[
E_{i,k} = \frac{\hbar^2 k^2}{2m^*} + E_i
\]

(1.15)

where \( k = (k_x, k_y) \) is the wave vector for free motion in the \( x, y \) plane. The energies \( E_i \) are the quantum well subband energies (Fig. 16) and may be found by direct numerical computation.
When only the lowest energy subband $E_0$ is beneath the substrate chemical potential the device will have a single dynamically two-dimensional electron gas at the oxide-semiconductor interface. Having calculated $E_0$ quantum mechanically it may then be regarded as the zero of potential energy for the two-dimensional gas and therefore the two-dimensional carrier density.
1.8 The GaAs-Al\(_x\)Ga\(_{1-x}\)As heterostructure

\(n_{2D}\) (m\(^{-2}\)) may be calculated from

\[ n_{2D} = \int_{E_0}^{\infty} g_{2D}(\varepsilon) f(\varepsilon) d\varepsilon \]  

where \(g_{2D}\) is the two-dimensional density of states (see Appendix IV). The spatial dependence of the well carrier density \(n_C(x, y, z)\) (m\(^{-3}\)) can then be calculated from \(n_{2D}\) by noting that this two-dimensional density is distributed in the interface region according to the probability density of the bound state \(E_0\). A two-dimensional hole gas may be made by applying a negative voltage to the metal gate of a metal-oxide-n-type junction Fig. 1.14.

1.8 The GaAs-Al\(_x\)Ga\(_{1-x}\)As heterostructure

In order to study quantum-mechanical effects in a two-dimensional electron gas it is necessary to make them as free from unintentional potential modulation as possible. The drawback with the metal-oxide-semiconductor junction is that there are occupied dopant states in the oxide-semiconductor interface region exactly where the two-dimensional system forms. In addition, the oxide barrier is not smooth and contains a high density of trapped charge since it is amorphous. The GaAs-Al\(_x\)Ga\(_{1-x}\) heterostructure, shown in Fig. 17a,b, partly solves these problems: these heterostructures are grown by molecular beam epitaxy [Kelly] and the interface region where the two-dimensional electron gas forms is flat to within a mono-layer with no trapped charge; and modulation doping is used to place dopants hundreds of Angstroms away from the interface to reduce their ability to scatter. Electrons in the donor region in Fig. 1.17(a) are well above the bulk chemical potential and therefore flow out to the surface and interface regions during equilibration. As with the metal-oxide-semiconductor junction a dynamically two-dimensional electron gas can form at the interface, here at the GaAs-AlGaAs interface, if the doping concentration and length scales are chosen correctly. Fig. 1.17(b) shows the badstructure after equilibration.

1.9 Capacitor model

The dependence of the carrier density \(n_{2D}\) on surface-gate voltage for both the metal-oxide-semiconductor junction and the GaAs-AlGaAs heterostructure may be calculated from a simple capacitor model. One plate of the capacitor is a metal surface gate and the other the two-dimensional electron gas. The capacitance of the device is approximately, \(C = \varepsilon A/d\) where \(d\) is
the distance from the surface gate to the two-dimensional electron gas and \( A \) is the area of the device. If we substitute this into the capacitor equation \( Q = CV_g \) we have

\[
e \Delta n_{2D} A = \frac{\varepsilon A}{d} \Delta V_g \quad (1.17)
\]

for an electron gas, where \( \Delta n_{2D} \) is the change in carrier density of the two-dimensional electron gas caused by a change in gate voltage \( \Delta V_g \), this simplifies to

\[
\Delta n_{2D} = \frac{\varepsilon}{ed} \Delta V_g \quad (1.18)
\]

For a hole gas we have

\[
\Delta p_{2D} = -\frac{\varepsilon}{ed} \Delta V_g \quad (1.19)
\]

Figure 1.18 shows a typical GaAs-Al\(_{0.3}\)Ga\(_{0.7}\)As heterostructure device [Kardynal] and Fig. 1.19 shows how its carrier density \( n_{2D} \) varies as a function of both front and back gate voltages \( V_{FG}, V_{BG} \). Figure 1.20 shows how the two-dimensional subband wave function and band edge vary as a function of back-gate voltage.
1.10 Bi-layer heterostructures: Weak coupling

In the GaAs-Al$_x$Ga$_{1-x}$As heterostructure system band engineering can be used to produce any number of parallel quantum wells. Figure 1.21 shows a heterostructure with two quantum wells each of which has a single occupied two-dimensional subband. The spacing between wells is 125Å and the distance from the centre of the double well to the surface gate is 300Å.

Figure 1.22 shows the carrier densities in the two subbands as a function of a metallic surface gate on the left-hand side of the device. In region I
Fig. 1.20. Wave functions in quantum well in Fig. 1.18 at three different back-gate voltages.

Fig. 1.21. (a) Schematic of double well wafer structure (b) Conduction band edge profile of a typical double quantum well system in a GaAs-Al$_x$Ga$_{1-x}$As heterostructure.

only the right quantum well is occupied. At point II the left quantum well becomes occupied and screens the effect of the surface gate on the right well so that its carrier density remains approximately constant over the rest of the gate voltage range. In this range the carrier density in the front well increases approximately linearly.
1.10 Bi-layer heterostructures: Weak coupling

Fig. 1.22. Carrier densities in the two quantum wells of Fig. 1.21 as a function of the left surface-gate voltage.

Fig. 1.23. Double quantum well wave functions and conduction band edge at different surface gate voltages $d = 125\,\text{Å}$.

Figure 1.23 shows the double well region and its wave functions at the gate voltages marked by stars in Fig. 1.22. Note how the bending of the conduction-band edge becomes more pronounced as more carriers enter the quantum wells. At the voltage where the two densities are equal, no anticrossing is seen. This is because the barrier is sufficiently high/wide that coherence is not preserved in multiple scattering of electrons between wells. Conservation of momentum in passing between quantum wells is observed though and can be used to carry out a kind of electron spectroscopy.
1.10.1 Magnetotunnelling spectroscopy [Eisenstein]

Electron tunnelling spectroscopy can be used to investigate the Fermi surfaces of two-parallel two-dimensional electron gases if they are sufficiently weakly coupled but close enough to allow tunnelling. The upper inset in Fig. 1.24 shows an electronic device that can be used for this purpose. The basic device consists of two parallel two-dimensional electron systems separated by a 125Å AlGaAs barrier in a GaAs-AlGaAs heterostructure. Metallic contacts connect to the upper and lower two-dimensional systems independently so that current passed from the left contact to the right contact must pass through the upper two-dimensional system, tunnel across the barrier and pass out through the lower two-dimensional system. Independent contact to the two layers is achieved by applying a negative bias to narrow top and bottom gate. They are shown as black rectangles on the upper and lower surfaces of the device and marked ‘a’ and ‘b’ in the inset figure. Negative voltages applied to these gates can be tuned so that they cut off the two-dimensional electron system closest to them.
A central metallic gate on the upper surface of the device can be used to change the relative carrier densities in the two quantum wells and therefore control tunnelling between the two-dimensional systems. The magnetotunnelning experiment consists of measuring the equilibrium tunnelling differential conductance $dI/dV$ ($I$ is the current passing between layers and $V$ the potential difference between layers) as a function of an externally applied magnetic field placed parallel to the two two-dimensional planes and also as a function of the relative carrier densities of the two systems. The term ‘equilibrium’ here refers to the fact that the voltage $V$ is made sufficiently small that it can be ignored in considering the individual properties of the two electron systems.

In equilibrium the two electron systems have the same chemical potential because the barrier between the two quantum wells is sufficiently narrow to allow electrons to tunnel between the two quantum wells. This is shown in the inset to figure [1.24]. At low temperatures the tunneling is restricted, by Fermi statistics, so that it only occurs at the device chemical potential and therefore between the Fermi surfaces of the two two-dimensional electron systems. Tunnelling is further restricted by conservation of momentum and therefore conservation of wavevector. The equilibrium tunnelling conductance of the device is then given by the overlap integral of the Fermi-surface spectral densities of the two-two-dimensional electron systems.

The wave functions in a two-dimensional system are plane waves

$$\Psi_{k_x,k_y} = e^{ik_xx}e^{ik_yy}$$  \hspace{1cm} (1.20)

and have a dispersion of the form

$$E_{k_x,k_y} = \frac{\hbar^2}{2m^*}(k_x^2 + k_y^2)$$  \hspace{1cm} (1.21)

Hence, at the Fermi surface where $E_{k_x,k_y} = E_F$ there is only finite spectral density on a ring with $k_F^2 = k_x^2 + k_y^2$, where $k_F$ is the Fermi wave vector, as shown in Fig. 1.25.

The effect of a parallel magnetic field is to shift the two spectral functions relative to one-another in $k$-space. Schrödinger’s equation in a magnetic field has the form

$$\frac{(p + eA)^2}{2m^*} \Psi + V \Psi = E \Psi$$  \hspace{1cm} (1.22)

In a parallel magnetic field $B = (B_x, B_y, 0)$ we can use the Landau gauge for the vector potential $A = (B_y z, -B_x z, 0)$. Substituting this into Eqn. 1.22
Fig. 1.25. Spectral density of a two-dimensional electron gas at its Fermi energy - the Fermi circle.

we get

\[ \left( \frac{(p_x + eB_y z)^2}{2m^*} + \frac{(p_y - eB_x z)^2}{2m^*} + \frac{p_z^2}{2m^*} \right) \Psi + V\Psi = E\Psi \quad (1.23) \]

For the device in Fig. 1.24, centering the upper two-dimensional system at \( z = 0 \) and the lower two-dimensional system at \( z = -d \) the respective Schrödinger equations have the form

\[ \left( \frac{p_x^2}{2m^*} + \frac{p_y^2}{2m^*} + \frac{p_z^2}{2m^*} \right) \Psi + V\Psi = E\Psi \quad (1.24) \]

\[ \left( \frac{(p_x - eB_y d)^2}{2m^*} + \frac{(p_y + eB_x d)^2}{2m^*} + \frac{p_z^2}{2m^*} \right) \Psi + V\Psi = E\Psi \quad (1.25) \]

so that their \( k \)-space origins are offset by

\[ p_x \rightarrow p_x - eB_y d \Rightarrow k_x \rightarrow k_x - \frac{eB_y d}{\hbar} \quad (1.26) \]

\[ p_y \rightarrow p_y + eB_x d \Rightarrow k_y \rightarrow k_y + \frac{eB_x d}{\hbar} \quad (1.27) \]

The form of the traces in figures [1.24] and [1.26] can readily understood in terms of a model consisting of the overlap between the two Fermi surface
1.11 Bi-layer heterostructures: Strong coupling [Davies]

Fig. 1.26. Magnetotunnelling conductance (a) experiment (b) theory for tunnelling between two parallel two-dimensional electron systems as a function of their relative carrier densities at different parallel field strengths.

rings of two two-dimensional electron gases. The theory (b) parts to each of these figures show examples of the relative positions of the two Fermi surfaces.

1.11 Bi-layer heterostructures: Strong coupling [Davies]

Figure 1.27 shows the dependence of the quantum well carrier densities at three different barrier thicknesses $d = 125\text{Å}$, $d = 45\text{Å}$ and $d = 25\text{Å}$. As the barrier thicknesses are made smaller a clear anti-crossing is observed. This is because multiple scattering begins to become coherent and the states in the two quantum wells begin to hybridize as they are brought closer together and electrons are able to tunnel faster.

The wave functions and eigenenergies for the 25Å case are shown in Fig. 1.28 at a series of different surface gate voltages.
Fig. 1.27. Carrier densities of front and back quantum wells in three different bi-layer heterostructure plotted as a function of surface gate voltage.

Fig. 1.28. Strongly coupled double quantum-well wave-functions and energies as a function of surface-gate voltage $d = 25\AA$. 
1.12 Exercises

1.1 In what way is motion of an electron in free space similar/different from motion of conduction electrons or holes in intrinsic semiconductors.

1.2 Describe qualitatively how the chemical potential, carrier densities, unionized donor and acceptor concentrations vary as a function of \(N_D\), \(N_A\) and temperature in a semiconductor.

1.3 How is equilibrium established when a \(p-n\) junction is formed.

1.4 What is meant by band bending in semiconductors and why does it occur.

1.5 Describe how a two-dimensional hole gas may be created in a Si MOSFET.

1.6 In what way is a two-dimensional electron or hole gas in a MOSFET dynamically two-dimensional.

1.7 Describe the different aspects of band engineering that go into making a GaAs-AlGaAs heterostructure.

1.8 If a square quantum well in a GaAs-AlGaAs heterostructure has a width of 150\(\AA\), what is the difference in energy between the lowest state of the well and the next.

1.9 If a quantum well in a GaAs-AlGaAs heterostructure, containing a single two-dimensional electron gas, has a Fermi energy of 10meV what is the carrier density. If the distance between the two-dimensional electron gas and a metal surface gate is 300nm what would the change in carrier density be for a change in gate voltage of 1 Volt.

1.10 The device shown in Fig. 1.17(b) has a problem for practical applications since the AlGaAs at the left surface would be likely to oxidise and therefore destroy the device. Design an alternative device that would be more practical.
2

Semi-classical electron transport.

2.1 Reference material


2.2 Introduction

Now that we have understood some of the general physics of two-dimensional systems in semiconductors it is natural to ask about their general transport characteristics. We will begin this process by looking at the predictions of a powerful semiclassical theory: Boltzmann transport theory.

In this section we derive the semi-classical Boltzmann equation for the two-dimensional case and use it to understand electronic transport in two-dimensional electron systems containing many scattering centers. The Boltzmann equation expresses transport dynamics in these systems in terms of a balance between acceleration, due to the Lorentz force, and deceleration, due to collisions with the scattering centers.

2.3 The non-linear Boltzmann equation

We consider a two-dimensional electron system confined to the $x - y$ plane. If this system is in equilibrium and there are no electric or magnetic fields present the electrons will be distributed according to the Fermi-Dirac function:

$$f^0(k) = \frac{1}{1 + \exp \left( \frac{\epsilon_k - \mu}{k_B T} \right)}.$$
At low temperature the occupied states fill the Fermi circle, which is centred around $\mathbf{k} = (k_x, k_y) = 0$, as shown Fig. 2.1. This distribution has no net momentum and therefore no net current flows.

When electric and magnetic fields $E, B$ are switched on, electrons passing through the system will be accelerated by the Lorentz force. An electron initially in a state $\mathbf{k}$ with velocity $v_k$ will receive an acceleration

$$\frac{dv_k}{dt} = \frac{\hbar}{m^*} \frac{d\mathbf{k}}{dt} = -\frac{e}{m^*} (E + v_k \wedge B)$$ (2.2)

This acceleration will cause the electron distribution $f$ to evolve in time $f \rightarrow f(\mathbf{k}, t)$. The distribution function $f(\mathbf{k}, t)d^2k/(2\pi)^2$ is defined as, the average number of electrons occupying the infinitesimal volume of $\mathbf{k}$-space $d^2k$ around wave vector $\mathbf{k}$ at time $t$. We can quantify the evolution of $f$ if we think of the acceleration $d\mathbf{k}/dt$ as a velocity in $\mathbf{k}$-space. This concept is illustrated in Fig. 2.2. Any electrons occupying a small volume of $\mathbf{k}$-space around wave vector $\mathbf{k}$, box 1 in Fig. 2.2, with velocity $\dot{\mathbf{k}}$ at time $t$ will move so that they are in a small volume of $\mathbf{k}$-space around wave vector $\mathbf{k}' = \mathbf{k} + \mathbf{k}dt$, box 2 in Fig. 2.2, at time $t + dt$. This implies that

$$f(\mathbf{k}, t) = f(\mathbf{k} + \mathbf{k}dt, t + dt)$$ (2.3)

If we now assume: that the change in $f$ is a small perturbation on $f^0$; and that the application of small magnetic and electric fields will not alter the
FIG. 2.2. Motion in \( k \)-space due to the Lorentz Force.

eigen spectrum of the two-dimensional system significantly - we will come back to this in a later lecture - we can make a Taylor expansion of Eqn. 2.3. To first order in \( dt \) this is:

\[
\approx f(k, t) + \nabla_k f \cdot \dot{k} \, dt + \frac{\partial f}{\partial t} \, dt
\]

\[
0 = \nabla_k f \cdot \dot{k} + \frac{\partial f}{\partial t}
\]

(2.4)

so that the rate of change of \( f \) with respect to time from the Lorentz force has the form

\[
\frac{\partial f}{\partial t} \bigg|_{\text{Lorentz force}} = -\nabla_k f \cdot \dot{k}
\]

(2.5)

If the acceleration of electrons in the system ran unchecked by any mechanism, the assumption that the new distribution function is a small perturbation on the zero field distribution would certainly be wrong. In real systems there are many compensating factors: (1) Electrons may scatter from impurities or crystal defects. At low temperatures this is typically the dominant mechanism. (2) They can scatter from phonons. This mechanism becomes dominant at higher temperatures. (3) They can collide with each other since they are charged particles. We will discuss this in more detail later in lectures.

After switching on the electric and magnetic fields for a short time an
2.4 The linear Boltzmann equation

The excess distribution

\[ g(k, t) = f(k, t) - f^0(k) \quad (2.6) \]

will build up. If the fields are then switched off we would expect that this excess distribution would decay away through collision processes at a rate proportional to the excess

\[ \left. \frac{\partial g(k, t)}{\partial t} \right|_{\text{Collisions}} = -\frac{g(k, t)}{\tau_k} \quad (2.7) \]

The characteristic time for this decay to occur, \( \tau_k \), is referred to as the relaxation time. Since the initial distribution function \( f^0 \) is independent of time, the time derivative of Eqn. 2.6 gives

\[ \left. \frac{\partial g(k, t)}{\partial t} \right|_{\text{Collisions}} = \left. \frac{\partial f(k, t)}{\partial t} \right|_{\text{Collisions}} \quad (2.8) \]

Equation 2.7 describes what is known as the relaxation time approximation. If one thinks of the detailed quantum mechanical multiple scattering between impurities that must occur at low temperatures, where phase coherence becomes important, this approximation will clearly break down.

We now have two different mechanisms that cause the electron distribution to change: the Lorentz force which tends to accelerate conduction electrons; and collisions which tend to decelerate them. Under the application of small electric fields the two-dimensional electron gas can therefore find a new dynamic equilibrium

\[ \left. \frac{\partial f(k, t)}{\partial t} \right|_{\text{Lorentzforce}} + \left. \frac{\partial f(k, t)}{\partial t} \right|_{\text{Collisions}} = 0 \quad (2.9) \]

substituting Eqns. 2.2,2.5-2.7 into 2.9 we have the non-linear Boltzmann equation.

\[ \nabla_k \left( g + f^0 \right) \cdot \frac{e}{\hbar} (E + v_k \wedge B) = \frac{g}{\tau_k} \quad (2.10) \]

It is a non-linear first order partial differential equation in \( g \).

2.4 The linear Boltzmann equation

If we expand the bracket in Eqn. 2.10 the first term is

\[ \nabla_k f^0 \cdot \frac{e}{\hbar} v_k \wedge B = 0 \quad (2.11) \]

It is equal to zero because

\[ \nabla_k f^0 = \frac{\partial f^0}{\partial \epsilon} \hbar v_k \quad (2.12) \]
Fig. 2.3. Shift of the Fermi circle owing to external electric field $\Delta k = -e\tau E/\hbar$.

and $v_k \cdot B = 0$. The second term is the non-linear term

$$\nabla_k g \cdot \frac{e}{\hbar} E$$  \hspace{1cm} (2.13)

It is non-linear because it is a product of the electric field $E$ and the excess distribution $g$ that the electric field produces. Removing these terms results in the linear Boltzmann equation:

$$\nabla_k g \cdot \frac{e}{\hbar} v_k \wedge B + \nabla_k f_0 \cdot \frac{e}{\hbar} E = \frac{g}{\tau_k}$$  \hspace{1cm} (2.14)

It will be valid if the excess distribution $g$ is small.

In zero magnetic field the solution to the linear Boltzmann equation is trivial since the only term containing a derivative is multiplied by the magnetic field $B$. The solution is:

$$g(k) = \frac{e\tau_k}{\hbar} \nabla_k f_0 \cdot E$$  \hspace{1cm} (2.15)

substituting from Eqn. 2.6 we find

$$f = f^0 + \frac{e\tau_k}{\hbar} \nabla_k f^0 \cdot E$$  \hspace{1cm} (2.16)

If we assume that the relaxation time $\tau_k$ is constant $\tau_k = \tau$ equation 2.16 yields a very simple picture of transport in a two-dimensional system. We may rewrite Eqn. 2.16 to the same order of approximation as

$$f = f^0 \left( k + \frac{e\tau}{\hbar} E \right)$$  \hspace{1cm} (2.17)

Equation 2.17 indicates that under the influence of an electric field the new
dynamic equilibrium is represented by shifting the zero-field Fermi circle in the opposite direction to the applied electric field (opposite since electronic charge is negative) by an amount \( \Delta \mathbf{k} = -e \tau E / \hbar \) in \( \mathbf{k} \)-space. This is shown Fig. 2.3. This shift may be viewed as each electron picking up an additional drift velocity \( v_d \) owing to the influence of the electric field. This drift velocity is defined by

\[
\frac{m^* v_d}{\hbar} = \Delta \mathbf{k} = -e \tau E / \hbar
\]  

(2.18)

From this equation it is clear that the longer the relaxation time \( \tau \) the larger the drift velocity \( v_d \) will be for a given electric field \( E \). Using equation [2.18] the electron mobility \( \mu \) can be written:

\[
\mu = \frac{|v_d|}{|E|} = \frac{e \tau}{m^*}
\]  

(2.19)

(note that the same symbol \( \mu \) is also used both for the chemical potential Eqn. 2.1 and for microns \( \mu m \)). Mobility is a measure of the purity of a system. Good electron mobilities in GaAs-AlGaAs heterostructures are \( \mu = 50 \text{m}^2/\text{Vs} \rightarrow 100 \text{m}^2/\text{Vs} \). This implies relaxation times of \( \tau = 19 \text{ps} \rightarrow 380 \text{ps} \) and typical drift velocities in a low temperature measurement of \( |v_d| = 0.6 \text{ms}^{-1} \rightarrow 12.5 \text{ms}^{-1} \) assuming a 1mV potential difference across an 80\( \mu \text{m} \) sample.

### 2.5 Semi-classical conductivity

We can use the functional form of the equilibrium excess distribution Eqn. 2.15 to calculate the conductivity of a two-dimensional system containing scattering centers. The equilibrium distribution function \( f(\mathbf{k}) \) represents a dynamic equilibrium in which a current is flowing through the system. This current is given by

\[
j = 2 \int \frac{d^2 \mathbf{k}}{(2\pi)^2} \cdot - e v_d f
\]  

(2.20)

We include a factor of two for spin degeneracy. In Si this would be a factor of four owing to the extra valley degeneracy. The twelvefold bulk degeneracy is reduced to four at the SiO2 Si interface. The term \( d^2 \mathbf{k}/(2\pi)^2 \) in Eqn. 2.20 is the density of states within an infinitesimal volume of \( \mathbf{k} \)-space.

Since there is no current flow in zero field Eqn. 2.20 may be written

\[
j = 2 \int \frac{d^2 \mathbf{k}}{(2\pi)^2} \cdot - e v_d g
\]  

(2.21)
substituting for $g$ from Eqn. 2.15 we find

$$j = \frac{2e^2}{(2\pi)^2} \int d^2k \left(- \frac{\partial f^0}{\partial \epsilon}\right) \tau_k v_k (v_k.E)$$

(2.22)

which may be rewritten as

$$j = \sigma E$$

(2.23)

where $\sigma$ is the conductivity tensor (a 2x2 matrix)

$$\sigma = \frac{2e^2}{(2\pi)^2} \int d^2k \left(- \frac{\partial f^0}{\partial \epsilon}\right) \tau_k v_k \otimes v_k$$

(2.24)

(Mathematical notation - outer product of two vectors $a$, $b$ is written $(a \otimes b)_{ij} = a_i b_j$. In our case $v_k$ is a 1x2 vector and so $v_k \otimes v_k$ is a 2x2 matrix).

If the functional form for $\tau_k$ is known Eqn. 2.24 can be integrated to find the conductivity of a twodimensional system with arbitrary shape Fermi surface.

For this lecture course, the important thing about Eqn. 2.24 is that the derivative of the Fermi-Dirac distribution $-\partial f / \partial \epsilon$ appears in the integrand.

Its presence implies that, even though we know that every electron has picked up the same additional drift velocity $v_d$ irrespective of its energy or wave vector, at low temperatures it is only those states that are at the Fermi-surface that contribute to the conductivity.

Electrons at the Fermi-surface of a two-dimensional electron gas travel at phenomenal velocities. For example with a typical Fermi energy of $E_F = 10$meV in a GaAs-AlGaAs heterostructure the Fermi velocity will be $v_F = 280000\text{ms}^{-1}$. For the relaxation times $\tau = 19\text{ps} \rightarrow 380\text{ps}$ this implies the mean free path for electrons at the Fermi surface is $l_F = v_F t = 5.3\mu\text{m}$. With modern lithographic techniques making electronic devices smaller than this has become standard. We will look at ballistic motion of electrons in such devices in the next part of the notes.

### 2.6 Drude conductivity

Equation Eqn. 2.24 reduces to the Drude formula for conductivity if we assume a constant relaxation time, zero temperature and a parabolic dispersion relation $\epsilon = \hbar^2k^2/2m^*$. Rewriting the integral Eqn. 2.22 in cylindrical co-ordinates in $k$-space we have

$$j = \frac{2e^2\tau}{(2\pi)^2} \int k dk d\phi \left(- \frac{\partial f^0}{\partial \epsilon}\right) (v_k \cos(\phi) k_x + v_k \sin(\phi) k_y) (v_k \cos(\phi) E_x + v_k \sin(\phi) E_y)$$

(2.25)
2.6 Drude conductivity

Fig. 2.4. Heterojunction interface with sub-band energy $E_0$ and Fermi energy $E_F$.

which expands to

$$j = \frac{2e^2 \tau}{(2\pi)^2} \int kdkd\phi \left( -\frac{\partial f^0}{\partial \epsilon} \right) \left( v_k^2 \cos^2(\phi)E_xk_x + v_k^2 \sin^2(\phi)E_yk_y + \cos(\phi) \sin(\phi) (E_xk_x + E_yk_y) \right)$$

performing the integral over $\phi$ we have

$$j = \frac{2e^2 \tau}{(2\pi)^2} \int kdk \left( -\frac{\partial f^0}{\partial \epsilon} \right) v_k^2 \pi E$$

(2.26)

In the limit of zero temperature the derivative of the Fermi-Dirac function becomes a delta function at the chemical potential $\mu$. If we choose the zero of potential energy to be at the two-dimensional sub-band energy $E_0$ (shown in Fig. 2.4) then

$$\lim T \to 0 \left( -\frac{\partial f^0}{\partial \epsilon} \right) = \delta(\epsilon - E_F)$$

(2.27)

(2.28)

The dispersion relation for the system is

$$\epsilon = \frac{\hbar^2 k^2}{2m^*} = \frac{1}{2} m^* v^2_k$$

(2.29)

$$d\epsilon = \frac{\hbar^2 kd\kappa}{2m^*}$$

(2.30)

Substituting Eqns. 2.28-2.30 into Eqn. 2.27 we have

$$j = \frac{2e^2 \tau}{(2\pi)^2} \frac{m^*}{\hbar^2} \int d\epsilon \delta(\epsilon - E_F) \frac{2\epsilon}{m^*} \pi E$$

(2.31)
Semi-classical electron transport.

Fig. 2.5. Elastic scattering from an impurity (a) $k \rightarrow k'$ (b) $k' \rightarrow k$.

which after some algebra gives

$$\sigma = \frac{e^2 \tau}{m^* n_{2D}} = e \mu n_{2D}$$

(2.32)

where $n_{2D}$ is the carrier density of the two-dimensional electron gas. This expression is often misinterpreted as implying that conductivity derives from all occupied states. As we have seen, this is not the case at low temperature, only states at the Fermi surface contribute.

2.7 Impurity scattering

As we have said, at low temperatures, the dominant scattering mechanism in a two-dimensional electron system is impurity scattering. We have used the constant relaxation time approximation to estimate the size that devices must be in order to be free from scattering. However, significant scattering is still seen in systems smaller than $l_F$. In order to understand why this is, we will derive the relaxation time $\tau_k$ from first principles and see how it relates to the time between scattering events.

The rate at which $g(k)$ diminishes by impurity scattering processes is an integral over all possible ways in which a particle may scatter out of a state $k$ into a state $k'$, minus an integral over all possible ways in which a particle may scatter from the state $k'$ into the state $k$ obeying Fermi statistics Fig. 2.5.

Thus the rate of change of the distribution function $f(k,t)$ with respect
2.7 Impurity scattering

to time can be written as
\[
\frac{\partial f(k,t)}{\partial t} \bigg|_{\text{Collisions}} = - \int \frac{d^2k}{(2\pi)^2} W_{k,k'} f(k)(1 - f(k')) - W_{k',k} f(k') (1 - f(k))
\] (2.33)

The factor \( f(k)(1 - f(k')) \) is the probability, that the state \( k \) is occupied and that the state \( k' \) is unoccupied - the necessary condition for an electron in state \( k \) to be able to scatter into a state \( k' \). The factor \( f(k')(1 - f(k)) \) is the probability for the reverse process Fig. 2.5(b). At low temperature these factors restrict scattering to the vicinity of the Fermi surface.

\( W_{k,k'} \frac{d^2k'}{(2\pi)^2} \) is the probability per unit time that an electron in a state \( k \) will be scattered into an infinitesimal volume \( d^2k' \) around the state \( k' \). This probability may be derived from Fermi’s Golden rule. The principal results of such a derivation are that: the scattering event is symmetric \( W_{k,k'} = W_{k',k} \); that energy is conserved in the collision \( |k|^2 = |k'|^2 \); that \( W_{k,k'} \) is an even function of only the angle between \( k \) and \( k' \); and that \( W_{k,k'} \) is proportional to the density of impurity centers. Under these assumptions Eqn. 2.33 simplifies to

\[
\frac{g(k)}{\tau_k} = \int \frac{d^2k}{(2\pi)^2} W_{k,k'} (g(k) - g(k'))
\] (2.34)

after substituting for the left-hand side from Eqn. 2.6. For a large diffusive system we can assume that the relaxation time is isotropic \( \tau_k = \tau_{|k|} \). Substituting this into Eqn. 2.15 and rearranging we have

\[
g = e\tau_{|k|} \left( \frac{\partial f_0}{\partial \epsilon} \right) v_k . E = \hbar e\tau_{|k|} \left( \frac{\partial f_0}{\partial \epsilon} \right) \frac{m^*}{e} E . k = a(\epsilon) . k
\] (2.35)

If we now substitute Eqn. 2.35 into Eqn. 2.34 we have

\[
a(\epsilon) . k = a(\epsilon) . \int \frac{d^2k}{(2\pi)^2} W_{k,k'} (k - k')
\] (2.36)

Next we resolve \( k' \) into components parallel \( k \) to \( k \) and perpendicular \( k_{\perp} \) to \( k \) Fig. 2.5.

\[
k' = |k'| \cos(\theta) k + |k'| \sin(\theta) k_{\perp}
\] (2.37)

which, since energy is conserved in the collision, becomes

\[
k' = \cos(\theta) k + |k| \sin(\theta) k_{\perp}
\] (2.38)

If we substitute Eqn. 2.38 into Eqn. 2.36, noting that \( W_{k,k'} \) is an even function of the angle \( \theta \) between \( k \) and \( k' \), we find an expression for the
relaxation rate

\[ \frac{1}{\tau_k} = \int \frac{d^2k}{(2\pi)^2} W_{k,k'} (1 - \cos(\theta)) \]

This result shows that the relaxation rate is a type of average over scattering events that gives little weight to forward scattering since the factor \(1 - \cos(\theta)\) is small for small \(\theta\). Therefore, if each scattering event itself is dominated by forward scattering, as is the case for electrons in a two-dimensional electron system scattering from impurity centers, the relaxation rate will remain small even if the rate at which scattering events occur is high. Hence, the mean free paths \(l_F\) calculated from \(l_F = v_F t\) above should only be taken as upper limits and not the distance between scattering events. They are measures of the distance through which electrons travel before suffering a collision that affects the bulk conductivity - large angle scattering. Equation 2.39 demonstrates two further features of impurity scattering. The relaxation rate is independent of temperature. It was eliminated in canceling \(a(\epsilon)\) from both sides of Eqn. 2.36. This explains why it becomes dominant at low temperatures, since all other scattering mechanisms diminish with decreasing temperature. We also find that the relaxation rate is proportional to the density of scattering centers.

2.8 Exercises

2.1 Write a list of the assumptions that have gone into the derivation of the linear Boltzmann Eqn. 2.14. Suggest ways in which a fully quantum mechanical formulation of the transport problem would have to be different.

2.2 The expression for the Boltzmann conductivity Eqn. 2.24 contains the derivative of the Fermi function in its integrand. How do you reconcile this with the fact that all electrons are equally affected by the application of an external electric field. What are the implications of this for low temperature electron transport.

2.3 Explain the distinction between the relaxation time and the time between scattering events within the Boltzmann formalism. What is the distinction between the mean free path and the distance between scattering events.
3

Particle-like motion of electrons

3.1 Sources

3.2 Introduction
These notes look at particle-like (ballistic) motion of electrons in high-mobility two-dimensional electron systems. In particular they concentrate on experiments that demonstrate electrons behaving in a similar way to photons in geometric optics. They cover: sources of electrons - the ohmic contact; the split-gate electron aperture; cyclotron motion; the detection of classical skipping orbits; the observation of chaotic motion; classical refraction; and the operation of a ballistic prism and a ballistic lens.

3.3 The Ohmic contact
In order to measure the transport properties of an electron system in a semiconductor device some way of connecting it to macroscopic measurement equipment must be found. For a two-dimensional electron gas this is typically achieved by creating a local conducting $n^{++}$ region that extends from the device surface, where a wire may be connected, to the interface where the electron system forms. For GaAs-AlGaAs heterostructures conducting contacts are made by evaporating a patch of metal alloy, sufficiently
large to be seen easily through a low power optical microscope, onto the surface of the device and then heating it until the alloy melts and diffuses down to the interface. This is shown in Fig. 3.1a. For a Si metal-oxide-semiconductor junction device the contacts are made by firing ion dopants into the surface: Fig. 3.1b. Typically, these contacts will make electrical connection to a large two-dimensional region that then feeds into the low-dimensional system being measured. Such contacts are referred to as ohmic contacts if their resistance is constant for small applied potential biases i.e. if they obey Ohm’s law $I = V/R$. Good ohmic contacts will have a low a resistance to prevent power dissipation and unknown voltage drops.

An ohmic contact is typically large in comparison to the mean free path of the two-dimensional system it contacts, its presence creates a large number of crystal defects and it is threedimensional. The exact way in which ohmic contacts achieve electrical contact is a complicated issue and is still in debate. However, an ideal ohmic contact should act like a Fermionic version of a black-body radiator, sending and receiving electrons equally at all energies and in all directions according to Fermi statistics. For this reason, an ohmic contact can be thought of as analogous to a diffuse light source (light bulb) in geometric optics.

### 3.4 The electron aperture

The geometric motion of photons in an optical experiment is immediately apparent if one places a small aperture in front of a strong light source. The linear path of the light is seen by the eye where dust scatters it. But how could such an experiment be performed on electrons in a two-dimensional electron system? The ohmic contacts take the place of both sources and detectors but how can we create an aperture?

As we have discussed in the first lecture, the carrier density in a two-
3.5 Cyclotron motion

dimensional electron system can be controlled by applying a potential difference between a metallic surface gate and the system itself. It is a small step from this concept to the realization that if the surface gate is made small it can control the electron density locally, either by depleting or enhancing the electron system directly under the gate. Fig. 3.2a shows a schematic diagram of the effect of applying a negative potential to a metallic split gate on the surface of a GaAs-AlGaAs heterostructure. The narrow channel that forms acts like a small aperture connecting two two-dimensional regions.

3.4.1 The effective potential

The general effect of applying potentials to arbitrary surface-gates patterned on a two-dimensional electron system can be summed up in terms of the so-called position-dependent two-dimensional ‘effective potential’. Fig. 3.2b shows the conduction-band edge both, under a metal surface gate (upper trace), and away from it when a negative potential has been applied to the gate. Under the gate, the lowest two-dimensional subband energy $E_0$ is above the chemical potential $\mu = 0$ and so no electrons will occupy the quantum well. Away from the gate $E_0$ is below the chemical potential and so the quantum well will be occupied. In general the subband energy $E_0$ will be a function of position $E_0(x, y)$ and is referred to as the two-dimensional ‘effective potential’ it is actually the position-dependent effective conduction band edge in region of the quantum well. Its position dependence is determined by the pattern of surface gates and the potentials applied to them and may be calculated by solving Schrödinger’s equation selfconsistently with Poisson’s equation. We will come back to this when we consider one and zero-dimensional systems in later lectures.

3.5 Cyclotron motion

The electron aperture that we have just described can be used as part of an electronic device to show that electrons in a two-dimensional electron gas travel ballistically but it cannot be done as simply as it can in an optical experiment. There are no useful analogies to dust or the human eye. The electron apertures are also stuck to the surface of the device and cannot be moved during an experiment. This means that although apertures can be made so that they sit on a straight line between a source and a detector, they cannot be moved so that it can be demonstrated that electrons only travel in a straight line. Electrons do however respond to external fields,
Particle-like motion of electrons

Fig. 3.2. (a) Schematic of a metallic split gate on the surface of a heterostructure. (b) The conduction-band edge under a metallic gate (upper) and away from a metallic gate (lower) when a negative potential is applied to the gate.

through the Lorentz force, and an external magnetic field can be used to bend the path electrons take.

In the absence of an electric field the Lorentz force on an electron in a two-dimensional system is \( F = -e v \times B \). Where \( v \) is the electron velocity and \( B \) the external magnetic field. Any component of the magnetic field parallel to the two-dimensional system will have no effect on the motion of the electron in the \( x - y \) plane since the resulting Lorentz force acts in the direction of confinement. In order to study the effect of a magnetic field on the motion of an electron forced to move in a two-dimensional plane, we need only consider the effect of a perpendicular magnetic field \( B = B \hat{z} \).

From Newton’s second law its motion described by

\[
F = -eBv_y \hat{x} + eBv_x \hat{y} \tag{3.1}
\]

Within the effective mass approximation, we can ignore the underlying crystal lattice in a semiconductor electron system if we take its effect into account via an effective mass \( m^* \). Using this effective mass in equation [3.1] we find

\[
\begin{align*}
\dot{v}_x &= -\frac{eB}{m^*} v_y = -\omega_c v_y \tag{3.2} \\
\dot{v}_y &= +\frac{eB}{m^*} v_x = +\omega_c v_x \tag{3.3}
\end{align*}
\]

\( \omega_c \) is referred to as the cyclotron frequency. Substituting [3.2] into [3.3] we find two independent differential equations that show that the electrons
velocities execute simple harmonic motion.

\[ \ddot{v}_x = -\omega_c^2 v_x \]  
\[ \ddot{v}_y = -\omega_c^2 v_y \] (3.4) (3.5)

As we saw in the previous section of the notes, the only electrons that are detected in the conduction process are at the Fermi surface. So, for the device shown in Fig. 3.3, the boundary conditions at the source aperture are \( v_x = V_F \) and \( v_y = 0 \). \( v_F \) is the Fermi velocity. With these boundary conditions the solution for the position of an electron from [3.2.5] as a function of time \( t \) will be

\[ x = \frac{v_F}{\omega_c} \sin(\omega_c t) \] (3.6)
\[ y = \frac{v_F}{\omega_c} (1 - \cos(\omega_c t)) \] (3.7)

This motion describes a circle \( x^2 + (y - R)^2 = R^2 \) where \( R = v_F/\omega_c \) is the cyclotron radius - this motion is known as 'cyclotron motion'.

Fig. 3.3. Cyclotron motion of electrons injected into a two-dimensional system.
3.6 Experiments

We will now look in detail at a series of experiments that demonstrate ballistic motion of electrons. All of the experiments were performed at low temperatures 1K on devices patterned on GaAs-AlGaAs heterostructures with a single two-dimensional electron system at the interface between the GaAs and the AlGaAs. The electron systems were of high quality with mobilities $\mu > 50 \text{ m}^2/\text{Vs}$. Surface gates are numbered 1,2,3... etc. They have negative voltages applied to them that are sufficiently large to deplete the electron gas under them and therefore define their shape as barriers in the two-dimensional system. Ohmic contacts are lettered A,B,C... etc and are either connected to earth, a current source, a voltage source or a volt meter.

3.7 Detection of ballistic motion  L. W. Molenkamp [1]

Fig. 3.4a shows a schematic diagram of the device used in an experiment by L. W. Molenkamp to detect ballistic motion. There is a source and detector aperture set opposite each other, gates 1,2 and gates 3,4. Behind the source aperture is the source ohmic contact A and behind the detector aperture is the detector ohmic contact C. A current $I_A$ is injected into the system from A. Ohmic contact D is connected to ground and acts as a sink to all electrons irrespective of their ballistic path. The voltage $V_C$ on ohmic contact C is then measured relative to that on B. The measured resistance $(V_C - V_B)/I_A$ is shown in Fig. 3.4b as a function of external magnetic field.

Electrons that travel ballistically from A to C cause an over pressure of electrons at contact C with respect to contact B where electrons must arrive by very complicated diffusive paths. When a magnetic field is applied ballistic electrons no longer pass directly from A to C and therefore the potential difference between C and B diminishes: Fig. 3.4b. This experiment, is one of the most clear demonstrations of ballistic electron motion in a two-dimensional system. In the referenced paper [1] the authors also perform numerical simulations using a ray tracing program to find the expected resistance for different shape apertures: Fig. 3.4b insets. The lower dashed line shows the expected ballistic peak for a square aperture exit. Such an aperture sprays electrons in all directions and therefore shows a lower ballistic peak than experiment. A series of dots that approximately follow the experimental trace show the result for a smoothly varying exit.
3.8 Collimation

When electrons pass through a narrow split-gate aperture they scatter from the electrostatically defined edges. These edges are typically very smooth and the reflection from them is elastic with angle of incidence equal to the angle of exit $\theta = \theta'$: Fig. 3.5.

Such reflection leads directly to collimation through two processes. (1) As electrons exit the aperture the effective potential reduces in value until it reaches the bulk potential. This causes the electrons to accelerate. Fig. 3.6a shows how such acceleration leads to collimation. (2) The exit of an aperture widens as it reaches the exit. Fig. 3.6b shows how this leads to collimation.

3.9 Skipping orbits J. Spector [2]

Fig. 3.3 shows that for a sufficiently strong magnetic field ballistic electrons in the Molenkamp experiment [1] would not reach the opposite ohmic contact but would end up skipping along the source aperture electrostatic boundary. Fig. 3.7 shows an experimental device used to demonstrate this.

Ohmic contacts E and F are earthed and current is injected from ohmic contact B. When a negative magnetic field is applied to the device electrons from the source aperture in front of B skip along the boundary between B and A. If a multiple of the cyclotron diameter $2R$ is equal to the distance
between the two apertures electrons pass directly from B into A and an enhanced voltage is measured. If the diameter is too large or too small electrons miss the aperture in front of A and only the diffuse background will be measured. This process appears as the oscillatory signal in Fig. 3.7b (trace (a)) as a function of magnetic field. The last peak results from a path with nine reflections from the electrostatic edge and clearly demonstrates the purity of the device being measured.

When the magnetic field is applied in the positive direction electrons are deflected to the left and a single peak is seen, trace (b), in Fig. 3.7b. It corresponds to an orbit without any reflection. Orbits that cause electrons to reflect from ohmic contact C do not give rise to peaks in the voltage at D since they are either absorbed or scattered randomly.

### 3.10 The Cross Junction

C. J. B. Ford [3]

The inset to Fig. 3.8(a) shows a schematic view of a cross surface-gate structure made by Christopher Ford when a postdoc at IBM.

A current $I_A$ is injected into the cross from contact A and a potential difference $V_{BC}$ between contacts B and C is measured. The experimental trace in Fig. 3.8a shows $R_B = V_{BC}/I_A$ as a function of magnetic field. At low magnetic fields cyclotron motion causes enhanced scattering from the
exit boundary walls but no electrons are directed to contact B. At some critical field the electrons are directly focused into contact B and a positive $R_H$ is measured. Fig. 3.8b shows the result of removing the corners of the cross. In this device, at small fields the electrons can bounce from the flat surface on gate 2 and give an enhanced voltage at C i.e. a negative $R_H$. $R_H$ becomes increasingly negative with increasing magnetic field until the cyclotron radius is sufficiently small that electrons can pass ballistically directly from A to B, $R_H$ then becomes positive. Fig. 3.8c shows the result of introducing an artificial impurity into the centre of the cross structure. The dashed line shows the result with no voltage on gate 5, giving a result identical to Fig. 3.8b, and the solid line with a negative voltage on gate 5. The deflection of the ballistic electrons by the impurity causes them to pass preferentially into contact B even at low fields.

3.11 Chaotic Motion  C. M. Marcus [4]

Figure 3.9 shows two gate patterns for two different devices used by Charles Marcus to investigate chaotic motion. In the upper panel of Fig. 3.9 the electron micro-graph shows a stadium shaped device defined by surface gates (gates in White, electron gas in black). The lower panel shows a circular shaped device. Each shape has an entrance and exit aperture. The resistance
traces shown in Fig. 3.9 are a measure of whether electrons are reflected from the structure (high resistance) or pass through (low resistance). We will show the direct relationship between resistance and reflection coefficients in subsequent lectures. The resistance is shown as a function of magnetic field. Each trace consists of a large number of peaks and dips. Peaks occur at magnetic fields where electrons are, eventually, reflected from the device and dips occur where electrons are eventually passed through the device. This type of structure is a signature of chaotic ballistic motion. It is a simple exercise in ray tracing to convince yourself that for a complicated scattering path around either object a slight change in magnetic field can make the difference between a particle exiting from the left or right aperture. Figs. 3.10a,b show a refinement to the experiment in Fig. 3.9, for which, in addition to changing the magnetic field it was possible to alter the shape of the confined region with the gate $V_g$ Fig. 3.10a.

The grey scale plot in Fig. 3.10b shows random fluctuation in the conductance of the device as a function of both parameters. It has the signature of chaotic motion—extreme sensitivity to boundary conditions.
3.12 Classical refraction

We have been discussing a series of experiments performed on ballistic electrons in which they appear to behave quite classically. We might therefore not expect to see anything like refraction in the same systems. However,

![Fig. 3.8. Hall resistance $R_H$ of (a) Cross gate pattern. (b) Corners removed (c) Artificial impurity.](image-url)
Particle-like motion of electrons

Fig. 3.9. (a),(b) shows two gate patterns in white. Resistance of (a) Stadium (b) Circle showing extreme sensitivity to boundary conditions.

there is a classical analogue to optical refraction. Consider a beam of electrons incident on a graded potential step: Fig. 3.11. When an electron passes through the graded region in Fig. 3.11 a normal force acts only in the \( y \) direction and therefore momentum in the \( x \) direction is conserved.

\[
|v_F| \sin(\theta) = |v'_F| \sin(\theta')
\]

(3.8)

Energy is conserved as the collision of the electron with the step is elastic. The electron travels across the step at the chemical potential and the two speeds \( v_F \) and \( v'_F \) are equal to the Fermi velocities in the two regions. Thus
3.12 Classical refraction

[3.6] may be rewritten as

$$\frac{\sin(\theta)}{\sin(\theta')} = \sqrt{\frac{n'_{2D}}{n_{2D}}}$$  \hspace{1cm} (3.9)

where $n_{2D}$ is the carrier density in the $V_g = 0$ region and $n'_{2D}$ in the $V_g < 0$ region.

Refraction in geometric optics is shown in Fig. 3.12. The wavefront of a beam of photons is always perpendicular to the direction of propagation so that in place of [3.6] we have

$$|v'| \sin(\theta) = |v| \sin(\theta')$$  \hspace{1cm} (3.10)

Substituting for the refractive indices $n = c/v, n' = c/v'$ in equation [3.8]
we find

\[
\frac{\sin(\theta)}{\sin(\theta')} = \frac{n'}{n}
\]  

(3.11)

The important difference between [3.7] and [3.9] is that for electrons under a metal surface gate with a negative voltage on it \( n_{2D} < n'_{2D} \) and therefore \( \theta' > \theta \), Fig. 3.11, whereas for light passing from air into glass \( n' > n \) and therefore \( \theta' < \theta \), Fig. 3.12. We will see how this affects the design of a ballistic lens in the last section of these notes.

### 3.13 Refraction J. Spector [5,6]

Fig. 3.13 shows an electron micrograph of a device that Joe Spector used to demonstrate classical refraction of electrons.

Electrons are injected from ohmic contact F and pass through a source aperture defined by gates 7,8. They then pass through a second aperture
Fig. 3.12. Refraction in optics. The wavefront is always perpendicular to the direction of motion.

defined by 6,9 and stray electrons are caught by ohmic contacts G,E. Additional stray electrons are removed after 6,9 with ohmic contacts W. Gate 5 is the active refractive gate, the refractive index of which is altered by changing the voltage on it. The three traces (a), (b), (c) in Fig. 3.13b show the potential measured on ohmic contacts A,B,C as a function of the voltage on gate 5. For the lowest density, most negative gate voltage, the angle of deflection is the greatest, ballistic electrons pass preferentially into contact A. Traces (a), (b) and (c) show the results of numerical simulations using a ray tracing program.

We noted in the derivation of classical refraction that it was in the opposite sense to optical refraction. This has its ultimate application in Joe Spector’s ballistic electron lens shown in Fig. 3.14 where a converging lens has the same shape as an optical diverging lens.

3.14 Exercises

3.1 Explain the electrical principles behind the detection of ballistic electron motion in a two-dimensional system. A magnetic focusing ex-
Particle-like motion of electrons

Fig. 3.13. (a) Pattern of device to demonstrate classical refraction (b) Solid line experiment, dashed theory.

The experiment is performed in two-dimensional electron gas with mobility $100 \text{m}^2/\text{Vs}$ and carrier density $1.0 \times 10^{15} \text{m}^{-2}$ using two split-gate apertures $4 \mu \text{m}$ apart along the same edge. At what magnetic fields do the first four focusing peaks occur? What might limit the number of peaks seen? How might you account for uneven peak heights. Assume $m^*=0.067m_e$ in all questions.

3.2 Explain how a surface gate can be used to cause a local modulation in the charge density of a two-dimensional system. Explain how a collimated beam of electrons can be created. How are optical and classical refraction different. A collimated beam of ballistic electrons in a two-dimensional electron gas of Fermi energy $10 \text{meV}$ is normally incident on one surface of a wedge shaped region of low density with Fermi energy $5 \text{meV}$ and internal angle $\theta=30^\circ$ caused by a surface gate. What will the angle of exit be referred to the angle of incidence. What application might such a device have and what factors might limit the speed of its operation.
Fig. 3.14. (a) Ballistic electron lens (b) Ratio of detected $I_d$ to injected current $I_e$: solid line experiment, dashed line theory.
4

The quantum Hall and Shubnikov de Haas effects

4.1 Sources

4.2 Introduction
These notes look at the effect of a perpendicular magnetic field on the electronic structure and transport properties of a two-dimensional electron system. We show that the semiclassical Boltzmann formalism fails to account for high magnetic field behaviour owing to the formation of a new set of states: the Landau levels. We show that their existence is consistent with a number of experimental observations.

The notes cover: the prediction of the Boltzmann formalism in finite magnetic field; the eigenstates of a two-dimensional electron system in a magnetic field; the density of electrons in a Landau level; disorder broadening of a Landau level; the oscillation of the Fermi energy; the oscillation of the capacitance; and the basic features of the Shubnikov de Haas and Quantum Hall effects.

4.3 Boltzmann prediction
As was shown in the second lecture, the linear Boltzmann equation at finite magnetic field has the form:

$$\nabla_k g \cdot \frac{e}{\hbar} v_k \wedge B + \nabla_k f^0 \cdot \frac{e}{\hbar} E = \frac{q}{\hbar}$$

(4.1)
4.3 Boltzmann prediction

In the constant relaxation time approximation $\tau_k = \tau$ its solution can be written in the form:

$$g(k) = \frac{e\tau}{\hbar} \nabla_k f_0 \cdot Z$$

(4.2)

where $Z$ is the so called Hall vector:

$$Z = \frac{E + \left(\frac{e\tau}{m^*}\right) B \wedge E + \left(\frac{e\tau}{m^*}\right)^2 (B \cdot E) B}{1 + \left(\frac{e\tau}{m^*}\right)^2 B \cdot B}$$

(4.3)

This form for $Z$ is derived by substituting Eqn. 4.2 into Eqn. 4.1 (see appendix for details) to find the following relationship between $E$, $B$, and $Z$:

$$E = Z - \left(\frac{e\tau}{m^*}\right) B \wedge Z$$

(4.4)

and then inverting it to find $Z$.

The solution Eqn. 4.2 for the distribution $g(k)$ has the same functional form as the zero field case, Eqn. 2.15, but with $E$ replaced by $Z$. The equilibrium occupation probability $f(k) = g(k) + f_0(k)$ is therefore also a shifted Fermi circle with $\Delta k = -(e\tau/\hbar)Z$:

$$f = f_0 \left(k + \frac{e\tau}{\hbar} Z\right)$$

(4.5)
Fig. 4.1 is a graphical representation of Eqn. 4.4. It shows that the locus of the Hall vector \( Z \) with magnetic field is a circle. It also shows that the drift of the Fermi circle is made smaller in magnitude and increasingly perpendicular to the electric field \( E \) with increasing magnetic field.

### 4.4 Conductivity

The derivation of the conductivity for the finite field case follows the same set of equations as given in section 2.5 of lecture 2 but where relevant with \( E \) replaced by \( Z \). Equation 2.32 becomes

\[
j = \sigma_0 Z = n_{2D} e \mu Z = n_{2D} \frac{e^2 \tau}{\hbar} Z \tag{4.6}
\]

Substituting Eqn. 4.6 into Eqn. 4.4 we find

\[
E = j \sigma_0 - \frac{e \tau}{m^*} B \wedge j \sigma_0 \tag{4.7}
\]

Since we are considering a two-dimensional system in a perpendicular magnetic field: \( j = j_x \hat{x} + j_y \hat{y}, E = E_x \hat{x} + E_y \hat{y} \) and \( B = B \hat{z} \). With these definitions, Eqn. 4.6 takes the form

\[
\begin{pmatrix}
E_x \\
E_y
\end{pmatrix}
= \begin{pmatrix}
\frac{1}{\sigma_0} & \frac{e \tau}{m^*} \frac{1}{\sigma_0} B \\
-\frac{e \tau}{m^*} \frac{1}{\sigma_0} B & \frac{1}{\sigma_0}
\end{pmatrix}
\begin{pmatrix}
j_x \\
j_y
\end{pmatrix}
\]

so that

\[
E = \hat{\rho} j = \begin{pmatrix}
\frac{1}{\sigma_0} & \frac{e \tau}{m^*} \frac{1}{\sigma_0} B \\
-\frac{e \tau}{m^*} \frac{1}{\sigma_0} B & \frac{1}{\sigma_0}
\end{pmatrix}
j \tag{4.9}
\]

Where \( \hat{\rho} \) is the resistivity tensor. Thus within the Drude formalism the diagonal resistivity \( \rho_{xx} = 1/\sigma_0 \) is independent of magnetic field whereas, the off-diagonal resistivity \( \rho_{x,y} = B/e n_{2D} \) or Hall resistivity increases linearly with magnetic field.

### 4.5 Experiment

The resistivity tensor of a two-dimensional electron system can be measured using a Hall bar. Fig. 4.2 is a schematic diagram showing a Hall bar patterned into a GaAs-AlGaAs heterostructure. Its shape is defined as a mesa by etching away the heterostructure wafer containing the two-dimensional electron gas in the regions where it is not wanted. In measurements of the resistivity tensor, a current \( I \) is passed between ohmic contacts S and D giving a current density \( j_x = I/W \). For the diagonal resistivity \( \rho_{xx} = E_x/j_x \) the voltage between ohmic contacts A and B, \( V_{AB} \),
4.5 Experiment

Fig. 4.2. Hall bar mesa patterned in a heterostructure wafer. Ohmic contacts A-D,S

is measured to find $E_x = V_{AB}/L$. For the Hall resistivity $\rho_{x,y} = E_y/j_x$ the voltage between ohmic contacts A and C, $V_{AC}$, is measured to find $E_y = V_{AC}/W$. Note that the Hall resistivity and Hall resistance are equal $\rho_{xy} = (V_{AB}/W) / (I/W) = V_{AC}/I = R_{xy}$. Fig. 4.3 shows a measurement of $\rho_{xx}$ and $\rho_{xy}$ for a two-dimensional electron system, with carrier density $n_{2D} = 5.6 \times 10^{15} \text{ m}^{-2}$, in a GaAs-AlGaAs heterostructure.

At magnetic fields below $B \sim 0.5\text{T}$ it is clear that the assumptions of the Boltzmann formalism work well: $\rho_{xx}$ is constant and $\rho_{x,y}$ is linear in
The quantum Hall and Shubnikov de Haas effects

magnetic field. However, above that magnetic field the diagonal resistivity begins to show oscillations that are periodic in $1/B$ and the Hall resistivity begins to show plateaux at integer multiples of $h/e^2$. The oscillations in $\rho_{xx}$ are referred to as the Shubnikov de Haas effect and the plateaux in $\rho_{xy}$ are referred to as the quantum Hall effect.

An important point to note here is that in the regions where $\rho_{xx} = 0$ this implies directly that the diagonal conductivity $\sigma_{xx} = 0$ since:

$$\hat{\sigma} = \hat{\rho}^{-1} = \begin{pmatrix} 0 & \rho \\ \rho & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1/\rho \\ 1/\rho & 0 \end{pmatrix}$$

The experimental result in Fig. 4.3 represents a breakdown of the Boltzmann formalism that is quantum mechanical in origin. The central assumption of the Boltzmann formalism is that when the Lorentz force acts it simply redistributes electrons among the fixed set of zero field $k-$ states. It is this assumption that has broken down in the experiment. At sufficiently high magnetic fields, when the cyclotron orbit is smaller than the sample size, the eigenstates of a two-dimensional system are qualitatively different from their zero field form.

### 4.6 Eigenstates in a Magnetic Field

We consider a two-dimensional electron system confined to the $xy$ plane and in a perpendicular magnetic field, $B = B \hat{z}$. Schrödinger’s equation for this system has the form

$$\frac{1}{2m^*} (p + eA)^2 \Psi = E \Psi$$

(4.11)

The vector potential $A$ must be chosen to satisfy the constraints

$$\nabla \cdot A = 0$$

$$\nabla \times A = B \hat{z}$$

(4.12)

Gauge invariance gives us the freedom to choose the exact form of the vector potential $A$ in a way that suits the symmetry of our system. The solutions to Schrödinger’s equation in one gauge can be found from those in another gauge by taking linear combinations of degenerate states, an operation that has no observable consequences. Here, we are considering an infinite two-dimensional electron system and so it is convenient to use the symmetric gauge:

$$A = -\frac{B}{2} y \hat{x} + \frac{B}{2} x \hat{y}$$

(4.13)
Substituting Eqn. 4.13 into Eqn. 4.11 we find

\[
\frac{1}{2m^*} \left( p_x - eB \frac{x}{2} \right)^2 \Psi + \frac{1}{2m^*} \left( p_y + eB \frac{y}{2} \right)^2 \Psi = E \Psi
\]  

(4.14)

Expanding both brackets we find

\[
\frac{1}{2m^*} \left( p_x^2 + p_y^2 \right) \Psi + \frac{eB}{2m^*} (xp_y - yp_x) \Psi + \frac{1}{2m^*} \left( eB \frac{x}{2} \right)^2 (x^2 + y^2) \Psi = E \Psi
\]  

(4.15)

which may be rewritten as

\[
\frac{p}{2m^*} \Psi + \frac{eB}{2m^*} L_z \Psi + \frac{1}{2m^*} \left( eB \frac{x}{2} \right)^2 r^2 \Psi = E \Psi
\]  

(4.16)

where \( L_z = (r \wedge p)_z \) is the \( z \)-component of the angular momentum. This equation has cylindrical symmetry and its solution in cylindrical co-ordinates \((r, \phi)\) therefore has the form

\[
\Psi = e^{il\phi} f(r)
\]

(4.17)

where \( l \) is the azimuthal quantum number. The radial part of the solution \( f(r) \) is best expressed in terms of the dimensionless parameter

\[
z = r^2 m^* \omega_c \frac{2\hbar}{eB}
\]

(4.18)

where \( l_B = \sqrt{\hbar/eB} \) is the magnetic length and \( \omega_c = eB/m^* \) is the cyclotron frequency. It may be found by writing it in the form \( f(r) = e^{-z/2} g(z) \) and solving the Eqn. 4.16 for \( g(z) \) using a power series expansion. As is the case for the solutions to Hermite’s equation (you will have seen this in 1B maths) this series truncates at finite order to give

\[
f(r) = z^{|l|} e^{\frac{1}{2} \frac{d}{dz}} \left( \frac{d}{dz} \right)^{n+|l|} z^n e^{-z}
\]

(4.19)

where \( n \) is a positive integer. The energy eigenvalues of Eqn. 4.16 are given by

\[
E_{n,l} = \left( n + \frac{|l| + l}{2} + \frac{1}{2} \right)
\]

(4.20)

which, since both \( n \) and \( l \) are integers, simplifies to

\[
E_{n,l} = \left( n + \frac{1}{2} \right) \hbar \omega_c
\]

(4.21)

Equation 4.21 shows that all the eigenvalues of a two-dimensional system lie on odd integer multiples of \( \hbar \omega_c/2 \). Each of these energies is infinitely
degenerate since $l$ may take any integer value. The density of states of a perfect two-dimensional system is therefore a set of delta functions as shown in Fig. 4.4. These delta functions are referred to as Landau levels.

If we had included electron spin in the initial Schrödinger equation 4.11 the eigen spectrum would have had the form

$$E_{n,l,\sigma} = \left( n + \frac{1}{2} \right) \hbar \omega_c + \frac{\sigma}{2} g \mu_B B$$

where $\sigma = \pm 1$. Valley splitting in Si further splits each spin-split Landau level in two.

**4.7 Density of electrons in a Landau level**

The radial solution to Schrödinger’s equation Eqn. 4.19 shows that the magnetic eigenstates lie in concentric rings about $r = 0$. The larger $|l|$, the larger the area enclosed by the ring. The product $z|l| e^{-z^2/2}$ defines an over all modulation for the shape of each state in all Landau levels. This is shown in Fig. 4.5. It has a maximum at $z = |l|$ and a half-height width of approximately $l_B$. If we consider a single Landau level, the number of states enclosed by a state with angular momentum $|l|$ is approximately equal to $|l|$. The radius of the state is $z = |l|$ which from Eqn. 4.18 gives $r_{|l|}^2 = 2l_B^2 |l|$. The area enclosed by this state is $\pi r^2 = 2\pi l_B^2 |l|$ and hence the density of
4.8 Disorder broadening of Landau levels

The effective potential $E_0(x, y)$ of a two-dimensional system in a semiconductor device will be modulated by electric fields from ionized impurities, interface defects and crystal defects. These fields add a random component to $E_0(x, y)$ that has the effect of broadening the delta-function Landau levels as shown in Fig. 4.6.

Schrödinger’s equation in the symmetric gauge Eqn. 4.16 has two terms

$$n_L = \frac{|l|}{2\pi l_B^2} = \frac{1}{2\pi l_B^2 B} = \frac{eB}{\hbar} \quad (4.23)$$

This result is not exact for a small system since the finite width of the wave functions ensures that not all electrons are within the defined circle $\pi r^2 = 2\pi l_B^2 |l|$. However for an infinite system Eqn. 4.23 is exact and gives the electron density of a Landau level.

The carrier density $n_{2D}$ in a two-dimensional system in a perpendicular magnetic field is therefore composed of a set of Landau levels so that

$$n_{2D} = \nu \frac{eB}{\hbar} \quad (4.24)$$

where $\nu$ is the number of filled Landau levels. The variable $\nu$ is usually referred to as the filling factor.
that depend on magnetic field. Firstly there is a term that is linear in the angular momentum and secondly there is a parabolic effective potential

$$V_B = \frac{1}{2m^*} \left( \frac{e B}{2} \right)^2 r^2$$  \hspace{1cm} (4.25)

The confinement from this potential becomes increasingly strong, and low angular-momentum states become increasingly localized to the point \( r = 0 \), as the magnetic field increased. We can use this fact to understand the way in which disorder broadens a Landau level. If we choose an arbitrary origin for the symmetric gauge \((x_0, y_0)\), Fig. 4.6, at high magnetic field the eigenstates with zero angular momentum will be localized to an area \( 2\pi l_B^2 \) around this origin and will have approximate energies given by

$$E_{n,x_0,y_0} \simeq \left( n + \frac{1}{2} \right) \hbar \omega_c + E_0(x_0, y_0)$$  \hspace{1cm} (4.26)

For larger values of \(|l|\) the energy of the states will be increasingly affected by the curvature of the potential \( E_0(x_0, y_0) \) and therefore Eqn. 4.26 will not hold. However if we had chosen a different location in the random potential to perform this analysis we would have found a different set of energies localized at the new point. As long as the wave functions at the two points we
4.9 Oscillation of the Fermi energy

Choose are sufficiently far apart that the states are independent \( d \simeq \sqrt{2\pi l_B^2} \) both sets of states will be valid eigenstates of Schrödinger's equation in any gauge. Thus we can write the eigenstates of a disordered two-dimensional system in a strong magnetic field as

\[
E_{n,x,y} \simeq \left( n + \frac{1}{2} \right) \hbar \omega_c + E_0(x,y)
\]  

(4.27)

where on average there is one state in an area \( 2\pi l_B^2 \). The dotted lines in Fig. 4.6 show the localized states in each Landau level following the effective potential \( E_0(x_0,y_0) \). The density of states is therefore found by creating a histogram of the states in each energy range Fig. 4.6b. The density of states in a single Landau level is a histogram of the amplitude variation of the effective potential normalized to give a density of \( 2\pi l_B^2 \). We will find that this picture, in which all the states in a two-dimensional system at high field are localized, will have to be modified but for now it serves as a qualitative basis for further analysis.

4.9 Oscillation of the Fermi energy

In the first lecture we showed that the density in a two-dimensional electron gas could be controlled by a surface gate. We found that the relationship between change in density and change in a surface-gate voltage was linear:

\[
\Delta n_{2D} = \frac{e}{ed} \Delta V_g
\]

(4.28)

In deriving this relationship we did not assume any particular shape for the density of states of the electron gas. The only assumption was that the system behaved like an ideal capacitor. Therefore Eqn. 4.28 is as valid in a finite magnetic field as it is in zero field.

Figs. 4.7a,b show two different situations that arise when the density of states is formed from Landau levels. In Fig. 4.7a a Landau level is at the chemical potential. If in this situation we use a surface gate to remove a small number of electrons from the electron system the Fermi energy \( E_F \) will not change much because in the centre of a disorder broadened Landau level the states are close together in energy. Removing one electron only changes the energy by an amount equal to the next energy spacing. The Landau level is said to be pinned at the chemical potential.

In Fig. 4.7b, there are no states at the chemical potential and in order to remove an electron, the Fermi energy must shrink rapidly, bringing occupied states in the next Landau level to the chemical potential. When the \( n^{th} \) Landau level is pinned at the chemical potential the Fermi energy will be
The quantum Hall and Shubnikov de Haas effects

Fig. 4.7. Heterojunction interface occupied by a single two-dimensional electron gas. (a) Landau level at the chemical potential $\mu$. (b) Chemical potential $\mu$ between Landau levels.

Fig. 4.8. Fermi energy in a two-dimensional system as a function of carrier density at fixed magnetic field $B$.

$E_F \approx (n + 1/2) \hbar \omega_c$, to within the width of the disorder broadened Landau level, and it will remain at this value until the density has been reduced by an amount $n_L = eB/h$, the density of electrons in a Landau level, the Fermi energy will then jump to the next value $E_F = (n - 1 + 1/2) \hbar \omega_c$ and so on. This process is shown schematically in Fig. 4.8.

When we increase the magnetic field at fixed carrier density two things
4.9 Oscillation of the Fermi energy

Fig. 4.9. Grey lines: Landau Fan $E_n = (n + 1/2)\hbar\omega_C$ as a function of $B$ for $n = 0..14$. Black line: The Fermi energy of a two-dimensional system as a function of magnetic field at constant density. Horizontal dark grey line: Zero field Fermi energy.

happen. Firstly the density in each Landau level $n_L = eB/h$ increases and therefore in order for the carrier density $n_{2D} = \nu eB/h$ to remain constant the filling factor $\nu$ must gradually reduce and successive Landau levels with lower index $n$ will pass through the chemical potential. Secondly the spacing between Landau levels increases and therefore whilst a particular Landau level is pinned at the chemical potential the Fermi energy will increase to accommodate this energy. These factors result in the Fermi energy oscillating as a function of magnetic field around its zero field value. This is shown schematically in Fig. 4.9.

The first experiments to detect the oscillation of the Fermi energy in a two-dimensional system were performed by Vladimir Pudalov [1] using a Si-MOSFET. His experimental circuit is shown in Fig. 4.10. When the switch in Fig. 4.10 is connected, a voltage $V_g$ is applied between the surface gate and the two-dimensional system. This induces a carrier density $n_{2D}$ in the two-dimensional system which remains when the switch is disconnected. In Pudalovs experiments this charge remained constant practically indefinitely since there was no current path between the two-dimensional system and the surface gate in the experimental configuration and the device was of extremely high quality. An electrometer with high input impedance, $Z > 10^{14}\Omega$, was used to detect the potential difference between the surface gate
and the two-dimensional system as a function of perpendicular magnetic field.

The potential difference detected in this way is directly proportional to the oscillation of the Fermi energy in the two-dimensional electron gas since although the Fermi energy in the metallic surface gate also oscillates it has an amplitude a factor of $10^{-4}$ times smaller than for the two-dimensional system. Fig. 4.11 shows the results from Pudalovs experiments for two different two-dimensional carrier densities. The shape of the oscillation in Fig. 4.11 is different from that in Fig. 4.9 because Fig. 4.9 does not include spin or valley splitting - both of which occur in Si device that Pudalov used.

### 4.10 Oscillation of the capacitance

In deriving the relationship between the change in density of a two-dimensional electron system with a change in surface-gate voltage, Eqn. 1.17, we assumed that the two-dimensional system acted like a perfect metal plate. This is actually an approximation because a perfect metal has an infinite density of states and a two-dimensional system does not. These two systems therefore screen external fields in different ways. This fact can be used to detect the oscillation of the density of states in a magnetic field directly. First we must re-derive the capacitance between a surface gate and a two-dimensional system in a heterostructure more carefully. Fig. 4.12a shows the band structure of a GaAs-AlGaAs heterostructure device.
4.10 Oscillation of the capacitance

Fig. 4.11. Examples of recordings of the oscillations of the potential difference between surface gate and two-dimensional electron gas at two different densities. Numbers above each trace record the filling factor of the electron gas.

Fig. 4.12. (a) Band edge in GaAs-AlGaAs heterojunction. (b) Model band edge used in capacitance calculation.
Capacitance is a measure of the change in charge density resulting from a change in voltage on a gate. In calculating the capacitance of the device in Fig. 4.12a we can therefore ignore the effect of the ionised donors since they form a static layer of charge that does not respond to changes in surface gate voltage. In addition, the presence of the band offset at the surface in Fig. 4.12a will not affect the capacitance. The straight band edge shown in Fig. 4.12b is therefore a sufficient model for use in calculating the capacitance of the device in Fig. 4.12a.

Using the analysis from the first lecture, appendix III, the relationship between the electric field in the AlGaAs barrier region \( E \) and the carrier density \( n_{2D} \) in the two-dimensional system is

\[
-E = -e \frac{n_{2D}}{\epsilon} \quad (4.29)
\]

We have assumed that sufficiently far to the right of the two-dimensional system the electric field will be zero. By definition \( E \) is related to the potentials \( V_1 \) and \( V_2 \) through

\[
V_1 - V_2 = dE \quad (4.30)
\]

The Fermi energy of the two-dimensional system is

\[
b - eV_2 = -E_F \quad (4.31)
\]

Since the band offset \( b \) is constant, for a small variation in \( V_2 \) we can rewrite Eqn. 4.31 as

\[
e\delta V_2 = \frac{\delta n_{2D}}{dn_{2D}/E_F} \quad (4.32)
\]

The denominator of Eqn. 4.32 is by definition the thermodynamic density of states: \( dn_{2D}/dE_F \). For our two-dimensional system it is similar to but not identical to the density of states: \( dn_{2D}/dE \) at fixed \( E_F \). This is because the Fermi energy is not a linear function of the carrier density.

The chemical potential of the surface gate is pinned at a fixed energy in the band gap of the surface GaAs cap layer and therefore there is a fixed offset between \( V_1 \) and the surface gate voltage \( V_g \). If we consider a small variation in \( V_g \) we have

\[
\delta V_1 = \delta V_g \quad (4.33)
\]

If we now substitute Eqns. 4.29,4.32,4.33 into equation Eqn. 4.30 we find

\[
\delta V_g = e\delta n_{2D} \left( \frac{d}{\epsilon} + \frac{1}{e^2dn_{2D}/dE_F} \right) \quad (4.34)
\]
The change in charge in the two-dimensional system will be \( \delta q = e \delta n A \), where \( A \) is the area of the device, and therefore we can write

\[
\delta V_g = \frac{\delta q}{C} = \delta q \left( \frac{1}{C_G} + \frac{1}{C_Q} \right)
\]

(4.35)

where \( C \) is the total capacitance and \( C_G \) and \( C_Q \) defined by

\[
C_G = \frac{\epsilon A}{d}
\]

(4.36)

\[
C_Q = e^2 \frac{dn_2 D}{dE_F} A
\]

(4.37)

are the geometric and quantum capacitances. Equation 4.35 therefore implies that we must think of the capacitance of a heterostructure device as being composed of a geometric capacitor \( C_G \) in series with a quantum capacitor \( C_Q \). Typically \( C_Q >> C_G \) so that the geometric part dominates any capacitance measurements and our analysis of the oscillation of the Fermi energy still stands. However, Smith et al [Smith] have performed accurate capacitance measurements on a two-dimensional electron gas in a GaAs-AlGaAs heterostructure that show the contribution of \( C_Q \) in \( C \). The results of this measurement are shown in Fig. 4.13a,b.

Figure 4.13b shows the thermodynamic density of states derived from the capacitance data using equations Eqn. 4.35,4.36,4.37. The dashed lines show the theoretical predictions for the corresponding experimental traces under the assumption that the density of states is a set of Gaussian peaks, one for each Landau level.

### 4.11 Conductivity and resistivity at high magnetic field

The two experiments described above demonstrate that the density of states, in a two-dimensional electron system in a strong magnetic field, is composed of disorder broadened Landau levels. When first considering electronic transport in a high magnetic field it is tempting to ask whether the diagonal conductivity \( \sigma_{xx} \) of a Hall bar is proportional to the density of states at the chemical potential, as it is in the Boltzmann formalism (c.f equation Eqn. 2.24):

\[
\sigma_{xx} \propto \left. \frac{dn_2 D}{dE} \right|_{E=\mu}
\]

(4.38)

An experimental trace of the diagonal conductivity of a two-dimensional
Fig. 4.13. The measured and calculated capacitance (a) and density of states (b) vs magnetic field. The dashed line is an approximation using Gaussian Landau levels. The arrow indicates the zero-magnetic field density of states.

system, taken from Smith [Smith], is shown in Fig.4.14. It consists of a series of peaks with decreasing height and in a small magnetic-field range around 1.6T the conductivity is equal to zero.

Neither of these facts is consistent with equation Eqn. 4.38 which would predict that the peak heights should increase with increasing magnetic field as the density in each Landau level increases (c.f Fig. 4.13(b)). In addition measuring a region of zero conductivity would be impossible since the Fermi energy must pass instantly through a region of zero density of states giving it no width in magnetic field. From a theoretical perspective equation Eqn. 4.38 is also wrong since in showing how Landau levels are broadened
by disorder we found that all the eigenstates in a random potential are localized. But localized states by their very nature do not conduct and therefore cannot give rise to finite conductivity.

The solution to the problem with equation Eqn. 4.38 is that close to the centre of each Landau level there are states that extend across the whole system that are responsible for giving it a finite diagonal conductivity. The standard way in which this idea is understood is to think of electrons filling the random effective potential $E_0(x,y)$ in the same way that water would. Such a random potential is shown in Fig. 4.15.

At low densities electrons fall into the deepest dips, black regions in Fig. 4.15, making small localized puddles, but as the carrier density rises each puddle begins to grow and puddles begin to link up. At some critical point the percolation threshold the puddles link one side of the system to the other. However, as the carrier density rises further disconnected islands form at the surface, light grey/white islands in Fig. 4.15, that become smaller and fragment into islands as the water level rises. Each side of the percolation threshold electron states are localized, either to dips or peaks in the random potential, and do not give rise to a diagonal conductivity but at the threshold some states are extended and give rise to a finite diagonal conductivity. We therefore imagine the total density of states $\rho_{tot}(E)$ in a
The quantum Hall and Shubnikov de Haas effects

Fig. 4.15. Grey scale plot of a random potential. Black represents low potential and white high potential. Contour lines outline the puddles that form at low density and islands that form at high density.

Fig. 4.16. Extended and localized states in a single Landau level. A simple model of resistivity and Landau levels (Zero temperature).

Landau level being a sum of localized states $\rho_{loc}(E)$ and extended states $\rho_{ext}(E)$. The extended states being distributed over a smaller energy range than the localized states, as shown in Fig. 4.16, giving rise to a conductivity $\sigma_{xx} \sim \rho_{ext}(\mu)$. This picture provides a model that allows us to understand the basic structure of the traces in Fig. 4.3.
4.12 A simple model

This model provides an understanding of some qualitative features of the diagonal and off–diagonal resistivities, shown in Fig. 4.3, and is based on the model of a Landau level shown in Fig. 4.16. It should not be taken as providing a full understanding of either the quantum Hall or Shubnikov de Haas effect. This will be provided in subsequent lectures.

We can approximate the functional form of disorder broadened Landau levels by assuming that they are Lorentzian in shape (the exact shape will of course depend on the nature of the disorder in the two-dimensional system as indicated by Eqn. 4.26) so that the total Landau–level density of states has the form:

\[ \rho_{tot}(E) = \frac{eB}{\hbar} \sum_{n,\sigma} \frac{1}{\pi \hbar^2} \frac{1}{1 + \left(\frac{E - E_0 - E_{n,\sigma}}{\Gamma_{tot}}\right)^2} \]  

(4.39)

where \( E_{n,\sigma} = (n + 1/2)\hbar\omega_C + \sigma g\mu_B B/2 \) is the Landau level spectrum, \( E_0 \) is the average subband energy, and \( \Gamma_{tot} \) is the width of each Landau level in energy. The carrier density of the two-dimensional system is then expressed as the following integral (at zero temperature):

\[ n_{2D} = \int_{-\infty}^{\mu} \rho_{tot}(E) dE \]  

(4.40)

where \( \mu \) is the chemical potential. By convention we set \( \mu = 0 \). Fig. 4.17 shows the oscillation of the Fermi energy \( E_F = \mu - E_0 \) as a function of magnetic field found by solving Eqn. 4.39,4.40 for \( E_F \) assuming \( n_{2D} \) is independent of magnetic field. This is easily done numerically using a bisection routine. Note that the integral of a Lorentzian is an arc-tangent function. The corners of the trace are rounded owing to the disorder broadening of the Landau levels. A carrier density of \( n_{2D} = 2.8 \times 10^{15} m^{-2} \), Landau level width \( \Gamma_{tot} = 0.06 \text{ meV} \) and Lande g-factor \( g = 0.44 \) was used.

The upper trace in Fig. 4.18 shows the oscillation of the total density of states \( \rho_{tot}(E) \) evaluated at the chemical potential corresponding to Fig. 4.17 evaluated using Eqn. 4.40 and using \( E_F \) as shown in Fig. 4.17. It is divided by the magnetic field \( B \) and normalized so that peaks have unit height at high field for clarity.

According to our model of the states in a Landau level, the extended states must have a narrower spread in energy than the total density of states and be centred around the centre of each Landau level, as in Fig. 4.16. We use
the following functional form for the extended density of states:

$$\rho_{\text{ext}}(E) = \sum_{n,\sigma} -\frac{\partial}{\partial E} f \left( \frac{E - E_0 - E_{n,\sigma}}{\Gamma_{\text{ext}}} \right)$$  \hspace{1cm} (4.41)

where \( f(E) = \frac{1}{1 + \exp(E)} \). The normalization for this function is such that the total integrated density of extended states in each Landau level is unity. This can be justified in terms of the percolation picture. It suggests that there should be only one energy per Landau level where percolation is possible for a sufficiently large system. In subsequent lectures this justification for Eqn. 4.41 will be strengthened and modified by the introduction of the edge-state picture. The lower trace in Fig. 4.18 shows the extended density of states at the chemical potential as a function of magnetic field using \( \Gamma_{\text{ext}} = 0.02 \text{meV} \) using \( E_F \) as shown in Fig. 4.17. This trace also has the peak heights normalized so that they tend to unity at high field.

There are two important things to notice about the traces in Fig. 4.18. Firstly, there are regions where the extended density of states is equal to zero and hence will give rise to zero diagonal conductivity. Secondly the peaks become wider at higher magnetic fields. This increase in width is not because the Landau levels are getting broader the calculation assumes that they are of constant width it is because, as the density in each Landau level
increases, the Landau levels become pinned at the chemical potential over greater magnetic field ranges, as can be seen in Fig. 4.17.

Since it is the extended states that are responsible for the diagonal resistivity we can also imagine that they must be responsible for the off-diagonal resistivity, the Hall resistivity. Motivated by this and the functional form of the off diagonal terms in Eqn. 4.9 the simple model calculates the Hall resistance from:

\[
\rho_{x,y} = \frac{h}{e^2} \int_{-\infty}^{\mu} \rho_{\text{ext}}(E) dE = \frac{h}{e^2} \sum_{n,\sigma} f \left( \frac{E_F - E_{n,\sigma}}{\Gamma_{\text{ext}}} \right) \]

The normalization \(h/e^2\) ensures that \(\rho_{xy}\) has the correct classical Hall slope. Fig. 4.19 shows the Hall resistivity corresponding to Fig. 4.17 calculated using Eqn. 4.42. It shows that the Hall plateaux occur when the density of extended states is essentially zero. These regions centre on magnetic fields where the filling factor is an integer and at these points the Hall resistivity has its classical value \(\rho_{xy} = B/en_{2D}\) which on substitution for the density \(n_{2D} = \nu eB/h\) becomes \(\rho_{xy} = h/\nu e^2\).
The quantum Hall and Shubnikov de Haas effects

Fig. 4.19. Hall resistance as a function of magnetic field. Parameters, as in Fig. 4.17.

4.13 Appendix I

As was shown in the second lecture, the linear Boltzmann equation at finite magnetic field has the form:
\[ \nabla_k g \frac{e}{\hbar} v_k \wedge B + \nabla_k f^0 \frac{e}{\hbar} E = \frac{g}{\tau_k} \] (4.43)

In the constant relaxation time approximation \( \tau_k = \tau \) its solution can be written in the form:
\[ g(k) = \frac{e\tau}{\hbar} \nabla_k f^0 . Z \] (4.44)

where \( Z \) is the so called Hall vector:

Substituting Eqn. 4.43 into Eqn. 4.44 we have
\[ \nabla_k \left( \frac{e\tau}{\hbar} \nabla_k f^0 . Z \right) \frac{e}{\hbar} v_k \wedge B + \nabla_k f^0 \frac{e}{\hbar} E = \frac{e}{\hbar} \nabla_k f^0 . Z \] (4.45)

Substituting the identity \( \nabla_k f^0 = (\partial f^0 / \partial \epsilon) \hbar v_k \) into equation Eqn. 4.45 we find
\[ \nabla_k \left( e\tau \frac{\partial f^0}{\partial \epsilon} v_k . Z \right) \frac{e}{\hbar} v_k \wedge B + \frac{\partial f^0}{\partial \epsilon} v_k . E \frac{\partial f^0}{\partial \epsilon} v_k . eZ = \frac{e}{\hbar} \nabla_k f^0 . eZ \] (4.46)

Expanding the grad term \( \nabla_k \left( e\tau \frac{\partial f^0}{\partial \epsilon} v_k . Z \right) \) in Eqn. 4.46 gives two terms.
The first term is zero since $\nabla_k \left( \frac{\partial f^0}{\partial \epsilon} \right) \propto v_k$ and $v_k \cdot v_k \wedge B = 0$. Eqn. 4.46 therefore reduces to

$$\frac{\partial f^0}{\partial \epsilon} \frac{e^2 \tau}{\hbar} \nabla_k (v_k \cdot Z) \cdot v_k \wedge B + \frac{\partial f^0}{\partial \epsilon} v_k \cdot eE = \frac{\partial f^0}{\partial \epsilon} v_k \cdot eZ \quad (4.47)$$

Using $m^* v_k = \hbar k$ the grad term in Eqn. 4.47 simplifies to

$$\nabla_k (v_k \cdot Z) = \frac{\hbar}{m^*} Z \quad (4.48)$$

Substituting Eqn. 4.48 into Eqn. 4.47 and using the vector identity $a \cdot b \wedge c = b \cdot c \wedge a = c \cdot a \wedge b$ equation Eqn. 4.47 becomes

$$\frac{\partial f^0}{\partial \epsilon} \frac{e^2 \tau}{m^*} v_k \cdot B \wedge Z + \frac{\partial f^0}{\partial \epsilon} v_k \cdot eE = \frac{\partial f^0}{\partial \epsilon} v_k \cdot eZ \quad (4.49)$$

Dividing Eqn. 4.49 through by $e \partial f^0 / \partial \epsilon$ and noting that the resulting expression is true for all $v_k$ we find

$$E = Z - \frac{e \tau}{m^*} B \wedge Z \quad (4.50)$$

4.14 Exercises

4.1 What are the quantized values of the Hall resistance in Fig. 4.3 (measure voltages from the left axis).

4.2 Use the Hall slope and the Shubnikov de Haas minima in Fig. 4.3 to make two different estimates of the carrier density of the two-dimensional system.

4.3 What are the principle features of the Shubnikov de Haas and Quantum Hall effects? Which aspects of these features are at odds with the assumptions of the Boltzmann transport formalism and why? Which aspects are consistent the existence of Landau levels and why?

4.4 In the limit of zero disorder there will be no localised states in the Landau levels of a two-dimensional system. What consequence would this have on the observation of the quantum Hall and Shubnikov de Haas effects?

4.5 How can you reconcile the difference between the solutions of the zero field Schrödinger equation $E_k = \hbar^2 k^2 / 2m^*$ and the solutions of Eqn. 4.16 at zero$^+ \text{ magnetic field } E_n = (n + 1/2)\hbar \omega_C$. 
Quantum transport in one dimension.

5.1 Sources


5.2 Introduction

These notes take a detailed look at the quasi-one-dimensional system and its experimental realization in the split-gate device. We show magnetotunnelling and capacitance data that demonstrate features of its eigenspectrum. We derive the Landauer transport formalism and use it to find the conductance of a saddle point potential. The topics covered in these notes are: the experimental realization of a quasi-one-dimensional system; its eigenstates; its density of states; its conductance; the Landauer formalism; and the saddle point potential.
5.3 Experimental realization of the quasi-one-dimensional system [1]

Figure 5.1 shows a schematic diagram of a split-gate device patterned on the surface of a GaAs-AlGaAs heterostructure containing a single two-dimensional-electron gas. When a sufficiently large negative voltage is applied to these gates the electron gas beneath them is depleted leaving behind a narrow quasi-one-dimensional electron gas. In the limit where the width of the channel is only a few Fermi wavelengths, the transport properties of such systems are determined quantum mechanically and transverse confinement leads to them having many of the properties of true one-dimensional systems.

5.4 Eigenstates of an infinitely long quasi-one-dimensional system

Schrödinger’s equation for a quasi-one-dimensional system with transverse confining potential $V(y)$ has the form:

$$\left(\frac{p^2}{2m^*} + V(y)\right)\Psi = E\Psi \quad (5.1)$$

It has translational invariance in the $x$ - direction and therefore solutions of the form

$$\Psi_{k_x,n} = \exp(ik_xx)\phi_n(y) \quad (5.2)$$
Substituting 5.2 into 5.1 we have

\[
\left( \frac{p_y^2}{2m^*} + V(y) \right) \phi_n = \left( E_{k_x,n} - \frac{\hbar^2 k_x^2}{2m^*} \right) \phi_n
\]

(5.3)

Equation 5.3 is the one-dimensional Schrödinger equation familiar from any quantum mechanics text book. The simplest example of its solution is for the Harmonic oscillator potential:

\[ V(y) = \frac{1}{2} m^* \omega_0^2 y^2 \]

(5.4)

The solutions to Eqn 5.3 with the potential in Eqn 5.4 are expressed in terms of the variable

\[ z = \left( \frac{m^* \omega_0}{\hbar} \right)^{\frac{1}{2}} y \]

(5.5)

and have the form

\[ \phi_n(y) = e^{-z^2/2} H_n(z) \]

(5.6)

where \( H_n \) is the \( n^{th} \) Hermite polynomial. These solutions are shown in Fig 5.2.

The corresponding eigenvalues have the form

\[ E_{k_x,n} = (n + \frac{1}{2}) \hbar \omega_0 + \frac{\hbar^2 k_x^2}{2m^*} \]

(5.7)

and are plotted as a function of \( k_x \) in Fig. 5.3.

In deriving the eigen functions and values for a parabolic confining potential, we have chosen a special case of a class of functions \( V(y) = A |y|^\alpha \).
5.5 Kardynal et al [2]

They have the property that for $\alpha > 2$ the eigenvalues of increasing energy become increasingly far apart in energy. The typical example is $\alpha \to \infty$, the square well. For $\alpha = 2$ eigenvalues are equally spaced and for $\alpha < 2$ the spacing between eigenvalues become smaller with increasing energy. The typical example is $\alpha = 1$, the triangular well. For an arbitrary potential $V(y)$ the eigenvalues will have the form

$$E_{k_x,n} = E_{0,n} + \frac{\hbar^2 k_x^2}{2m^*}$$

(5.8)

We refer to the dispersion curve for each value of $n$ as a one-dimensional subband and $E_{0,n}$ as the subband energy.

5.5 Kardynal et al [2]

Magnetotunnelling is an ideal probe of the quantum mechanical structure just described. Fig. 5.4 shows an electronic device used to observe the subband energies and wave functions of a quasi-one-dimensional electron system. The basic device consists of two parallel two-dimensional electron systems (light grey) separated by a 125 Å AlGaAs barrier in a GaAs-AlGaAs heterostructure. Ohmic contacts (dark grey) contact the upper and lower two-dimensional systems independently so that current passed from the left contact to the right contact must pass through the upper two-dimensional system, tunnel across the barrier and pass out through the lower two-dimensional system. Independent contact to the two layers is achieved.
Quantum transport in one dimension.

Fig. 5.4. Schematic diagram of the patterned double quantum well device used by Kardynal et al [2] to observe subband energies and wave functions of quasi-one-dimensional electron systems.

via front and back gates $V_a$ and $V_b$. Large negative voltages applied to them deplete the two-dimensional electron system closest to them.

Above the upper two-dimensional system is a set of 51 150$\mu$m long metallic gates. When a negative voltage $V_g$ is applied to these gates the electron gas is depleted under them (white pattern on grey upper two-dimensional system) leaving behind a set of 50 identical long quasi-one-dimensional systems. The experiment consists of measuring the equilibrium tunnelling differential conductance $dI/dV$, where $I$ is the current passing between layers and $V$ the potential difference between layers, as a function of an externally applied magnetic field placed parallel to the two two-dimensional planes. The term equilibrium refers to the fact that the voltage $V$ is made sufficiently small that it can be ignored in considering the individual properties of the two electron systems.

Since each quasi-one-dimensional system in Fig. 5.4 is nominally identical, it is sufficient to consider tunnelling between a single one-dimensional system and a two-dimensional system in a parallel magnetic field in order to understand the results of tunnelling conductance measurements.

At low temperatures, the tunnelling between the two electron systems can be taken to occur only at their mutual chemical potential, that is, between the Fermi surfaces. The two layers have the same chemical potential in equilibrium because the barrier width is sufficiently small to allow tunnelling. Electrons tunnelling from one layer to the other conserve momentum. Conservation of momentum is equivalent to conservation of wave vector. We therefore need to know the probability for finding an electron of a given wave vector (the spectral density) on the Fermi surface of both the one-
dimensional and the two-dimensional system. The tunnelling conductance of the device is then the overlap integral of these two spectral functions.

The wave functions in a two-dimensional system are plane waves
\[
\Psi_{k_x, k_y} = e^{ik_x x} e^{ik_y y}
\]
(5.9)
and have a dispersion of the form
\[
E_{k_x, k_y} = \frac{\hbar^2}{2m^*} \left( k_x^2 + k_y^2 \right)
\]
(5.10)
Hence, at the Fermi surface where \( E_{k_x, k_y} = E_F \) there is only finite spectral density on a ring with \( k_F^2 = k_x^2 + k_y^2 \), where \( k_F \) is the Fermi wave vector, as shown in Fig. 5.5.

The spectral density in a quasi-one-dimensional system is complicated (fortunately) by the fact that the wave function is only plane wave like in the \( x \) - direction and composed of the non-plane wave like subband wave functions in the \( y \) - direction
\[
\Psi_{k_x, n} = \exp(ik_x x)\phi_n(y)
\]
(5.11)
The dispersion relation for each subband \( n \) has the form
\[
E_{k_x, n} = \frac{\hbar^2 k_x^2}{2m^*} + E_{0,n}
\]
(5.12)
so that, at the Fermi surface the values of $k_x$ are quantised (see Fig. 5.3) with

$$k_x^n = \pm \sqrt{\frac{2m^*(E_F - E_{0,n})}{\hbar^2}} \quad (5.13)$$

The two solutions in Eqn. 5.13 represent forward and reverse moving electrons. The spectral density for the $n^{th}$ subband at the chemical potential has the form

$$A_{n}^{1D}(k_x, k_y) \propto \left| \text{FourierTransform} \left[ \Psi_{k_x^n, n}(x, y; \mu) \right] \right|^2$$

$$\propto \left| \int dx dy e^{ik_x x} e^{ik_y y} \Psi_{k_x^n, n}(x, y; \mu) \right|^2$$

$$\propto \left| \int dx dy \left( e^{ik_x x} + e^{-ik_x x} \right) e^{ik_y y} \phi_n(y) e^{ik_y y} \right|^2$$

$$\propto \left( \delta (k_x^n - k_x) + \delta (-k_x^n - k_x) \right) \left| \int dy \phi_n(y) e^{ik_y y} \right|^2 \quad (5.14)$$

The Fourier transforms of the subband wave functions that appear in equation Eqn. 5.14 are easily evaluated for the case of the Harmonic potential since

$$\int dz e^{ik_z z} e^{-z^2/2} H_n(z) \propto e^{-k_z^2/2} H_n(k) \quad (5.15)$$

For this special case the $k_y$ dependence of the spectral density is exactly the same as the $y$ - dependence of the subband wave functions (these are shown in Fig. 5.2). In general, the two are not the same but the Fourier transform of an arbitrary subband wave function always has exactly the same number of maxima and minima as the wave function itself. Fig. 5.6 is a schematic representation of the spectral function of a one-dimensional system at its chemical potential with three occupied subbands.

The six vertical lines have $k_x$ values given by equation Eqn. 5.13 and are modulated by the modulus squared of the Fourier transform of the respective subband wave functions as indicated by equation Eqn. 5.14. As we have said, the tunnelling conductance of the device in Fig. 5.4 is measured as a function of an external magnetic field placed parallel to the two-dimensional layers. The reason for this is that when an electron leaves the upper layer and tunnels to the lower layer in the presence of such an external field its path is bent by the Lorentz force which alters the condition for conservation of momentum. Fig. 5.7(a) is a schematic diagram of the tunnelling process for the case where the magnetic field is applied both parallel to the two-dimensional system and to the one-dimensional systems $B = B_x$. 
Fig. 5.6. Schematic diagram of the spectral density of a quasi-one-dimensional electron gas with three occupied subbands.

Fig. 5.7. (a) Effect on a tunnelling electron of a magnetic field in the $x$ - direction (b) As for (a), for $y$ - direction.

For this case momentum is conserved in tunnelling if

$$k_{y}^{2D} \rightarrow k_{y}^{1D} + \frac{eB_{x}d}{\hbar}$$  \hspace{1cm} (5.16)

Equation 5.16 is justified in lecture notes 1 and strictly only applies to the case where the one-dimensional and two-dimensional systems can be treated as completely independent. The variable $d$ is the distance between the centres to the two quantum wells. Fig. 5.8(a) shows the spectral density for a one-dimensional system with six occupied subbands and the spectral density for a two-dimensional system on the same graph. The one-dimensional
Fig. 5.8. (a) Relative displacement of the spectral functions of the two-dimensional system and the quasi-one-dimensional system for a magnetic field applied in the x-direction (b) for a magnetic field applied in the y-direction.

and the two-dimensional spectral densities are offset by $\Delta k_y = eB_x d/\hbar$ indicating the overlap between equivalent $k$-states in the two systems at an arbitrarily chosen magnetic field $B_x$. The points where both spectral densities have finite weight are the points in $k$-space where tunnelling occurs. The differential tunnelling conductance is given by the overlap of the one-dimensional and the two-dimensional spectral densities:

$$\frac{dI}{dV} \propto \int \int dk_x dk_y A_{2D} \sum_{n_{occ}} A_{1D}$$

(5.17)

Fig. 5.7b is a schematic representation of the tunnelling process for the case where the magnetic field is applied parallel to the two-dimensional system and perpendicular to the one-dimensional systems $B = B_y$. For this case, momentum is conserved in tunnelling if

$$k_{x2D}^{2D} \rightarrow k_{x1D}^{1D} - \frac{eB_y d}{\hbar}$$

(5.18)

Fig. 5.8(b) shows the equivalent of Fig.5.8(a) for this case.

If we look again at Fig. 5.8(a) and ask how the overlap between the two-dimensional and one-dimensional spectral densities varies as a function of $B_x$, we see that at zero magnetic field the two densities have no overlap and therefore we expect zero $dI/dV$. As the magnetic field is increased the Fermi circle of the two-dimensional system passes across the spectral function of the one-dimensional system and a modulated $dI/dV$ is observed. It is found
that $dI/dV$ is most strongly modulated by the spectral density variations
due to the highest occupied subband with the contributions due to lower
energy subbands tending to be smeared out giving an average background.
Fig. 5.9 shows the tunnelling differential conductance of the device in Fig. 5.4
at a series of different surface gate voltages ($V_g$) for a magnetic field swept
in the $B_x$ direction.

Each trace in Fig. 5.9 is an image of the highest occupied one-dimensional
subband wave function in $k$ - space. One can see how as $V_g$ becomes increas-
ingly negative the number of oscillations decreases and the symmetry of the
oscillations varies from odd to even, indicating the sequence of subband
depopulations.

If we now look again at Fig. 5.8(b) and ask how the overlap between Fermi
circle and the one-dimensional spectral density varies as a function of $B_y$, we
see that as the magnetic field is increased, the tunnelling conductance is a
maximum each time the trailing edge of the Fermi circle intersects with one
of the fixed $k_x$ values given by Eqn. 5.13. Fig. 5.10 shows the corresponding
experimental traces at a series of different gate voltages.

In Fig. 5.10 one can see how as $V_g$ becomes increasingly negative, the
number of oscillations decreases as subbands depopulate. From the $k_x$ peak
positions in Fig. 5.10, one can determine the subband energies using equation
Eqn. 5.13. These are shown in Fig. 5.11 as a function of surface gate voltage.
Quantum transport in one dimension.

Fig. 5.10. Tunnelling differential conductance of the device shown in Fig. 5.4 as a function of $B_y$ taken at surface gate voltages $V_g=-0.48,-0.50,...,-0.62$V - upper trace to lower trace. The centre to centre distance between surface gates was 350nm and the gap between gates was 170nm.

Fig. 5.11. Subband energies taken from Fig. 5.10 as a function of surface gate voltage.

The subband energies in Fig. 5.11 show a tendency to become larger as $V_g$ is made more negative. This trend can be understood once the shape of the effective potential $V(y)$ as a function of the surface gate voltage is understood. Each of the traces in Figs. 5.9,5.10 are a cross-section of the one-dimensional spectral density at each gate voltage and the positions of
Fig. 5.12. Experimental and calculated tunnelling differential conductance for $V_g = -0.48$V for $B_x$ a,b and $B_y$ c,d for the 350nm device a,c and the 570nm device b,d. (e) shows the best fitting potential profile for the 350nm device and (f) that for the 570nm period device.

Peaks in each trace are sufficiently unique that they can be used to determine the shape of the effective potential ($V(y)$ in Eqn. 5.1). Figures 5.12a,c show single tunnelling conductance traces (solid lines) for $B_x$ and $B_y$ for a device with ‘narrow’ quasi-one dimensional systems $W \sim 170$nm.

Dotted lines in Figs.5.12a,c show the corresponding theoretical traces for the effective potential shape, Fig. 5.12e, which gave the best match for peak positions. In the narrow wire at the gate voltage shown, the electron density is low and therefore cannot screen the external field from the surface gates which gives an effective potential of approximately parabolic shape and quite a strong curvature. Figures 5.12b,d,f show the equivalent to Figs. 5.12a,c,e for a ‘wide’ quasi-one dimensional system $W \sim 230$nm, measured at the same gate voltage. In the wide wire the electron density is sufficiently high that it can screen the external electric field from the surface gates to give the effective potential a flat centre region.
Quantum transport in one dimension.

Fig. 5.13. (a) Device geometry. The layer thicknesses are 24nm for the undoped GaAs top layer, 36nm for the doped AlGaAs layer, 10nm for the undoped AlGaAs spacer layer and 6.6μm for the GaAs substrate. The donor binding energy and heterojunction conduction band edge offset are taken to be 0.05 and 0.12eV respectively. The calculation assumes a uniform surface charge density in the gate opening. (b) Calculated transverse potential profiles at 4.2K along a line 5.6nm from the GaAs-AlGaAs interface, near the mean of the vertical electron charge distribution for four values of gate voltage.

Calculation of the shape of a quasi-one dimensional effective potential is a complicated numerical task. It requires the simultaneous solution of Schrödinger’s equation and the Poisson equation. Fig. 5.13a shows a device configuration for which a calculation was performed [3] and Fig. 5.13b shows the resulting effective potential, $V(y)$, at four different gate voltages. At -1.52V the potential is clearly well approximated by a parabola but at higher gate voltages the screening by electrons gives it a flat centre region.

From the shape of these effective potentials we can understand the voltage dependence of the subband energies in Fig. 5.11. For the wide flat centred effective potential found at low gate voltages the subband energies are closely spaced and for the tight parabolic effective potential at high gate voltages the subband energies are large.

5.6 Density of States in a quasi-one-dimensional system

Each subband of a quasi-one-dimensional system defined by the Schrödinger equation 5.1 is an independent dynamically one-dimensional system. The
density of states in a quasi-one-dimensional system is therefore a sum of the densities of states for each one-dimensional subband.

If we consider a one-dimensional system of length $L$ with periodic boundary conditions then $k_x L = 2\pi n$, where $n$ is an integer, so that the spacing between states in one-dimensional $k$-space is $\Delta k_x = 2\pi / L$. If the highest occupied state has wave vector $k$, then the density of occupied states will be

$$n_1 = \frac{(2k/\Delta k_x)}{L} = \frac{k}{\pi}$$  \hfill (5.19)

If we now substitute for $k$ from Eqn. 5.13 the density in the $n^{th}$ subband for electrons filled to energy $E$ is for

$$n_n = \frac{1}{\pi \hbar} \sqrt{2m^*(E - E_{0,n})} \quad \text{for} \quad E_{0,n} \leq E$$

$$= 0 \quad \text{for} \quad E < E_{0,n}$$  \hfill (5.20)

so that the one-dimensional carrier density at zero temperature is

$$n_c = g \sum_{n, E_{0,n} \leq E} \frac{1}{\pi \hbar} \sqrt{2m^*(E - E_{0,n})}$$  \hfill (5.21)

$g$ is the degeneracy of each one-dimensional subband $g = 2$ in GaAs and $g = 4$ in Si. Fig. 5.14 shows the carrier density of a square well potential as a function of energy evaluated using Eqn. 5.21.

If we now differentiate Eqn. 5.21 with respect to energy $E$ the density of
Quantum transport in one dimension.

Fig. 5.15. One-dimensional density of states as a function of energy for a square well potential.

The density of states has the form
\[
\rho_{1D}(E) = g \sum_{n, E_{0,n} \leq E} \frac{\sqrt{2m^*}}{2\pi\hbar} \left( E - E_{0,n} \right)^{-\frac{1}{2}}
\]  
(5.22)

Fig. 5.15 shows the density of states, evaluated using Eqn. 5.22, corresponding to Fig. 5.14.

5.7 Oscillation of the Fermi energy and Capacitance

The density of states modulation of a quasi-one-dimensional system is very similar to that of a two-dimensional system in a perpendicular magnetic field owing to the singularity at each subband energy. By comparison with the high magnetic field case, it is clear that this density of states modulation will cause oscillations in both the Fermi energy and the capacitance of a quasi-one-dimensional system. This section shows two experiments. The first is by Linda Macks and shows the oscillation of the Fermi energy through magneto-tunnelling measurements. The second is by Drexler and is a capacitance measurement of a set of induced quasi-one-dimensional systems.

5.7.1 Macks et al [4]

Figure. 5.16 shows the result of a magneto-tunnelling measurement on a set of 200 quasi-one-dimensional systems. The device structure was the same as that shown in Fig. 5.4. The surface-gate array had a period of
5.7 Oscillation of the Fermi energy and Capacitance

250nm and each gate had a width of 100nm and length 45µm. The quasi-one-dimensional systems were sufficiently narrow that by applying increasingly negative split-gate voltages the lowest five subbands could be observed: Figs. 5.16(a,b). A grey-scale image of the $B_y$ dependence of the magneto-tunnelling conductance, Fig. 5.16(a), shows kinks in the trajectories of the 1st and 2nd subband tunnelling peaks as the 3rd and 4th subbands pin at the device chemical potential. Fig. 5.16(d) shows a selfconsistent Poisson and Schrödinger calculation of the 1D subband energies as a function of surface-gate voltage. It shows how the subband energies pin in regions of high 1D density of states and jump in regions of low 1D density of states. It also shows the effect of this pinning on the effective potential minimum. At the most negative gate voltages it can be seen that the quantum well actually deepens as the last subband depletes.
Quantum transport in one dimension.

Fig. 5.17. (a) Sketch of the interdigitated gate geometry of the top of the device and voltage sources applied at the gate electrodes. (b) Cross-section of the sample. The electron systems are generated beneath the centre gate biased at voltage $V_g$. The side gates are biased with respect to this gate by a voltage $V_d$. (c) The calculated bare potential induced by the gate configuration in (b) at the location of the electron systems with gate voltages indicated.

5.7.2 Drexler et al [5]

There have been a number of measurements of the oscillation of the capacitance of a quasi-one-dimensional system but the clearest were performed by Drexler et al [5] using a GaAs-AlGaAs MISFET (Metal insulator semiconductor field effect transistor). This device is similar to the Si MOSFET in that there is no electron gas at the heterojunction interface until a potential is applied to a surface gate. A set of parallel quasi-one-dimensional systems are induced by placing potential differences between two sets of interdigitated surface gates and a back gate - shown in Fig. 5.17.

The interdigitated surface gate geometry is formed by 300 metal strips of 100nm width, 500nm period, and with 180µm length. The capacitance between the set of gates above the one-dimensional systems and a back gate is measured. Fig. 5.18 shows the measured capacitance.

Drexler et al assume that the backgate and the set of quasi-one-dimensional systems are at the same chemical potential. This gives the same series capacitance relation as for the 2D case $1/C = 1/C_G + 1/C_Q$. Where $C_G$ is the geometric capacitance and $C_Q$ is the quantum capacitance which is proportional to the thermodynamic density of states. Thus in Fig. 5.18 the larger the capacitance the larger the thermodynamic density of states. The difference between Fig. 5.15 and Fig. 5.18 - quite apart from Fig. 5.15 being the density of states and Fig. 5.18 being a measure of the thermodynamic density of states - indicates that for these wires, disorder and averaging over 300 slightly different wires has nearly removed all modulation in the density of states.
5.8 The Conductance of a quasi-one-dimensional system

In a perfect quasi-one-dimensional system a current will travel in any one-dimensional subband without scattering into any other. If we consider current passing through a single subband labelled $i$ between two perfect Ohmic contacts. The left contact is at chemical potential $\mu_L = \mu - eV$ and the right contact at chemical potential $\mu_R = \mu$ where $V$ is an applied potential difference between the two contacts. The current passing to the right in a small energy range $d\epsilon$ is then

$$dI^+_i = -e v^i_\epsilon f(\epsilon + eV) \frac{1}{2} \frac{dn_i}{d\epsilon} d\epsilon$$  \hspace{1cm} (5.23)

The term $-e v^i_\epsilon$ is the current in a single state in subband $i$ at energy $\epsilon$. $v^i_\epsilon$ is the group velocity at energy $\epsilon$ in subband $i$ defined by

$$v^i_\epsilon = \left. \frac{1}{\hbar} \frac{dE_{kz,i}}{dk_x} \right|_\epsilon$$ \hspace{1cm} (5.24)

The second term $f(\epsilon + eV) \frac{1}{2} \frac{dn_i}{d\epsilon} d\epsilon$ is the number of electrons passing to the right through subband $i$ in the energy range $d\epsilon$ around energy $\epsilon$ (note
Quantum transport in one dimension.

that the inclusion of the factor 1/2 is because the definition of \( n_i \) in equation Eqn. 5.19 includes electrons propagating in both directions but here we only wish to include electrons passing to the right. Substituting Eqn. 5.19 and Eqn. 5.24 into Eqn. 5.23 we find that the energy dependence of the group velocity and density of states cancel and we are left with

\[
dI_i^+ = -\frac{e}{\hbar} f(\epsilon + eV)d\epsilon
\]

That is, the current in any subband is independent of its index. A similar expression is derived for states moving to the left in subband \( i \)

\[
dI_i^- = -\frac{e}{\hbar} f(\epsilon)d\epsilon
\]

The total current flowing through subband \( i \) is then

\[
I_i = \int (dI_i^+ - dI_i^-) = \frac{e}{\hbar} \int_{-\infty}^{\infty} (f(\epsilon) - f(\epsilon + eV))d\epsilon
\]

In the limit of zero temperature this becomes

\[
I_i = \frac{e^2}{\hbar} V
\]

If there are \( N \) subbands occupied in the quasi-one-dimensional system then the total current will be

\[
I = \sum_i I_i = N \frac{e^2}{\hbar} V
\]

which implies a conductance of

\[
G = N \frac{e^2}{\hbar}
\]

This result is quite remarkable, the conductance of a perfect quasi-one-dimensional system is equal to an integer multiple of \( e^2/\hbar \). The experimental confirmation of this result [6,7] resolved a theoretical debate spanning thirty years [8].

5.9 Thomas et al [9]

Recent developments in growth and fabrication technology have produced quasi-one-dimensional systems of astounding purity. It is interesting to quote from Thomas et al [9]:

"The samples used in this study were made from two modulation doped GaAs-Al_{0.33}Ga_{0.67}As heterostructures, T119 and T166, grown by molecular beam epitaxy. Both wafers have the same structure above the plane"
of the two-dimensional system: 600Å undoped Al$_{0.33}$Ga$_{0.67}$As spacer layer, 2000Å Si-doped (1.2x10$^{17}$cm$^{-3}$) Al$_{0.33}$Ga$_{0.67}$As layer, and 170Å undoped GaAs cap layer. Split gates of width 0.75µm and length 0.4µm were defined on a Hall bar using electron beam lithography. Ohmic contacts were made by thermal evaporation of Au/Ge/Ni alloys. After 20s of illumination with a red light-emitting diode positioned 1cm away from the sample, the saturated values of the low temperature mobility and carrier concentration of sample T119 were 3x10$^6$cm$^2$/Vs and 1.9x10$^{11}$cm$^{-2}$, while those of sample T166 were 1.5x10$^6$cm$^2$/Vs and 1.9x10$^{11}$cm$^{-2}$. Prior to illumination the mobility and carrier concentrations were found to be less by 35% and 50% respectively. Although both heterostructures have the same growth sequence, the electron mobility is higher in T119, as it was grown towards the end of a growth run. Two-terminal ac conductance measurements of the illuminated samples were carried out at the base temperature (less than 100mK) of a dilution refrigerator using an excitation of 1µV at a frequency of 85Hz. Curves (a) and (b) in Fig. 5.19 show the differential conductance $G(V_g) = dI/dV_{sd}$ of two one dimensional constrictions in samples T166 and T119 respectively. Curve (a) shows 25 conductance plateaux and curve (b) shows 21 conductance plateaux quantised in units of 2$e^2/h$, after subtraction of a series resistance $R_s$ which ensures that the lower plateaux align with the exact quantised values 2$N e^2/h$. Where $N$ is the subband index. Upon application of a negative voltage $V_g$=-0.6V to the split gate, electrons are depleted from beneath both gates and the channel is defined; the conductance (see Fig.5.8 inset) drops dramatically at this definition voltage. As $V_g$ is made more negative, the constriction is narrowed and the carrier density is reduced, and $G$ decreases in units of 2$e^2/h$ as the one-dimensional subbands in the constriction are depopulated."

The above passage mentions a series resistance that occurs in addition to the resistance of the split gate structure. The device is not a perfect manifestation of a quasi-one-dimensional system connected to perfect Ohmic contacts and therefore not all the voltage that is applied is dropped across the device. The conductance is measured in units of 2$e^2/h$ because in zero magnetic field GaAs has a spin degeneracy of 2.

Perhaps the most important observation for us to make in Fig. 5.19 is that the conductance steps are not the perfect step functions predicted by Eqn. 5.30 but have a finite spread to them that is much wider than the temperature or excitation voltage used. One is tempted to say immediately that disorder broadens the steps and the result is therefore trivial. However the way in which disorder or any feature in the split gate effective potential
Quantum transport in one dimension.

Fig. 5.19. Differential conductance $G(V_g) = dI/dV_{sd}$ measured at a temperature of $T=50\text{mK}$, quantised in units of $2e^2/h$ after a series resistance $R_S$ has been subtracted: (a) sample T166 showing 25 quantised plateaux ($R_S=345\Omega$). (b) sample T119 showing 21 quantised plateaux ($R_S=230\Omega$). Inset: raw data showing the definition and pinch off characteristics of: (1) sample T166 and (2) sample T119 (data offset by 500$\mu$S).

broadens the ballistic steps is not at all trivial and requires a quantum-mechanical transport formalism to account for it.

5.10 The Landauer Formalism

If we wish to calculate the two-terminal conductance of a device such as the split gate device described above, there are two routes that are mathemati-
5.10 The Landauer Formalism

cally equivalent: the Kubo formalism and the Landauer formalism [8]. We chose the latter.

The first step in deriving the Landauer formalism is to convert the quantum mechanical problem into a scattering matrix problem for electrons incident on an electronic device at energy \( \epsilon \). We consider the situation shown in Fig. 5.20 in which two semi-infinite quasi-one-dimensional systems are connected by a region of electron gas with arbitrary effective potential \( V(x, y) \).

The corresponding two-dimensional Schrödinger equation has the form

\[
\left( \frac{p^2}{2m^*} + V(x, y) \right) \Psi = E \Psi \tag{5.31}
\]

To the left of the region the wave function has the form

\[
\Psi_L(x, y) = \sum_i e^{ik_i^L x} \phi_i^+(y) \frac{a_i}{j_i} + \sum_i e^{-ik_i^L x} \phi_i^-(y) \frac{a_i^*}{j_i} \tag{5.32}
\]

and to the right it has the form

\[
\Psi_R(x, y) = \sum_i e^{ik_i^R x} \phi_i^+(y) \frac{b_i}{j_i} + \sum_i e^{-ik_i^R x} \phi_i^-(y) \frac{b_i^*}{j_i} \tag{5.33}
\]

The quantities \( \phi_i^+(y), \phi_i^-(y) \) are the transverse eigenfunctions in the quasi-one-dimensional systems evaluated at energy \( \epsilon \). They form a complete basis at any chosen energy \( \epsilon \) only if all subbands are included, irrespective of whether for that energy their wave vector \( k_i \) is imaginary. States with imaginary \( k_i \) are referred to as ‘evanescent modes’ in analogy with their counterparts in optics. \( j_i \) is the current in mode \( i \). \( a_i^+, a_i^-, b_i^+, b_i^- \) are the current amplitudes for forward and reverse going states on the left and right of the connecting region.

If we substitute Eqn. 5.32 and Eqn. 5.33 into Eqn. 5.31 we can solve for the relationship between the current amplitudes on the left and right of the inserted region. Typically this is done by discretising Eqn. 5.31 to achieve a finite transverse basis set and then solving numerically. The result may be
written in matrix form as

$$
\begin{bmatrix}
b^+ \\
a^-
\end{bmatrix} =
\begin{bmatrix}
t^+ & r^+ \\
r^- & t^-
\end{bmatrix}
\begin{bmatrix}
a^+ \\
b^-
\end{bmatrix} =
S
\begin{bmatrix}
a^+ \\
b^-
\end{bmatrix}
$$

(5.34)

where \(a^+\), \(a^-\), \(b^+\), \(b^-\) written now without subscripts are vectors containing the amplitudes \(a_i^+, a_i^-, b_i^+, b_i^-\). The matrices \(t^+, r^+, t^-, r^-\) are the transmission and reflection matrices of the arbitrary region. The elements \(t_{i,j}^+, r_{i,j}^+, t_{i,j}^-, r_{i,j}^-\) are the current transmission and reflection probability amplitudes. The elements of the matrix \(S\) must be such that the current passing into the system via propagating modes is equal to the current passing out of the system via propagating modes.

We now connect the quasi-one-dimensional leads to perfect Ohmic contacts with, as before, chemical potential \(\mu_L = \mu - eV\) on the left and chemical potential \(\mu_R = \mu\) on the right. On a microscopic level, the currents flowing into each of the subbands of the quasi-one-dimensional systems from the Ohmic contacts are random owing to the discreteness of electronic charge. In a simple transport measurement we detect the time averaged current. The average current flowing to the right in Fig. 5.20 in a small energy range \(d\epsilon\) about energy \(\epsilon\) is

$$
dI^+ = \sum_i \langle |b_i^+|^2 \rangle_{\text{time}} = \langle (b^+) (b^+) \rangle_{\text{time}} = \langle a^+ t^+ t^+ a^+ \rangle_{\text{time}}
$$

$$
= \sum_{i,j} (t^+ t^+)^{i,j} \langle a_i^+ a_j^+ \rangle_{\text{time}}
$$

(5.35)

The time average \(\langle a_i^+ a_j^+ \rangle_{\text{time}}\) simplifies because the amplitudes \(a_i^+, a_j^+\) for \(i \neq j\) for are uncorrelated in time so that

$$
\langle a_i^+ a_j^+ \rangle_{\text{time}} = \delta_{i,j} \langle |a_i^+|^2 \rangle_{\text{time}}
$$

(5.36)

The current entering each subband \(\langle |a_i^+|^2 \rangle_{\text{time}}\) will be exactly the same as was derived for the perfect case above

$$
\langle |a_i^+|^2 \rangle_{\text{time}} = -\frac{e}{\hbar} f(\epsilon + eV) d\epsilon
$$

(5.37)

Unless it is an evanescent mode, for which the current will be zero. Substituting Eqn. 5.37 and Eqn. 5.36 into Eqn. 5.35 we have

$$
dI^+ = -\frac{e}{\hbar} \sum_i \langle t^+ t^+ \rangle_{\text{Occ}} f(\epsilon + eV) d\epsilon = -\frac{e}{\hbar} \text{Trace} \left[ t^+ t^+ \right] f(\epsilon + eV) d\epsilon
$$

(5.38)
5.11 The saddle point potential

where the trace is over occupied states only. By a similar argument we derive

\[ dI^- = \frac{e}{\hbar} \text{Trace} \left[ t^{+\dagger}t^+ \right] f(\epsilon) d\epsilon \]  

(5.39)

Therefore the total current will be

\[ I = \frac{e}{\hbar} \int_{-\infty}^{\infty} \text{Trace} \left[ t^{+\dagger}t^+ \right] \left( f(\epsilon + eV) - f(\epsilon) \right) d\epsilon \]  

(5.40)

The differential conductance \( G \) then has the form

\[ G = \frac{dI}{dV} = \frac{e^2}{\hbar} \int \text{Trace} \left[ t^{+\dagger}t^+ \right] \left( -\frac{df}{d\epsilon} \right) d\epsilon \]  

(5.41)

In the limit of zero temperature this becomes

\[ G = \frac{e^2}{\hbar} \text{Trace} \left[ t^{+\dagger}t^+ \right] \]  

(5.42)

where the expression is evaluated at the chemical potential.

As we have derived them, Eqns. 5.41,5.42 give the conductance for an arbitrary effective potential sandwiched between two semi-infinite quasi-one-dimensional systems attached to perfect Ohmic contacts. One might reasonably ask what bearing this has on the experimental system described above. The answer is that in calculating the conductance of an electronic device using Eqns. 5.41,5.42 the connecting wires should be made as wide as possible and be joined smoothly to the arbitrary region so that any effect due to their resistance or multiple scattering at the join is minimised.

5.11 The saddle point potential

We used the result of the calculation on the conductance of a perfect quasi-one-dimensional system to understand ballistic quantisation in an experimental split-gate device. However since the constrictions in these experiments are electrostatically induced by a pair of split gates, the potential is a smoothly varying function of position and therefore we must address the question of whether Eqn. 5.30 remains valid under these circumstances. We follow a paper by Büttiker [10] in describing this situation.

The bottle neck of a constriction forms a saddle point Fig. 5.21 which may be expanded to second order as

\[ V(x, y) = V_0 - \frac{1}{2} m^* \omega_x^2 x^2 + \frac{1}{2} m^* \omega_y^2 y^2 \]  

(5.43)
Quantum transport in one dimension.

Fig. 5.21. Schematic diagram of a saddle point potential.

Here $V_0$ is the electrostatic potential in the middle of the saddle point, and the curvatures of the potential are expressed in terms of the harmonic oscillator energies $E_x = \hbar \omega_x$, $E_y = \hbar \omega_y$. The total energy is given by the potential in equation 5.43 supplemented by the kinetic energy $p^2/2m^*$. The Hamiltonian is separable into a transverse wave function associated with energies $(n + 1/2)E_y$, $n = 0,1,2,3...$ and a wave function for motion along the $x$-direction in an effective potential $V_0 + (n + 1/2)E_y - m^*\omega_x^2x^2/2$. This effective potential can be viewed as the band bottom of the $n^{th}$ one-dimensional subband in the region of the saddle point. In the absence of quantum tunnelling the subbands with threshold energy

$$E_n = V_0 + (n + \frac{1}{2})E_y$$

(5.44)

below the Fermi energy are open (fully transmitted) and the subbands with threshold energy $E_n$ above the Fermi energy are closed (perfectly reflected). Quantum mechanically transmission and reflection at the saddle allows for subbands which are neither completely open nor completely closed, but which permit transmission with probability $T_{n,m} = |t_{n,m}|^2$. Here the index $n$ refers to the incident subband, and $m$ refers to the outgoing subband. The transmission probabilities for this case are calculated in references [11] and can be expressed with the help of the variable

$$\epsilon_n = 2 \left[ E - E_y \left( n + \frac{1}{2} \right) - V_0 \right] / E_x$$

(5.45)
5.11 The saddle point potential

Fig. 5.22. Single subband transmission probabilities $T_{n,n}$ and the total transmission probability (conductance) $T$ as a function of $(E - V_0)/E_x$ for the case where $E_y/E_x = 3$. The opening of successive one-dimensional subbands over narrow energy intervals leads to the quantisation of the conductance.

in the simple form

$$T_{n,m} = \delta_{n,m} \frac{1}{1 + e^{-\pi \epsilon_n}}$$  \hspace{1cm} (5.46)

Only transmission probabilities for which the incident subband and the outgoing subband are the same are non-zero. Because of the quadratic form of the saddle-point there is no subband mixing. The transmission probabilities $T_{n,n}$ for $n = 0,1,2,3,\ldots$, for the case $E_y/E_x = 3$ are shown in Fig. 5.22 as a function of $(E - V_0)/E_x$.

For $\epsilon_n \ll 0$ the transmission probability is exponentially small: $T_{n,n} \rightarrow \exp(-\pi \epsilon_n)$. For $\epsilon_n \gg 0$ the transmission probability is close to one: $T_{n,n} \rightarrow 1 - \exp(-\pi \epsilon_n)$. The transition from zero transmission to a transmission probability close to unity occurs in the neighbourhood of the classical threshold energy $E_n$ in Eqn. 5.44. The size of the energy interval needed for the
Quantum transport in one dimension.

Fig. 5.23. Constriction conductance as a function of \((E-V_0)/E_x\) for differing saddle point potentials characterised by a ratio \(E_y/E_x\) for ratios in an interval from 0 to 5 in increments of 0.25.

The transition is determined by \(E_x\). The conductance from Eqn. 5.42 is

\[ G = \frac{e^2}{\hbar} \sum_n T_{n,n} = \frac{e^2}{\hbar} T \]  

The total transmission probability \(T\) (or the two-terminal conductance) therefore shows a series of well developed steps if the transition region for the opening of a one-dimensional subband is small compared to the subband separation. Since the width of the transition region is \(E_x\) and the separation of subbands at the saddle point is determined by \(E_y\) well-pronounced steps occur if

\[ E_y \geq E_x \]  

Fig. 5.23 shows a series of conductance traces with ratios of \(E_y/E_x\) increased in increments of 0.25 in the interval 0 to 5. It is seen that already for ratios which are moderately larger than unity the conductance shows considerable structure, and if the ratio approached the maximum value shown the plateaux are entirely flat.

5.12 Determination of saddle point potential shape

The analysis in the previous section explains the broadening of the transition regions between the ballistic plateaux of a split gate device in terms of the
5.12 Determination of saddle point potential shape

ratio $E_y/E_x$ and it is natural to ask how the curves in Fig. 5.23 can be matched with those in Fig. 5.19 to determine this ratio for an experimental device. This is typically done by assuming that in an experimental split gate device $V_0$ depends linearly on the gate voltage $V_g$ and that both $E_x$ and $E_y$ are constant over a small range of transitions. The ratio $E_y/E_x$ is therefore found by scaling the gate voltage axis of an experimental conductance trace linearly until the plateau and transition widths can be matched with a trace from Fig. 5.23.

It is not possible to determine $E_x, E_y$ individually from differential conductance traces such as those in Fig.5.19. However, if in addition to an a.c excitation a d.c bias is applied across the device the Fermi surfaces on each side of it can be used as probes of the subband energies. Fig. 5.24 shows a cross-section of the potential in a split gate device taken at the centre of the constriction in the $x$ - direction with no d.c bias applied.

The numbered lines drawn above the potential edge are the one-dimensional subband minima. The chemical potential is constant across the device and no current is flowing.

When a d.c bias is applied across the device capacitive coupling between the electrons in the quasi-one-dimensional system and the surface gate will cause the electron density at each point in the device to change from its zero d.c bias/equilibrium value. However, this coupling is typically so small that even if a d.c bias of several milli volts were applied the density at any point would only be different from its equilibrium value by 0.1%. The application of the d.c bias therefore distorts the effective potential in such a way as to keep the local density and therefore local Fermi energy constant ($E_F(x) \propto n_{2D}(x)$). Fig.5.25 is a schematic representation of this where for
Quantum transport in one dimension.

\[ \mu_L = \mu_0 - eV/2 \]
\[ \mu_R = \mu_0 + eV/2 \]

Fig. 5.25. Distortion of split gate effective potential under the application of a d.c. bias. \( \mu_L = \mu_0 - eV/2 \) chemical potential on left hand side of device. \( \mu_R = \mu_0 + eV/2 \) chemical potential on right hand side of device. \( \mu_0 \) equilibrium chemical potential, kept constant as an energy reference.

\[ \mu_L = \mu_0 - eV/2 \]
\[ \mu_R = \mu_0 + eV/2 \]

Fig. 5.26. (a) Subband minima and left and right chemical potentials for the case where the applied d.c bias is smaller than the subband spacing. (b) case where the d.c. bias is larger than one subband spacing.

Convenience we have chosen to make our energy reference point the equilibrium chemical potential \( \mu_0 \) so that for a potential difference \( V \) across the device the chemical potential on the right is \( \mu_R = \mu_0 + eV/2 \) and the chemical potential on the left is \( \mu_L = \mu_0 - eV/2 \).

By using Fig. 5.25 as our model for the application of a d.c bias across a split gate device we may think of the subband energies in the centre of the saddle point as being unaffected by the application of the d.c bias but capable of being moved up an down by the application of different split gate voltages.

If, as in Fig.5.26a, both the left and right chemical potentials lie between the subbands \( N \) and \( N + 1 \) then at zero temperature we may write the total current as

\[ I = I^+ - I^- = -N e^2 h/2 \mu_L + N e^2 h/2 \mu_R = N e^2 h/2 2 eV = N e^2 h V \] \hspace{1cm} (5.49)
5.12 Determination of saddle point potential shape

Fig. 5.27. Transconductance of a split gate device. The bottom trace corresponds to zero d.c bias and successive traces which are vertically off-set are different in applied d.c bias by 0.08mV.

so that the differential conductance is

\[ G = \frac{dI}{dV} = Ne^2/h \quad (5.50) \]

If, as in Fig. 5.26b, the left chemical potential lies between the subbands \( N \) and \( N + 1 \) and the right chemical potential lies between the subbands \( N + 1 \) and \( N + 2 \) then at zero temperature we may write the total current as

\[ I = I^+ - I^- = -N e/h \mu_L + (N + 1) e/h \mu_R = (N + 1/2) e^2/h V + e/h \mu_0 \quad (5.51) \]

so that

\[ G = \frac{dI}{dV} = \left( N + \frac{1}{2} \right) \frac{e^2}{h} \quad (5.52) \]

In general we find that whenever a subband minimum at the centre of the saddle point passes either the left or the right chemical potential the differential conductance changes by an amount \( \Delta G = e^2/2h \). This change in \( G \) implies a peak in the transconductance \( dG/dVg \).

Fig. 5.27 shows a series of transconductance traces as a function of split gate voltage for a split gate from reference [9].
Quantum transport in one dimension.

Fig. 5.28. (a) Dark lines - transconductance of a split gate. Numbers - two terminal conductance. (b) Differential conductance within the saddle point model.

In Fig. 5.27 at zero d.c bias as the split gate voltage is made more negative successive subbands are depopulated at the centre of the saddle point. Each time this happens there is a peak in the transconductance. As the d.c bias increases each of these peaks slits into two. The peak at less negative gate voltage corresponds to the subband minimum passing across the left chemical potential and the peak at more negative voltage corresponds to the subband minimum passing across the right chemical potential. The trajectories of these peaks in the gate voltage $V_g$ vs. d.c bias $V$ plane are straight lines: if we change the gate voltage and move a particular subband minimum by an energy $E$ then in order to keep it aligned with a particular chemical potential we must change that chemical potential by the same energy. Fig. 5.28a is a schematic diagram of the transconductance peaks of a split gate device in the $V_g$ vs $V$ plane. Numbers indicate the value the differential conductance will have, in units of $e^2/h$, in each diamond shaped region defined by the peaks in the transconductance. Fig. 5.28b shows the differential conductance calculated within the saddle point model at a series of different d.c. biases.

In Fig. 5.28a lines with positive gradients correspond to subband minima aligning with the right chemical potential and lines with negative gradient correspond to subband minima aligning with the left chemical potential.
Grey circles mark the points where transconductance peaks cross. They appear at set of gate voltages where at zero d.c bias two subband minima are equally close to the chemical potential. The d.c bias at which they occur is equal to the spacing in energy between these two subbands. Therefore by accurately measuring the d.c bias at which these crossings occur the corresponding subband spacings can be obtained. Other crossings of transconductance peaks allow the determination of other subband spacings. However it is clear from Fig. 5.25 that the effective potential in the constriction becomes increasingly distorted by the d.c bias and therefore the energy differences measured have no bearing on the zero d.c bias effective potential.

Fig. 5.28 has deliberately been drawn to indicate the experimental observation that the measured subband spacings decrease with increasing gate voltage. That is, if the constriction were a true parabolic saddle point potential then $E_y$ would increase with increasingly negative gate voltage.

5.13 Exercises

5.1 A GaAs-AlGaAs heterostructure has a surface, a split-gate of width 700nm and length 700nm patterned on its surface and a two-dimensional electron gas of density carrier density $1.0 \times 10^{15} \text{m}^{-2}$ 250nm below its surface.

(a) Calculate, assuming a simple capacitance model, the split-gate voltage at which a one-dimensional channel would be defined by the surface gates. Estimate the voltage at which the channel would pinch off.

(b) Describe an experiment that would allow a determination of the dependence of the one-dimensional subband energies on split-gate voltage. Describe and account for the structure that would be seen.

(c) The two-terminal conductance of the first plateau has a value $G = 1.95e^2/h$. Assuming that a series resistance is responsible for the deficit, what is the value of this resistance? Why would you expect a series resistance?

(d) Why are the transition regions between conductance plateaux broadened.
Quantum transport in one dimension.

(e) The curvature of the split-gate potential causes some inter-subband scattering. Will this necessarily destroy the ballistic quantisation?

(f) A d.c bias experiment is performed and it is found that peaks in the transconductance cross at the following points

\[ n=1, 2: V_g = -5.0V, V_{dc} = 2.5mV \]
\[ n=4, 5: V_g = -3.0V, V_{dc} = 1.0mV \]
\[ n=10, 11: V_g = -1.0V, V_{dc} = 0.5mV \]

- amongst other points. Calculate the corresponding subband energies assuming that some of the applied d.c bias is dropped across the series resistance (use the series resistance calculated in (c)). Account for the trend you find in the subband energies with decreasing gate voltage. Approximately what would the width of the channel be at the chemical potential at the three gate voltages given assuming that the potential is (i) parabolic (ii) square?

5.2 Why is there a singularity in the quasi-one-dimensional density of states? For finite size systems, is the one-dimensional density of states ever higher than the two-dimensional density of states? Describe the principle behind an experiments to observe the Fermi energy modulation and the density of states modulation and account for the fact that singular behaviour is likely to be strongly damped.

Assume \( m^* = 0.067m_e \) and \( \epsilon = 13.0 \).
General quantum transport theory

6.1 Sources

6.2 Introduction
This section of the notes extends the Landauer quantum transport theory to the general case for finite perpendicular field with multiple Ohmic contacts. This will allow us to make a comprehensive study of the quantum Hall and Shubnikov de Haas effects as measured on a standard Hall bar. The notes cover: the quasi-one-dimensional system in finite perpendicular magnetic field, its eigen states, density of states, effective potential, and two-terminal conductance; the saddle point potential; the multi-probe Landauer-Büttiker formalism; edge states, disorder and the quantum Hall and Shubnikov de Haas effects.

6.3 Eigenstates of an infinite quasi-one-dimensional system
Schrödinger’s equation for a quasi-one-dimensional electron system with a transverse confining potential $V(y)$ in a perpendicular magnetic field $B = B\hat{z}$ has the form:

$$\left(\frac{1}{2m} (p + eA)^2 + V(y)\right) \Psi = E \Psi$$

(6.1)
Although the physical system is translationally invariant in the $x-$direction, Eqn. 6.1 is only translationally invariant in the $x-$direction if we make a suitable gauge choice for $A$. Here we use the so called Landau gauge:

$$A = -By\hat{x} \tag{6.2}$$

Substituting Eqn. 6.2 into Eqn. 6.1 and expanding the dot product we have

$$\left(\frac{p_y^2}{2m^*} + \frac{1}{2m^*}(p_x - eBy)^2 + V(y)\right)\Psi = E\Psi \tag{6.3}$$

The translational invariance of the system implies that the solutions to Eqn. 6.3 have the form

$$\Psi_{k_x,n} = \exp(i k_x x) \phi_{k_x,n}(y) \tag{6.4}$$

Substituting Eqn. 6.4 into Eqn. 6.3 we find

$$\left(\frac{p_y^2}{2m^*} + \frac{1}{2}m^*\omega_c^2 \left( y - \frac{\hbar k_x}{eB} \right)^2 + V(y) \right) \phi_{k_x,n} = E_{k_x,n}\phi_{k_x,n} \tag{6.5}$$

Hence, in this gauge, the effect of the magnetic field is simply to add a parabolic potential

$$V_B(y) = \frac{1}{2}m^*\omega_c^2 (y - y_0)^2 \tag{6.6}$$

where

$$y_0 = \frac{\hbar k_x}{eB} \tag{6.7}$$

to the effective potential $V(y)$. This parabolic potential represents the effect of the Lorentz force. $y_0$ is the classical centre of motion for electrons passing along the quasi-one-dimensional system. It is referred to as the guiding centre.

If we find the solutions to equation Eqn. 6.5 on the assumption that the effective potential is parabolic

$$V(y) = \frac{1}{2}m^*\omega_0^2 y_0^2 \tag{6.8}$$

they will give us sufficient insight to understand the nature of the solutions for an arbitrary confining potential. The sum of Eqn. 6.6 and Eqn. 6.8 in Eqn. 6.5 results in an confining potential that is also parabolic. Equation 6.5 can therefore be written in the form

$$\left(\frac{p_y^2}{2m^*} + \frac{1}{2}m^*\omega_c^2 b^2 \left( y - \frac{1}{b^2}y_0 \right)^2 + \frac{1}{2}m^*\omega_0^2 \frac{y_0^2}{b^2} \right) \phi_{k_x,n} = E_{k_x,n}\phi_{k_x,n} \tag{6.9}$$
6.3 Eigenstates of an infinite quasi-one-dimensional system

where

\[ b^2 = 1 + \frac{\omega_0^2}{\omega_c^2} \]  \hspace{1cm} (6.10)

The solutions to Eqn. 6.9 are the same as those for Eqn. 5.3 with confining potential Eqn. 5.4 but with an origin offset by \( y_0/b^2 \)

\[ \phi_{y_0,n}(y) = e^{-(y-y_0/b^2)^2/2l_B^2} H_n \left( \frac{(y-y_0/b^2)}{\sqrt{2}l_B} \right) \]  \hspace{1cm} (6.11)

where

\[ l_B = \left( \frac{m^*\omega_c b}{\hbar} \right)^{-\frac{1}{2}} \]  \hspace{1cm} (6.12)

The eigenvalues have the form

\[ E_{y_0,n} = \left( n + \frac{1}{2} \right) \frac{\hbar \omega_c b}{2} + \frac{1}{2} m^* \frac{\omega_0^2}{b^2} y_0^2 \]  \hspace{1cm} (6.13)

An example of the wave functions given by Eqn. 6.11 with energies given by Eqn. 6.13 is shown in Fig. 6.1 for a given value of \( y_0 \). The wave functions in Eqn. 6.11 are localised around classical cyclotron orbits that have radii \( R_n = l_B \sqrt{2n+1} \) and are centred on position \( y = y_0 \). States that carry momentum in opposite directions are forced close to opposite walls of the quasi-one-dimensional system even if they are at the same energy and have the same Landau level index. The larger the momentum the further out from the centre \( y = 0 \) the states sit. If \( V(y) \) is a broad parabolic confinement

---

Fig. 6.1. Schematic diagram showing effective potential \( V(y) \) and magnetic effective potential \( V_B(y) \) at an arbitrary position \( y = y_0 \) together with the wave functions and eigenenergies at this position.
such that $\omega_c \gg \omega_0$ then $b \to 1$ and Eqn. 6.13 takes the form

$$E_{y_0,n} = \left( n + \frac{1}{2} \right) \hbar \omega_c + V(y_0)$$  \hspace{1cm} (6.14)$$

This equation is the quasi-one-dimensional equivalent of Eqn.4.25. Since at high field $l_B$ is sufficiently small that at any position $y_0$ an arbitrary effective potential may be approximated locally by a parabola the high field eigenstates of an arbitrary effective potential will also be given by Eqn. 6.14.

Fig. 6.2 shows the subband energies $E_{y_0,n}$ of a quasi-one-dimensional system with a square-well effective potential $V(y)$, of width $W = 150\text{nm}$, as a function of magnetic field. It can be seen that all subband energies increase with increasing magnetic field and that they fall onto the Landau levels $E_{0,n} \to (n + 1/2) \hbar \omega_c$ when their respective cyclotron orbits are smaller in diameter than the width of the system.

Figs. 6.3a,b,c,d are the dispersion curves for this system at four different magnetic fields. At zero magnetic field, Fig. 6.3a, the dispersion is parabolic and subband energies are proportional to the square of the subband index.
6.3 Eigenstates of an infinite quasi-one-dimensional system

At $B = 0.3$, Fig. 6.3b, the effect of the magnetic field is noticeable in the spacing between the first and second subbands. At this field the cyclotron energy is comparable to the zero field energy of the first subband. At $B = 1.0$T the cyclotron diameters for the first two subbands are much smaller than the width of the system. For guiding centre positions where the cyclotron orbits do not intersect with the potential boundaries ($y = \pm 75$nm) the subband energies are dispersionless with energies equal to the cyclotron energies. For guiding centre positions where the cyclotron orbit does intersect the potential boundaries the energies are no longer degenerate with the cyclotron energies and rise in energy the more tightly they are pushed against the boundaries. Fig. 6.4a shows the wave function of the lowest subband, in this system, for propagation in the positive direction, at an energy of 10meV, for a series of different magnetic fields. Fig. 6.4b shows the corresponding wave functions for propagation in the negative direction.

Fig. 6.3. Dispersion curves for a 150nm wide square well at (a) 0.0T (b) 0.3T (c) 1.0 T (d) 2.0T.
Fig. 6.4. $n=0$ wave function at 10meV in a 150nm square well potential five magnetic fields (a) negative momentum (b) positive momentum.

These two sequences show how the Lorentz force acts to push wave functions against the walls of the system with increasing magnetic field.

### 6.4 Group Velocity

The group velocity of a state with wave vector $k_x$ in the $n$th Landau level is given by the gradient of the dispersion curve

$$v_{k_x,n} = \frac{1}{\hbar} \frac{\partial E_{k_x,n}}{\partial k_x}$$

(6.15)

If we consider Fig. 6.3d in light of this expression we see that states with a guiding centre position such that their cyclotron orbit does not touch the walls of the wire have zero group velocity and are therefore stationary cyclotron orbits. Those states with orbits that touch the walls have a group velocity that increases the closer they are forced to the walls. This reflects the fact that electrons can only pass through the system by skipping along the edges of the wire these skipping orbits were discussed in notes 3. The density of states therefore naturally splits into two parts: bulk states that
are stationary cyclotron orbits and edge states that skip along the walls of the wire.

6.5 Density of states

The density of states in each subband is inversely proportional to the group velocity at any magnetic field. If we consider a finite wire with periodic boundary conditions we may write

\[
dE_n = \frac{\partial E_n}{\partial k_x} dk_x = \frac{\partial E_n}{\partial k_x} \pi dn
\]

so that

\[
\frac{dn_n}{dE} = \frac{2}{2\pi \frac{\partial E_n}{\partial k_x}} = \frac{2}{\hbar v_n}
\]

(6.16)

Figs. 6.5a,b show the density of states of the quantum wire at zero magnetic field and at one Tesla taken from Fig. 6.3a,c. They both have the same basic shape but it can be seen that at high field, states are more localised in energy around the subband energies.

Figure 6.6 shows data from the capacitance experiment by Drexler (notes 5) performed at series of perpendicular magnetic fields between zero and seven Tesla. The traces at high field show the tendency towards greater degeneracy at subband energies. The experiment averages over 300 wires each containing different disorder potentials. This averaging blurs the sharp peaks expected in the density of states.

6.6 Effective potential at high magnetic field [1]

Fig. 6.7a shows a smoothedged effective potential with occupied states marked by black dots of width \(l_B\). Fig. 6.7b shows the corresponding distribution of charge density from these occupied states. Where only the first Landau level is occupied the density is \(n(x) = eB/h\) and where the first and second Landau levels are occupied the density is \(n(x) = 2eB/h\). Fig. 6.7c shows a step-like effective potential edge and below it in Fig. 6.7d the corresponding density distribution. The two situations in Figs. 6.7a,b and Figs. 6.7c,d show that owing to the nature of the eigenstates in a magnetic field a device could have a smooth effective potential shape with a stepped density distribution or a stepped potential shape with a smooth density distribution. The physical choice for an experimental system is the latter since the density distribution is determined by a smooth distribution of background ionised donors and remote surface gate potentials and is almost independent
of magnetic field. The discontinuities in the density distribution in Fig. 6.7d arise from the fact that it is impossible to achieve the minimum energy density distribution when $l_B$ is too small. The density distribution in Fig. 6.7d is therefore that which is closest to the minimum energy density distribution given the constraints that the eigenspectrum imposes.

### 6.7 Two-terminal conductance in a magnetic field

A quasi-one-dimensional system in a perpendicular magnetic field has two important features in common with the zero field case. (1) Irrespective of whether a magnetic field is present or not the product of the group velocity $v_n$ and the density of states $n_n$ for each one-dimensional subband is a constant, independent of both energy and subband index

$$ n_n v_n = \frac{1}{\hbar} $$

(6.18)

(2) Since all subband wave functions are eigenstates of the current opera-
Fig. 6.6. Capacitance of the wire array recorded as a function of surface–gate voltage $V_g$ at different magnetic fields. The magnetic field increases in steps of 1T from the top to the bottom trace. The scale refers to the top trace ($B=0$T) and a constant offset is subtracted from the traces at finite magnetic field for clarity. The values $n=1,2,3,4$ indicate the subbands that start to be filled in the corresponding gate voltage regimes.

The only difference being that at zero magnetic field the transverse wave functions are independent of both energy and direction of propagation. As we have...
Seen this is not the case for finite magnetic field since states propagating in different directions tend to be pushed towards opposite walls of the potential. The two-terminal conductance of an arbitrary effective potential in a magnetic field and zero temperature may therefore be written

\[ G = \frac{e^2}{h} \text{Trace} (t^+ t^+) \]  

(6.19)

Whilst the form of Eqn. 6.19 is independent of magnetic field, the transmission coefficients are not. This is because they depend upon the functional forms of the wave functions and eigenenergies.

**6.8 The saddle point in a finite magnetic field [2]**

We have shown that the effective potential in a quasi-one-dimensional system changes as a function of magnetic field — acquiring sharp steps in the high field limit. However, if we consider a split-gate device at magnetic fields away from this limit we can approximate its effective potential by a parabolic
The saddle point in a finite magnetic field [2]

saddle point - just as we did for the zero field case. In the Landau gauge, Schrödinger’s equation for the saddle point potential may be written

\[ \left( \frac{p_y^2}{2m^*} + \frac{1}{2m^*} (p_y - eBy)^2 + V_0 - \frac{1}{2} m^* \omega_c^2 x^2 + \frac{1}{2} m^* \omega_y^2 y^2 \right) \Psi = E \Psi \] (6.20)

Unlike the zero field case, this equation has no simple separation of variables. However we can gain an idea of what its transmission and reflection coefficients will be like if we consider the potential as a series of narrow sections in which \( V(x) \) is constant – Fig. 6.8.

Schrödinger’s equation for the section centred on the point \( x \) has the form

\[ \left( \frac{p_y^2}{2m^*} + \frac{1}{2} m^* \omega_c^2 \left( y - \frac{h k_x}{eB} \right)^2 + \frac{1}{2} m^* \omega_y^2 \right) \phi_{k_x,n} = E' \phi_{k_x,n} = \left( E - V_0 + \frac{1}{2} m^* \omega_c^2 x^2 \right) \phi_{k_x,n} \] (6.21)

where we have made the substitution

\[ \Psi_{k_x,n} = \exp(i k_x x) \phi_{k_x,n}(y) \] (6.22)

Equation Eqn. 6.21 has the same form as equation Eqn. 6.9 and therefore has wavefunctions of the form Eqn. 6.11 and eigenenergies of the form Eqn. 6.13. These solutions consists of a set of independent subbands through which current flows without scattering. If we now connect together two sections and ask what happens at their boundary then it is clear that if they join sufficiently smoothly the join will not cause any scattering between subbands. Thus if the whole saddle-point potential is sufficiently smooth there will be no scattering between subbands and the transmission matrix will again be diagonal, as it was for the zero field case – Eqn.5.46. This is the case for the parabolic-saddle-point potential. Using the definition...
Fig. 6.9. Conductance of a saddle-point potential, with $\omega_y/\omega_x = 1$, as a function of $(E - V_0)/E_x$ at a series of different magnetic fields $\omega_c/\omega_x = 0, 0.25, \ldots 5$.

$\Omega^2 = \omega_c^2 + \omega_x^2 - \omega_y^2$ the energies $E_1, E_2$ that govern transmission and reflection at a saddle-point potential in a magnetic field are

\[ E_1 = \frac{1}{2} \frac{\hbar}{\sqrt{2}} \left( \left( \Omega^4 + 4\omega_x^2\omega_y^2 \right)^{1/2} - \Omega^2 \right)^{1/2} \]  

\[ E_2 = \frac{\hbar}{\sqrt{2}} \left( \left( \Omega^4 + 4\omega_x^2\omega_y^2 \right)^{1/2} + \Omega^2 \right)^{1/2} \]  

(6.23)  

(6.24)

The transmission probability is a function of the variable

\[ \epsilon_n = \frac{E - E_2 (n + 1/2) - V_0}{E_1} \]  

(6.25)

and is, as in the absence of a magnetic field, given by

\[ T_{n,m} = \delta_{n,m} \frac{1}{1 + e^{-\pi \epsilon}} \]  

(6.26)

Fig. 6.9 shows the conductance of a saddle-point potential with $\omega_y/\omega_x = 1$ at a series of magnetic fields. From the traces in this figure it is clear that the quantisation of the ballistic plateaux become increasingly accurate with
increasing magnetic field. The reason for this is that the magnetic field suppresses backscattering by physically separating modes that propagate in opposite directions. Fig. 6.10 shows ballistic quantisation in a magnetic field taken from experiment [3].
6.9 Magnetic Depopulation [3,4]

The experimental traces in Fig. 6.10 and theoretical traces in Fig. 6.9 show a remarkable similarity that conceals an important difference between them. In the theoretical traces, at a particular value of the parameter \((E - V_0)/E_x\), the number of transmitted subbands reduces with increasing magnetic field because the cyclotron energy increases and therefore at constant local Fermi energy \(E_{F0} = E - V_0\) in the centre of the constriction subbands must depopulate. However, in the experiment at a fixed gate voltage it is the distribution of density that is constant and not the local Fermi energy. Subbands depopulate because the degeneracy of each Landau level increases with increasing magnetic field. Figs. 6.11a,b give a detailed picture of how subbands depopulate in experiment [4].

It is interesting to note that the gate voltage at which the first subband in a split-gate potential populates is independent of magnetic field. The reason for this is that the gate voltage at which the first electrons arrive in the device, and therefore the subband energy, is independent of magnetic field to within small corrections.

6.10 Multi-probe Landauer-Büttiker formalism [5]

Ballistic quantisation in a split-gate device and the quantum Hall effect are intimately related. However, ballistic quantisation derives from a two-terminal measurement, where the same Ohmic contacts are used to inject current as are used to measure voltage, and the Hall effect derives from a four-probe measurement, in which current is passed between one pair of Ohmic contacts and voltage measured across another pair. The Shubnikov de Haas measurement is also a four-probe measurement. We have described these measurements in notes 4. In order to understand these measurements better, we will derive a multi-probe version of the Landauer formula for conductance.

Fig. 6.12 shows a device structure with a number of Ohmic contacts. If we consider the \(i\)th Ohmic contact then the total current \(I_i\) that this Ohmic contact injects into the bulk of the device is made from three parts:

1. Current injected directly:

\[
I_{i,\text{injected}} = -\frac{e}{h} N_i \mu_i \tag{6.27}
\]

where \(N_i\) is the number of occupied propagating modes passing out from Ohmic contact \(i\) and \(\mu_i\) is the chemical potential in lead \(i\).
Fig. 6.11. (a) Conductance of a split gate of lithographic width 450nm as a function of split gate voltage measured at different magnetic fields. (b) A plot of the split gate voltages at which subbands depopulate as a function of magnetic field. Lines are drawn as a guide to the eye. Note that the density in the wire was different in (a) and (b) and therefore they cannot be compared directly [4].

(2) Current which is reflected current back after passing into the bulk:

\[ I_{i,\text{reflected}} = \frac{e}{\hbar} R_{i,i} \mu_i \]  

(6.28)
The reflection probability $R_{i,i}$ has the form

$$R_{i,i} = \sum_{\nu_1,\nu_2=0}^{N_i} |r_{\nu_1,\nu_2}^{i,i}|^2$$  \hspace{1cm} (6.29)

where $r_{\nu_1,\nu_2}^{i,i}$ are the quantum-mechanical reflection amplitudes from subband $\nu_2$ in Ohmic contact $i$ to subband $\nu_1$ in Ohmic contact $i$. They obey the symmetry relation $R_{i,j}(B) = R_{j,i}(-B)$ under reversal of the external magnetic field $B$.

(3) Current which is transmitted from other contacts:

$$I_{i,\text{transmitted}} = -\frac{e}{h} \sum_{i \neq j} T_{i,j} \mu_j$$  \hspace{1cm} (6.30)

The transmission probabilities $T_{i,j}$ have the form

$$T_{i,j} = \sum_{\nu_1=1}^{N_i} \sum_{\nu_2=1}^{N_j} |t_{\nu_1,\nu_2}^{i,j}|^2$$  \hspace{1cm} (6.31)

where $t_{\nu_1,\nu_2}^{i,j}$ are the quantum mechanical transmission amplitudes from subband $\nu_2$ in Ohmic contact $j$ to subband $\nu_1$ in Ohmic contact $i$. They obey the symmetry relation $T_{i,j}(B) = T_{j,i}(-B)$ under reversal of the external magnetic field $B$.

The total current that Ohmic contact $i$ injects into the bulk is then

$$I_i = I_{i,\text{injected}} - I_{i,\text{reflected}} - I_{i,\text{transmitted}}$$  \hspace{1cm} (6.32)
6.11 Edge states and obstacles

\[ I_i = -\frac{e}{\hbar} [N_i - R_{ii}] \mu_i - \frac{e}{\hbar} \sum_{i \neq j} T_{i,j} \mu_j \] (6.33)

were \( V_i \) is the voltage on Ohmic contact \( i \).

### 6.11 Edge states and obstacles

The problem with equation Eqn. 6.33 is that in order to find the output potentials \( \mu_i \) and currents \( I_i \) for a given set of input potentials \( \mu'_i \) and currents \( I'_i \) one must evaluate all of the reflection and transmission coefficients for the system. If done exactly, this is a task that must be done numerically and can take hours of computational time. If one uses effective potentials calculated selfconsistently, such as that shown in Fig. 6.7c, then the time for this calculation must be added and it is easy to spend months calculating a single magnetic field sweep or gate voltage sweep for comparison with a particular experimental device. However, the existence of edge states simplifies the situation very considerably, offering a framework within which many high-field transport phenomena can be understood.

As we have seen, at high field, edge states can pass through a parabolic saddle point potential without scattering. It is only when forward and reverse travelling edge states are brought within a magnetic length of each other that backscattering occurs and the transmission coefficient is reduced. Figs. 6.13b,c,d show three examples of edge states negotiating the random potential shown in Fig. 6.13a. They are grey-scale representations of the wave function at the chemical potential. In Fig. 6.13b the two-terminal conductance was close to \( 3e^2/h \) and one can make out the three bumps of the wave function passing from one Ohmic contact to the other. In Fig. 6.13c,d the two-terminal conductances were close to \( 2e^2/h \) and \( e^2/h \) respectively. In general edge states behave like their classical counterparts - the skipping orbits. Edge states follow the equipotentials of an effective potential, however contorted the path, and bound states form where the paths close on themselves.

The picture of a random potential being carpeted by localised states that we used in notes 4 is now replaced by a picture in which complicated networks of edge states follow every contour in the random potential at every energy. The puddles and peaks are still filled/covered with/by localised states but now they are bound edge states. At the percolation threshold when the puddles link to form a continuous equipotential though the device a single edge state is transmitted. This single state then remains at the
edge of the device as states are filled beyond the percolation point. When $N$ Landau levels are filled there will be $N$ edge states passing from one end of the Hall bar to the other in addition to the localised states. This situation is shown schematically in Fig. 6.14.

6.12 The Quantum Hall Effect

Fig. 6.15 shows a standard Hall bar, which we imagine to be very large in comparison to the magnetic length. We consider the situation where the chemical potential is in the localised states between Landau levels and $N$ edge channels pass into and out of each Ohmic contact. The Hall resistance is found by passing a current $I$ between Ohmic contact $S$ and $D$ and measuring
6.12 The Quantum Hall Effect

Fig. 6.14. (a) Cross-section of a Hall bar showing edge states and bulk states (b) Corresponding wave function at chemical potential c.f Fig. 6.13

Fig. 6.15. Hall bar with $N$ edge states passing between each Ohmic contact and a current $I$ passing between $S$ and $D$.

the voltage difference between Ohmic contacts $A$ and $C$

$$R_{x,y} = \frac{V_A - V_C}{I}$$  (6.34)
The corresponding diagonal resistance is found from

$$R_{x,x} = \frac{V_A - V_B}{I} \quad (6.35)$$

Under these conditions equation Eqn. 6.33 gives five coupled linear equations

\[
\begin{align*}
I &= \frac{e^2}{h} NV_S - \frac{e^2}{h} NV_C \quad (6.36) \\
-I &= \frac{e^2}{h} NV_D - \frac{e^2}{h} NV_B \\
0 &= \frac{e^2}{h} NV_A - \frac{e^2}{h} NV_S \\
0 &= \frac{e^2}{h} NV_B - \frac{e^2}{h} NV_A \\
0 &= \frac{e^2}{h} NV_C - \frac{e^2}{h} NV_D \\
\end{align*}
\]

Equation 6.36 and Eqn. 6.37 describe current flowing out of $S$ and into $D$. Equations 6.38-6.40 describe voltage probes - for which no total current flows in or out. We can immediately evaluate the diagonal resistivity in Eqn. 6.35 since from Eqn. 6.39 $V_A = V_B$ and therefore

$$R_{x,x} = 0 \quad (6.41)$$

Using Eqn. 6.38 we find $V_A = V_S$ substituting this into Eqn. 6.36 we find

$$I = N \frac{e^2}{h} (V_A - V_C) \quad (6.42)$$

which gives

$$R_{x,y} = \frac{h}{Ne^2} \quad (6.43)$$

The experimental confirmation of the results in Eqn. 6.41 and Eqn. 6.43 have been shown in Fig.4.3.

The quantum Hall effect derived in equations Eqs. 6.36-6.43 is extremely robust. Fig. 6.16 is a schematic representation of an elastic scattering event.
Fig. 6.17. Hall bar with many Ohmic contacts along each edge. Note that interaction with an Ohmic contact can be thought of as an inelastic scattering event.

along the edge of a Hall bar. If all \( N \) edge states originate from the same Ohmic contact they will all be at the same potential. The scattering event simply redistributes current between states. The total current passing into the event must equal the total current passing out and therefore equations Eqns. 6.36-6.40 remain valid in the presence of such scattering events and we would measure the same Hall resistance Eqn. 6.43 and Shubnikov de Haas effect Eqn. 6.41.

Fig. 6.17 shows a Hall bar with many Ohmic contacts along each edge. If a current is passed between \( S \) and \( D \) then equations Eqns. 6.33 take the form

\[
I = N \frac{e^2}{h} V_S - N \frac{e^2}{h} V_{B_1} \quad (6.44)
\]
\[
0 = N \frac{e^2}{h} V_{A_1} - N \frac{e^2}{h} V_S \quad (6.45)
\]
\[
0 = N \frac{e^2}{h} V_{A_i} - N \frac{e^2}{h} V_{A_{i-1}} \quad (6.46)
\]
\[
-I = N \frac{e^2}{h} V_D - N \frac{e^2}{h} V_{A_n} \quad (6.47)
\]
\[
0 = N \frac{e^2}{h} V_{B_n} - N \frac{e^2}{h} V_D \quad (6.48)
\]
\[
0 = N \frac{e^2}{h} V_{B_i} - N \frac{e^2}{h} V_{B_{i+1}} \quad (6.49)
\]

From Eqns. 6.48-6.49 we find \( V_{B_i} = V_D \) and from Eqns. 6.45-6.46 we find
Fig. 6.18. Reflection of edge states from a split-gate potential in a Hall bar. $M$ edge states pass through the split gate and $N$ edge states are occupied in the bulk two-dimensional regions either side of the split gate.

$V_{A_i} = V_S$ therefore from Eqns. 6.44 we have

$$I = N \frac{e^2}{h} (V_{A_i} - V_{B_j}) \quad (6.50)$$

so that

$$R_{SD,ij} = \frac{h}{Ne^2} \quad (6.51)$$

which is independent of the contacts the voltage is measured between. Thus even if edge states encounter phase-randomising events in Ohmic contacts or elsewhere the value of the quantum Hall effect is unaltered.

The theory of the quantum Hall effect predicts that there should be as many edge states as there are occupied Landau levels. This prediction may be tested directly by reflecting edge states from a split gate using the Hall bar configuration shown in Fig. 6.18.

The bulk two-dimensional system has $N$ occupied Landau levels and the split gate reflects $N-M$ edge states so that, at its centre only $M$ Landau levels are occupied. A current is passed between Ohmic contacts 1,4 and voltages are measured between other contacts. Equations 6.33 give

$$I = N \frac{e^2}{h} V_1 - N \frac{e^2}{h} V_6 \quad (6.52)$$

$$0 = N \frac{e^2}{h} V_2 - N \frac{e^2}{h} V_1 \quad (6.53)$$
6.12 The Quantum Hall Effect

\[ 0 = N \frac{e^2}{h} V_3 - M \frac{e^2}{h} V_2 - (N - M) \frac{e^2}{h} V_5 \] (6.54)

\[ -I = N \frac{e^2}{h} V_4 - N \frac{e^2}{h} V_3 \] (6.55)

\[ 0 = N \frac{e^2}{h} V_5 - N \frac{e^2}{h} V_4 \] (6.56)

\[ 0 = N \frac{e^2}{h} V_6 - M \frac{e^2}{h} V_5 - (N - M) \frac{e^2}{h} V_2 \] (6.57)

These equations are simplified, by noting that Eqns. 6.53 and 6.56 imply that \( V_2 = V_1 \) and \( V_5 = V_4 \), and by choosing Ohmic contact 1 to be at zero potential. Thus we have

\[ I = -N \frac{e^2}{h} V_6 \] (6.58)

\[ 0 = N \frac{e^2}{h} V_3 - (N - M) \frac{e^2}{h} V_4 \] (6.59)

\[ -I = N \frac{e^2}{h} V_4 - N \frac{e^2}{h} V_3 \] (6.60)

\[ 0 = N \frac{e^2}{h} V_6 - M \frac{e^2}{h} V_4 \] (6.61)

which give resistivities of the form

\[ R_{14,26} = \frac{V_2 - V_6}{I} = \frac{h}{N e^2} \] (6.62)

\[ R_{14,25} = \frac{V_2 - V_5}{I} = \frac{h}{M e^2} \] (6.63)

\[ R_{14,23} = \frac{V_2 - V_3}{I} = \left( \frac{1}{M} - \frac{1}{N} \right) \frac{h}{e^2} \] (6.64)

The Fractional quantisation predicted by Eqn. 6.64, that demonstrates the successive reflection of edge states from a split gate, is readily observable experimentally [6] Fig. 6.19.

The use of split-gates also allows the detection of inter-edge state scattering directly. The Hall bar shown in Fig. 6.20 has a Hall resistance that is sensitive to the current scattering probability \( S \), between edge states originating from Ohmic contacts 4 and 1, along the upper edge of the Hall bar.

A current \( I \) is passed between Ohmic contacts 1,3 and a voltage measured between Ohmic contacts 2,4. Equations 6.33 have the form

\[ I = M \frac{e^2}{h} V_1 - M \frac{e^2}{h} V_4 \] (6.65)

\[ 0 = M \frac{e^2}{h} V_2 - (M - S) \frac{e^2}{h} V_1 - S \frac{e^2}{h} V_4 \] (6.66)
Fig. 6.19. Fractional quantisation in the integer quantum Hall regime of the four-terminal longitudinal conductance $R_{14,23}^{-1}$ of a split gate in a magnetic field of 1.4T at $T=0.6$K. The solid horizontal lines indicate the quantised plateaux predicted by Eqn. 6.64 with $N=5$ and $M=1,2,3,4$. The dashed lines give the location of the spin-split plateaux, which are not well resolved at this magnetic field [6].

![Image](image.png)

Fig. 6.20. Hall bar used for the detection of inter-edge state scattering.

Rearranging Eqn. 6.66 we have

$$M \frac{e^2}{h} V_1 = \frac{M^2}{M - S} \frac{e^2}{h} V_2 - \frac{M S}{M - S} \frac{e^2}{h} V_4$$  \hspace{1cm} (6.67)

so that

$$R_{13,24} = \frac{V_2 - V_1}{I} = \left( \frac{M - S}{M^2} \right) \frac{h}{e^2}$$  \hspace{1cm} (6.68)

Thus the Hall resistance is reduced by inter-edge-state scattering. In high mobility two-dimensional systems such experiments have shown that
electrons can travel in edge states without scattering for 1000µm or more.

6.13 Exercises

6.1 How does a perpendicular magnetic field affect the eigenstates of quasi-one-dimensional system? What is the distinction between edge states and bulk states? Why does the effective potential in a quantum wire acquire steps at high magnetic field?

6.2 Describe an experiment to determine the density of states modulation in a quasi-one-dimensional system at finite perpendicular magnetic field.

6.3 What are Landau levels and how does their existence account for the quantum Hall effect in two-dimensional systems?

6.4 What is the nature of states in a Landau level in a disordered two-dimensional system? Why is disorder necessary to observe the quantum Hall effect?

6.5 Why is Landauer’s formula still valid in a magnetic field?

6.6 Describe how a perpendicular magnetic field affects the shapes of steps in the two-terminal conductance of a saddle point potential.

6.7 How do edge states negotiate obstacles. Why is the quantum Hall effect insensitive to scattering between edge states in a standard Hall bar geometry. Describe an experiment to detect scattering between edge states.
7
Quasi-particles in two dimensions

7.1 Sources

7.2 Introduction
In these notes we introduce the quasi-particle concept. This will give some insight into why electron-electron interactions can frequently be ignored when considering the transport properties of low-dimensional systems. The topics covered are: electrical forces; Landau’s theory of quasi-particles;
quasi-particle decay; the temperature dependence of quasi-particle decay; the density-of-states dependence of quasi-particle decay; the formation of new quasi-particles under the influence of a quantizing magnetic field.

7.3 Electrical forces

Electrical forces are immensely strong. As pointed out by Feynman in his lecture notes [1], if two people stand at arms length from each other, and each has one percent more electrons than protons, then the repelling force between them is sufficient to lift a weight equal to that of the earth [1]! Electron-electron interactions in a two-dimensional electron system can also be very large. Consider a two-dimensional electron system in a GaAs/AlGaAs heterostructure with a carrier density of \( n_{2D} = 1.0 \times 10^{-15} \text{m}^{-2} \). On average electrons will be 320\text{Å} apart so that the Coulomb force between adjacent pairs of electrons will be \( F = \frac{e^2}{4\pi\epsilon r^2} = 2.0 \times 10^{-14} \text{N} \). This is quite a small force from the perspective of a human but from the perspective of an electron of effective mass \( m^* = 0.067 m_e \) it is a very large force. It is sufficient to cause two electrons to accelerate away from each other at approximately a billion-billion meters per second squared. In a GaAs-AlGaAs heterostructure it is the electrostatic attraction of the ionised donors that prevents electrons from evacuating the system. In this case they provide a force of attraction per electron that is an order of magnitude larger than the repulsion between electrons \( F = eE = e^2 n_{2D}/\epsilon = 2.2 \times 10^{-13} \text{N} \). The ionised donors are therefore able to force the strongly interacting electrons together to maintain the two-dimensional system. In this system, given the strength of the electron-electron interactions, one would expect the transport properties to be dominated by inter-particle interaction effects. It is interesting to consider what the experiments on ballistic electron transport would yield in this system or how the strong interaction might affect the observation of ballistic quantization or the quantum Hall effect. The answer to these questions is simple yet quite surprising. Every experiment we have described to date has been performed in two-dimensional systems with densities similar to the case described above and yet we have been able to interpret each new phenomenon as if there is no interaction between electrons at all! The strong interactions appear to change nothing. The reason for this was proposed by Landau in his theory of quasi-particles [2-3].
Fig. 7.1. Development of eigenstates of an electron system as the interaction between electrons is slowly switched on.

### 7.4 Landau’s theory of quasi-particles [2-4]

Consider an electron system in which the interaction between electrons is switched on very slowly such that every state of the non-interacting system develops smoothly and adiabatically into a state of the interacting electron system Fig. 7.1.

If the interacting electron system is still in its ground state after this process then it is referred to as a normal Fermi liquid. The two-dimensional electron system that forms at a heterojunction interface is an example of a normal Fermi liquid. There are many counter examples. Perhaps the best known are the superconductors in which the state achieved by adiabatically switching on the interaction between particles is an excited state that can lose energy by forming Cooper-pairs.

The ground state of a normal Fermi liquid will have states filled according to Fermi-Dirac statistics and therefore will have a well defined Fermi surface at low temperature. The ground state is an equilibrium state with zero momentum. If we consider an excited state of the non-interacting system that consists of the filled Fermi circle together with an additional particle of momentum $p$ (Fig. 7.2a) then, due to conservation of momentum in particle collisions, when the interaction is switched on we will generate an excited state of the interacting system also of momentum $p$ (Fig. 7.2b).
Fig. 7.2. Excited state of (a) a non-interacting system (b) an interacting system. The interacting system and the non-interacting system have exactly the same momentum due to conservation of momentum in particle collisions.

Fig. 7.3. R. D. Mattucks horse (a) and quasi-horse (b) [4].

This additional particle slowly perturbs the distribution of other particles in its vicinity as the interaction is increased. Once the interaction is completely switched on the additional particle will move around the system surrounded by the perturbation it produces. This particle, together with its perturbation, is called a quasi-particle and may be considered to be an independent entity that is a composite consisting of contributions from many particles. R. D. Mattuck in his consideration of quasi-particles [4] uses the idea of a horse galloping across a dusty plain to picture this concept. As the horse moves along it kicks up a dust cloud that follows it producing thereby a quasi-horse consisting of a horse and a dust cloud: Fig. 7.3a,b.

This analogy emphasizes the fact that the perturbation around the additional particle tends to screen the quasi-particles interaction with other quasi-electrons just as the dust cloud tends to obscure the horse. This screening happens because the perturbation tends to consist of the removal
of other particles from the additional particles vicinity leaving behind a positive background of ionised donors. This positive background will partially neutralise the charge seen by other quasi-electrons.

The problem with the quasi-horse analogy is that it does not emphasize the fact that electrons are indistinguishable. The cloud around the additional electron is also made from electrons and it is not correct think of a single bare electron entering the system creating a distortion in the particle distribution which then propagates together with the particle through the system. Within the quasi-horse analogy it would be better to think of a horse running into a field of horses, causing a great deal of confusion, and some time later another horse with the same momentum running out of the field. An analogy that emphasizes this aspect of quasi-particle nature is the Newtons Cradle analogy. Newtons cradle consists of a row of metal balls hanging on strings Fig. 7.4.

The first ball is injected into the row with some momentum $p$ a quasi-ball of momentum $p$ then passes through the row of balls mediated by a string of collisions and the final ball leaves the system with the same momentum. Depending upon the length of the row of balls several quasi-balls may be injected at one time and quasi-balls travelling in opposite directions can be made to collide and interact with each other.

Thus we have the answer to the question of why experiments see only single-particle behaviour. Although an electron system at a heterojunction interface is composed of very strongly interacting electrons if we add a particle with a particular momentum from a reservoir, as we do in transport measurements, conservation of momentum requires that a quasi-particle of the same momentum propagate across the system. Quasi-particles are weakly interacting due to the screening that their particle perturbation provides.
7.5 Quasi-particles

The Hamiltonian for an interacting many-particle system can be written in the form

\[ H = \sum_{i=1}^{N} \left( \frac{p_i + eA_i}{2m^*} + V_i(r_i) \right) + \sum_{i \neq j=1}^{N} V(r_i, r_j) \]  \tag{7.1}

The first term represents a set of independent particles of mass \( m \) each in their own effective potential \( V_i(r_i) \). The second term represents interactions between these particles. As we have seen, the interaction term for the two-dimensional electron system, is typically very large. For the case we considered in the section on electrical forces the interaction was the same order of magnitude as the Fermi energy for each pair interaction. Thus there is no sense in which we can think of the distribution of bare particles as being perturbed by the interaction term for this case. It must absolutely dominate the solutions to Schrödinger’s equation. However, if we recall the two-body problem, a change of co-ordinates to a centre of mass co-ordinate and a relative co-ordinate reduces the interacting problem to two independent single particle problems. Most many-particle problems are reducible in exactly the same way, but through more general co-ordinate transformations and gauge transformations:

\[
\begin{align*}
R_i &= R_i(r_1, ..., r_N, p_1, ..., p_N) \\
P_i &= P_i(r_1, ..., r_N, p_1, ..., p_N) \\
A'_i &= A_i + \nabla \phi_i
\end{align*}
\]  \tag{7.2}

These transformations are referred to as canonical transformations if the new position and momentum co-ordinates obey the same commutation relations as the original co-ordinates. Under these transformations the Hamiltonian can be re-written in the form

\[ H = \sum_{i=1}^{N} \left( \frac{P_i + eA_i}{2m^*} + V_i(R_i) \right) + \sum_{i \neq j=1}^{N} V(R_i, R_j, P_i, P_j) \]  \tag{7.3}

The first term represents a set of independent quasi-particles each with their own vector potential and electrostatic potential and the second term represents the interaction between these quasi-particles. If, as for the two-body problem, the transformation eliminates the interaction term then the Hamiltonian describes a set of independent non-interacting quasi-particles. If the interaction term has been minimised by the transformation then we have, in some sense, the best set of quasi-particles. As external conditions
Quasi-particles in two dimensions

Fig. 7.5. Decay of a quasi-electron through the creation of a particle hole pair.

change, e.g magnetic and electric fields, what constitutes the best set of quasi-particles will change.

7.6 Quasi-particle decay

As we have seen above, quasi-particles are only approximate eigenstates of a real interacting system if the interaction term is not zero and therefore an excited quasi-particle will decay through collisions with other quasi-particles. We can use this fact to probe the validity of Landauls Fermi liquid theory by making predictions about the temperature and density of states dependence of the relaxation rate due to quasi-particle decay (c.f notes 2).

Fig. 7.5 shows an example of a process in which a quasi-electron of energy \( \epsilon \) decays due to a collision with a quasi-electron of energy \( \epsilon_2 \).

The products of the decay are a quasi-electron of energy \( \epsilon'_1 \) and a quasi-electron quasi-hole pair with energies \( \epsilon_2 \) and \( \epsilon'_2 \). Higher order decay processes, in which several quasi-electron quasi-hole pairs are created, are possible but occur at a much smaller rate. The relaxation rate for quasi-electrons of energy \( \epsilon \) due to this type of particle decay may be written qualitatively as

\[
\frac{1}{\tau_{\epsilon}} \propto \int \delta(\epsilon+\epsilon_2-\epsilon'_1-\epsilon'_2) \left\{ \rho_2(\epsilon_2)f^0_2 d\epsilon_2 \right\} \left\{ \rho_1(\epsilon'_1)(1-f^0_1)d\epsilon'_1 \right\} \left\{ \rho_2(\epsilon'_2)(1-f^0_2)d\epsilon'_2 \right\}
\]

(7.4)

The delta function in Eqn. 7.4 is included to ensure conservation of energy in the collision. Other conservation rules and the detailed nature of the quasi-particle collision are ignored since they do not affect the result we are trying to derive. \( \rho(\epsilon) \) is the quasi-particle density of states and \( f^0(\epsilon) \) is the Fermi Dirac distribution function. The three terms in curly brackets
represent the availability of final states for the decay process.

\[
\begin{align*}
\rho_2(\epsilon_2) f_2^0 d\epsilon_2 &= \text{[average no. of quasi-electrons occupying } d\epsilon \text{ around } \epsilon_2] \\
\rho_1(\epsilon_1')(1 - f_1^0') d\epsilon_1' &= \text{[average no. of quasi-holes occupying } d\epsilon \text{ around } \epsilon_1] \\
\rho_2(\epsilon_2')(1 - f_2^0') d\epsilon_2' &= \text{[average no. of quasi-holes occupying } d\epsilon \text{ around } \epsilon_2']
\end{align*}
\]

(7.5)

The integral in Eqn. 7.4 has most of its weight within an energy range \(\Delta \epsilon = \pm \max[\epsilon, k_B T]\) of the chemical potential \(\mu\) and therefore may be simplified if we consider a density of states that varies slowly over this energy range. We may then write Eqn. 7.4 in the form

\[
\frac{1}{\tau_\epsilon} \propto \rho^3(\mu) \int \delta(\epsilon + \epsilon_2 - \epsilon_1' - \epsilon_2') f_2^0(1 - f_1^0')(1 - f_2^0') d\epsilon_2 d\epsilon_1' d\epsilon_2'
\]

(7.6)

Changing variables in Eqn. 7.6 to

\[
x_i = \frac{\epsilon_i - \mu}{k_B T}
\]

(7.7)

we have

\[
\frac{1}{\tau_\epsilon} \propto \rho^3(\mu) \frac{(k_B T)^3}{k_B T} \int \delta(x + x_2 - x_1' - x_2') f_2^0(1 - f_1^0')(1 - f_2^0') dx_2 dx_1' dx_2'
\]

\[
\propto \rho^3(\mu) (k_B T)^2 \pi^2 + x^2(1 - f_1^0(x))
\]

\[
\propto \rho^3(\pi k_B T)^2 + (\epsilon - \mu)^2 \frac{1 + \exp \left( \frac{\epsilon - \mu}{k_B T} \right)}
\]

(7.8)

The form of Eqn. 7.8 implies that at low temperature, quasi-particles with energies close to the chemical potential will propagate through an interacting electron system over long distances without any collision affecting their motion. The fundamental reason for this is that at low temperature there are no available states for the products of a decay process. At high temperature or at energies away from the chemical potential, quasi-particles will have shorter life-times and the concept of a stable quasi-particle begins to break down.

7.7 Resistivity

The decay of quasi-particles gives rise to a resistivity that adds to that due to impurity scattering and phonon scattering. We can derive its approximate dependence on density of states and temperature using Eqn. 7.8. The
Fig. 7.6. Drag experiment. Drive and drag voltages have opposite sign.

Resistivity is proportional to the average scattering rate

$$\rho_{xx}^{\text{QP}} \propto \langle \frac{1}{\tau} \rangle = \int \frac{1}{\tau} \left( -\frac{\partial f^0}{\partial \epsilon} \right) \rho(\epsilon) d\epsilon$$

$$\propto \rho^4(\mu)(k_B T)^2 \int \frac{(x^2 + \pi^2) \exp(x)}{(1 + \exp(x))^3}$$

$$\propto \rho^4(\mu)(k_B T)^2$$

This temperature dependence is difficult to observe experimentally in a single two-dimensional system because at high temperature the resistivity is dominated by phonon scattering $\rho_{xx} T^5$ and at low temperature by impurity scattering $\rho_{xx} \text{ const}$ (c.f notes 2).

### 7.8 Temperature dependence of quasi-particle decay [5]

As we have said, the temperature dependence of quasi-particle decay is difficult to see in a single two-dimensional system. However, Gramilla et al [5] have found a way to catch some of the quasi-electron quasi-hole pairs that are created when quasi-electrons decay by collision with other quasi-electrons. In their experiment a current $I$ is passed through a two-dimensional electron system (the drive layer) that is in close proximity to, but electrically isolated from, another two-dimensional electron system (the drive layer) as shown in Fig. 7.6[6].

Quasi-electrons passing between Ohmic contacts in the drive layer occasionally 'collide' with quasi-electrons in the drag layer thereby imparting momentum to them. Many such events cause a build up of electrons at one
end of the drag layer and therefore a chemical potential difference between different ends. The voltage difference across the drag layer $V_D$ is of opposite sign to that in the drive layer. The drag resistivity is defined as the ratio of the drag electric field to the drive current density

$$\rho_D = \frac{V_D}{I/W}$$  \hspace{1cm} (7.10)

By analogy with the Drude formalism we may write the drag resistivity in terms of a momentum relaxation rate

$$\rho_D = \frac{m^*}{n_1e^2\tau_D}$$  \hspace{1cm} (7.11)

where $\tau_D$ is then the average time a quasi-electron propagates in the drive layer before colliding with a quasi-electron in the drag layer. We may derive the temperature and density of states dependence of the drag resistivity following the analysis in Eqn. 7.4-7.10. Fig. 7.7 shows the ‘collision’ process between quasi-electrons in the drive and drag layers.

This process has the same form as that shown in Fig. 7.5 except that two of the final products of the decay process are in the drive layer and two in the drag layer and therefore the drag resistivity will have the form

$$\rho_D \propto \frac{1}{\tau_D} \propto \rho_1^2(\mu_1)\rho_2^2(\mu_2)T^2$$  \hspace{1cm} (7.12)
where $\rho_1(\mu_1)$ and $\rho_2(\mu_2)$ are the densities of states at the chemical potentials in the drive and drag layers respectively.

Fig. 7.8a shows the wafer structure used in [5] and Fig. 7.8b shows the conduction band edge for this wafer. Two parallel two-dimensional electron systems form in the two quantum wells with nearly identical carrier densities $n_{2D}=1.5\times10^{15}\text{m}^{-2}$ and mobilities $350\text{m}^2/\text{Vs}$. The distance between the wells is such that their wave functions are completely isolated from each other. The resistance of the barrier region between the wells was found to be $2\Omega$ compared to a sheet resistance of $10\Omega/\square$. The wafer was patterned into a Hall bar geometry using a mesa etch as shown Fig. 7.9.

If an Ohmic contact is required to contact only the top well then the back gate closest to that Ohmic contact is used to deplete the back two-dimensional electron system. If contact to the back two-dimensional system is require the closest front gate is used. Fig. 7.10 shows the measured drag resistivity. It shows the expected $T^2$ dependence for the quasi-particle decay rate. This remarkable result confirms that at low temperatures even in a strongly interacting electron system quasi-particles interact only weakly. We
have therefore been justified in using single electron theory to predict the transport properties of two-dimensional electron systems in notes 1-6.

7.9 Density of states dependence of quasi-particle decay

N. P. R. Hill et al [7] have performed a drag experiment using a bi-layer device similar to that of [5] but in the presence of a quantizing magnetic field. From equation Eqn. 7.12 the formation of an oscillating Landau level density of states should lead to an oscillating drag resistivity in phase with the single layer longitudinal resistivity. Fig. 7.11 shows the experimental confirmation of this result taken from [7].

7.10 Formation of new quasi-particles in a quantising magnetic field

We have shown that the quasi-particle scattering rate in a two-dimensional electron system increases as the fourth power of the density of states at the chemical potential. The data from the N. P. R. Hill et al [7] is consistent with this prediction showing a strong increase in the quasi-particle scattering rate with increasing density of states. From our derivation of
Fig. 7.10. Drag resistivity as a function of temperature demonstrating the $T^2$ law for decay of quasi-particles [6].

quasi-particles it is clear that their form will change as the interaction strength/type changes since in general a co-ordinate transformation that works for one strength/type of interaction will not work for another. This change is typically in the form of a change in effective mass but in a quantizing magnetic field a change in effective mass only affects the shape of the Landau levels and therefore cannot eliminate the interaction between quasi-particles. However, if at the peak of a Landau level (half integer filling factors $\nu = 1/2, 3/2, 5/2...$) the interaction between quasi-particles becomes larger than the cyclotron energy it is clear that the system will have sufficient energy to readjust itself in some way to screen this interaction. The way in which this is achieved is one of the most remarkable facts of nature.
Figs. 7.12-7.16 show the Shubnikov de Haas and Hall resistivities of four different two-dimensional electron systems over different magnetic field ranges. Fig. 7.12 shows the result from the lowest mobility device. It has smooth Shubnikov de Haas maxima and is a demonstration of text book' integer Shubnikov de Haas effect. Figs. 7.13-7.16 however show a great deal of additional structure in their Shubnikov de Haas oscillations. The basic pattern of the Shubnikov de Haas oscillations around zero magnetic field, shown in Fig. 7.12, appears to repeat itself around each half integer filling factor 1/2, 3/2, 5/2, 7/2, ... and then again around each quarter integer filling factor 1/4, 3/4, 5/4, ... . If we interpret this structure as the Landau level
structure of a new type of quasi-particle then these quasi-particles must be such that they feel no magnetic field at even denominator filling factors and an effective magnetic field less than the external magnetic field away from these points. The symmetrical form of the oscillations around even denominator filling factors implies that the effective magnetic field felt by these new quasi-particles can be either positive or negative with respect to the external magnetic field.

A detailed study of the way in which Shubnikov de Haas minima change as a function of the orientation of the external magnetic field with respect to the two-dimensional system allows one to determine the number of full quasi-particle Landau levels at each minimum [12]. Fig. 7.16 shows the result of such an investigation around filling factors 1/2 and 3/2.

The first row of Fig. 7.16 shows filling factors at which there is one fully occupied Landau level of any type. The second row shows the filling factors at which there are two fully occupied Landau levels and so on. The differences between the three columns are that, in column 1 the Landau levels are formed from quasi-electrons, in column 2 the Landau levels are formed from quasi-electron like quasi-particles that feel zero effective magnetic field at $\nu=1/2$, and in column 3 the Landau levels are formed from quasi-hole
like quasi-particles (holes in two filled column 1 type Landau levels) that feel zero effective magnetic field at $\nu=3/2$.

The particle density of each quasi-particle Landau level is taken to be $eB_{\text{eff}}/h$ where $B_{\text{eff}}$ is the effective magnetic field. For filling factors around $\nu = 1/2$ we can therefore write

$$n_{2D} = \nu \frac{eB}{h} = \nu_{\text{eff}} \frac{eB_{\text{eff}}}{h}$$

(7.13)

where $\nu_{\text{eff}}$ is the effective filling factor from Fig. 7.16 column 1 and $\nu$ is the filling factor from column 2. From the relationship between columns 1 and 2 we find

$$\nu = \frac{\nu_{\text{eff}}}{2\nu_{\text{eff}} + 1}$$

(7.14)

Substituting Eqn. 7.14 into Eqn. 7.13 we find the relationship between the effective magnetic field felt by the new quasi-particles and the external magnetic field to be

$$B_{\text{eff}} = B - B_{1/2} = B - \frac{2n_{2D}h}{e}$$

(7.15)
Quasi-particles in two dimensions

Fig. 7.14. (c) same as (b) but taken to higher field data taken from Goldman published by Jain [8].

Around filling factor $\nu = 3/2$ experiment suggests that quasi-particle Landau levels with density $eB_{eff}/h$ are actually formed as Landau levels of holes in two filled non-composite fermion Landau levels, that is

$$n_{2D} = \frac{\nu eB}{h} = 2 \frac{eB}{h} - \nu_{eff} \frac{eB_{eff}}{h}$$

(7.16)

Fig. 7.16 column 1 and $\nu$ is the filling factor from column 3. From the relationship between columns 1 and 3 we find

$$\nu = \frac{3\nu_{eff} + 2}{2\nu_{eff} + 1}$$

(7.17)

Substituting Eqn. 7.17 into Eqn. 7.16 we find the relationship between the effective magnetic field and the external magnetic field for these quasi-particles to be

$$B_{eff} = -3(B - B_{3/2}) = -3B + \frac{2n_{2D}h}{e}$$

(7.18)

Various tests for low field behaviour around $\nu = 1/2$ have been performed
Fig. 7.15. (d) A detail around filling factor $\nu=3/2$ data taken from Du [14].

Fig. 7.17 shows a comparison between a magnetic focusing experiment performed around zero magnetic field (c.f notes 3) and around $\nu = 1/2$. One can see that they contain similar oscillatory structure demonstrating cyclotron motion of composite fermions characteristic of the expected effective magnetic field Eqn. 7.15.

7.12 Composite fermions [16-17]

We have been discussing a new type of quasi-particle of a two-dimensional system that feels an effective magnetic field rather than the actual external
Quasi-particles in two dimensions

<table>
<thead>
<tr>
<th>Minima near $\nu \to \infty$</th>
<th>Willet: minima near $\nu = 1/2$</th>
<th>Du: minima near $\nu = 3/2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/3</td>
<td>5/3</td>
</tr>
<tr>
<td>2</td>
<td>2/5</td>
<td>8/5</td>
</tr>
<tr>
<td>3</td>
<td>3/7</td>
<td>11/7</td>
</tr>
<tr>
<td>4</td>
<td>4/9</td>
<td>14/9</td>
</tr>
<tr>
<td>5</td>
<td>5/11</td>
<td>17/11</td>
</tr>
<tr>
<td>6</td>
<td>6/13</td>
<td>*</td>
</tr>
<tr>
<td>7</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>−7</td>
<td>7/13</td>
<td>19/13</td>
</tr>
<tr>
<td>−6</td>
<td>6/11</td>
<td>16/11</td>
</tr>
<tr>
<td>−5</td>
<td>5/9</td>
<td>13/9</td>
</tr>
<tr>
<td>−4</td>
<td>4/7</td>
<td>10/7</td>
</tr>
<tr>
<td>−3</td>
<td>3/5</td>
<td>7/5</td>
</tr>
<tr>
<td>−2</td>
<td>2/3</td>
<td>4/3</td>
</tr>
<tr>
<td>−1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 7.16. Equivalent $\rho_{xx}$ minima around $B = 0$, $\nu = 1/2$ and $\nu = 3/2$

magnetic field. To find such quasi-particles requires a gauge transformation of the two-dimensional many-particle Hamiltonian

$$H_0 = \frac{1}{2m^*} \sum_i \left( (p_i + eA(r_i))^2 + V(r_i) \right) + \sum_{i<j} u(r_i - r_j) \tag{7.19}$$

The relevant transformation for column 2 table 1 has the form

$$\chi_0^{-1} H_0 \chi_0 \Psi_0 = E \chi_0^{-1} \Psi_0 \tag{7.21}$$

where $z_i = x_i + iy_i$ and $(x_i, y_i)$ are the co-ordinates of the $i$th particle and $\tilde{\phi}$ is a real number. Under the transformation Schrödinger’s equation has the form

$$\chi_0^{-1} H_0 \chi_0 \Psi_0 = E \chi_0^{-1} \Psi_0 \tag{7.21}$$

and the new Hamiltonian has the form

$$\chi_0^{-1} H_0 \chi_0 = \frac{1}{2m^*} \sum_i \left( \left( p_i + eA(r_i) - e\tilde{\phi} \sum_{j \neq i} a(r_i - r_j) \right)^2 + V(r_i) \right) + \sum_{i<j} u(r_i - r_j) \tag{7.22}$$
Eqn. 7.22 contains a gauge potential of the form

$$ a(r) = \frac{\hbar}{2\pi e} \frac{\hat{z} \wedge r}{r^2} \quad (7.23) $$

taking the curl of Eqn. 7.23 we find

$$ \nabla \wedge a = \frac{\hbar}{e} \delta(r) \hat{z} \quad (7.24) $$

The vector potential in Eqn. 7.23 therefore represents adding a tube of magnetic flux

$$ \Phi = -\frac{\hbar}{e} \hat{z} \quad (7.25) $$

to the centre of an electron and the term $e\phi \sum_{i \neq j} a(r_i - r_j)$ in Eqn. 7.22 represents adding such a flux tube to each electron in the system. The gauge
transformation Eqn. 7.21 has not changed the system in any way. The two
Hamiltonians $H_0$ and $\tilde{H}_{\phi}$ represent exactly the same physical system but in
terms of two different types of quasi-particle. The new quasi-particles with
their attached flux are fermions if $\tilde{\phi}$ is an even integer and bosons if $\tilde{\phi}$ is
an odd integer. This can be seen by calculating the Aharonov-Bohm flux
incurred in exchanging two of these quasi-particles. The new quasi-particles
are referred to respectively as composite fermions and composite bosons.
The variety of interest to us here are the composite fermions.

### 7.13 Composite fermion effective magnetic field

A composite fermion propagating through a two-dimensional system may
be viewed as passing through a forest of flux tubes due to other composite
ermions, Fig. 7.13, its dynamics being determined by the external vector
potential plus the vector potential due to the flux tubes.

The vector potential the particle feels from the flux tubes fluctuates about
a mean background due to the constant motion of other composite fermions.
This mean background is equivalent to a constant background magnetic field
$B_g$. If we consider some area $A$ of a two-dimensional system containing $N$
composite fermions then the total flux enclosed will be

$$B_gA = -\frac{\hbar}{e} N$$

(7.26)
Fig. 7.19. (a) Drift velocity, external magnetic field and electric field for bare electrons in a wire (b) as for (a) but for composite fermions.

If the area taken is sufficiently large then

$$B_g = -\tilde{\phi} \frac{h}{e} n_{2D}$$

(7.27)

where \(n_{2D}\) is the carrier density of the two-dimensional system. This background magnetic field adds to the external magnetic field to give an effective magnetic field felt by composite fermions

$$B_{\text{eff}} = B - \tilde{\phi} \frac{h}{e} n_{2D}$$

(7.28)

This is exactly equal to the effective magnetic field found experimentally in Eqn. 7.15 if \(\tilde{\phi} = 2\). The same value of \(\tilde{\phi}\) also gives the effective magnetic field in Eqn. 7.18 and the correct effective magnetic fields around all half integer filling factors. Taking \(\tilde{\phi} = 4\) in Eqn. 7.28 we find the correct effective magnetic fields around all quarter integer filling factors and so on.

The fractional quantum Hall effect may therefore be thought of as an integer quantum Hall effect of composite fermions.

### 7.14 Composite fermion effective electric field

When a current passes along a wire in the presence of an external magnetic field \(B\) an electric field \(E\) builds up such that

$$0 = E + v_d \wedge B$$

(7.29)

where \(v_d\) is the drift velocity Fig.7.19a (c.f notes 2).

The amplitude of the drift velocity is given by

$$v_d = -\frac{E}{B}$$

(7.30)

In the composite fermion gauge the quasi-particles respond to an effective
magnetic field $B_{\text{eff}}$ but, since the drift velocity is an average quantity, that
is independent of any gauge choice they must also respond to an effective
electric field $E_{\text{eff}}$ such that
\[ \nu_d = -\frac{E_{\text{eff}}}{B_{\text{eff}}} \] (7.31)

Combining equations Eqn. 7.30, 7.31 and 7.15 the effective electric field
for the sequence of filling factors in column 2 of table 1 has the form
\[ E_{\text{eff}} = \frac{B_{\text{eff}}}{B} E = (1 - \tilde{\phi} \nu) E \] (7.32)

This represents the production of a gauge electric field $E_g$ of the form
\[ E_g = -\tilde{\phi} \nu E \] (7.33)

which results from the motion of the composite fermion flux tubes and
can be determined directly by consideration of Faraday’s law.

7.15 Composite fermion Hamiltonian

We have established that if the fluctuations in the gauge magnetic field $B_g$
may be ignored then, in the composite fermions, we have a set of quasi-
particles that behave in agreement with experiment. The single-particle
Hamiltonian for these composite fermions has the form
\[ H_S = \frac{1}{2m^*} (p + eA_{\text{eff}})^2 + V \] (7.34)

where
\[ \nabla \wedge A_{\text{eff}} = B_{\text{eff}} \] (7.35)

The potential term $V$ includes the effective electric field.

7.16 Fractional quantum Hall effect

We found for the integer quantum Hall effect that when the chemical poten-
tial was midway between two Landau levels the Hall effect took its classical
value. In the fractional quantum Hall regime the Hall effect takes its classi-
ical value when the chemical potential is between composite fermion Landau
levels
\[ \rho_{xy} = \frac{B}{en_{2D}} = \frac{h}{\nu e^2} \] (7.36)

See Figs. 7.13-7.16.
7.17 Fractional edge-state model [18]

In order to understand quantum transport for our new set of quasi-particles acting in effective magnetic and electric fields, we must ask about the nature of their eigenstates. Equation 7.34 is identical to Schrödinger’s equation in a magnetic field $B = B_{\text{eff}}$ and therefore it must also have eigenstates that are edge states at energies between Landau levels (c.f notes 6). In each edge state, we must also find a cancellation of the group velocity and density of states so that each edge carries the same current. This fact then allows us to use the Landauer-Büttiker formalism to predict the transport properties of mesoscopic systems in the fractional quantum Hall regime. However, there is a difference. Since bare electrons and composite fermions see different electric fields, the chemical potential difference for composite fermions will be different from the chemical potential difference for the bare electrons. It is the chemical potential difference for bare electrons that experiment measures, since composite fermions are only good quasi-particles in a two-dimensional system at high magnetic field and not in Ohmic contacts or measuring equipment.

7.17.1 Fractional quantum Hall effect

Fig. 7.16 shows a standard Hall bar with $N$ composite fermion edge states passing between each Ohmic contact and corresponding bulk filling factor $\nu = N/2N + 1$.
\( \nu = N/(2N+1) \). A current is passed between Ohmic contacts \( S \) and \( D \) and a voltage measured between Ohmic contacts 1 and 2. Using the Landauer-Büttiker formalism we find

\[
I = N \frac{e^2}{h} V_s^* - N \frac{e^2}{h} V_2^* \\
0 = N \frac{e^2}{h} V_1^* - N \frac{e^2}{h} V_s^* \\
0 = N \frac{e^2}{h} V_2^* - N \frac{e^2}{h} V_D^* \\
-I = N \frac{e^2}{h} V_D^* - N \frac{e^2}{h} V_1^* \tag{7.37}
\]

We have added stars to the potentials to indicate that they are the potentials that the composite fermions feel not those that are measured. Equations 7.38 and 7.39 imply that \( V_s^* = V_1^* \) and \( V_D^* = V_2^* \). Therefore from equation 7.37 we find

\[
I = N \frac{e^2}{h} (V_1^* - V_2^*) \tag{7.41}
\]

Since a current of magnetic flux is passing between ohmic contacts 1,2 we can use Eqn. 7.32 to find the relationship between the composite fermion potential difference and the bare electron potential difference measured in an experiment.

\[
V_1^* - V_2^* = E_{eff} W = EW (1 - 2\nu) = (V_1 - V_2)(1 - 2\nu) \tag{7.42}
\]

where we have used \( \tilde{\phi} = 2 \). Substituting Eqn. 7.42 into Eqn. 7.41 we find

\[
I = N \frac{e^2}{h} \nu (V_1 - V_2) = \frac{N}{2N+1} \frac{e^2}{h} (V_1 - V_2) = \nu \frac{e^2}{h} (V_1 - V_2) \tag{7.43}
\]

Hence

\[
R_{SD,12} = \frac{h}{\nu e^2} \tag{7.44}
\]

7.17.2 Reflection of composite fermion edge states

The fractional edge state model has withstood every experimental test that the integer edge state model has passed. A good example is the reflection
Fig. 7.21. Selective reflection of edge states in the fractional quantum Hall regime.

of edge states from a split gate structure. Equations 6.58-6.61 from notes 6 give

\[ I = N \frac{e^2}{h} (V_2^* - V_6^*) \]  
\[ I = M \frac{e^2}{h} (V_2^* - V_5^*) \]  
\[ I = \left( \frac{1}{M} - \frac{1}{N} \right) \frac{e^2}{h} (V_2^* - V_3^*) \]

Using Eqn. 7.32 we find

\[ V_2^* - V_6^* = \left( 1 - \frac{N}{2N+1} \right) (V_2 - V_6) = \frac{V_2 - V_6}{2N+1} \]  
\[ V_2^* - V_5^* = \left( 1 - \frac{M}{2M+1} \right) (V_2 - V_5) = \frac{V_2 - V_5}{2N+1} \]  
\[ V_2^* - V_3^* = V_2 - V_3 \]

Therefore we have

\[ R_{14,26} = \frac{V_2 - V_6}{I} = \frac{2N+1}{N} \frac{h}{e^2} = \frac{h}{\nu_b e^2} \]
Quasi-particles in two dimensions

\[ R_{14,25} = \frac{V_2 - V_5}{I} = \frac{2M + 1}{M} \frac{h}{e^2} = \frac{h}{\nu_{sg} e^2} \]  \hspace{1cm} (7.52)

\[ R_{14,23} = \frac{V_2 - V_3}{I} = \left( \frac{1}{M} - \frac{1}{M} \right) \frac{h}{e^2} = \left( \frac{1}{\nu_{sg}} - \frac{1}{\nu_b} \right) \frac{h}{e^2} \]  \hspace{1cm} (7.53)

where \( \nu_{sg} \) and \( \nu_b \) are respectively the filling factors in the split gate and in the bulk. The results in Eqns. 7.51-7.53 are what are observed experimentally - Fig. 7.21 [19].

7.18 Exercises

7.1 What is a quasi-particle and why is this concept necessary to account for single-particle explanations of mesoscopic phenomena?

7.2 What are the fractional quantum Hall and Shubnikov effects and how may they be explained as an integer quantum Hall effect of composite fermions? Why is disorder necessary to observe the fractional quantum Hall effect?

7.3 Account for the temperature and energy dependence of the life time of quasi-particles in a degenerate electron gas. Describe experiments to probe each of these dependencies.

7.4 How can quasi-particles form in a quasi-one dimensional channel? Describe the passage of a quasi-particle across a saddle point potential.

7.5 Account for the effective magnetic field which composite fermions respond to. Plot the effective magnetic field for the \( n=1/2 \) series composite fermions as a function of position across a parabolic channel at a series of characteristic densities at the channel centre.
Quantum dots

8.1 Sources

8.2 Introduction
In the first section of these notes we introduce the quantum dot - a quasi-zero-dimensional system confined in all three dimensions. We look at its spectrum and conductance and discuss the role of the Coulomb interaction in determining the physics of both. In the second section we derive the Aharonov-Bohm effect and look at its manifestation at low field and at high-field in edge-state networks. These notes cover: the experimental realization of a quasi-zero-dimensional system; the single-particle eigen-spectrum of a quasi-zero-dimensional system; single-particle conductance; experimental conductance; classical Coulomb blockade; quantum Coulomb blockade; the artificial atom; the Aharonov-Bohm effect; and the Aharonov-Bohm effect with edge states.

8.3 Quasi-zero dimensional systems [1,2]
Figure 8.1 shows two different types of quasi-zero-dimensional electron system, so called 'quantum dots', fabricated in semiconductor materials. The
first, Fig. 8.1a,b, shows a vertical resonant tunnelling diode configuration. The quantum dot forms between the two AlGaAs barriers. Conduction occurs vertically via heavily doped GaAs contact regions above and below the quantum dot. The number of electrons trapped in the quantum dot is controlled by the metal side gate. Fig. 8.1(c) shows a lateral quantum dot formed using metal surface gates on a GaAs/AlGaAs heterostructure and Fig. 8.1(d) shows the corresponding effective potential in the device two-dimensional electron gas. The number of electrons in this dot can be controlled via the numbered surface gates.
8.4 The single-particle eigen-spectrum of a quantum dot [3]

Using Eqn. 4.15, Schrödinger’s equation, in the symmetric gauge \( A = (-By/2, Bx/2, 0) \), for a quantum dot with an effective potential \( V(r, \phi) \) has the form

\[
\frac{p^2}{2m^*} \Psi + \frac{\omega_c}{2} L_2 \Psi + \frac{1}{2} m^* \omega_c^2 \frac{r^2}{4} \Psi + V(r, \phi) \Psi = E \Psi \tag{8.1}
\]

Equation 8.1 can be solved exactly for a parabolic confining potential

\[
V(r, \phi) = \frac{1}{2} m^* \omega_c^2 r^2 \tag{8.2}
\]

The eigen-energies of this potential are

\[
E_{n,l} = \hbar \omega_c \left( b \left( n + \frac{1}{2} \right) + \frac{b |l| - l}{2} \right) \tag{8.3}
\]

where

\[
b = \sqrt{1 + 4 \frac{\omega_c^2}{\omega_0^2}} \tag{8.4}
\]

Equation 8.3 has essentially the same form as Eqn. 4.20. The minus sign that appears in Eqn. 8.3 is a convention choice since, \( l \), the azimuthal quantum number, may be either positive or negative. Fig. 8.2 is an evaluation of Eqn. 8.3 showing the generic form of the single-particle eigen-spectrum of a parabolic quantum dot.

8.4.1 Zero field limit

Equation 8.3 can be rearranged in the following way

\[
E_{n,l} = \hbar \omega_c b \left( n + \frac{|l|}{2} \right) - \frac{\hbar \omega_c l}{2} \tag{8.5}
\]

Taking the zero field limit of \( \omega_c b \)

\[
\lim_{B \to 0} \omega_c b = \lim_{B \to 0} \sqrt{\omega_c^2 + 4 \omega_0^2} = 2 \omega_0 \tag{8.6}
\]

equation Eqn. 8.5 then becomes

\[
\lim_{B \to 0} E_{n,l} \tag{8.7}
\]

Equation 8.7 implies that at zero field the eigen-energies of a parabolic quantum dot are degenerate on integer multiples of \( \hbar \omega_0 \) since the term \( 2n + |l| + 1 \) is an integer. The form of this term implies that an energy \( m \hbar \omega_0 \) will have an \( m \) fold degeneracy.
8.4.2 High field limit

In the high field limit

\[
\lim_{B \to \infty} b = 1
\]

so that

\[
\lim_{B \to \infty} E_{n,l} = \hbar \omega_c \left( n + \frac{1}{2} + \frac{|l| - l}{2} \right)
\]

(8.9)

For \( l \geq 0 \)

\[
\lim_{B \to \infty} E_{n,l} = \hbar \omega_c \left( n + \frac{1}{2} \right)
\]

(8.10)

For

\[
\lim_{B \to \infty} E_{n,l} = \hbar \omega_c \left( n + |l| + \frac{1}{2} \right)
\]

(8.11)

All energies therefore converge on one or other Landau level at high field.

Fig. 8.2. Single-particle eigen-spectrum of a parabolic quantum dot.
8.5 Conductance of a quantum dot

The conductance of any mesoscopic system, including a quantum dot, may be calculated using the Landauer formalism (for two-terminal measurements) or the Landauer-Büttiker formalism (for multi-terminal measurements). The problem of calculating the conductance of a quantum dot therefore reduces to finding its transmission coefficients as a function of external parameters. This section investigates the relationship between the energy dependence of the transmission coefficients of a quantum dot and its eigen-spectrum.

Fig. 8.3(a) shows a one-dimensional cross-section of a quantum dot. The two barriers are assumed to be identical and to have individual reflection and transmission coefficients \( r \) and \( t \). For conservation of particle number, they are related through

\[
|t|^2 + |r|^2 = 1 \quad (8.12)
\]

In general these coefficients will be arbitrary complex numbers, but we will assume that the reflection coefficient is real - implying either zero or \( \pi \) phase change on reflection.
When an electron enters a quantum dot it will scatter backwards and forwards between the two barriers so that the total transmission and reflection coefficients are composed of a sum of all possible multiple scattering paths through the dot. The contribution from each path is shown in Fig. 8.3(b).

From Fig. 8.3(b) the total transmission coefficient of the quantum dot in Fig. 8.3(a) has the form

\[ T = t^2 e^{i\phi/2} \left( 1 + e^{i\phi r^2} + \left( e^{i\phi r^2} \right)^2 + \left( e^{i\phi r^2} \right)^3 + \ldots \right) \tag{8.13} \]

summing the series this gives

\[ T = \frac{t^2 e^{i\phi/2}}{1 - e^{i\phi r^2}} \tag{8.14} \]

The two terminal conductance of the quantum dot then has the form

\[ G = \frac{e^2}{h} |T|^2 = \frac{e^2}{h} \frac{t^4}{1 + r^2 - 2r^2 \cos(\phi)} \tag{8.15} \]

Fig. 8.4(a) shows $|T|^2$ as a function of $\phi$ for $t = 0.5$. The resonant peaks in this figure are of unit height because the two barriers of the quantum dot are assumed to be identical. Fig. 8.4(b) shows $|T|^2$ as a function of both $\phi$ and $t$. In this figure it can be seen that for large values of $t$ the transmission resonances are broad and for small values of $t$ they are narrow. On the left side of this plot the transmission resonances are so narrow that they are not visible within the resolution of the figure.

From Eqn. 8.15, it can be seen that the condition for a resonant peak in the conductance is

\[ \phi = 2n\pi \tag{8.16} \]
where \( n \) is an integer. This establishes the link between the eigen-spectrum of a quantum dot and its transmission coefficient since Eqn. 8.16 is exactly the condition for an eigenenergy of a quantum dot. Quantum dots therefore transmit electrons resonantly when an eigenenergy passes the reservoir chemical potentials. Interference in multiple scattering strongly suppresses tunnelling away from this condition.

8.6 McEuen et al [2]

Figure 8.5 shows the two-terminal conductance of the quantum dot in Fig. 8.1(b) as a function of the gate marked (1). Note that the vertical scale is logarithmic. The experiment shows temperature broadened resonant lines that are similar in shape to those in Fig. 8.4(b). The fact that the peak heights are an order of magnitude smaller between Figs. 8.4 and 8.6 could be accounted for by noting that there are many mechanisms that can reduce peak height for example: temperature; asymmetry between barriers; and inelastic scattering.

8.6.1 Magnetic field dependence of resonant peaks [1,2]

At low densities the effective potential of a quantum dot will be approximately parabolic, since screening is less efficient at low densities. We would therefore predict that the resonant peaks shown in Fig. 8.5 should oscillate in magnetic field, and trace out the spectrum shown in Fig. 8.2, because of link we have established between the transmission resonances and the eigenstates of a quantum dot. Fig. 8.6a,b show the trajectories of resonant peaks for both dots shown in Fig. 8.1. As can be seen, the resonant peaks in Fig. 8.6 oscillate as a function of magnetic field in a way that is similar to the spectrum shown in Fig. 8.2. However, it appears as if the spectrum of the parabolic dot has been taken apart section by section in such a way that each line in Fig. 8.2, which corresponds to a line of constant particle number in the dot, is displaced in gate voltage from each other by a similar amount. This is demonstrated in the lower panel of Fig. 8.6(b) where these displacements have been removed. It shows (all be it in a wobbly way) a nice detail from the spectrum shown in Fig. 8.2. This offset between resonant peaks derives from the Coulomb interaction. The effect is known as ‘the Coulomb blockade’ and is a mechanism which operates even in quantum dots that are so large that quantum mechanics is irrelevant.
Quantum dots

Fig. 8.5. Two-terminal conductance of the quantum dot shown in Fig. 8.1(b). Details show resonant line shapes.

8.7 Classical Coulomb blockade [4].

Consider a small metallic dot, Fig. 8.7, in weak electrical contact with two reservoirs via two high resistance connections and capacitively coupled to a nearby gate with a potential $V_g$ applied to it. We imagine that the dot is sufficiently large that its eigen-spectrum forms a continuum.
8.7 Classical Coulomb blockade [4].

Fig. 8.6. (a) Trajectories of resonant peaks for Fig. 8.1(a). (b) for Fig. 8.1(b).

\[
\phi(Q) = Q/C + \phi_{ext} \tag{8.17}
\]

where \( Q \) is the charge on the dot, \( C \) is the capacitance between the dot and the rest of the system, and \( \phi_{ext} \) is the external potential that arises from
the charge in the rest of the system

\[ Q_{\text{ext}} = C\phi_{\text{ext}} \]  

(8.18)

. This external charge is distributed between the surface gate, ionised donors and reservoirs. Integrating Eqn. 8.17 the electrostatic potential energy of the electrons on the dot is

\[ U(Q) = \int_0^Q \phi(Q)dQ \]  

(8.19)

Since the dot is only weakly coupled to the rest of the system the total
number of electrons on it must be an integer \( Q = Ne \) so that Eqn. 8.19 becomes

\[
U(N) = \frac{(Ne)^2}{2C} - Ne\phi_{\text{ext}} \quad (8.20)
\]

The external charge \( Q_{\text{ext}} \) may be varied continuously by means of the surface gate \( V_g \) (in contrast to \( Q \) which is restricted to integer multiples of \( e \)). Rearranging Eqn. 8.20 and substituting for \( \phi_{\text{ext}} \), \( U(N) \) takes the parabolic form

\[
U(N) = \frac{(Ne - Q_{\text{ext}})^2}{2C} - \frac{Q_{\text{ext}}^2}{2C} \quad (8.21)
\]

Fig. 8.8(a) shows \( U(N) \) from Eqn. 8.21 plotted for a series of different values of \( Q_{\text{ext}} \) as a function of \( N \). The values of \( U(N) \) for which \( N \) is an integer are marked with black dots and labelled by the number of electrons trapped in the dot. For \( Q_{\text{ext}} = Ne \) it can be seen from Fig. 8.8(a) and Eqn. 8.17 and Eqn. 8.21 that when the dot contains \( N \) electrons: the electrostatic potential is zero \( \phi(N) = 0 \); the electrostatic potential energy \( U(N) \) is a minimum; and the minimum energy to add an extra electron is

\[
U(N + 1) - U(N) = \frac{e^2}{2C} \quad (8.22)
\]

Unless this energy can be supplied by the reservoirs (by putting a sufficiently large d.c. bias across the dot) electrons cannot hop onto the dot and its resistance will be very large. This phenomenon is referred as the ‘Coulomb blockade’ because it is the Coulomb interaction that, under these circumstances, is preventing electrons from hopping onto the dot. The energy difference \( U(N + 1) - U(N) \) is referred to as the minimum addition energy. The addition spectrum of the dot is shown in Fig. 8.8(b) at the same series of \( Q_{\text{ext}} \) used Fig. 8.8(a).

As \( Q_{\text{ext}} \) increases, with more positive \( V_g \), it can be seen in Fig. 8.8(b) that the minimum addition energy decreases and becomes zero when:

\[
Q_{\text{ext}} = (N + 1/2)e \quad (8.23)
\]

At this point, it costs as much energy to have \( N \) electrons in the dot as it does to have \( N + 1 \) and electrons may therefore hop onto or off the dot without experiencing any Coulomb blockade. The resistance of the dot is then simply the sum of the resistances of the two contacts and the dot itself (this is its classical resistance value). A current will begin to flow through the dot once the d.c bias between reservoirs is larger than the minimum addition energy.
Quantum dots

Once \( Q_{ext} \) increases past \( Q_{ext} = (N + 1/2)e \) there will be \( N + 1 \) electrons trapped in the dot and again it will be Coulomb blockaded since there will be a gap to allowing the \( N + 2 \)nd electron through. The sequence then repeats itself and the resistance is very large until \( Q_{ext} = (N + 1 + 1/2)e \), where it is again equal to the classical resistance of the dot.

The conductance of a small metallic dot, as a function of an external gate voltage, therefore consists of a series of equally spaced isolated peaks occurring when there is no Coulomb blockade. The external potential \( \phi_{ext} \) has a linear dependence on the surface-gate potential \( V_g \) so that for some change in surface-gate voltage \( \Delta V_g \): \( \Delta \phi_{ext} = \alpha \Delta V_g \). Between two peaks in conductance, one extra electron enters the dot and \( \Delta Q_{ext} = e \). Therefore the spacing in gate voltage between conductance peaks is

\[
\Delta V_g = \frac{e}{\alpha C} = \frac{e}{C_g}
\]

where \( C_g \) is the capacitance between the gate and the dot.

8.8 Quantum Coulomb blockade

Even though the Coulomb blockade is a classical effect, it is still important in small quantum dots. In these systems, the potential energy \( U(N) \) is a sum of the Coulomb charging energy Eqn. 8.21 and the energies of discrete levels in the dot

\[
U(N) = (Ne - Q_{ext})^2/2C - Q_{ext}^2/2C + \sum_{i=1}^{N} \epsilon_i
\]  

(8.25)

Here \( \epsilon_i \) represents the single-particle energy of the \( i \)th state measured relative to the minimum of the quantum dot effective potential. It should be noted that, as the dot fills, its effective potential changes and the nature of the single-particle spectrum changes. The minimum addition energy \( E_{min} \) of a quantum dot containing \( N \) electrons then depends on the single-particle energy for the \( N + 1 \)st state

\[
E_{min} = U(N + 1) - U(N) = \frac{(N + 1/2)e^2 - eQ_{ext}}{C} + \epsilon_{N+1}
\]  

(8.26)

In the classical case the addition energies above \( E_{min} \) form a continuum. In the quantum case they form a discrete set of energies \( E_i = E_{min} + \epsilon_i - \epsilon_{N+1} \). Fig. 8.9 shows the addition spectrum at a series of different \( Q_{ext} \). Once the d.c. bias between the reservoirs either side of a quantum dot is larger than \( E_{min} \) electrons may tunnel through the dot. As the d.c. bias is increased
and successive single-particle states pass beneath the chemical potential of either reservoir the conductance increases in a series of steps.

At zero d.c. bias, a peak in the conductance of a quantum dot occurs when its addition energy is zero. The resistance of a quantum dot will be less than the classical resistance because of resonant multiple scattering. Away from a resonant peak interference in multiple scattering may aid the Coulomb blockade and make it more resistive than the classical case.

From Eqn. 8.26 the external charge at a conductance peak is

$$Q_{\text{ext}} = (N + 1/2)e + \frac{C}{e}\epsilon_{N+1}$$

(8.27)

so that, between the $N$th and $N+1$st peaks

$$\Delta Q_{\text{ext}} = e + \frac{C}{e}(\epsilon_{N+2} - \epsilon_{N+1})$$

(8.28)

This implies that either slightly more or slightly less external charge, than the classical case, is required to compensate for differences in the single-particle energies as each extra electron is added to the quantum dot. The spacings between peaks in gate voltage are then

$$\Delta V_g = \frac{e}{C_g} + \frac{1}{\alpha e}(\epsilon_{N+2} - \epsilon_{N+1})$$

(8.29)

The single-particle energies therefore modulate the resonant peak positions.
confirming that the oscillatory structure in the spectra shown in Fig. 8.6 does in fact derive from the single-particle spectrum of each quantum dot.

8.9 Artificial atoms

Figure 8.10(a) shows a grey-scale plot of the transconductance $dG/dV_g$ ($G$ is the differential conductance of the quantum dot $dI/dV_{dc}$) as a function of dot gate voltage $V_g$ and d.c bias $V_{dc}$ for the quantum dot shown in Fig. 8.1(a). Just as for d.c bias measurements of a split-gate structures, see notes 6, each transconductance peak splits into two peaks in d.c. bias, one for each chemical potential. The width of each diamond (coloured black for clarity in Fig. 8.10(a)) gives the magnitude of the minimum addition energy for each integer number of electrons filling the dot. These addition energies are plotted in Fig. 8.10(b). The number next to each peak indicates the number of electrons occupying the quantum dot. Configurations with 2, 6, 12 electrons have particularly high minimum addition energies. These fillings correspond to the zero field degeneracies of a parabolic quantum dot, once spin is included, and they therefore correspond to complete two-dimensional shells: nobel artificial atoms. This is shown schematically in Fig. 8.11. The other peaks correspond to half-filled shells indicating a two-dimensional
Fig. 8.11. Shell structure of a two-dimensional artificial atom. Black and Grey represent the two spin species.

Fig. 8.12. Arbitrary line in a two-dimensional plane.

Hunds rule: in a partially filled shell the minimum energy configuration has maximum spin.

8.10 The Aharonov-Bohm effect [5]

The Aharonov-Bohm effect is a truly quantum-mechanical phenomenon. It is an interference effect that arises from the vector-potential dependence of a wave function. The first step in understanding it is to find an expression
for the way in which a wave function changes along an arbitrary path in a magnetic field.

In zero magnetic field Schrödinger’s equation, in a two-dimensional plane with an effective potential $V(x, y)$, has the form

$$\frac{p^2}{2m^*} \Psi + V(x, y) \Psi = E \Psi$$  \hspace{1cm} (8.30)

If at a some point $l$ in the plane, the wave function is $\Psi(l)$ then some small distance $dl$ away it will be

$$\Psi(l + dl) = e^{ip(l) \hbar/\hbar} \Psi(l)$$  \hspace{1cm} (8.31)

This assumes that the variation of the wave function is much faster than the variation of the effective potential - allowing it to be treated locally as if it is constant. Iterating Eqn. 8.31 along the line in Fig. 8.12 we find

$$\Psi(l_N) = e^{ip(N) \hbar/\hbar} \Psi(l_0)$$  \hspace{1cm} (8.32)

This can then be written as

$$\Psi(l_N) = e^{i \int_{l_0}^{l_N} \frac{p(l)}{\hbar} \hbar} \Psi(l_0)$$  \hspace{1cm} (8.33)

At finite magnetic field we must subtract the field momentum from the canonical momentum to get the correct kinetic energy in Schrödinger’s equation Eqn. 8.30. This gives the following replacement in Eqn. 8.33: $p \rightarrow p + eA$ so that

$$\Psi(l_N) = e^{i \int_{l_0}^{l_N} \frac{p(l) + A(l)}{\hbar} \hbar} \Psi(l_0)$$  \hspace{1cm} (8.34)

This equation shows that the phase accumulated in passing along an arbitrary path in a two-dimensional plane depends both on the canonical momentum $p$ and the field momentum $eA$.

The Aharonov-Bohm effect is most strikingly seen in the situation shown in Fig. 8.13. It consists of a wire that constrains electrons to move in a circle around a region that has a finite magnetic field $B$ in an area $S$ somewhere enclosed within it. From a classical perspective, we would not expect the electrons to have any part of their motion affected by the enclosed magnetic field since they will not actually move through any region where the magnetic field is finite. However, the particle’s wave function is affected by the existence of the enclosed flux. From Eqn. 8.34 the total phase accumulated in passing round the loop will be

$$\Delta \phi = \oint \frac{p + eA}{\hbar} \hbar dl = \frac{2\pi rp}{\hbar} + \frac{e}{\hbar} \int \nabla \times A \cdot dS = 2\frac{\pi rp}{\hbar} + \frac{e}{\hbar} \int B \cdot dS$$  \hspace{1cm} (8.35)
where $r$ is the radius of the electron loop so that

$$\Delta \phi = \frac{2\pi rp}{\hbar} + 2\pi \frac{e}{h}BS$$

The field momentum part of this phase change $eBS/h$ is called the Aharonov-Bohm phase. We will now show how this phase affects the conductance of a ring.

**8.11 Webb et al [6]**

Figure 8.14 (inset) shows an experimental device that consists of a small gold ring patterned on the surface of a semiconductor. The upper oscillatory trace (a) shows the resistance of this ring as a function of magnetic field and the lower trace (b) its power spectrum. The ring was sufficiently pure that its phase-coherence length was greater than twice the circumference of the ring. Its conductance can therefore be calculated using the Landauer formalism. The transmission coefficient of the loop should be calculated by summing all possible paths that take electrons from left to right. They form a very complicated set since any one path can alternate randomly between propagation clockwise or anticlockwise around the loop. Summing all paths is therefore not a particularly easy job. Fortunately, in order to understand the data in Fig. 8.14 it is not necessary - we can get away with just looking at the two principal paths shown in Fig. 8.14(c). Considering only these paths, and using equation Eqn. 8.34, the transmission coefficient of the loop has the form

$$t \propto t_{OA} + t_{OB} = \left( e^{i\int_{0}^{A} \frac{p(l)+eA(l)}{\hbar} dl} + e^{i\int_{0}^{B} \frac{p(l)+eA(l)}{\hbar} dl} \right) = e^{i\phi_{A}} + e^{i\phi_{B}}$$

(8.37)
Fig. 8.14. (a) Magnetoresistance of the (inset) ring measured at 0.01K. (b) Fourier power spectrum in arbitrary units containing peaks at h/e and h/2e. The inside diameter of the ring was 784nm, and the width of the wire was 41nm. (c) Schematic of ring.

so that the conductance is

\[ G \propto |t|^2 \propto (1 + \cos(\phi_A - \phi_B)) \]  

where

\[ \phi_A = \int_0^A \frac{p + eA}{\hbar} \cdot dl + \frac{e}{\hbar} \int_0^A A \cdot dl \]  

Using the symmetric gauge \( A = (-Bx/2, By/2, 0) = Br/2\phi \) in cylindrical coordinates this becomes

\[ \phi_A = 2\pi pR/\hbar + \frac{e}{\hbar} \int_0^\pi BR/2 = 2\pi pR/\hbar + \frac{e}{\hbar} \frac{B\pi R^2}{2} \]  

and

\[ \phi_B = 2\pi pR/\hbar + \frac{e}{\hbar} \int_0^{-\pi} BR/2 = 2\pi pR/\hbar - \frac{e}{\hbar} \frac{B\pi R^2}{2} \]  

so that

\[ \phi_A - \phi_B = \frac{e}{\hbar} B\pi R^2 = 2\pi \frac{e}{\hbar} BS \]  

This implies that as the external magnetic field increases, for each flux quantum \( h/e \) added to the ring, the conductance oscillates through one period.
8.12 Highfield Aharonov-Bohm effect.

In looking at edge states in the quantum Hall regime, notes 6, we found that they follow equipotentials in the selfconsistent effective potential of a device. The majority of these states in a Hall bar form closed loops either in dips (quantum dots) as shown in Fig. 8.15(a) or around peaks (quantum anti-dots) as shown in Fig. 8.15(b).

Any state on a closed loop must have an accumulated phase that is equal to an integer multiple of $2\pi$: $\Delta \phi = 2n\pi$. This causes the states on these loops to have a discrete energy spectrum.

At high magnetic fields the Aharonov-Bohm phase dominates the phase shift around any closed loop so that for any high field eigenstate of a dot or an anti-dot

$$2\pi \frac{e}{\hbar} BS = 2m\pi \Rightarrow BS = m\frac{\hbar}{e}$$

where $S$ is the area enclosed by the state. The eigen-spectrum for a symmetric potential, such as a parabolic quantum dot, can then be written

$$E_{n,m} = (n + 1/2)\hbar \omega_c \pm \frac{1}{2} g\mu_B B + V(r_m)$$
The second term in Eqn. 8.44 is Zeeman spin splitting. Equations Eqn. 8.44-8.45 can be compared directly with Eqn. 8.3 for a parabolic effective potential, if spin-splitting is ignored.

\[
B \pi r_m^2 = \frac{m}{e} \tag{8.45}
\]

The second term in Eqn. 8.44 is Zeeman spin splitting. Equations Eqn. 8.44-8.45 can be compared directly with Eqn. 8.3 for a parabolic effective potential, if spin-splitting is ignored.

### 8.13 Edge-state networks, dots and anti-dots.

Figs. 8.16(a),(b) show generic patterns of edge states forming around a dot (a) and an anti-dot (b) at high magnetic field. Arrows indicate the direction in which the edge states propagate and horizontal and vertical dashed lines indicate possible points where tunnelling between edge states may occur. The grey shapes show split-gate patterns that could be used to create such dots and anti-dots artificially.

As can be seen, the two edge-state networks for the two cases are very similar in form, except for the direction in which the bound states propagate. The conductance of these devices can be calculated from the Landauer-Büttiker formalism and therefore depends on the network transmission coefficients. Evaluating all the multiple scattering paths that contribute to the transmission coefficient for this case is much easier than for the gold rings we considered earlier because electrons propagate in only one direction along an edge state.
There are two basic resonant processes that operate simultaneously in these networks: resonant transmission through bound states from left to right (horizontal tunnelling); and resonant reflection by tunnelling top to bottom (vertical tunnelling). It is convenient to consider the effect of each process separately as an initial step to understanding the transmission properties of the complete networks. Figs. 8.17a,b show edge state networks for the horizontal and vertical tunnelling processes individually. For network (a) in Fig. 8.17a the transmission coefficient from left to right $t_{AB}$ is a sum of the paths shown in Fig. 8.17(c). In terms of the individual tunnel barrier
reflection and transmission coefficients $r, t$ it has the form

$$t_{AB} = t^2 e^{i\phi/2} \left(1 + r^2 e^{i\phi} + (r^2 e^{i\phi})^2 + (r^2 e^{i\phi})^3 + \ldots\right) \quad (8.46)$$

$$= \frac{t^2 e^{i\phi/2}}{1 - r^2 e^{i\phi}} \quad (8.47)$$

For network (b) in Fig. 8.17b the transmission coefficient has the form

$$t_{AB} = r + r t^2 e^{i\phi} \left(1 + r^2 e^{i\phi} + (r^2 e^{i\phi})^2 + (r^2 e^{i\phi})^3 + \ldots\right) \quad (8.48)$$

$$= r + \frac{r t^2 e^{i\phi/2}}{1 - r^2 e^{i\phi}} \quad (8.49)$$

The corresponding conductances for (a) and (b) are shown in Figs. 8.17e,f as a function of the phase shift $\phi$ and transmission coefficient $t$. As would be expected (a) leads to resonant peaks in conductance with increasing $\phi$ and (b) lead to resonant dips in conductance with increasing $\phi$. In combination, these two processes compete with each other so that when the barrier transmission coefficient for vertical tunnelling is greater than the barrier tunnelling coefficient for horizontal tunnelling resonant dips down are seen and when the barrier transmission coefficient for horizontal tunnelling is greater than the barrier tunnelling coefficient for vertical tunnelling resonant peaks are seen.

8.14 Mace et al [7]

Figure. 8.18(a) shows a quantum anti-dot device, for which the split gate potentials have been adjusted to pass the two edge states of the $n=0$ Landau level (spin up and spin down). These edge states form a set of bound states around the anti-dot. States in the $n=1$ Landau level (spin up and spin down) are reflected from the anti-dot. Figs. 8.18(b),(c) show cross-sections through the device. Bound edge states are shown as dotted lines, continuous edge states as solid lines, the effective potential as a thick line and the chemical potential as a dashed line.

Figure 8.19 shows a series of conductance traces from the anti-dot as a function of magnetic field. Between different traces the voltage on the anti-dot gate has been changed. Trace (1) has the most negative anti-dot gate voltage and trace (8) the least negative. In trace (8) the anti-dot is small and the gaps between the upper and lower split gates are large (Fig. 8.18(c)) so that vertical tunnelling is suppressed. This results in horizontal tunnelling dominating and therefore resonant transmission peaks appear in the conductance. The peaks form spin-split pairs since spin is preserved in tunnelling.
The Aharonov-Bohm period should be taken between peaks deriving from like spins. As can be seen in Figs. 8.18(b),(c) and equations Eqn. 8.44-8.45, each spin forms its own separate set of bound states. At the chemical potential the spin up states sit outside the spin down states. For horizontal tunnelling this results in equal tunnelling distances between states of like spin (Fig. 8.18(b)) and therefore equal resonant tunnelling peak heights.

As the anti-dot gate voltage is made more negative the anti-dot diameter increases and so the gaps between the split gates get smaller (Fig. 8.18(c)). This strengthens the vertical tunnelling and therefore diminishes the resonant transmission peak heights, eventually, in trace (1), turning them both into resonant dips. The left peak of a spin-split pair derives from the spin-up bound states and resonant transmission peaks from them diminish and turn into resonant dips faster than peaks deriving from the spin-down elec-
Quantum dots

Fig. 8.19. Conductance traces from the anti-dot shown in Figs 19. Anti-dot gate voltage is more negative in trace (1) than in trace (8). Traces show spin split Aharonov-Bohm oscillations.

electrons because at any gate voltage the vertical tunnelling distance for spin up electrons is always less than that for spin down electrons (Fig. 8.18(c)).

8.15 Exercises

8.1 Plot the high field eigen-spectrum of a quantum anti-dot using Eqn. 8.44-8.45.

8.2 What could account for the extra structure outside the principal diamonds in Fig. 8.10(a)?

8.3 In the experiment by Webb [6], demonstrating the Aharonov-Bohm
effect, the external magnetic field was not zero inside the ring. What effect might this have?

8.4 Why do you think the lines in Fig. 8.6(a) appear in pairs?
9
Quantum computation

9.1 Sources
[1]

9.2 Introduction
In these notes we introduce the

9.3 Classical vs Quantum computation
9.4 The DiVincenzo rules
9.5 The charge qubit
9.6 Fujisawa
9.7 g-factor in a semiconductor
9.8 Hanson
9.9 Single spin detection: ...
9.10 Interaction in the Hubbard model
9.11 Single-spin rotation
9.12 Surface acoustic wave current quantisation
9.13 The surface acoustic wave quantum processor
9.14 Exercises

9.1
Appendix 1

Occupation probabilities.

A1.1 Derivation of occupation probabilities:

The ratio of the probability that an \( n \)-electron state of energy \( \epsilon \) is occupied to the probability that it is unoccupied, when it is bought into contact with an electron reservoir at temperature \( T \) with chemical potential \( \mu \), is given by the Gibbs factor: [Kittel]

\[
\frac{p(n, \epsilon)}{p(0, 0)} = \exp\left(\frac{n\mu - \epsilon}{k_B T}\right) \tag{1.1}
\]

If we consider a state in the conduction band of a semiconductor we would expect to find that its occupation probability is the Fermi-Dirac distribution \( f(\epsilon) \). Conduction-band states can be occupied in four different ways (1) with zero electrons, (2) with an electron of spin up, energy \( \epsilon \), (3) with an electron of spin down, energy \( \epsilon \), (4) with two electrons of opposite spin, energy \( 2\epsilon \).

The average number of electrons occupying the state is then

\[
f(\epsilon) = \frac{\sum_i n_i p(n_i, \epsilon_i)}{\sum_i p(n_i, \epsilon_i)} = \frac{2p(1, \epsilon) + 2p(2, 2\epsilon)}{p(0, 0) + 2p(1, \epsilon) + p(2, 2\epsilon)} = \frac{2}{1 + \exp\left(\frac{\epsilon - \mu}{k_B T}\right)} \tag{1.2}
\]

the factor of two results from the spin degeneracy.

If we now consider a single localized donor state. These states can also be occupied in four different ways (1) with zero electrons (2) with an electron of spin up, energy \( \epsilon \) (3) with an electron of spin down, energy \( \epsilon \) (4) with two electrons of opposite spin. But now, since the donor wave function is localized to a single bond the energy for (4) is much larger than \( 2\epsilon \) owing to the Coulomb interaction and the corresponding probability for occupation will be zero so that

\[
\langle n \rangle = \frac{2p(1, \epsilon)}{p(0, 0) + 2p(1, \epsilon)} = \frac{2}{2 + \exp\left(\frac{\epsilon - \mu}{k_B T}\right)} \tag{1.3}
\]
Occupation probabilities.

the corresponding average for an acceptor state of energy $\epsilon$ is

$$
\langle p \rangle = \frac{2}{2 + \exp\left(\frac{\epsilon - \mu}{k_B T}\right)}
$$

(1.4)
Appendix 2

Density of states in three dimensions.

We assume that the Fermi surface is spherical. The solution to the effective Schrödinger equation has the form $\psi = \exp(i \mathbf{k} \cdot \mathbf{r})$. Periodic boundary conditions on a cube of side $L$ gives, $\exp(i k_x (x + L)) = \exp(i k_x x)$ for the $x$-direction, this implies that the wave vector in the $x$-direction has the following quantized values $k_x = 2n\pi/L$ similar results are obtained for the $y, z$ directions. The total number of occupied states in the cube is then

$$N = \frac{4\pi k_F^3}{(2\pi/L)^3} \quad (2.1)$$

so that the carrier density becomes

$$n = \frac{N}{L^3} = \frac{k_F^3}{3\pi^2} \quad (2.2)$$

The Fermi energy is related to the Fermi wave vector through

$$E_F = \frac{\hbar^2 k_F^2}{2m^*} \quad (2.3)$$

so that the density of states takes the form

$$g(E) = \frac{dn}{dE} = \frac{m^* \sqrt{2E}}{\hbar^3 \pi^2} \quad (2.4)$$
Appendix 3

Band bending from a delta layer

Consider a delta function layer of charge with surface density $\sigma$ (m$^{-2}$). If the system is homogeneous Poisson's equation takes the form $\nabla.E = \rho/\varepsilon$. Integrating this expression over the pill box in Fig. A3.1

$$\int \nabla.E d\tau = \int \frac{\rho}{\varepsilon} d\tau$$

(3.1)

gives

$$\oint E.dS = \frac{\sigma A}{\varepsilon}$$

(3.2)

which, if the interfaces are infinite gives

$$E_2 - E_1 = \frac{\sigma}{\varepsilon}$$

(3.3)

Fig. A3.1. Sheet of charge.
Now $E = -\nabla \phi = \nabla \frac{E_C}{e}$, where $E_C$ is the conduction band edge therefore if $\sigma$ is positive the gradient of $E_C$ is larger on the right hand side of the figure and if $\sigma$ is negative the gradient is smaller.
Appendix 4
Density of states in two dimensions.

We assume the Fermi surface is circular. This gives
\[ n = \frac{N}{L^2} = \frac{\pi k_F^2}{(2\pi)^2} = \frac{m^*}{2\pi\hbar^2} E_F \]  
(4.1)
so that
\[ g_{2D}(\epsilon) = \frac{dn}{d\epsilon} = \frac{m^*}{2\pi\hbar^2}. \]  
(4.2)
For a single bound state in a well the carrier density will be
\[ n_{2D} = \int_{E_0}^{\infty} g_{2D}(\epsilon) f(\epsilon) d\epsilon \]  
(4.3)
We must multiply this by the degeneracy factor to find the total density. For GaAs each two-dimensional sub-band is doubly degenerate due to spin. For Si each sub-band is four fold degenerate. Only two of the Si valleys remain degenerate at the oxide-Si interface, the other factor of two comes from spin degeneracy. The carrier density in a two-dimensional sub-band is distributed according to the modulus squared of its wave function.