NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Magneto-tunneling of Holes in $Si/Si_{1-x}Ge_x/Si$ Double Barrier Structures

by R. C. D. Low

Thesis

submitted to the University of Ottawa

in partial fulfilment of the requirements for the degree of

Master's in Physics,

University of Ottawa, January 16, 1991.

ISBN 0-315-68061-X

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

Abstract

Resonant and non-resonant tunneling processes in p-type $Si/Si_{1-x}Ge_x/Si$ double barrier structures were investigated at cryogenic temperatures in both planar and perpendicular magnetic fields. A differential conductance measurement technique was devised to observe non-resonant tunneling features superimposed on a rapidly increasing current background. An additional resonant feature was observed at high bias voltage and is attributed to tunneling of holes from the two-dimensional heavy hole subband in the emitter to a third quasi-bound state in the quantum well. Experimental evidence for non-resonant tunneling processes is presented for the first time in p-type double barrier structures. Inter-Landau level transitions between the emitter and the quantum well occur either elastically by scattering from impurities or inelastically with the emission of optical phonons. Both processes have been observed in our samples. Data above 4.2 K suggest that tunneling processes also take place from the three-dimensional valence band edge states in the emitter. A novel spectroscopy of the quasi-bound states in the well was performed by measuring the shift in energy of the resonant tunneling features as a function of applied planar magnetic field.

Dedicated to my Mother and
my Father, who made this
all possible.

Acknowledgments

The production of this thesis is of course indebted to the assistance of numerous people and the greatest danger in the writing of these pages is the failure to acknowledge the contributions of each of them.

I would like to thank the members of the Quantum Physics Group, in the Institute for Microstructural Sciences at the National Research Council Canada. Since the whole of the research on which this thesis is based was conducted in this location, it is impossible to express my gratitude sufficiently for their kind acceptance of my ignorance, their subsequent gentle attempts to set me on the correct path, and their good humour in all my silly requests for help.

In particular I would like to thank Don Brown who tended to the very impressive magnet and cryostat facility with utmost competency, and handling with aplomb the times when I would test that competency with an equal amount of incompetency (i.e. quench). The afternoons spent on the golf course were a blessing.

I am indebted to Jeff Young who is a living reminder that purple herbies do indeed mix with serious physics, and that fun and physics are not mutually exclusive terms. He is like the son I never had, and he treated me likewise: for that I am entirely grateful.

To Andy Sachrajda for teaching me the wonderful phrases which are hurled between supporters of competing football sides. I'll miss the repartee with him.

To Bruce Kettles for sharing his office and company with me. To Paul Kelly and Sylvain Charbonneau for their help. To all of them for leaving the cafeteria each time I arrived.

To Richard Taylor, who put me up in my hour of need. For his companionship, his poor taste in music and for doing the seal whilst I did the penguin. Simp, a man of eclectic ways and strange behaviour, God bless him.

To H. C. Liu for his understanding and discussions regarding not only my research but my future. For setting me straight on many occasions.

To Margaret Buchanan for her infinite help and wonderful company during my work in her lab: a rebel who operated several sports cars in a career at which I can only marvel. Her understanding of physics is only excelled by her better understanding of human nature and the importance of tea-time. The only way I can conceive of returning her generosity is to complete the carrot experiment someday on her favourite subject.

I would also like to acknowledge the friends in Optoelectronics, with whom I partook in lunch and coffee and beer in my stay here.

To Richard Normandin for his help in ridding the building of that dreaded virus and for his patience in explaining his work.

To Douglas Halliday, a fellow of wit and prudence, a pal, a Scot, for his encouragement in my most disheartened moments of this writing. I shall never forget the athletic prowess he so brilliantly demonstrated in the art of pebble throwing.

In the rare moments in which I was at the University I had the pleasure of receiving help from some very good people in the Department of Physics.

To Prof. Benson for his help in setting up my teaching duties, which afforded me the pleasures both of working with bright young kids and of having funds with which to . To Prof. Smith who was responsible for giving me the opportunity to come to Ottawa and pursue my research. To Mike Jackson, who was irascible only because he did like me. To Roger Chagnon and the rest of the members of the teaching laboratories for their help and kindness. To the secretaries, Lorraine, Line, Carole and Bonnie for taking care of the administrative details in which I was so loath to be involved, and for which my memory was none too good.

To James Carolan at the University of British Columbia, my mentor and friend, it was he who inspired me to contemplate a future in physics.

Lastly, and most importantly, to my supervisors.

To Francis Guillon, for being my point man at the University of Ottawa. He ensured that I was imformed of such trivialities as deadlines, important guest speakers, financing of my education, etcetera. I would be hard pressed to imagine my years here without his support and generosity.

To Marie D'Iorio who took me in and guided me to this conclusion. To forgive me my errors and ignorance throughout my two years was a most gracious thing. To persist in leading me on the right path despite all my shortcomings was the act of a saint. To her I am indebted for financial support, academic teaching and supervision. But if all this were set aside, I have an even greater obligation for her friendship which has made the entire experience so much more the enjoyable. She allowed me the freedom to pursue my many other interests, had the good humour to dismiss my failures, and gave me continual encouragement at all points. For all the time she devoted to my work, including midnight hours, and for the patience she showed, I thank her from the bottom of my heart.

Table of Contents

List of Figures

List of Tables

I. Introduction

During the past three decades, the field of solid state electronics has undergone a revolution due to devices based on electron transport in inversion layers in semiconductors. One particularly important semiconductor structure is the field effect transistor (FET) which forms the basic component of computer logic circuits. The operation of this device is controlled by an electric field applied at the surface of the semiconductor causing charge carriers to accumulate in a thin layer at the semiconductor-insulator interface. More recently the interest in **III-V** semiconductors has flourished due to the advent of material deposition techniques such as molecular beam epitaxy (MBE), which have made possible the growth of thin layers of different semiconductors with extremely clean interfaces. These structures are known as heterostructures, multiple quantum well or superlattices depending on the number and thickness of the constituent layers; they generally have a band structure which differs from that of the constituent layers depending on parameters such as doping, layer thickness etc. which can be set by the growth. The novel properties of these heterostructures stem from the confinement of the charge carriers in one or more dimensions; they are the focus of both fundamental and applied solid state physics research (*cf.* Stiles, 1987 for a review of this subject).

In particular, silicon MBE has focused on the growth of strained layer superlattices (SLS) of Si and Ge. The strain arises from the lattice constant mismatch between Si and Ge and it can be used to change the bandgap of the superlattice. The compatibility of Si-Ge SLS with existing very large scale integrated (VLSI) circuits of silicon makes it a particularly attractive and relevant structure to study (*cf.* People, 1986). In this thesis, an experimental study of hole transport in double barrier tunneling structures (DBTS) of $Si/Si_{1-x}Ge_x/Si$ is reported. Whereas

resonant tunneling of holes has been demonstrated experimentally in these structures, the non-resonant tunneling processes of holes has not been studied in any p-type system be it group IV or III-V. In this report the transport properties of holes are examined as a function of magnetic field applied either parallel or perpendicular to the plane of the layers. The evolution of features in the I-V characteristic with the applied magnetic field provides spectroscopic information on inter-subband transitions of holes from the contact to the quantum well. Non-resonant tunneling processes are reported for the first time in p-type DBTS and attributed to various scattering mechanisms.

An introduction to two-dimensional (2D) electron systems will be given in Chapter II and in particular, the confinement of charge carriers by an electric potential and the quantization of their motion by a magnetic field will be discussed. Theoretical aspects of resonant tunneling are treated in Chapter III and applied to the study of double barrier structures in the presence of a magnetic field. Chapter IV describes the properties of the Si/SiGe strained layer heterostructure beginning with a brief review of the basic properties of these materials and proceeding to a description of the Si/SiGe/Si DBTS. The fabrication and processing of the samples used in this work is outlined in Chapter V and the details of the experimental set-up are presented. Experimental results are presented and discussed in Chapter VI. Conclusions are given in Chapter VII.

This thesis presents the first observation of non-resonant tunneling of holes in $Si/Si_{1-x}Ge_x/Si$ double barrier tunneling structures. The spectroscopy of the quasi-bound states in the quantum well is performed by the application of a magnetic field at low temperature. The interpretation of the non-resonant tunneling features is corroborated by experimental results taken at temperatures up to 77 K.

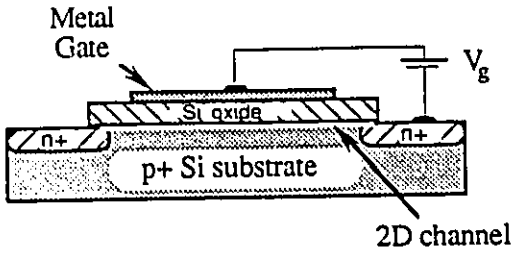
II. Two-dimensional Electronic Systems in Semiconductors

2.1 Introduction

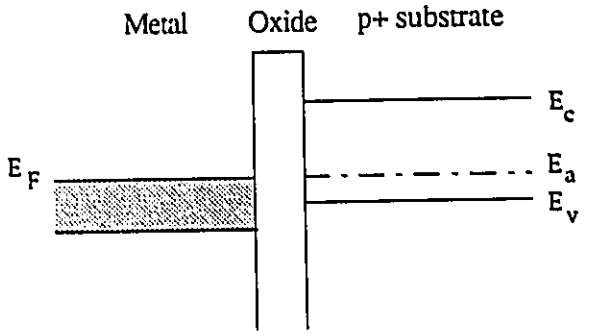
A two-dimensional electron gas (2DEG) is formed at any semiconductor interface where the motion of the charge carriers is free in two spatial dimensions but confined in the third. This confinement arises due to spatial variations in the potential energy of the conduction and valence bands which result in the formation of a potential well. The pioneering work on 2D electron systems was performed by Fowler *et al.* (1966) on Si-MOSFETs (Silicon Metal-Oxide Field Effect Transistor) which are easily manufactured and allow for a wide range of the electron concentration by variation of the gate voltage. However, the interface roughness between the silicon and oxide layers and conduction via several partially degenerate valleys result in poor electron mobility which limited their speed of operation. These limitations spurred research into advanced growth techniques such as MBE and MOCVD (metal-organic chemical vapour deposition) which allow materials of differing composition to be deposited epitaxially on a substrate (*cf.* Chang and Ploog, 1985 for a treatise on growth techniques). The resultant high-quality interfaces and spatially-controlled doping of these heterostructures has led to phenomenal increases in electron mobilities, and the study of these structures has become one of the most exciting fields in solid-state physics.

2.2 The Silicon MOSFET

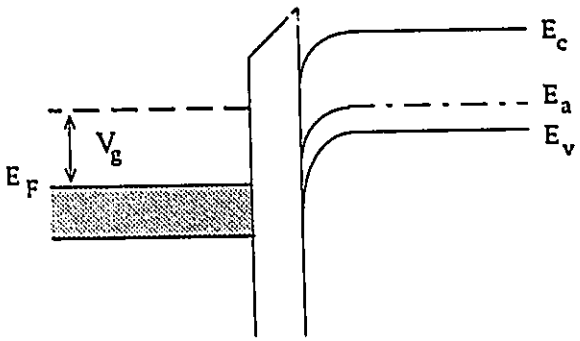
The initial experimental study of two-dimensional electronic systems began with the Si-MOSFET. The electronic properties of this structure (*cf.* Ando *et al.*,1982 for a definitive review) form the foundation of all 2D electronic systems. As illustrated in the schematic of Figure 2.1a, the n-channel Si-MOSFET consists



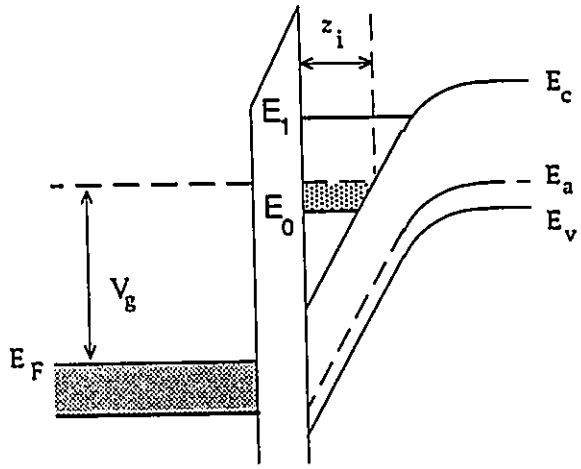
a) the Si-MOSFET. Current flows between the n+ contacts when a 2D electron gas forms in a narrow channel at the interface.



b) band edge profile with the gate voltage Vg = 0. The Fermi level is pinned at the acceptor level Ea.



c) the application of a positive voltage at the gate results in the bands bending downward at the interface



d) for sufficiently large Vg the conduction band edge lowered beyond the Fermi level and electrons accumulate at the interface. The spatial confinement results in the formation of a 2DEG.

Fig. 2.1 A schematic of the Si-MOSFET and the band edge profiles under various bias conditions.

of a p^+ doped silicon substrate separating two rectifying n^+ doped contacts. The gate electrode and substrate form the plates of a capacitor where the intermediate oxide layer serves as an insulating dielectric, and the voltage applied across this capacitor determines the operation of the device. As shown in Figure 2.1b, the band edge profiles in the absence of a gate voltage are flat and no current can flow between the contacts. When a positive bias is applied to the gate, electrons induced to the Si/SiO_2 interface recombine with holes, forming a depletion layer where the bands bend downward (Figure 2.1c). As the bias is further increased the conduction band edge is bent below the Fermi level E_F, and electrons accumulate in the triangular potential well at the interface (Figure 2.1d). This results in the formation of an inversion layer where the electron density is greater than the hole density in the bulk. Since the spatial extent of this layer ($z_i \sim 50$ Å) is comparable to the de Broglie wavelength of the electrons at the Fermi energy, the motion of the electrons normal to the interface is quantized into energies E_0, E_1, ... which are the eigenstates of the potential well. In general only the lowest level is occupied and the electrons have the normal free dispersion of Si in the interface plane; therefore the charges in the accumulation layer form a two-dimensional electron gas. Because of the multiple valley band structure of Si and the high longitudinal and transverse effective masses ($m_L = 0.9\, m_o$ and $m_T = 0.2\, m_o$), commercial production Si-MOSFETs generally have a poor mobility and a slow saturation speed which precludes high frequency oscillator and fast circuitry applications.

2.3 Semiconductor Heterojunctions

The quest for faster, higher-performance devices led to the proposal by R. L. Anderson (1960) of fabricating accumulation layers at the interface between two semiconductors of differing composition. The formation of a 2DEG is then

dependent on the parameters of the constituent materials such as the work functions, electronic affinities and band gap energies. As illustrated in Figure 2.2 for an interface between two different semiconductors, the difference in electron affinities χ_1 and χ_2 results in offsets in the conduction ΔE_c and/or valence bands ΔE_v. Additionally, if the work functions of the two materials W_1 and W_2 are not the same, charge transfer occurs across the interface so as to maintain a constant Fermi level. If the transfer of charge is substantial the band bending may be sufficiently strong to result in quantization of the electronic states in the well (*e.g.* E_0 and E_1 subbands).

Whereas in Si MOSFETs the carrier density in the inversion layer can be varied by the bias applied to the gate, in heterojunctions the carrier concentration is predetermined during the growth process by the level of doping in the layers. A spatial separation of the dopant atoms from the 2DEG can be effected during growth by means of modulation doping, *i.e.* doping only the larger bandgap material. Generally an undoped spacer layer of large bandgap material is grown next to the interface to separate the charge carriers from the tail of the long range impurity potential of the dopant atoms, as illustrated in Figure 2.3. Here the potential well is not formed at the interface but is due to the actual compositional change of one thin layer between others of larger bandgap. The phenomenal increase of the mobility at low temperature is linked to the reduced diffusion of ionized impurities with respect to the bulk. In this manner the 2DEG in heterostructures is free of the mobility-limiting interface roughness and impurity scattering encountered in the Si-MOSFET.

A wide assortment of semiconductor alloys can be selected for the formation of a heterojunction, resulting in a variety of heterostructure combinations. As illustrated in Figure 2.4, these compounds fall into sets characterized by their lattice constant and bandgap energies. A heterojunction between a given pair of alloys may

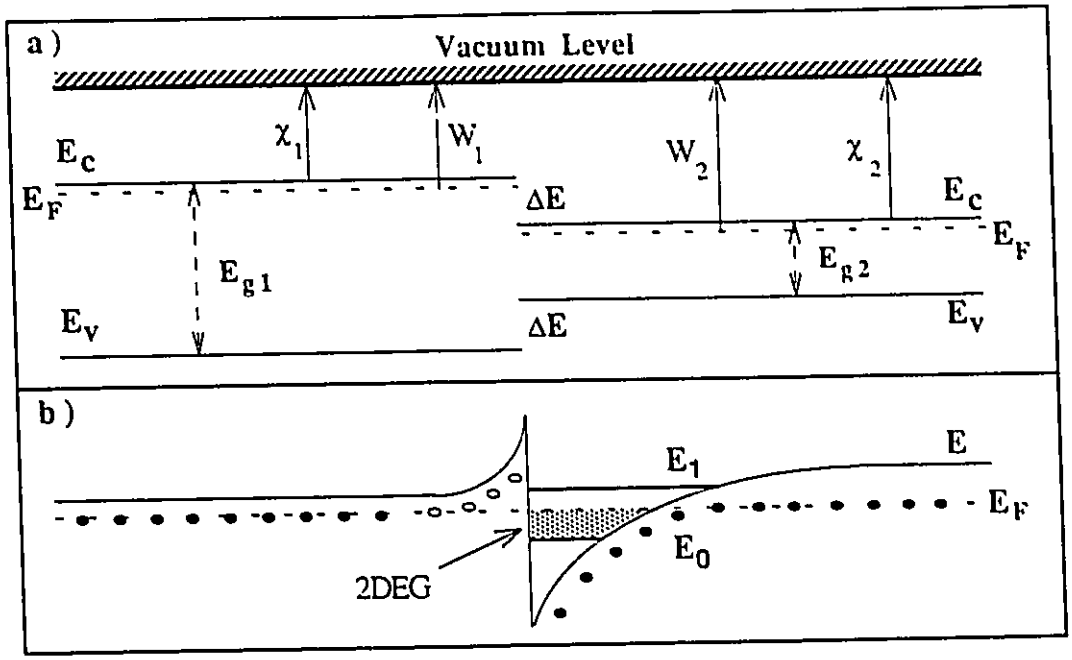


Fig. 2.2 a) The relationship between the electron affinities χ, work function W, Fermi level E_F, and band edges E_c, E_v is shown prior to the formation of the interface.

b) The energy levels near the conduction band are shown after the formation of a heterojunction. The dots represent donor levels and the hollow circles denote the ionized impurities. Electrons occupy states up to E_F with quantized momenta in the z-direction of energy E_0.

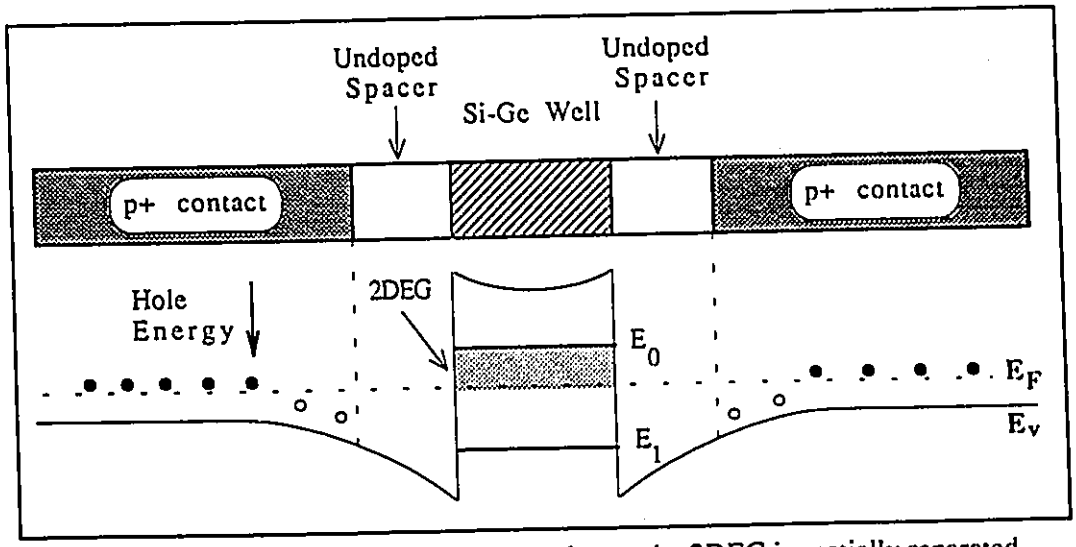


Fig. 2.3 By introducing undoped spacer layers the 2DEG is spatially separated from the impurities. Here the system is p+ doped, and the charge carriers are holes. The convention is to plot the energy of the holes as increasing downwards, and the energy zero is usually taken at the valence band edge E_v.

then be classified into four categories according to their band edge alignments, as illustrated in Figure 2.5. A Type I heterojunction (*e.g.* GaAs/AlGaAs, Si/SiGe), sometimes referred to as normal, is formed when the bandgap of one material lies entirely within that of the other. The electron or hole gases are confined in the lower bandgap material. In a Type II heterojunction both the conduction and valence band edges of the smaller bandgap material lie below their counterparts in the other material. Depending on the doping, the electron and hole gases are confined in adjacent layers. In Type II misaligned heterojunctions (*e.g.* InAs/GaSb) both the valence and conduction bands of one material lie above the conduction band of the other. An intrinsic charge transfer occurs across the interface (*i.e.* electrons from the GaSb valence band are transferred to the conduction band of InAs) leading tc semi-metallic behaviour. A Type III heterojunction (*e.g.* CdTe/HgCdTe) is formed when one of the materials forming the interface has a zero bandgap due to symmetry inversion at one of the symmetry points of the Brillouin zone.

2.4 Electronic Quantization

When particles are confined to within a spatial extent z_i which is comparable to their de Broglie wavelength, quantum effects become important. For electrons at the Fermi level in semiconductor heterostructures, this implies that $z_i \sim 100$ Å. The electronic states must then be described in quantum mechanical rather than classical terms. Assuming that the potential energy variation in the well is small within a de Broglie wavelength, the Schrödinger equation can be written in the effective mass approximation as

$$\left[\frac{p_x^2 + p_y^2}{2m^*} + \left\{\frac{p_z^2}{2m^*} + V(z)\right\}\right] \psi(\vec{r}) = E\,\psi(\vec{r}) \qquad (2.1)$$

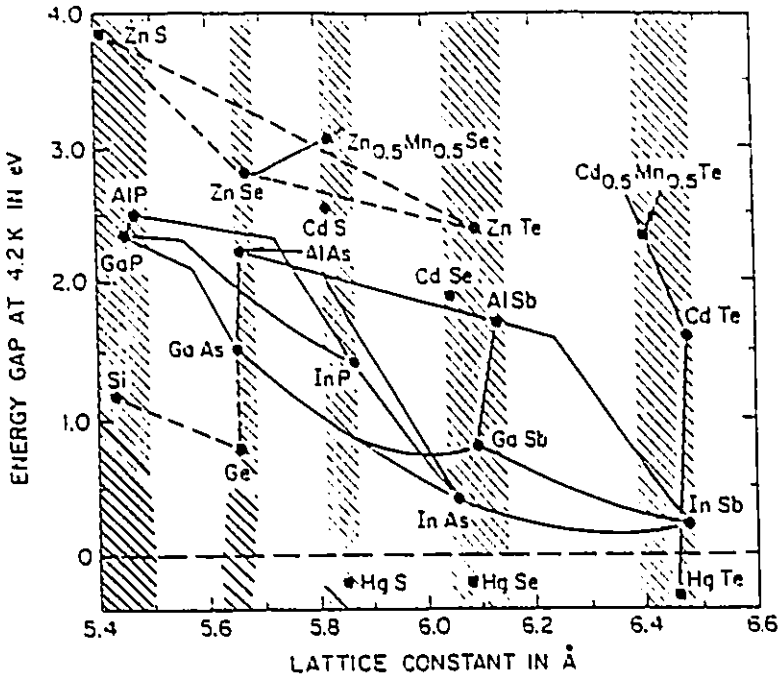


Fig 2.4 The bandgap energies of some semiconductor heterostructure alloys. The shaded regions indicate those compounds which are lattice-matched.

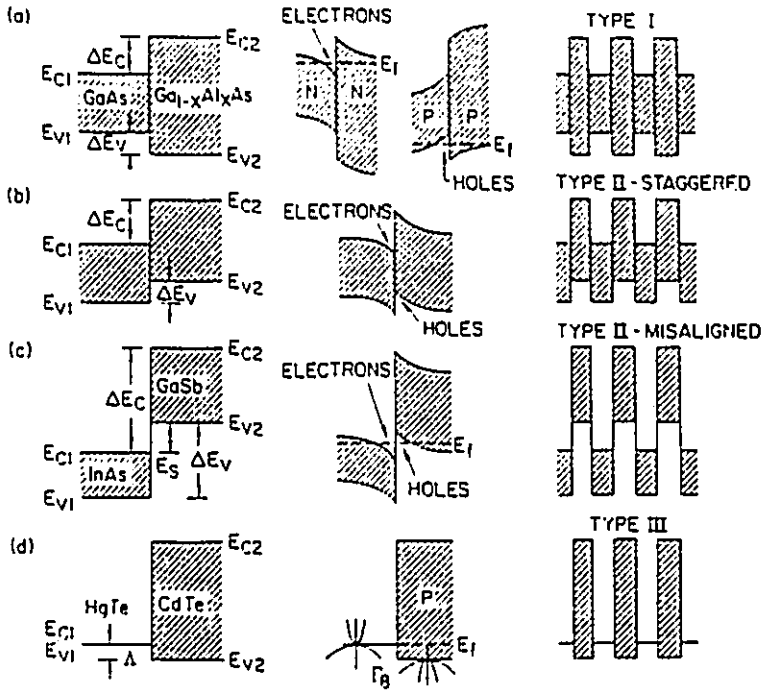


Fig. 2.5 The four classifications of heterojunctions showing, from left to right, the band edge alignments with undoped layers, doped layers and undoped superlattices. The bandgap energies are shown by the shaded regions and therefore indicate the potential barriers.

for an isotropic effective mass and a parabolic band. The periodic potential of the ionic cores and the screening due to electron-electron interactions is contained in the expression for the effective mass, and the confinement potential is given by $V(z)$. In the free electron limit (in the plane normal to the quantization axis), the electronic wavefunctions can thus be written as

$$\psi(\vec{\mathbf{r}}) = \varphi_i(z) \exp\left[i\left(k_x x + k_y y\right)\right] \tag{2.2}$$

where $\varphi_i(z)$ is the eigenfunction associated with the quantized level in the well, and the free motion in the x–y plane is described by plane waves. Since the spatial extent of the eigenfunctions $\varphi_i(z)$ is not zero, strictly one should refer to a *quasi* or *dynamically* two-dimensional electronic system.

The energy band is then quantized into a series of parabolic subbands, and the eigenenergy E_i associated with the confinement in the z direction defines the energy bottom of the ith subband. The solutions for the eigenenergies depend on the form of the confining potential. Two important cases are examined below:

2.4.1 Infinite square well

The simplest form of the potential for which analytical solutions are readily obtained is the infinite square well, which is the first approximation to the potential well formed in a semiconductor layer sandwiched between two layers of larger bandgap material. The potential is illustrated in Figure 2.6a and can be expressed in the form

$$V(z) = \begin{cases} 0 \ \textit{when} \ -\dfrac{d}{2} \le z \le \dfrac{d}{2} \\ \infty \ \textit{when} \ \ \dfrac{d}{2} \le |z| \end{cases} \tag{2.3}$$

The solutions to Equation (2.1) are plane waves with eigenenergies

$$E_i = \frac{(\pi\hbar)^2}{2m^*d}(i+1)^2 \qquad i = 0, 1, 2, \ldots \tag{2.4}$$

The solutions of the finite well are more complicated due to the penetration of the wavefunctions into the barrier, but the eigenenergies can be obtained by transcendental solutions of the Schödinger equation (*cf.* Cohen-Tannoudji *et al.*, 1976).

2.4.2 Asymmetric triangular well

As a first approximation, the potential well formed at a heterojunction interface can be modelled by an infinite triangular potential well where the electric field F_s can be expressed in the form

$$V(z) = \begin{cases} eF_s z & when\ z > 0 \\ \infty & when\ z \le 0 \end{cases} \tag{2.5}$$

and $F_s = \dfrac{e\left(N_d + fN_s\right)}{\varepsilon}$ *where* $\begin{cases} f = 1/2 \text{ is the average electric field} \\ f = 1 \text{ is the field at the interface} \\ f = 0 \text{ is the depletion charge term} \end{cases}$

where e is the charge on the electron, ε is the dielectric constant, N_s is the charge density in the well and N_d is the depletion charge density which takes into account the charge required to recombine with impurities near the well. The eigenenergies can be obtained from a solution of the semi-classical Bohr-Sommerfeld quantization condition (*cf.* Zawadzki, 1984)

$$E_i = \left(\frac{\hbar^2}{2m^*}\right)^{1/3}\left[\frac{3\pi}{2}eF_s\left(i+\frac{3}{4}\right)\right]^{2/3} \qquad i = 0, 1, 2, \ldots \tag{2.6}$$

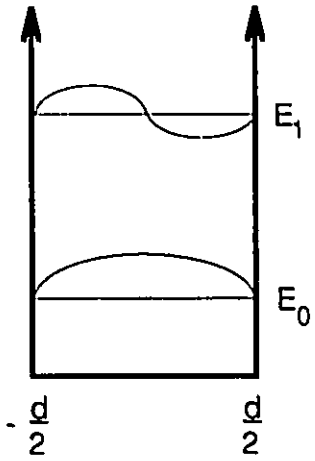
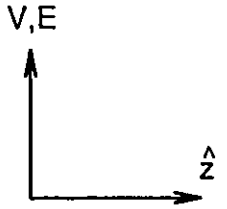


Fig. 2.6a The infinite square well potential, showing the first two bound states and the associated eigenfunctions.



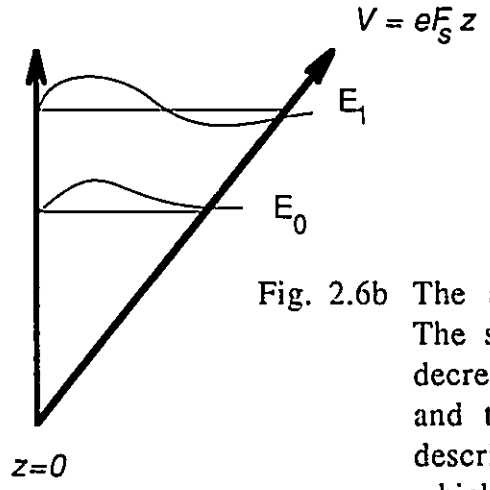


Fig. 2.6b The asymmetric triangular well. The spacing of the energy levels decreases with increasing energy, and the eigenfunctions are described by Airy functions which are 0 for z=0 and are finite for the linear potential forming the right wall of the well.

Fig. 2.6 The potential profiles of the square and asymmetric triangular wells.

The eigenfunctions can be described by taking appropriate parts of the Airy function A_i, which vanish for $z \leq 0$:

$$\varphi_i(z) = A_i\left[\left(\frac{2m^*}{\hbar^2}eF_s\right)^{1/3}\left(z - \frac{E_i}{eF_s}\right)\right] \qquad i = 0, 1, 2, \ldots \qquad (2.7)$$

The eigenenergies E_i in the Airy function are determined by the boundary condition $\varphi_i(0) = 0$ and only asymptotically approach Equation (2.6) for i >>0. Therefore, in practice it is easier to use variational functions with the proper symmetry to model the wavefunctions.

The average position of the 2DEG is then $z_i = 2E_i/3eF_s$. Although the ground state is approximated to within 10% by this calculation, the position of the second level depends on band bending via residual doping of the material which results in a potential that is not modelled properly by a simple triangular well. As well, a self-consistent calculation from an estimated potential energy must be used to account for the finite interface potential and the penetration of the wavefunctions into the barrier (Price and Stern 1983; Stern and Das Sarma, 1984).

2.4.3 Two-dimensional density of states

The 2D density of states $D_{2D}(E)$ is independent of the energy for a parabolic band and is written as

$$D_{2D}(E) = \frac{m^*}{\pi\hbar^2} \qquad (2.8)$$

for each subband. As illustrated in Figure 2.7, the total density of states for i subbands is a staircase function which tends towards the 3D density of states for a large number of subbands.

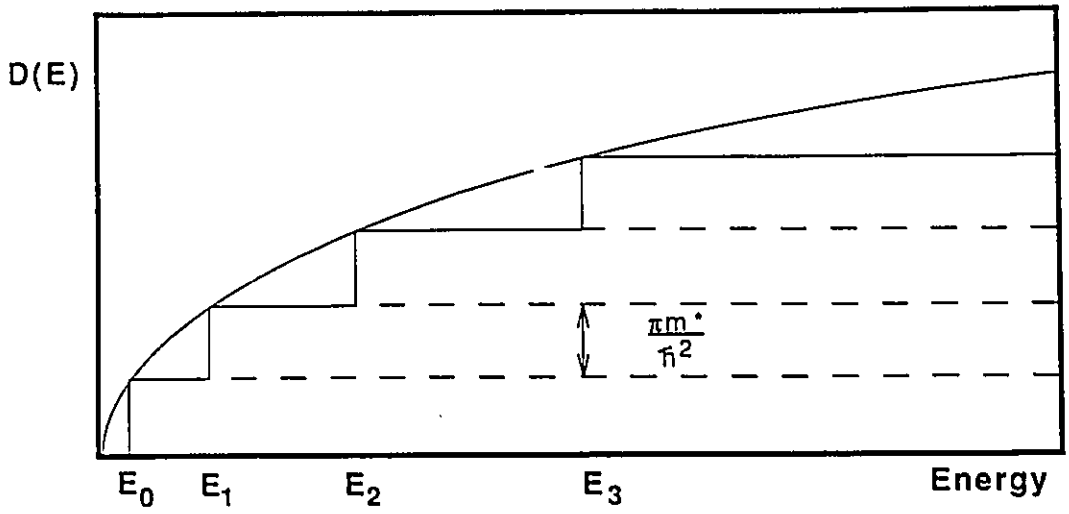


Fig. 2.7 The DOS of a 2D system is a series of steps at the quantized energies E_i. The 3D curve shown for comparison for a layer of width L and follows a $E^{1/2}$ dependence. Note that the 2D and 3D cases are identical at the quantized energies E_i.

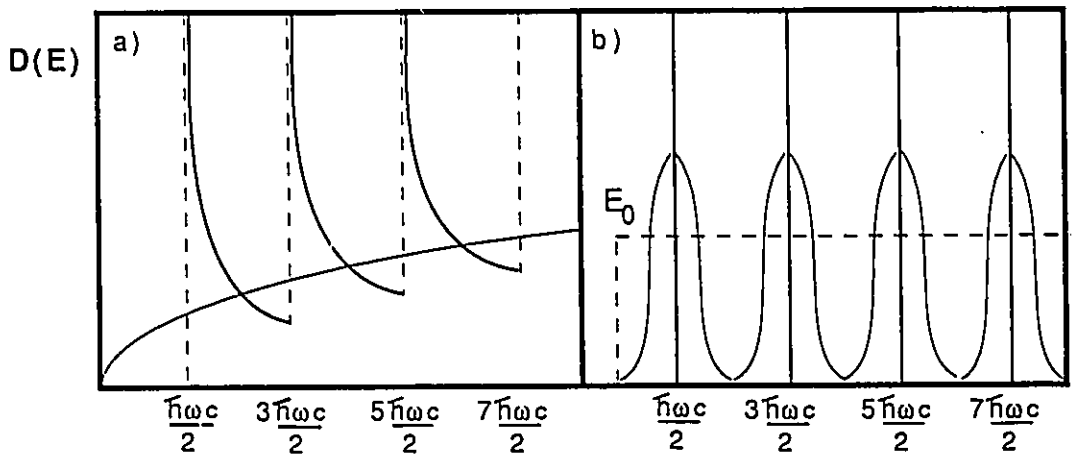


Fig. 2.8 The density of states under a normal magnetic field for $\frac{\hbar\omega_c}{2}$

a) the 3D system. The DOS is discontinuous at the Landau energies Shown for reference is the zero field DOS. Thermal broadening results in a rounding off of the discontinuities.

b) the 2D system with the field applied normal to the plane. The DOS is a series of delta functions for each subband Ei, and separated by the cyclotron energy. Spin-splitting would result in each delta function splitting into two delta functions centered at the original energy values. Thermal broadening results in the oscillatory form of the DOS which is shown as a Gaussian. The zero field DOS of a single subband is indicated by the dashed line.

$$D_{3D}(E) = \frac{1}{2\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{3/2} E^{1/2} \tag{2.9}$$

The steps occur at the quantized energies E_i, which define the bottom of each subband. In physical systems, disorder in the system results in level broadening which smooths out the discontinuities in the DOS.

2.5 Magnetic Quantization

2.5.1 *Three dimensions*

Under the influence of a strong magnetic field, electrons in a 3D volume will follow a helical trajectory whose axis is parallel to the field direction, and the motion of the charge carriers in the plane normal to the field is quantized into discrete cyclotron orbits. In the effective mass approximation the free electron Hamiltonian in a field **B** applied along the z direction is written as

$$\mathrm{H} = \left(\frac{1}{2m^*}\right)\left(\mathbf{p} + e\mathbf{A}\right)^2 \tag{2.10}$$

where **p** is the momentum and **A** the vector potential. It is convenient to choose the vector potential in the asymmetric Landau gauge as

$$\mathbf{A} = (-By, 0, 0) \tag{2.11}$$

such that $\mathbf{B} = \nabla \times \mathbf{A} = (0,0,B)$ as required. This choice of gauge results in the Hamiltonian taking the form of that for an off-centered harmonic oscillator:

$$\mathrm{H} = \left(\frac{1}{2m^*}\right)\left[\left(p_x - m^*\omega_c y\right)^2 + p_y^2 + p_z^2\right] \tag{2.12}$$

where $\omega_c = eB/m^*$ is the cyclotron frequency. In this form the variables can be separated and the solutions in the x and z directions are plane waves. For a system with areal dimensions L_x and L_z and infinite in the y direction, the eigenfunctions are therefore given by:

$$\psi_{n,k_x,k_z}(\vec{r}) = \left(L_x L_z\right)^{-1/2} \chi_n(y - y_o) \exp\left[i\left(k_x x + k_z z\right)\right] \qquad (2.13)$$

where the normalized wavefunctions χ_n are those of a harmonic oscillator centered about $y_o = \hbar k_x/(m^*\omega_c) = r_o^2 k_x$ where r_o is the classical radius of the cyclotron orbit in the ground state $r_o = (\hbar/eB)^{1/2}$. The eigenvalues are then

$$E_{n,k_z} = \left(n + \frac{1}{2}\right)\hbar\omega_c + \frac{\hbar^2 k_z^2}{2m^*} \qquad n = 0, 1, 2,... \qquad (2.14)$$

The circular motion in the x–y plane is described by the first term while the motion parallel to the magnetic field is given by the second term. In three dimensions, the dispersion relation will be given by a set of n parabolas equally spaced in energy by $\hbar\omega_c$ referred to as Landau levels. The degeneracy of each Landau level is the total number of states with a given n and k_z value and is denoted by $\nu = m^*\omega_c/2\pi\hbar = eB/h$ per unit area and per spin orientation.

The total density of states is therefore:

$$D(E) = \left(\frac{2m^*}{\hbar^2}\right)^{3/2} \frac{\hbar\omega_c}{4\pi^2} \sum_{n=0}^{m} \left[E - \left(n + \frac{1}{2}\right)\hbar\omega_c\right]^{-1/2} \qquad (2.15)$$

and in comparing this with the zero field density of states $D_o = \left(\frac{1}{2\pi^2}\right)\left(\frac{2m^*}{\hbar^2}\right)^{3/2} E^{1/2}$, it is observed that the magnetic field introduces a set of regularly spaced divergences as illustrated in Figure 2.8a.

2.5.2 *Two dimensions*

In a two-dimensional system limited to the x-y plane the motion of the charge carriers is affected only by the component of the field along the z direction. When a magnetic field is applied in the z direction the wavefunction becomes:

$$\psi_{n,i,k_y} = L_y^{-1/2} \chi_n(x - x_o)\, \varphi_i(z) \exp\left(ik_y y\right) \tag{2.16}$$

where n denotes the Landau level index and i the index of the quantized energy level in the well. The eigenfunctions $\varphi_i(z)$ depend on the form of the confining potential (*cf.* §2.4) and have associated eigenenergies labelled E_i. The eigenenergies of the total Hamiltonian are

$$E_{n,i} = \left(n + \frac{1}{2}\right)\hbar\omega_c + E_i \tag{2.17}$$

and the 2D density of states for each subband E_i is a series of delta functions separated by the cyclotron energy $\hbar\omega_c$,

$$D_{2D}(E) = \left(\frac{2eB}{\hbar}\right) \delta\left(E - E_{i,n}\right) \tag{2.18}$$

This is illustrated in Figure 2.8b for the lowest energy subband E_0.

In a strong magnetic field the spin degeneracy of the Landau levels will be lifted and the energy separation will be given by $g^*\mu_B B$ where g^* is the Landé factor of the charge carrier in the material and $\mu_B = e\hbar/m_o g_o$ is the Bohr magneton, with the free electron mass m_o and free electron g factor $g_o = 2$. The eigenenergies are therefore

$$E_{i,n,\sigma} = E_i + \left(n + \frac{1}{2}\right)\hbar\omega_c + \sigma g^* \mu_B B \tag{2.19}$$

where $\sigma = \pm 1/2$ corresponds to the two spin orientations. Each of the delta function peaks in Figure 2.8b is further split into two delta functions separated by $2\sigma g^*\mu_B B$. The 2D density of states can therefore be written in the form

$$D_{2D}(E) = \left(\frac{2eB}{\hbar}\right) \delta\left(E - E_{i,n,\sigma}\right) \tag{2.19}$$

In general, this spin-splitting is small compared to the cyclotron energy.

The above results were derived with the assumption that the system is in the limit of zero temperature. At finite temperatures however, the charge carriers can interact with impurities, defects and phonons of the material, thereby broadening the Landau levels and damping the delta function singularities of the density of states. As illustrated in Figure 2.8b this leads to an oscillatory form of the 2D DOS, and the amplitude of the oscillations is dependent on the amount of broadening in the system. For the system to remain in the quantum limit, the Landau level broadening must be smaller than the cyclotron energy and the quantization conditions are usually written as:

$$\begin{gathered} \omega_c\tau \ll 1 \quad or \quad \mu B \gg 1 \\ \hbar\omega_c \ll kT \end{gathered} \tag{2.20}$$

where τ is the time between scattering events, μ is the mobility and kT is the mean thermal energy.

III. Resonant Tunneling in Double Barriers

3.1 Introduction

Tunneling through a potential barrier is a fundamental result of quantum mechanics and its experimental observation was a confirmation of the wavelike properties of matter. The concept of tunneling has been used extensively in condensed matter physics: from field emission in metals by Fowler and Nordheim in 1928 to the scanning tunneling microscope by Binnig and Rohrer in 1982. In 1974, Chang, Esaki and Tsu were the first to observe resonant tunneling in a double barrier heterostructure. In this Chapter, an introduction to the physics of resonant tunneling will be given and tunneling processes in an applied magnetic field will be discussed. Excellent review articles have been written on tunneling by Duke (1969) and resonant tunneling by Mendez (1987).

3.2 Tunneling Current

Electrons tunneling in a semiconductor heterojunction are subject to the periodic potential of the ionic cores which is screened by electron-electron interactions. The Hamiltonian describing this system is complicated due to the many-body interactions, and is made manageable by using one-electron states which take into account the periodic component of the electron-ion potential. The potential barrier is described by a general potential $V(z)$ illustrated in Figure 3.1a and the charge distribution at the barrier interface is treated as a slowly varying potential governed by Poisson's equation. Further simplifications include:

- the expansion of the wavefunctions in terms of a single band on each side of the interface,

- the separation of Schrödinger's equation into components parallel and perpendicular to the plane of the interface, and

- specular tunneling meaning that both total energy and momentum parallel to the interface are conserved.

The problem is then described by the one-dimensional Schrödinger equation and the electrostatic potential near the interface $V(z)$:

$$\left\{ -\frac{\hbar^2}{2m^*}\frac{d^2}{dz^2} + V(z) \right\} \psi(z) = E\,\psi(z) \tag{3.1}$$

with boundary conditions set by the conservation of current at the interface. The current per unit area can be expressed in terms of the tunneling probability $T(E_z)$ and the Fermi-Dirac distribution $f(E)$ as follows:

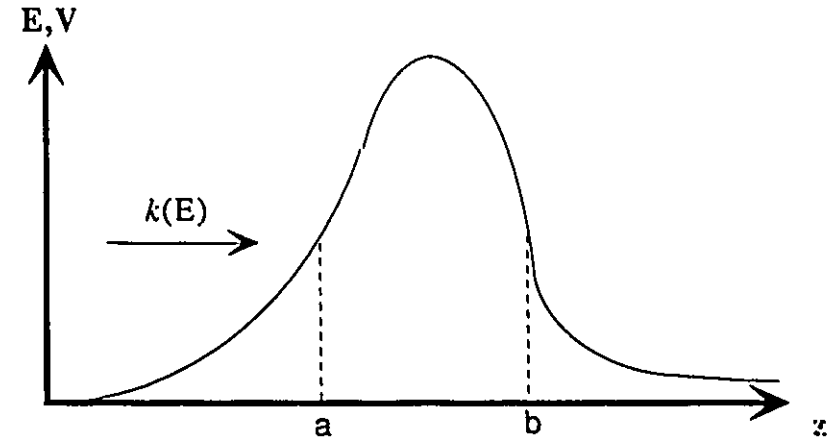
$$\mathrm{J} = \left(\frac{e}{4\pi^3\hbar}\right) \int dk_z\, d^2k_{//}\, f(E)\, T(E_z) \left(\frac{\partial E}{\partial k_z}\right) \tag{3.2}$$

When the barrier is subjected to an external bias, the net current through the barrier is the difference between currents flowing from left to right and from right to left:

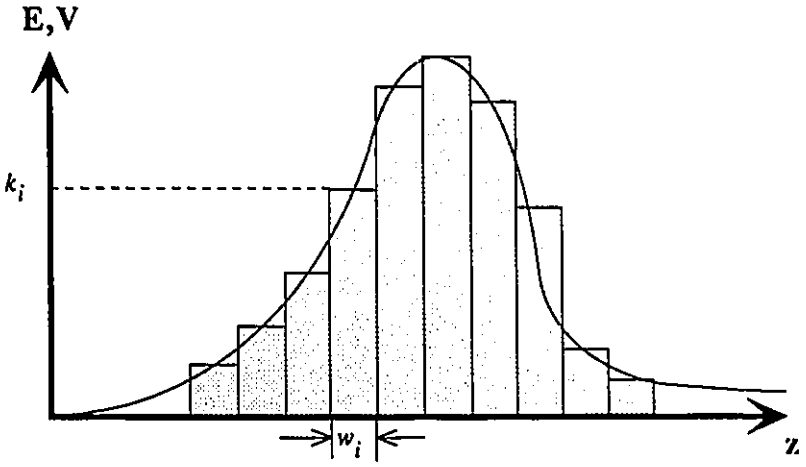
$$\begin{aligned} \mathrm{J} &= \mathrm{J}_{L\to R} - \mathrm{J}_{R\to L} \\ &= \left(\frac{e}{4\pi^3\hbar}\right) \int dE_z\, d^2k_{//} \Big[f(E) - f(E+eV) \Big]\, T(E_z) \end{aligned} \tag{3.3}$$

In the limit of zero temperature and the free electron approximation

$$d^2k_{//} = \frac{2\pi m^* dE_{//}}{\hbar^2}$$



a) a general potential barrier. For a particle of energy E, the barrier limits are a and b.



b) The WKB approximation to the barrier, consisting of several thin rectangular barriers of width w_i. The tunneling probability is then proportional to the total attenuation given by the product of each barrier. In the limit of infinitessimal widths, the product becomes an integral over the range a to b.

Fig. 3.1 A general potential barrier

and Equation (3.3) can therefore be rewritten as:

$$\mathbf{J} = \left(\frac{e}{2\pi^2\hbar^3}\right)\left[eV\int_0^{E_F-eV} dE_z\, T\left(E_z\right) + \int_{E_F-eV}^{E_F} dE_z\left(E_F - E_z\right)\, T\left(E_z\right)\right] \quad \text{for} \begin{cases} eV \le E_F \\ \\ eV \ge E_F \end{cases} \tag{3.4}$$

$$\mathbf{J} = \left(\frac{em^*}{2\pi^2\hbar^3}\right)\int_0^{E_F} dE_z\left(E_F - E_z\right)\, T\left(E_z\right)$$

and this can be readily evaluated so long as the tunneling probability $T(E_z)$ can be determined from the Schrödinger equation.

Often however, the one-dimensional Schrödinger equation cannot be solved analytically for the potential of interest $V(z)$, such as the general barrier illustrated in Figure 3.1a. In these cases approximation methods such as the semi-classical WKB approximation can be used (Franz, 1969). The WKB approximation can be used to describe tunneling through a potential barrier provided that i) the change in potential energy within a de Broglie wavelength is small in comparison to the kinetic energy of the particle, and ii) the kinetic energy is small in comparison to the barrier height. The exact barrier profile can then be approximated by several thin rectangular barriers of constant potential as illustrated in Figure 3.1b, and the tunneling probability is therefore (Roy, 1986)

$$T \approx \exp\left[-2\int_a^b k(z)dz\right] \tag{3.5}$$

where

$$k(z) = \left(\frac{2m^*}{\hbar^2}\right)^{1/2}\left[V(z) - E\right]^{1/2} \tag{3.6}$$

and a and b are the coordinates of the left and right edges of the barrier respectively. The exponential dependence of the transmission represents an attenuation by the barrier: for increasing barrier heights and widths the transmission is decreased. The implicit dependence of the transmission probability on the particle energy is incorporated in the wavevector k, and indicates a monotonic increase in transmission for increasing energy.

When the potential is of the form of several barriers separated by potential wells, the total attenuation is just the product of that due to each barrier: there is no dependence of the transmission on the dimensions of the well regions. However, because of the semi-classical nature of the WKB method, no account can be made for reflection of the particle by the barriers. The reflected wavefunctions may interfere constructively in the well and destructively in the barrier regions, leading to a build-up of the wavefunction in the well. For certain resonant energies, it is possible for there to be no attenuation of the transmission probability and therefore the barriers will be effectively transparent.

3.3 The Transfer Matrix Method

To include the effects of interference between the wavefunctions, a fully quantum mechanical formulation must be invoked. The transfer matrix (or transfer Hamiltonian) method developed by Kane (1969) is most often used as it yields an analytical expression for the tunneling probability through any number of rectangular barriers of arbitrary height. The general potential is described by a series of regions of constant potential separated by discontinuous potential steps. The method then involves finding the eigenfunctions in each of the regions. Since the potential is constant, the solutions of the one-dimensional Schrödinger equation (3.1) are easily obtained, and can be expressed in the general form

$$\varphi(z) = A\exp(ikz) + B\exp(-ikz) \tag{3.7}$$

where the wavevector is defined by

$$\frac{\hbar^2 k^2}{2m^*} = E - V(z) \tag{3.8}$$

When E – V > 0, the wavevector k is real and the wavefunctions are left and right-going plane waves. When E – V < 0, k is imaginary; by re-writing the wavevector as $k = i\kappa$ where κ is real it is seen that Equation (3.7) describes exponentially growing and decaying waves. The coefficients (A, B) are determined by matching the wavefunctions at the interfaces between each of the regions. The matching conditions require that both φ and $\partial\varphi/\partial z$ be continuous across the potential step at each interface. These boundary conditions can be expressed in the form of a 2 x 2 matrix **R** acting on the coefficients of the wavefunction on the left and right of the potential step. Thus we can write

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \mathbf{R} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{3.9}$$

where **R** is

$$\mathbf{R} = \frac{1}{2k_1} \begin{Bmatrix} (k_1 + \kappa_2)\exp\left[i(-k_1 + \kappa_2)z_1\right] & (k_1 - \kappa_2)\exp\left[i(-k_1 - \kappa_2)z_1\right] \\ (k_1 - \kappa_2)\exp\left[i(k_1 + \kappa_2)z_1\right] & (k_1 + \kappa_2)\exp\left[i(k_1 - \kappa_2)z_1\right] \end{Bmatrix} \tag{3.10}$$

In general, for a potential profile of n regions, there will be n-1 matrices acting on the coefficients. Using Equation (3.9) these relations can be expressed in the form

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \left(\mathbf{R}_1 \mathbf{R}_2 \ldots \mathbf{R}_{n-1}\right) \begin{pmatrix} A_n \\ B_n \end{pmatrix} \tag{3.11}$$

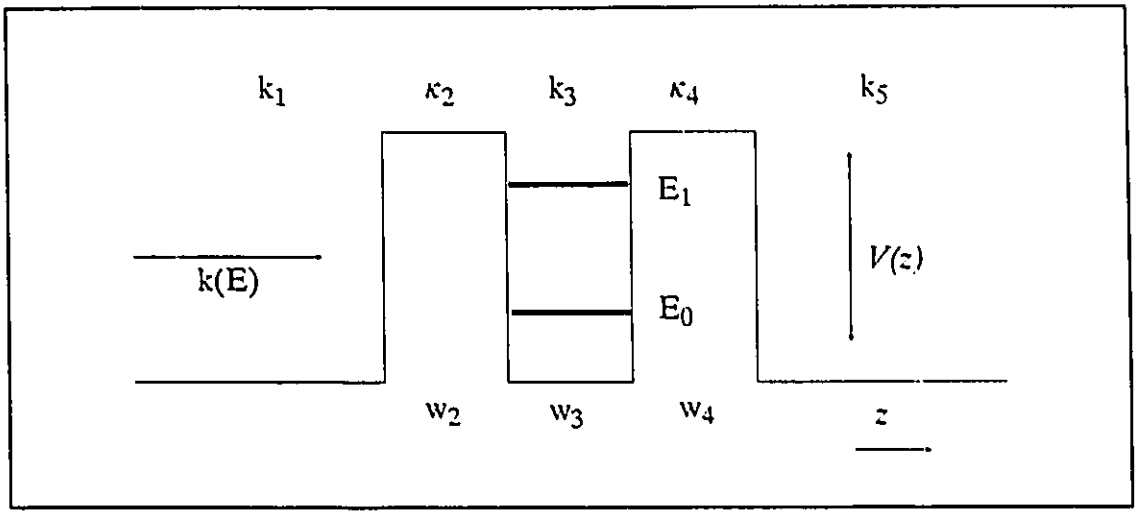


Fig. 3.2 The potential profile of the symmetric double barrier structure. The subcripts index the various regions, where w is the width of the region and k and κ are the wavevectors. E_0 and E_1 are the eigenenergies of the first two quasibound states of the well.

Consider the case of tunneling through the two symmetric barrier profile illustrated in Figure 3.2. There are 5 regions of constant potential and therefore four transfer matrices are required. Equation (3.11) is then written

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \left(\mathbf{R}_1 \mathbf{R}_2 \mathbf{R}_3 \mathbf{R}_4\right) \begin{pmatrix} A_5 \\ B_5 \end{pmatrix} \tag{3.12}$$

If the particle is incident from the left in region 1, only a transmitted wave will be observed on the right in region 5 and therefore $B_5 = 0$. The quantity of interest is the ratio of the transmitted current to the incident current, which is the transmission probability given by

$$T = \frac{j_5}{j_1} = \frac{k_5 m_1}{k_1 m_5} \cdot \left| \frac{A_5}{A_1} \right|^2 \tag{3.13}$$

where j_n is the current density, k_n is the wavevector and m_n is the effective mass in region n. By substituting Equation (3.10) for the explicit form of the matrices $\mathbf{R}_n$ in Equation (3.12), the transmission can be written in the form

$$\frac{j_t}{j_i} = \frac{2^8 k_1 \kappa_2^2 k_3^2 \kappa_4^2 k_5 / |K|^2}{(k_1^2 + \kappa_2^2)(\kappa_2^2 + k_3^2)(k_3^2 + \kappa_4^2)(\kappa_4^2 + k_5^2)} \tag{3.14}$$

and the expression is dominated by the exponential factor

$$\begin{aligned} K = {} & \exp(\kappa_2 w_2 + \kappa_4 w_4)\left\{e^{i(-\phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5)} - e^{i(\phi_1 + \phi_2 - \phi_3 - \phi_4 + \phi_5)}\right\} \\ & + \exp(\kappa_2 w_2 - \kappa_4 w_4)\left\{e^{i(\phi_1 + \phi_2 - \phi_3 + \phi_4 - \phi_5)} - e^{i(-\phi_1 + \phi_2 + \phi_3 - \phi_4 - \phi_5)}\right\} \\ & + \exp(-\kappa_2 w_2 + \kappa_4 w_4)\left\{e^{i(\phi_1 - \phi_2 + \phi_3 - \phi_4 + \phi_5)} - e^{i(-\phi_1 - \phi_2 - \phi_3 + \phi_4 + \phi_5)}\right\} \\ & + \exp(-\kappa_2 w_2 - \kappa_4 w_4)\left\{e^{i(\phi_1 - \phi_2 - \phi_3 - \phi_4 - \phi_5)} - e^{i(-\phi_1 - \phi_2 + \phi_3 + \phi_4 - \phi_5)}\right\} \end{aligned} \tag{3.15}$$

where w_n is the width of region n and the phase factors ϕ_i are

$$\begin{aligned} \phi_1 &= k_3 w_3 \\ \phi_2 &= \tan^{-1}(\kappa_2/k_1) \\ \phi_3 &= \tan^{-1}(\kappa_2/k_3) \\ \phi_4 &= \tan^{-1}(\kappa_4/k_3) \\ \phi_5 &= \tan^{-1}(\kappa_4/k_5) \end{aligned}$$

It is clear that the first term in Equation (3.15) is the largest and represents an attenuation due to the barriers (regions 2 and 4). This term can be made to vanish if the phase factors are such that

$$\phi_1 = \phi_3 + \phi_4 + r\pi \qquad r = 0, 1, 2, \ldots \tag{3.16}$$

in which case the terms involving exp ($\kappa_2 w_2 + \kappa_4 w_4$) and exp ($-\kappa_2 w_2 - \kappa_4 w_4$) go to zero. Since the barriers are symmetric the terms in exp ($\kappa_2 w_2 - \kappa_4 w_4$) are unity, and there is no attenuation of the transmission. To interpret the condition given in Equation (3.16), one notes that when the potential in regions 2 and 4 is infinite, the phase factors ϕ_3 and ϕ_4 are $\pi/2$, so then

$$n\pi = \phi_1 = k_3 w_3 \qquad n=1, 2, 3, \ldots \tag{3.17}$$

and by squaring both sides, it is seen that this is just an expression for an eigenstate of the infinite well. Similarly for finite barriers, the expression corresponds to an eigenenergy of the finite well (Roy, 1986). Since for the symmetric barrier profile $V_1 = V_3$ and therefore $k_1 = k_3$, Equation (3.17) implies that the transmission is maximized when the energy of the incident wavefunction coincides with an eigenenergy of a bound state of the well. This is the condition for resonant tunneling, where the transmission is enhanced due to constructive interference of the wavefunctions in the well.

The transmission probability for electrons tunneling through a 50 Å well is illustrated in Figure 3.3a and shows a general trend of increasing transmission for increasing particle energy, similar to the results from a WKB calculation. However, it is noted that the transmission probability approaches unity at two specific energies, and away from these values falls off rapidly. These sharp peaks correspond to resonant tunneling via eigenstates of the well. Other more broad maxima are observed at higher energies, and these represent resonances with virtual states whose energies lie above the well.

If the two barriers are of different widths, resonances are still observed but the transmission is less than unity. This attenuation is observed because the

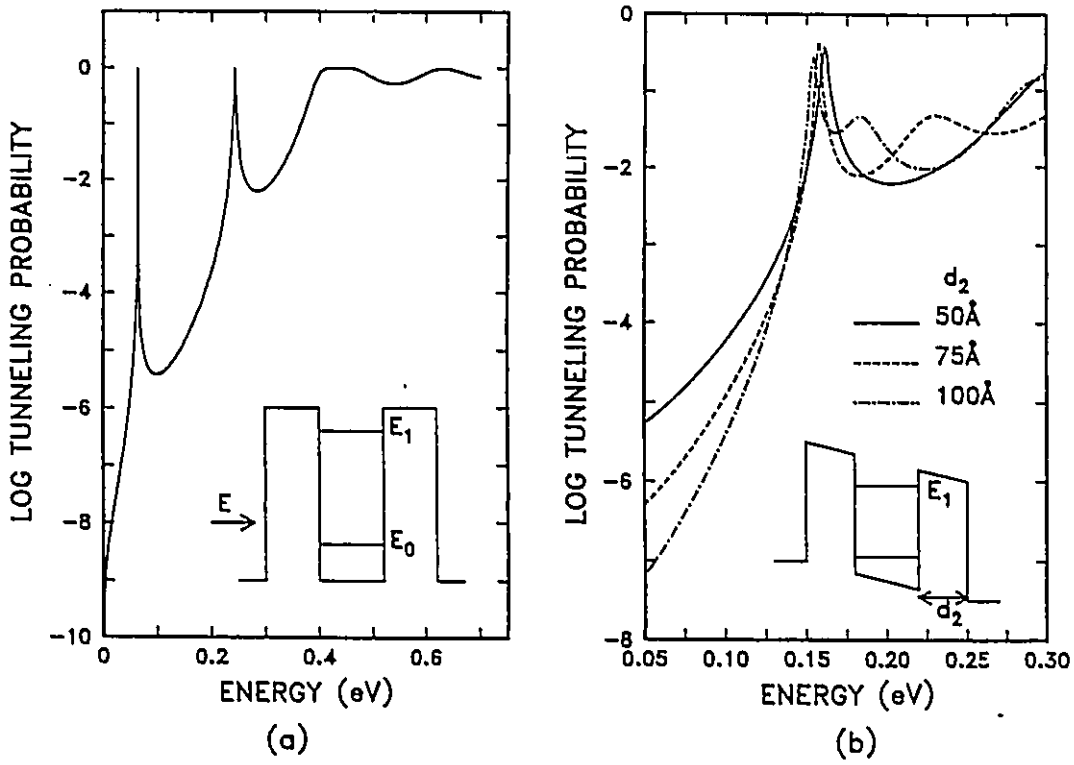


Fig. 3.3a A calculation of the transmission probability as a function of incident particle energy for a symmetric double barrier where the barrier height is 0.3 eV, barrier widths of 50 Å and a well width of 60 Å. The effective mass of the tunneling particle is taken to be 0.1 m_0 in the barrier regions and 0.067 m_0 in the well. Two resonances are clearly observed in the transmission probability, corresponding to tunneling via the quasi-bound states of the well with eigenenergies E_0 and E_1 The broad maxima at higher energies are due to tunneling resonances via virtual states which lie above the well.

Fig. 3.3b The tunneling probability through a double barrier subject to an electric field of 10^5 V/cm. The left barrier is set at a width of 50 Å and the transmission is calculated for exit barrier widths of 50, 75 and 100 Å. The other parameters are the same as in Fig. 3.3a. The peak at 0.16 eV corresponds to resonant tunneling via E_1 (E_0 is not shown). Optimum transmission is obtained for an exit barrier width of 75 Å, for which the system is effectively symmetric.

transparency of each barrier is different, and therefore the coefficient K is always greater than unity. Similar results can be obtained by the application of an electric field, which effectively destroys the symmetry of a system. The transfer matrix method, which was developed for regions of constant potential, can still be used when a potential gradient exists. In this case the true potential is approximated by a series of small potential steps, and the matrices are matched as previously. It has been shown that the method converges to the exact result very rapidly, so only a few steps are required (Mendez, 1986). The results of such a calculation are given in Figure 3.3b, which illustrates the effect of an electric field on the transmission.

3.4 Transit Time and Charge Accumulation

It was shown above that resonant tunneling occurs when the particle energy coincides with an eigenenergy of the well. If the barriers are infinitely wide, these energies correspond to the bound states of the well. For finite barriers however, the energies define only quasibound states: a particle localized in the well at time t = 0 will tunnel out of the well in some characteristic time τ, which is related to the tunneling probability at resonance. This transit time τ defines the lifetime of the quasibound state and, by the uncertainty principle, is also related to the energy width of the resonance ($\Delta E \sim \hbar/\tau$). From the WKB approximation, τ can be expressed in terms of the transmission probability by

$$\frac{1}{\tau} = \left(\frac{2E}{m^*}\right)^{1/2} \frac{T(E)}{2w} \tag{3.18}$$

where w is the width of the well. For typical well and barrier dimensions it has been shown that the transit time of a resonantly tunneling particle is of the order of 10^{-13} seconds (Mendez *et al.*, 1986). However, because the tunneling probability

varies exponentially with barrier dimension and particle energy, τ can also change over several orders of magnitude.

Within the transit time, a particle (represented by a wavepacket peaked at the resonant energy) will be reflected in phase several times within the well. The constructive interference between the waves leads to the build-up of the wavefunction in the well. If the tunneling particles are charged, this implies that an electrostatic charge will accumulate in the well. The charge density Q can be related to the tunneling time via

$$Q = \tau J \tag{3.19}$$

where J is the current density through the double barrier. In principle then, one could calculate the tunneling current density by knowing both the charge density in the well and the tunneling time.

3.5 Resonant Tunneling in Double Barrier Structures

The theoretical analysis above is quite general, and the results obtained are independent of the physical parameters of the materials forming the double barrier. In an experimental system however, it is well established that the details of the structure have a significant effect on the tunneling currents measured. The following analysis is specific to the tunneling current measured in a semiconductor double barrier structure. The well width is assumed to be sufficiently thin (of the order of less than 100 Å) so that the bound states are highly separated in energy, and the barrier heights, given by the conduction band edge discontinuities, are greater than 100 meV. The tunneling electrons are supplied from doped contact layers on either side of the barriers.

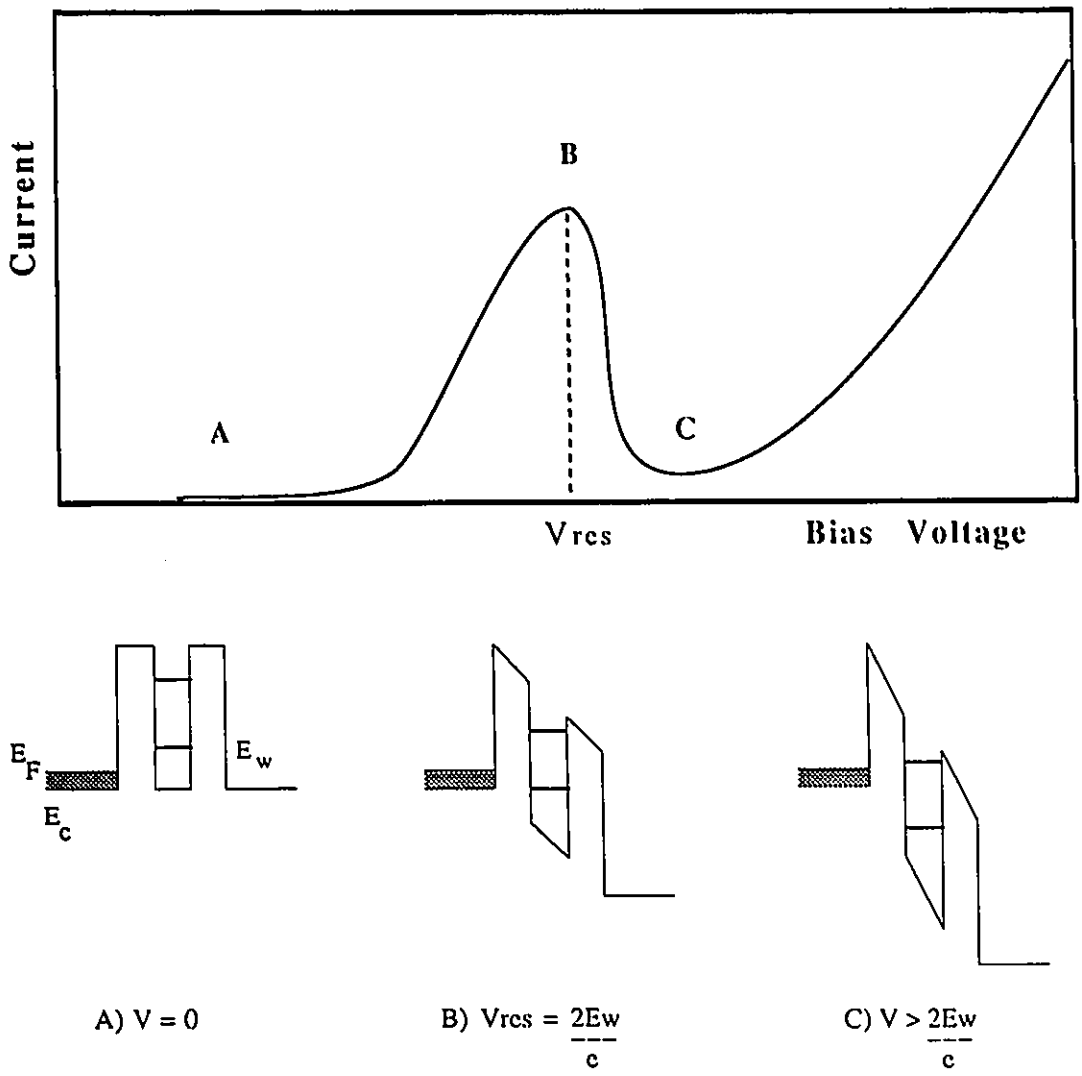


Fig. 3.4 A typical current-voltage characteristic of a double barrier tunneling structure. As the bias voltage brings the Fermi level into alignment with a bound state of the well, the current rapidly rises. The peak current is measured at point **B** where the conduction band edge is aligned with the bound state of the well. Beyond resonance the tunneling is limited by conservation of momentum. Successive peaks in the tunneling current are measured for each bound state of the well.

A typical current-voltage characteristic for a double barrier heterostructure is presented in Figure 3.4, along with potential profiles representative of various bias conditions. At zero bias, the bands are flat and no current is measured. Region A corresponds to non-resonant tunneling: electrons in the left electrode (emitter) tunnel through the structure and a small but finite current is measured. When the Fermi energy of the emitter is aligned with an eigenenergy of the well electrons are available to tunnel resonantly and a rapid increase in the current is measured. The current increases linearly until the conduction band edge is aligned with the eigenstate of the well (point B). If the potential is assumed to drop linearly across the barriers then the bias position of the resonant peak is

$$V_{res} = \frac{2E_w}{e} \tag{3.20}$$

where E_w is the energy of the eigenstate. Experimentally however, non-linear effects such as potential drops in the contact layers, space charge build-up at the interfaces and charge accumulation in the well all modify the potential. An extrapolation of the energy of the resonant state from the bias potential must then be made within an entirely self-consistent calculation which takes into account all these dynamic effects.

For an electron tunneling through the barrier without being scattered, planar momentum is conserved. When the bias raises the conduction band edge beyond an eigenstate, no electrons are available which can tunnel resonantly while conserving in-plane momentum (Ricco and Azbel, 1970) and therefore the current drops rapidly. This phenomenon is known as negative differential resistance (NDR), where an increase in bias voltage results in a decrease of the tunneling current. The peak-to-valley current ratio (points B to C on Figure 3.4) is a measure of the strength

of the resonance, and theoretical transmission probabilities predict extremely large peak-to-valley ratios (several orders of magnitude in Figure 3.3), with the valley current going to zero. In an experiment the peak-to-valley measured is much smaller, of the order of 10 depending on the material and other mechanisms must be considered along with resonant transmission. Thermal excitation of electrons across the structure and scattering of the electrons by impurities and interface roughness may result in non-resonant currents which contribute strongly to the background. The net effect of these contributions is a decrease in the peak-to-valley ratio.

The entire sequence detailed above is repeated for each of the quasibound states in the well. Thus the I-V characteristic is a series of triangular peaks in the current, followed by a final monotonic increase. The number of resonances observed in the current corresponds to the number of quasibound levels in the well, consistent with the barrier heights and well width.

3.6 Resonant Tunneling in a Magnetic Field Normal to the Layers

Most of the charge carriers contributing to the resonant current traverse the barriers without scattering and the planar component of the wavevector $k_{//}$ (perpendicular to the tunneling direction) is conserved. The energy of motion of charge carriers in the plane of the layers is quantized by a magnetic field **B** into discrete Landau levels (Stiles, 1984). The energies of the charge carriers in the contact and quantum well can be written as:

$$\begin{aligned} E_{em} &= E_0 + \left(m + \frac{1}{2}\right)\hbar\omega_c \\ E_{qw} &= E_1 + \left(n + \frac{1}{2}\right)\hbar\omega_c \end{aligned} \tag{3.21}$$

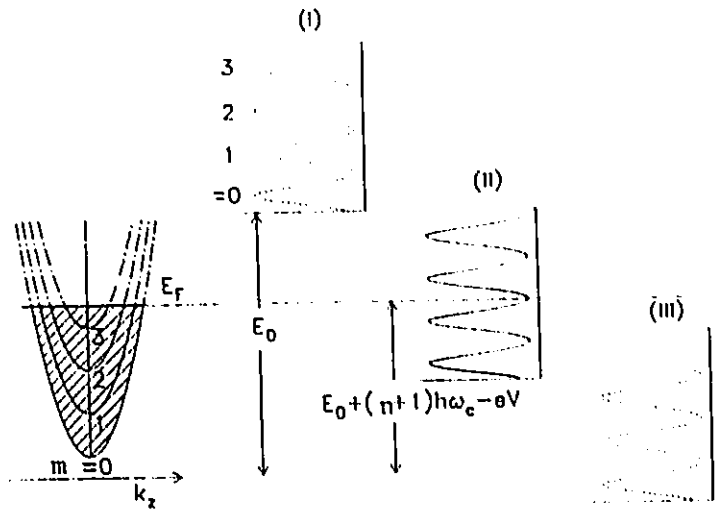
where E_0 and E_1 are the lowest bound state energies in the emitter and the quantum well respectively. The conservation of planar momentum $k_{//}$ requires that, in the presence of a magnetic field, transitions take place between Landau levels in the emitter and the quantum well for which $(m - n) = 0$.

Figure 3.5 shows the dispersion of the emitter subbands in the presence of a magnetic field in the growth direction, along with the density of states of the quantum well. The threshold voltage for resonance is then

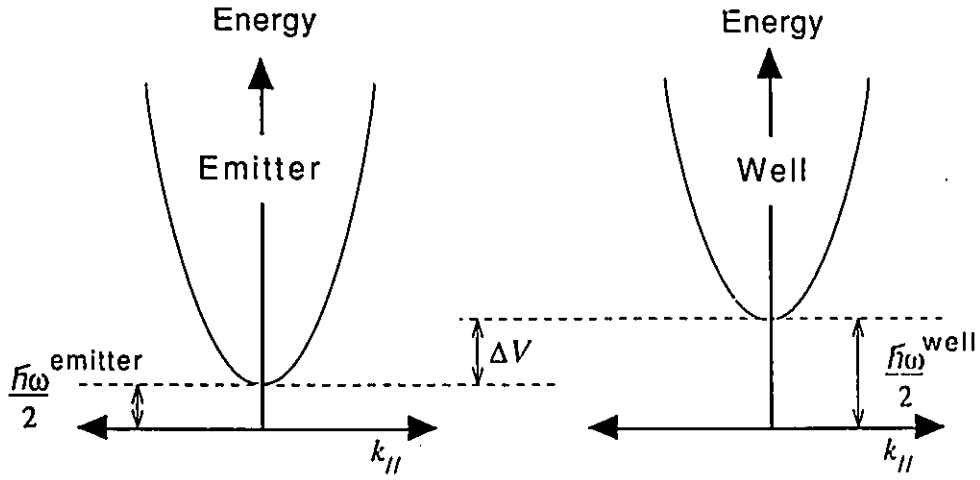
$$eV_{th} = E_1 - E_F + \frac{1}{2}\left(\hbar\omega_c^{qw} - \hbar\omega_c^{em}\right) \tag{3.22}$$

where E_1 is the bound state energy and E_F is the Fermi energy. Resonant tunneling via Landau levels begins at the threshold energy and continues until planar momentum conservation prohibits tunneling. The tunneling probability oscillates as the Fermi level is successively aligned with each Landau level and more states are made available for tunneling, leading to oscillations in the differential conductance. In general, thermal broadening of the Landau levels may smear out these oscillations at low fields.

An important consideration is the field dependent shift of the ground state energy due to the magnetic quantization. When the effective mass of the carriers in the emitter and quantum well are the same, the shifts of the energies in the emitter and quantum well due to the magnetic field will be identical and the resonance occurs at the same bias voltage independent of the field. However, when the effective masses are different, then Equation (3.22) describes a field dependent shift in the bias voltage of the resonance which can be expressed in the following form



The emitter subbands filled up to E_F in the emitter and the DOS of a quasi-bound subband in the well. Case (i) is for zero applied bias where there is no resonant tunneling. Resonant tunneling begins for a bias $eV_0 = E_0$. The tunneling probability oscillates as successive levels in the well are brought into resonance as shown in case (ii). Case (iii) corresponds to larger biases when resonant tunneling is forbidden by conservation of planar momentum.



The shifts in the ground state energy due to the magnetic field may be different for both the emitter and well subbands, leading to a shift in the bias position ΔV of the resonance.

Fig. 3.5 Resonant tunneling in a quantizing magnetic field B applied normal to the double barrier layers.

$$\Delta V = \frac{\hbar B}{2}\left(\frac{1}{m^*_{qw}} - \frac{1}{m^*_{em}}\right) \tag{3.23}$$

If the effective mass is larger in the emitter than in the quantum well, the application of a magnetic field will result in a shift of the resonance peak to higher bias voltage. Conversely, the resonance peak will be shifted to a lower bias when the effective mass in the emitter is less than that in the well. By measuring the size and direction of the shift, an estimate of the planar effective masses may be obtained.

3.7 Resonant Tunneling in a Planar Magnetic Field

The effect of a magnetic field applied parallel to the plane of the interface (along x) can be viewed classically in terms of a Lorentz force acting on the charge carriers impinging on the barriers and quantum well (along z). The force acts normal to the tunneling direction and deflects the charges from their original trajectory, resulting in a change in the transverse momentum of the carriers and a decrease in the kinetic energy in the tunneling direction (Ben Amor *et al.*, 1988). The motion parallel to the field (in the x direction) is unaffected and therefore the charge carriers tunnel to states with the same wavevector k_x in the well. For momentum to be conserved in the plane of the layers, electrons in the emitter with initial wavevector k_y must then tunnel to

$$k_y \rightarrow k_y + \frac{eBz}{\hbar}$$

where z is the distance from the emitter to the middle of the well. The parabolic dispersion relations of the transverse momentum of the well states are thus shifted in the k_y direction with respect to the emitter states. This is illustrated in Figure 3.6.

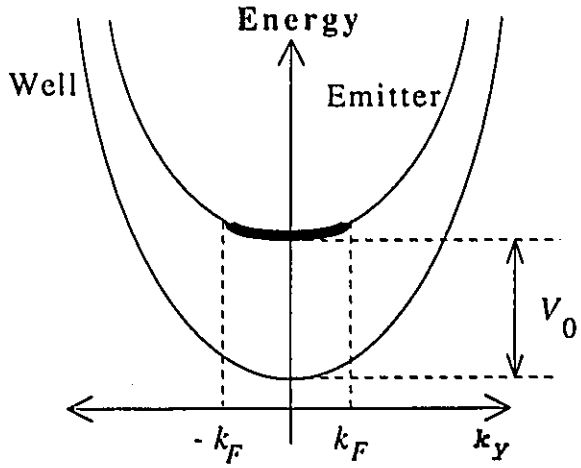
Because of the deflection of the trajectory, the electron experiences a potential decrease in the z-direction given by

$$\Delta V(z) = \frac{e^2 B^2 z^2}{2m^*} + \frac{\hbar e B k_{y0} z}{m^*} \tag{3.24}$$

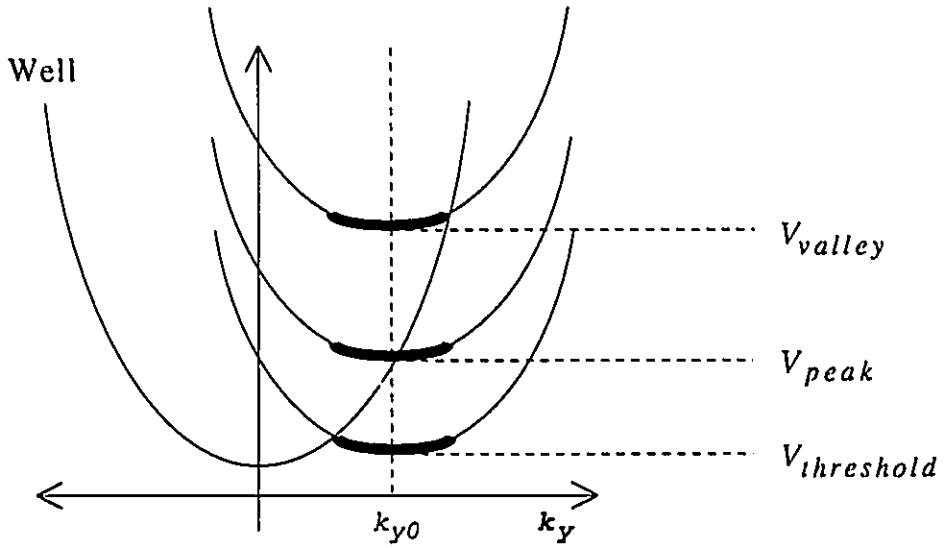
where $z = 0$ is taken from the average position of the accumulation layer in the emitter. An additional bias is thus required to overlap the emitter and quantum well states and according to Equation (3.24) results in a shift of the bias position of the resonance which is quadratic in field. As seen in Figure 3.6, the emitter and quantum well parabolas overlap over a greater range of biases with increasing magnetic field resulting in a smearing of the NDR regions. As the quantum well and barriers are narrow in double barrier tunneling structures, it is possible to use a perturbation technique to describe the energy levels in a transverse magnetic field (Eaves *et al.*, 1990). The zero of energy is taken with respect to the center of the quantum well and the results of the calculation are as follows:

$$\begin{aligned} E_{em} &= E_0 + eE\left(b_1 + \frac{w}{2}\right) + \frac{3m^*\omega_c^2}{2a^2} + \frac{\hbar^2 k_x^2}{2m^*} + \frac{\hbar^2\left(k_y - k_{y0}\right)^2}{2m^*} \\ E_{qw} &= E_1 + \frac{m^*\omega_c^2 w^2}{24} + \frac{\hbar^2 k_x^2}{2m^*} + \frac{\hbar^2 k_y^2}{2m^*} \end{aligned} \tag{3.25}$$

where a is the variational parameter of the Fang-Howard wavefunction, b_1 and w are the thicknesses of the emitter barrier and the quantum well, and E is the applied electric field. The terms involving ω_c^2 represent the diamagnetic energy shifts, which are small and can generally be neglected. The in-plane momentum k_{y0} in the quantum well corresponds to the change in the wavevector under the influence of



a) the planar dispersion of the emitter and well subbands in zero field. Resonant tunneling occurs when the occupied states in the emitter (dark curve) overlap with the states in the well, which occurs only for a single bias V_0.



b) for a magnetic field applied in the plane of the layers. The three curves for the well correspond to the biases when an overlap exists between occupied states in the emitter and quasibound states in the well. Resonant tunneling occurs at the threshold voltage up until the valley voltage where no overlap exists.

Fig. 3.6 Tunneling in a planar magnetic field.

the Lorentz force as the charge carrier tunnels from the emitter to the centre of the well:

$$k_{y0} = \frac{e\mathrm{B}}{\hbar}\left(\frac{3}{a} + b_1 + \frac{w}{2}\right) \tag{3.26}$$

Assuming that the effective mass is the same in both the well and the emitter so that the energy terms in the x direction cancel, the energy conservation rule for resonant tunneling may be written as:

$$E_0 + e\mathrm{E}\left(b_1 + \frac{w}{2}\right) + \Delta E + \frac{\hbar^2}{2m*}\left(k_y - k_{y0}\right)^2 = E_1 + \frac{\hbar^2 k_y^2}{2m^*} \tag{3.27}$$

where $\Delta E = m^*\omega_c^2\left(\frac{3}{a^2} - \frac{w^2}{12}\right)$ is the difference in diamagnetic shifts and is negligibly small. At low temperature the only states which are occupied in the emitter are those for which $\left|k_y - k_{y0}\right| \leq k_{\mathrm{F}}$. This is represented in Figure 3.6 as a parabola centered at $k_y = k_{y0}$ between $k_{y0} - k_F$ and $k_{y0} + k_F$ while the well states lie on a parabola centered at $k_y = 0$. Since $k_{y0} = 0$ at zero magnetic field, both parabolas are centered about the same axis. At zero bias or when $\mathrm{E} = 0$, the emitter curve is below the energy level in the well and no resonant tunneling can occur. As the bias voltage is increased, the emitter curve moves up relative to the well; resonant tunneling occurs when the voltage drop across the emitter and half of the well is such that $eV_{th} = E_1 + E_{\mathrm{F}}$. Neglecting charge accumulation in the well, V_{th} is the only bias at which tunneling can occur. The application of a transverse magnetic field shifts the center of the emitter parabola by k_{y0}. The threshold and cut-off for tunneling occurs when the wavevector is within a Fermi wavevector from the center of the parabola: the threshold at $k_y = k_{y0} - k_F$ and the cut-off at $k_y = k_{y0} + k_F$. When charge accumulation and variation of transmission coefficient with k_y are

neglected, the largest number of emitter states available for tunneling occurs at $k_y = k_{y0}$; the peak in the tunneling current will therefore occur at a bias voltage:

$$\begin{aligned} eV_{pk} &= E_1' + \frac{\hbar^2}{2m^*}\left(k_{y0}^2 + k_F^2\right) \\ &= eV_0 + \frac{\hbar^2 k_{y0}^2}{2m^*} \end{aligned} \tag{3.28}$$

where $k_{y0} = \frac{eB}{\hbar}\left(\frac{3}{a} + b_1 + \frac{w}{2}\right)$ and $E_1' = E_1 - \Delta E$.

A quadratic shift in the peak position and a linear increase of the width of the resonance is expected with magnetic field. The quadratic shift reflects the parabolic nature of the dispersion relation in the quantum well.

3.8 Non-Resonant Tunneling in Double Barrier Structures

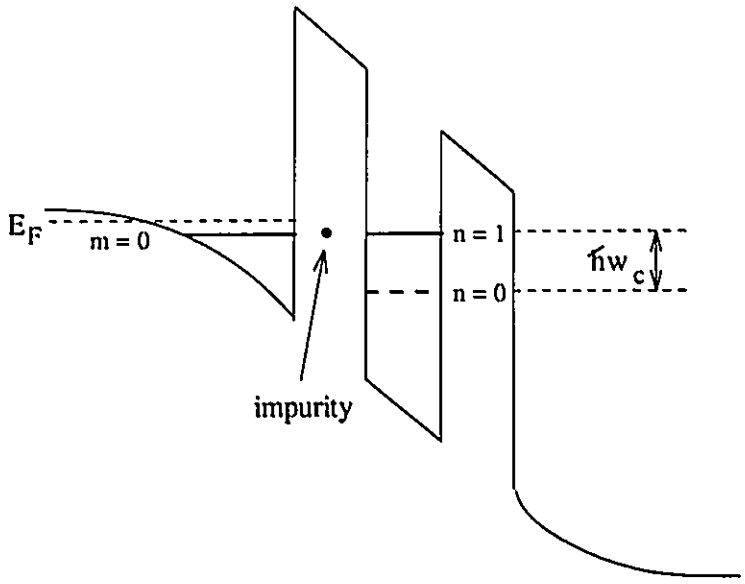
In the previous sections the tunneling was assumed to proceed resonantly via an eigenstate of the well with the conservation of both the total energy and planar momentum of the charge carrier. If the tunneling particle is scattered during its transit across the double barrier, the conservation rules are relaxed and tunneling may proceed non-resonantly. In a quantizing magnetic field normal to the layers of a double barrier structure three types of scattering transitions may be distinguished (Eaves *et al.*, 1988): (1) an elastic scattering induced transition from the $m = 0$ Landau level in the emitter to the $n = 1$ level in the quantum well via impurities or interface roughness, (2) quasi-elastic scattering transitions induced by the emission of an acoustic phonon and (3) inter-subband transitions due to inelastic scattering of particles tunneling from the $m = 0$ Landau level in the emitter to the $n = 0, 1, 2, \ldots$ levels in the well with the emission of an longitudinal optical phonon

of energy $\hbar\omega_L$. The energy selection rules for tunneling may then be expressed in the following form

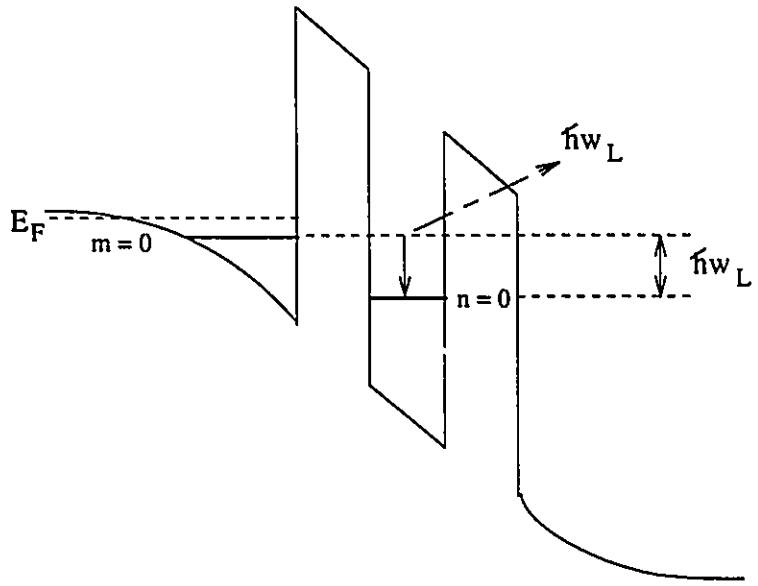
$$\begin{aligned} E_0 &= E_1 + \left(m - n\right)\frac{\hbar eB}{m^*} && \textit{elastic and quasi-elastic} \\ E_0 &= E_1 + \left(m - n\right)\frac{\hbar eB}{m^*} + \hbar\omega_L && \textit{inelastic} \end{aligned} \tag{3.29}$$

where E_0 and E_1 are the lowest bound state energies in the emitter and quantum well respectively.

The various scattering processes are illustrated in Figure 3.7 for electrons tunneling in a perpendicular magnetic field. Here the effective mass in the emitter and quantum well are assumed to be the same, and therefore the energy shifts of the bottoms of the subbands are identical. Figure 3.7a illustrates a bias voltage such that the $m = 0$ Landau level in the emitter is aligned with the $n = 1$ Landau level in the quantum well. There is no resonant enhancement of the tunneling probability due to tunneling through this level since planar momentum cannot be conserved ($m \neq n$). However, the particle may tunnel into the $n = 1$ level in the well if it is elastically scattered while traversing the left barrier, and the transmission probability is consequently enhanced. Figure 3.7b illustrates the energy levels at a bias which is $\hbar\omega_L$ beyond resonance. By emitting a LO phonon of this energy, the particle may tunnel via a quasi-bound state of the well with a corresponding increase in the tunneling probability. In the situation illustrated planar momentum is conserved ($m = n = 0$), but the particles may also be scattered into Landau levels $n \neq m$.



The elastic scattering process, displayed here for scattering off an impurity. The scattering changes the planar momentum of the tunneling particle, allowing tunneling via the n = 1 Landau level which is energetically aligned with the emitter states.



An inelastic scattering process with the emission of a phonon of energy $\hbar w_L$. The particle is shown here to conserve Landau level indices, but transitions are possible via Landau levels $p = n - m > 0$.

Fig. 3.7 Non-resonant tunneling processes.

IV. Si/SiGe Strained Layer Heterostructures

4.1 Introduction

In this thesis the tunneling characteristics of double barrier tunneling structures formed from layers of Si and $Si_{1-x}Ge_x$ alloy are investigated. The Si-Ge system is important due to its compatibility with current Si-based VLSI production techniques. Technologically Si-Ge heterostructures are of particular interest, as the bandgap can be modified by the strain in the layers. Basic properties of the Si-Ge system are reviewed in this chapter.

4.2 General Properties of Bulk Si and Ge

^{14}Si and ^{32}Ge are semiconductors belonging to the group IV elements of the periodic table. These materials crystallize in the diamond structure: as illustrated in Figure 4.1, the atoms are situated on two interpenetrating face-centered cubic (FCC) lattices displaced by one-quarter of the body diagonal in the (111) direction. In $Si_{1-x}Ge_x$ alloys atoms of either species will be interspersed on these sites. In the special case of $Si_{0.50}Ge_{0.50}$, the lattice describes the zincblende structure: the Si atoms occupy only the sites of one FCC lattice and the Ge atoms are situated on the sites of the other. The lattice constants of pure unstrained Si and Ge are 5.43 Å and 5.66 Å respectively, and the $Si_{1-x}Ge_x$ alloys will have an intermediate lattice constant which depends on the alloy composition x (the lattice constant increases for increasing compositions of Ge).

The energy band structures of bulk Si and Ge are displayed in Figure 4.2, for a temperature of 300 K. Both Si and Ge are indirect bandgap semiconductors, as are the alloys $Si_{1-x}Ge_x$ for all compositions x. At zero temperature the fundamental

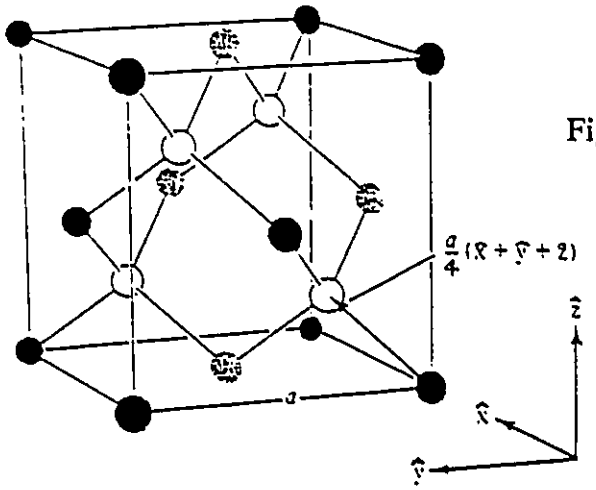
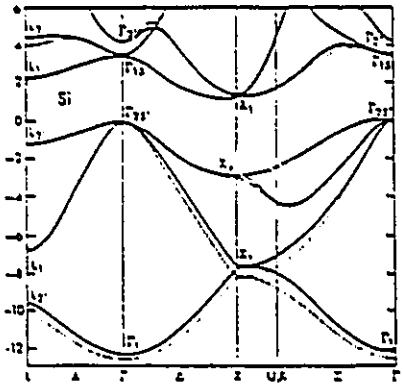


Fig. 4.1 The crystal structure of Si-Ge alloys. The dark spheres correspond to the sites of one FCC lattice and the hollow spheres to the other.



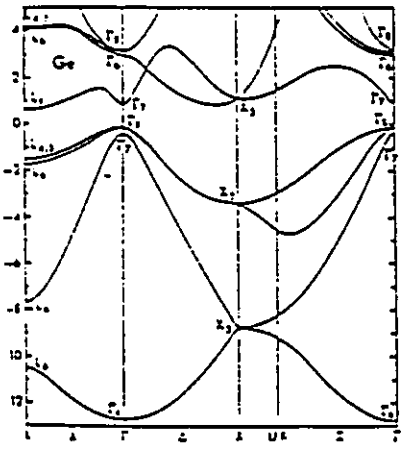


Fig. 4.2 The band structure of Si (above) and Ge.

bandgaps are 1.17 eV for Si and 0.75 eV for Ge. These bandgaps are highly temperature dependent and at 300 K, the values drop to 1.12 eV for Si and 0.67 eV for Ge (from Ashcroft and Mermin, 1976).

The electronic properties of materials formed from Si and Ge are largely determined by the type and level of impurity doping. In this study the excess charge carriers are holes and therefore the valence bands are of primary interest. In the absence of spin, these materials have six valence bands which coincide at the zone center ($k = 0$). Spin-spin interactions are negligibly small and therefore these bands may be considered as three sets of degenerate doublets. Spin-orbit coupling leads to the splitting of these bands into a four-fold set of states (angular momentum quantum number $J = 3/2$) towards higher energy, and a doublet ($J = 1/2$) towards lower energy. For all temperatures of interest, the lower doublet is split-off sufficiently far in energy (approximately 44 meV) to be inaccessible to charge carriers. The upper pair of doublets are distinguishable by the different curvatures in their dispersion: since the effective mass tensor is given by the reciprocal of the band curvature, this results in both heavy and light hole behaviour. Because of the close proximity of these bands to each other, band interactions are not negligible and are an important consideration in the calculation of the band structure.

4.3 The Role of Strain

Ideally, heterojunctions should be lattice matched at the interfaces to reduce scattering due to defects. Since there exists a 4% lattice mismatch between Si and Ge, there will be strain in some or all of the layers in Si-Ge heterostructures. By growing on thick Si substrates, the layers are all forced to assume the lattice constant of strain-free Si in the plane of the layers. This produces a bi-lateral compression

of the $Si_{1-x}Ge_x$ layers, and the lattice constant normal to the layers expands commensurate with minimizing the free energy.

The strain can be modelled by a tetragonal distortion of the lattice, which yields a very simple form for the strain tensor e_{ij}. If z is defined along the [001] growth direction then e_{ij} is non-zero only along the diagonal and $e_{xx} = e_{yy}$. The effect of the strain on the electronic bands may be estimated using phenomenological deformation potential theory (Kleiner and Roth, 1959). The following results were found by People (1986) for strained $Si_{1-x}Ge_x$ layers on unstrained Si:

- The previously degenerate valence bands at $k = 0$ are split because of the strain.
- As the molar fraction of Ge (or equivalently the strain) is increased, the bandgap decreases.
- The effective mass tensor is anisotropic and the effective mass in strained layers can vary dramatically from that in unstrained layers.

As illustrated in Figure 4.3, the heavy hole band is shifted up in energy for all Ge compositions x. The light hole band initially moves away from the heavy hole band, and may split-off far enough to be inaccessible to charge carriers at low temperatures. In Figure 4.4, the indirect bandgap of $Si_{1-x}Ge_x$ is shown as a function of the composition x, for growth on unstrained Si substrates. The bandgap decreases with increasing strain and is always lower than the unstrained alloy. With the knowledge of the bandgaps the band alignment at the interface can be determined; a means of estimating one of the band offsets ΔE_c or ΔE_v is required. Modelling performed by van de Walle and Martin (1985) indicates that the valence band edge of the $Si_{1-x}Ge_x$ alloy lies above that of Si for all values of x; therefore Si always acts as a barrier layer in the p-type system. It has been shown (People *et al.*, 1984) that

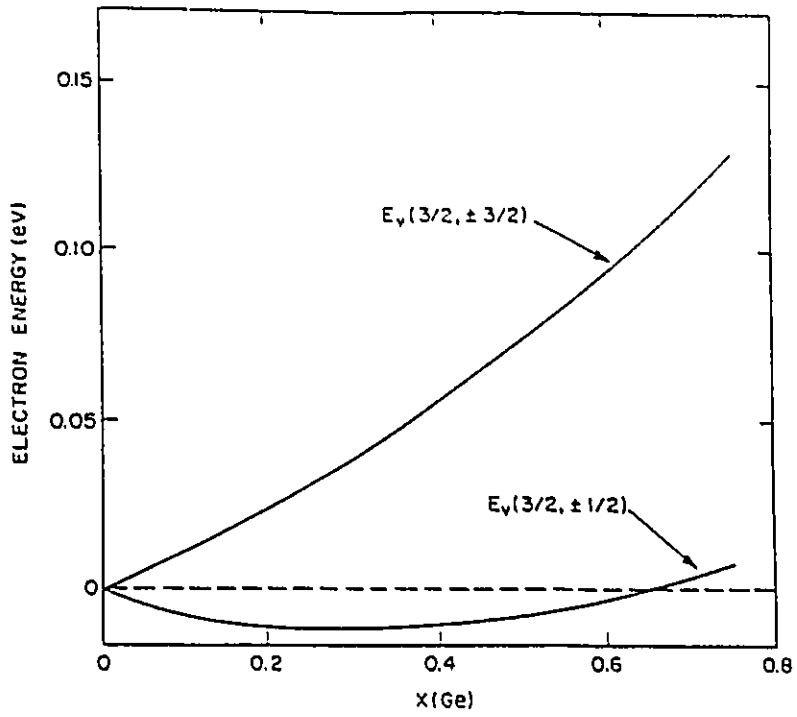


Fig. 4.3 The splitting of the heavy hole (J = 3/2) and light hole (J = 1/2) valence bands in $Si_{1-x}Ge_x$ alloys as a function of the composition x.

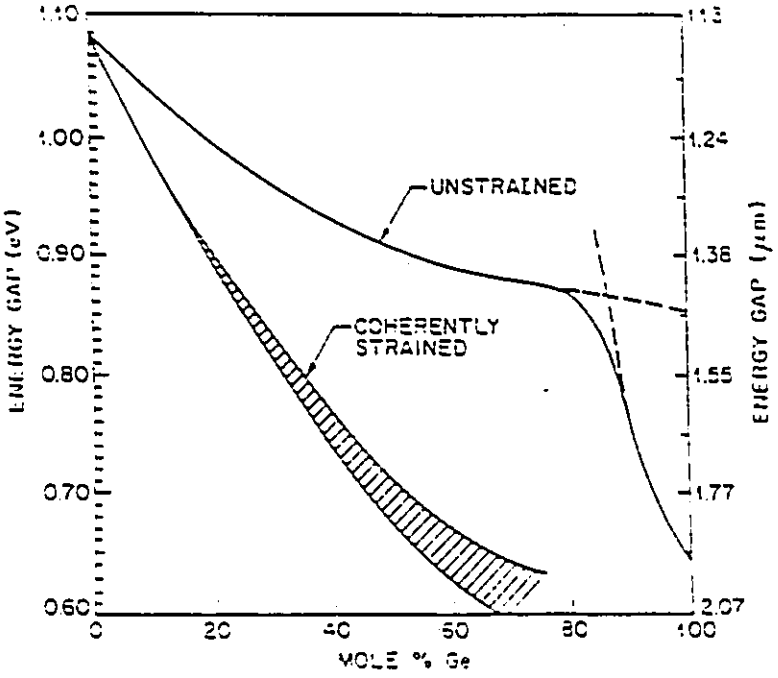


Fig. 4.4 The fundamental bandgap as a function of the percentage of Ge for unstrained and strained $Si_{1-x}Ge_x$ alloy on unstrained Si (001) substrates.

the alignment of the conduction band edge is dependent on the strain in the layers. For strained alloy layers on unstrained Si (001) substrates, the bandgap of the alloy falls within that of Si, and the alignment is Type I in character. The conduction band offset E_c is small (≈ 0.020 eV) and therefore the bandgap difference between the two materials is taken up mainly in the valence band offset ΔE_v. Type II alignments are possible when the alloy layers are deposited on Ge or strained Si substrates.

Because the electron dynamics are largely determined by the band structure near the Γ point ($k = 0$), it is important to know the dispersion of the bands. For the valence bands, the in-plane dispersions can be calculated using the Luttinger-Kohn (L-K) Hamiltonian in the spin $-\frac{3}{2}$ basis (Wessel and Altarelli, 1989). This model ignores the spin split-off and conduction bands, which have been shown to have a negligible effect on the hole transport. By using the spherical approximation for the band structure, the L-K Hamiltonian for a bulk strained semiconductor decouples into two 2×2 matrices. If an inversion symmetry plane exists in the structure, these matrices are equivalent and the complexity of the problem is greatly reduced.

In the quantum well, the confinement energy separates each band into a series of quantized subbands. In the L-K model, each subband may be considered separately and it is convenient to label them according to their zero wavevector properties: *e.g.* the lowest energy heavy hole and light hole subbands are denoted HH1 and LH1 respectively. As the band structure is symmetric about the x-y plane, the dispersion may be calculated along any arbitrary direction in this plane. From O'Reilly and Witchlow (1986, 1987), the following conclusions are drawn for the in-plane dispersion in strained and unstrained quantum wells:

- the lowest energy (ground) state is the HH1 heavy hole band which has a comparatively small effective mass for small in-plane wavevector.

- the most important interaction is of order k between LH1 and HH2, leading to non-parabolic band behaviour. In bulk alloys, HH2 is lower in energy than LH1 but for sufficiently thin strained layers, the strain lowers LH1 below HH2. Strain therefore serves to select the energy at which the crossover occurs.

4.4 The $Si/Si_{0.78}Ge_{0.22}/Si$ Quantum Well

The in-plane dispersion of these subbands was calculated by Liu *et al.* (1989) using the L-K Hamiltonian in the spherical approximation, taking a plane of symmetry at the middle of the well. The heavy hole barrier height was assumed to be 0.20 eV and a value of 0.043 eV was taken for the energy splitting between the heavy and light hole bands. The zero energy reference was taken from the top of the uppermost valence band edge. The results of this calculation are shown in Figure 4.5. It is noted that the basic features predicted by the model are observed: the LH1 subband has an electron-like dispersion at $k = 0$ with a local maximum occuring at ~ 0.02 $Å^{-1}$, and the interaction between the subbands results in the LH1 level becoming the second bound state of the well.

From the dispersion calculations, the effective masses of each of the subbands were determined from the subband curvature. The effective masses of the well subbands are given in Table 4.1 below, along with the effective masses of the emitter both parallel and perpendicular to the layers. The emitter layers were treated as bulk strained $Si_{0.78}Ge_{0.22}$ alloy where the strain results in an anisotropic heavy hole effective mass: 0.44 m_o in the tunneling direction and 0.16 m_o in the plane of the layers.

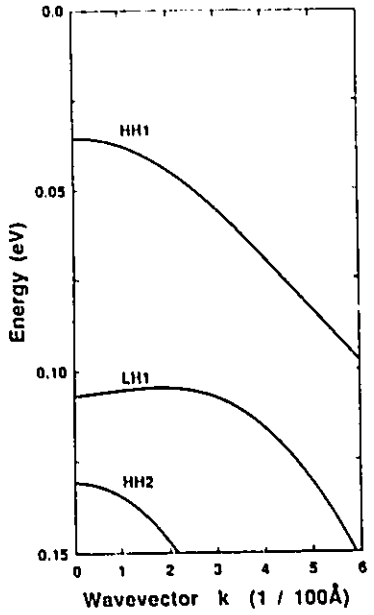


Fig. 4.5
The subband dispersion for the $Si_{0.78}Ge_{0.22}$ quantum well. Three bound states are calculated for a well depth of 0.2 eV.

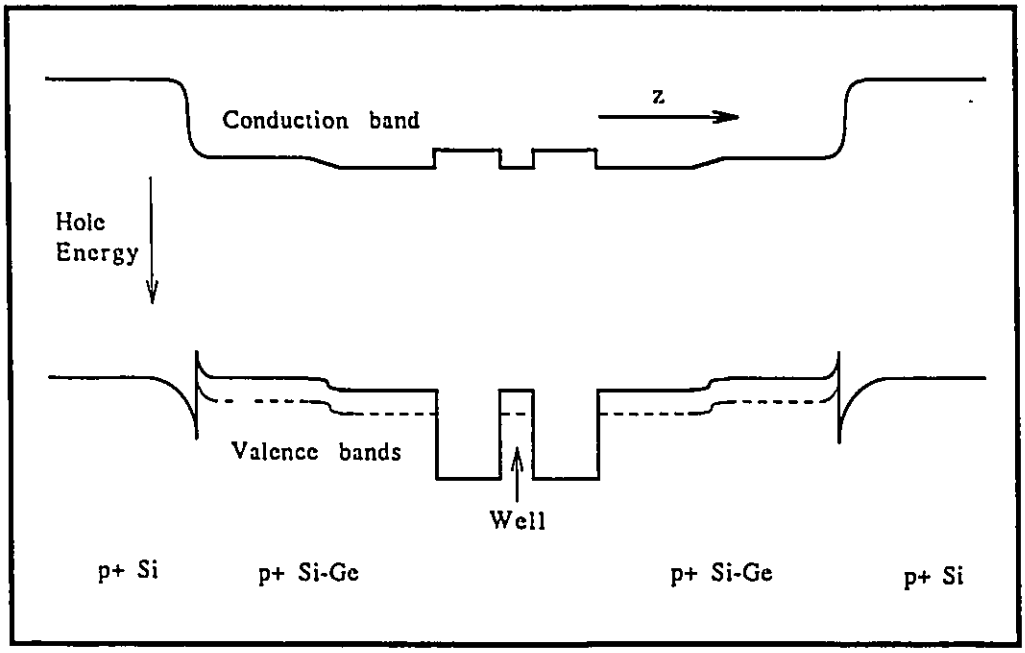


Fig. 4.6 The band edge of the double barrier tunneling structure. The heavy hole band edge is shown by the solid line, and the dashed line indicates the light hole valence band edge. The offset in the contact layers is due to different doping levels. Note that the layer thicknesses are not drawn to scale. Also, the bandgap energies (~ 1 eV) are much greater than shown.

A schematic of the band edge alignment is shown in Figure 4.6. Strain energy calculations for a Ge fraction of x = 0.22 yield an estimated valence band discontinuity of 0.23 – 0.27 eV. This rather large range of values indicates the uncertainty to which the valence band offset has been determined. Two heavy hole and one light hole subbands are expected to be bound in the well.

Quantum Well (and bound state energies)	In-plane effective masses in units of m_0
HH1 (–0.035 eV)	0.16
LH1 (–0.107 eV)	–0.14
HH2 (–0.130 eV)	0.06
Emitter	
Heavy hole band	0.16

Table 4.1 The planar effective masses in the emitter and quantum well.

V. Experimental Techniques

5.1 Samples

5.1.1 Sample Growth

The samples were processed from a single wafer grown by molecular beam epitaxy, an advanced method for the deposition of ultrathin layers of material originally developed from the study of chemical reactions on the surfaces of III-V compounds. The growth was done at the National Research Council Canada using a Vacuum Generators VG 80H MBE system and is detailed by Houghton *et al.* (1989) and Liu *et al.* (1989). A brief description of the method is presented here. A schematic of a typical MBE system is shown in Figure 6.1. The growth is performed in an ultra-high vacuum (UHV) environment to ensure that adsorption on the sample substrate is due almost solely to the source atoms. The substrate was a thick monolithic silicon wafer in the (001) orientation, which was rotated at approximately 20 revolutions per minute to improve the uniformity of growth. A heater mounted below the substrate sets the temperature (normally between 350 °C and 700 °C), which determines the chemi-sorption coefficient of the deposition. Effusion cells surrounded by a heating coil provide the Si and Ge atoms to be deposited. The temperature of the cells controls the vapour pressure of the sources, which in turn determines the flux directed towards the substrate. For a growth temperature of 500 °C, the deposition rate of Si + Ge was approximately 0.5 nm/s. A shutter on each of the effusion cells switches the flux on and off, and the shutters were programmed to open and close at the appropriate intervals for the deposition of layers of a desired thickness. Doping was achieved by simultaneously opening a shutter on the dopant material cell during the deposition. In this manner, the doping can be spatially controlled to as fine as one monolayer, a technique known as delta doping.

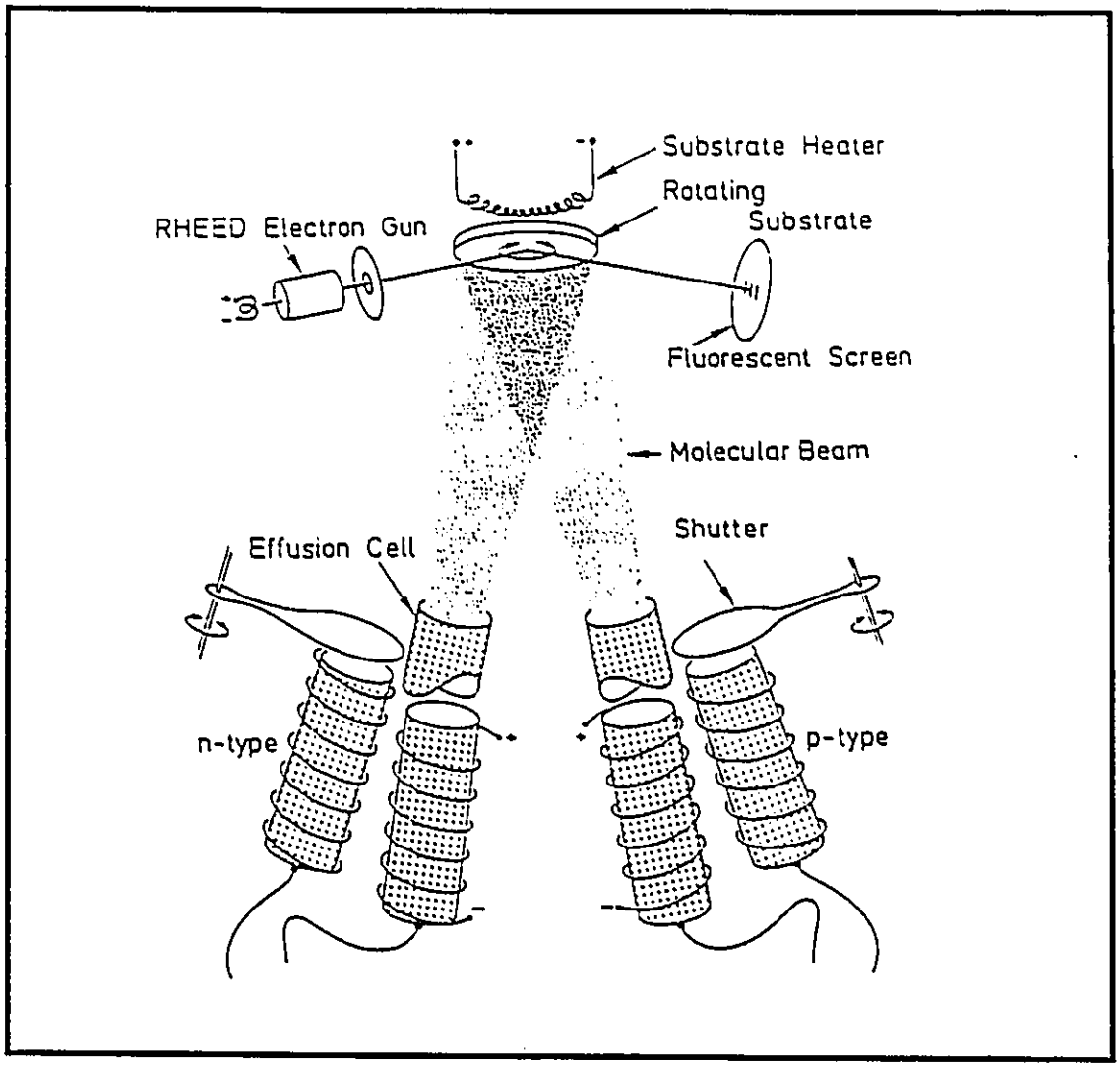


Fig. 5.1 A schematic of the MBE apparatus. The entire apparatus is set in an ultra-high vacuum environment. There is an effusion cell for each species of atoms to be deposited including dopant cells. There are heating coils wrapped around each cell to generate a flux of molecules. The substrate is rotated at a constant velocity and maintained at a set temperature during the growth. Shown near the substrate is a RHEED analyzer to monitor the deposition rate during the growth.

In practice however, diffusion of the dopant atoms often broadens the region over which the dopant is defined. Here the contact layers were p-type doped with boron by the co-evaporation of B_2O_3 during the growth. To prevent free charges from diffusing into the quantum well, undoped spacer layers were grown on both sides of the double barrier layers separating the contact and barrier layers.

An important aspect of the UHV technology is that it enables in-situ measurement of the physical and chemical properties of the films deposited. Some of the more important techniques used are reflection high energy electron diffraction (RHEED) and Auger electron spectroscopy (AES). In particular, the RHEED oscillation technique can be used to measure the growth rate during deposition, thereby enabling the real-time monitoring of the growth process (Ploog, 1987). It is this processing control which enables the deposition of single monolayers.

The layer sequence of the samples is given in Table 5.1, along with the layer thicknesses and doping levels. A transmission electron micrograph (TEM) of a cross-section of a typical sample is presented in Figure 5.2, which shows the high quality of the interfaces. As can be seen in the TEM, the double barrier structure is symmetric about the centre of the well. The dimensions of the quantum well and the barriers given in Table 5.1 are known to ± 0.3 nm and the alloy composition to ± 0.01. To aid in the identification of the double barrier structure, the layer sequence is also illustrated in an schematic expansion of the micrograph cross-section. The distribution of dopant was determined by electrochemical C-V profiling, and the alloy composition was determined by x-ray diffraction (XRD) measurements of the lattice parameter (Houghton *et al.*, 1989). The TEM and XRD measurements both confirmed coherent (lattice-matched) growth.

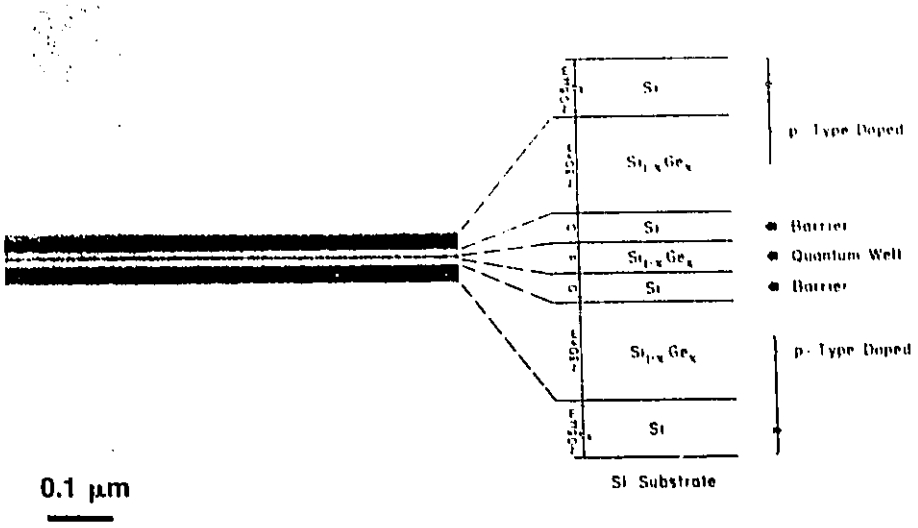


Fig. 5.2 A transmission electron micrograph of a cross-section of a typical double barrier tunneling structure. The dark areas correspond to the $Si_{1-x}Ge_x$ alloy layers . A diagram showing a blow-up of the layer sequence is given on the right to aid in the identification.

Fig. 5.3 A photograph of a typical mesa (7 μm) taken with a scanning electron microscope. The contact pad to the dot extends as a finger from the top right of the photo and the contact to the annulus can be seen in the upper left corner.

Table 5.1 The layer sequence of the samples grown on Si (001) substrates

Composition	Dimensions	Doping level
p^+ Si layer	500 nm	~ 10^{19} cm^{-3}
p^+ $Si_{0.78}$ $Ge_{0.22}$	15 nm	~ 10^{19} cm^{-3}
$Si_{0.78}$ $Ge_{0.22}$ spacer	15 nm	nominally zero
Si barrier	53 Å	undoped
$Si_{0.78}$ $Ge_{0.22}$ well	34 Å	undoped
Si barrier	53 Å	undoped
$Si_{0.78}$ $Ge_{0.22}$ spacer	15 nm	nominally zero
p^+ $Si_{0.78}$ $Ge_{0.22}$	15 nm	~ 10^{19} cm^{-3}
p^+ Si layer	500 nm	~ 10^{19} cm^{-3}

Individual tunneling structures were defined on the wafer by wet-etching to the bottom contact layer. For simplicity, a standard mask was used in the etching procedure to define several mesas per section of wafer. As shown in the scanning electron microscope (SEM) photograph of Figure 5.3 and illustrated in Figure 5.4a, the geometry of the structure is a cylindrical mesa whose axis is normal to the layer plane. The top contact is defined by a dot and the bottom contact as a concentric annulus. Additional contact pad layers were deposited on the smaller mesas using standard photolithographic techniques. For mesa diameters greater than 50 μm, contacts could be made directly to the mesa.

Each wafer section (approximately 1 cm^2) was then mounted onto a 12 pin ceramic header and secured with high vacuum grease (Figure 5.4b,c). An ultrasonic wirebonder (Westbond) was used to bring gold wires from the contacts on the sample to the gold pins of the header. The operation of this instrument relies on

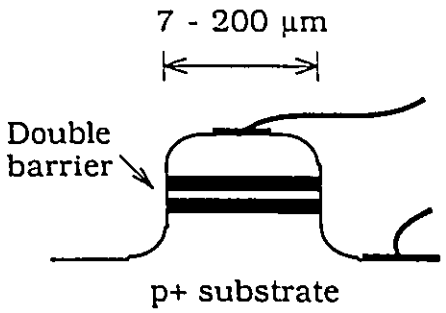


Fig. 5.4 A schematic of a typical mesa, not drawn to scale. Gold wires extend from the contact layers on top and to the substrate. The growth direction is upwards.

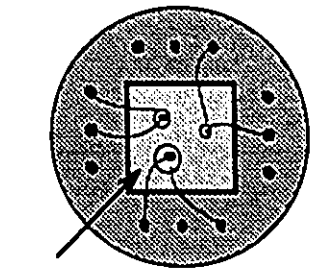


Fig. 5.5a A top view of a wafer mounted on the header. Three mesas are shown.

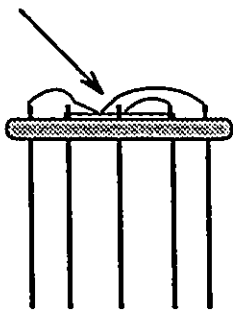


Fig. 5.5b A side view of the header, showing some of the contact wires and pins.

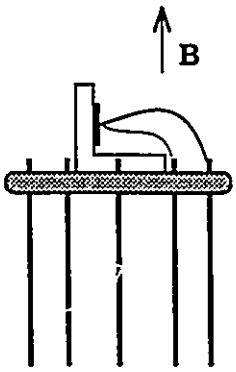


Fig. 5.5c The mounting configuration for the magnetic field in the plane of the layers

of this instrument relies on the surface adhesion of the wire to the sample. In most instances two wires were connected to both top and bottom contacts, allowing four-wire measurements to be made. In the planar field configuration this was not possible due to limited access to the header pins and therefore only single wires were bonded to top and bottom.

5.2 Cryogenic System

5.2.1 T = 1.7 K

The sample header was mounted on the probe of an Oxford Instruments model 400 top loading dilution refrigerator which incorporates a vacuum lock to insert the sample probe into the glass tail extension of the dilution unit centered in the bore of the superconducting magnet. The probe carries five twisted pairs of 40 gauge copper wire and four pairs of manganin wire for electrical measurements and thermometry as well as a flexible Lakeshore Cryogenics coaxial cable. The base temperature of the dilution refrigerator is 18 mK, but the measurements reported here were performed at 1.7 K. The bath temperature was obtained by measuring the resistance of a calibrated carbon resistor located near the sample. Since the sample was fully immersed in the ^{3}He-^{4}He mixture, the sample temperature could be assumed to be that of the bath for low input powers.

A maximum axial magnetic field of 15 Tesla at 4.2 K was generated by hybrid Nb_3Sn and NbTi coils manufactured by Oxford Instruments. The bore diameter is 44.5 mm and the field homogeneity at the centre of the magnet is better than four parts in 10^4 over a 10 mm diameter spherical volume. The inductance of the coil is approximately 170 Henry and a ± 10 Volt regulated power supply (Oxford Instruments model 2115A) provides a current of 100 Amperes at 15 Tesla. Sweep rates

ranging from 0.05 to 2.0 A/min were available, corresponding to field rates of 0.0075 to 0.30 Tesla/min.

Although the sample is usually mounted such that the magnetic field is perpendicular to the plane of the wafers, measurements can also be taken in the parallel configuration. This is accomplished by means of a special mount appended to the header, shown in Figure 5.5c. It consists of an L-shaped cylinder of ceramic, secured to the header using General Electric RTV silicone sealant. The sample is attached to the mount using high vacuum grease.

5.2.2 T = 4.2 K to 77 K

To investigate the temperature dependence of the tunneling characteristics of the samples, additional measurements were taken in an Oxford Instruments SM-2 cryostat with a maximum axial field of 8 Tesla. The system was fitted with a variable temperature insert (VTI) which allowed operation from the bath temperature up to 300 K. Temperature control and measurement is effected by the use of an Oxford Instruments ITC-4 temperature controller monitoring a calibrated platinum resistor situated on a copper block on the probe. Heat input was provided by supplying current to a 20 Ω wire-wound resistor situated at the needle valve of the pick-up line to the He bath. The maximum output voltage of the ITC-4 was 40 Volts, yielding a maximum output power of 80 Watts.

5.3 Electronics

5.3.1 Direct Current Measurements

A block diagram of the DC measuring circuit is presented in Figure 5.6 for a four-wire configuration. Whenever possible a four-wire measurement was made

in order to reduce the contribution of the lead resistance. The circuit consisted of a Keithley 230 voltage source controlling the voltage across the sample in a remote sensing mode. The bias potential across the sample and the potential drop across the standard series resistor were measured using Keithley 181 nanovoltmeters thus yielding the I-V characteristic curve for the sample. These measurements were performed as a function of magnetic field for different mesa diameter devices, from 7 μm up to 200 μm. The measurement of the tunneling current as a function of magnetic field at constant bias voltage is an effective way of tracking the displacement with field of a small feature in the current.

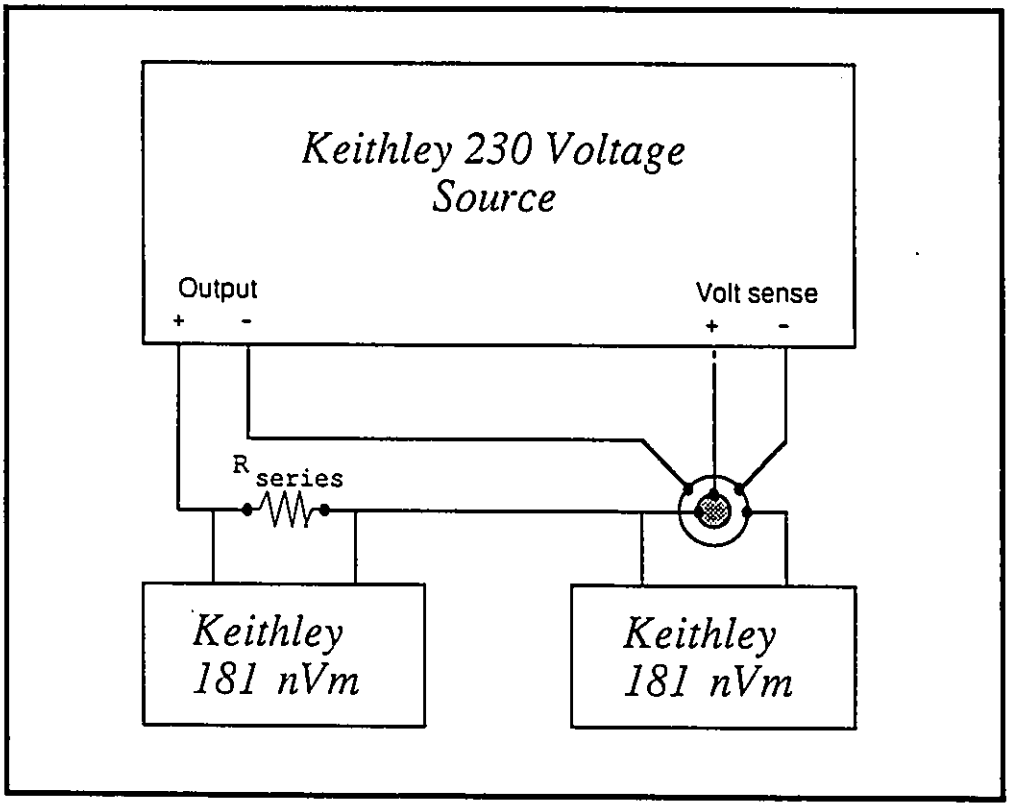


Fig 5.6 The DC measuring circuit. In this remote sensing mode, the output of the source is such as to maintain a fixed bias across the sample independent of any stray voltage drops along the rest of the circuit.

5.3.2 *Differential Conductance Measurements*

The differential conductance $\partial I/\partial V$ was also measured in order to resolve any fine features in the tunneling current as a function of bias voltage. The technique involves the superposition of a small AC signal on a DC bias current, and the measurement of the AC response of the sample using lock-in amplification (Thomas and Klein, 1963; Patterson and Shewchun, 1964). A block diagram of the circuit is illustrated in Figure 5.7a. The DC current was supplied by a Keithley 220 current source while the AC modulation was provided by a Hewlett-Packard 3325A synthesizer/function generator; Keithley 181 nanovoltmeters and Stanford SR 530 lock-in amplifiers were used to measure I, V and ∂I, ∂V respectively. In order to achieve the mixing of the signals without loss of sensitivity, a modulation circuit was designed by R. J. Douglas at the National Research Council, and the circuit diagram is given in Figure 5.7b. A low frequency (697 Hz) sinusoidal waveform is applied across a 1:100 voltage divider to a Burr-Brown INA 101AM instrumentational amplifier. This device is designed with a closed loop gain circuit; no gain resistor is used so that the amplifier acts as a buffering stage with unity gain. The output is then fed through a 10 kΩ resistor and summed with a DC signal applied to the negative input of an OP-07 operational amplifier. Because of negative feedback, this point is a virtual ground and therefore the output of the op-amp is a voltage across the feedback resistor which is sufficient to null the differential inputs. This output is applied to the gate of a FET and determines the current flowing in the transistor. The voltage dropped by the current across a 100 kΩ resistor is buffered and inverted by a second INA 101 AM and fed back to the positive input of the OP-07. The op-amp is operating in an open-loop gain mode, and consequently acts as a comparator. The output of the op amp will be such that the FET will supply a

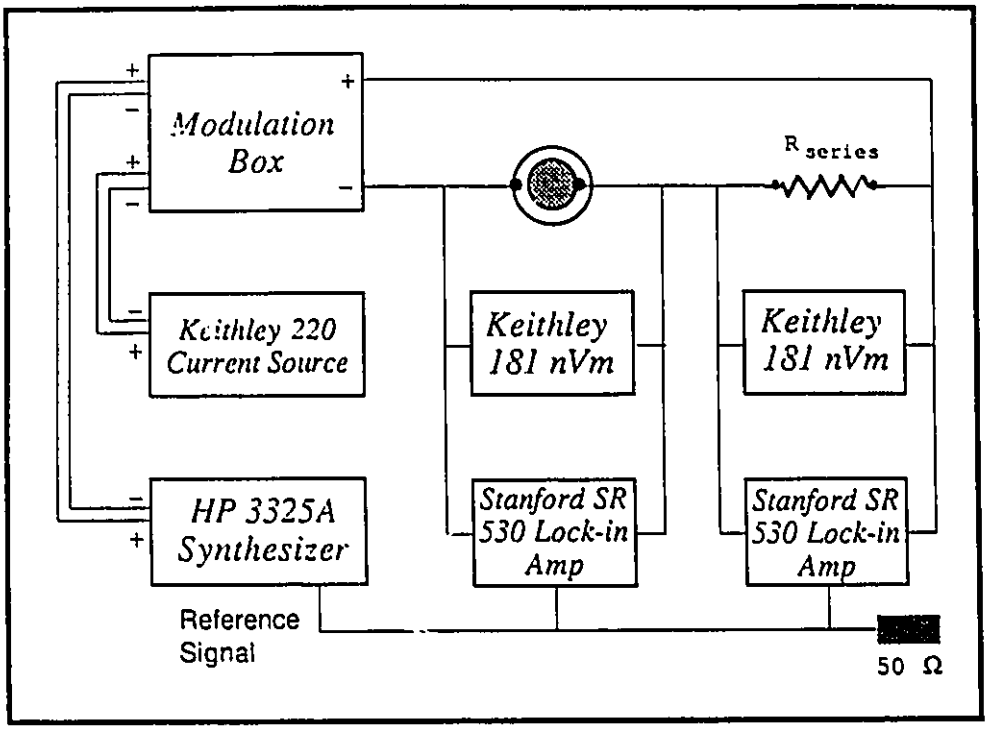


Fig. 5.7a A block diagram of the AC measuring circuit.

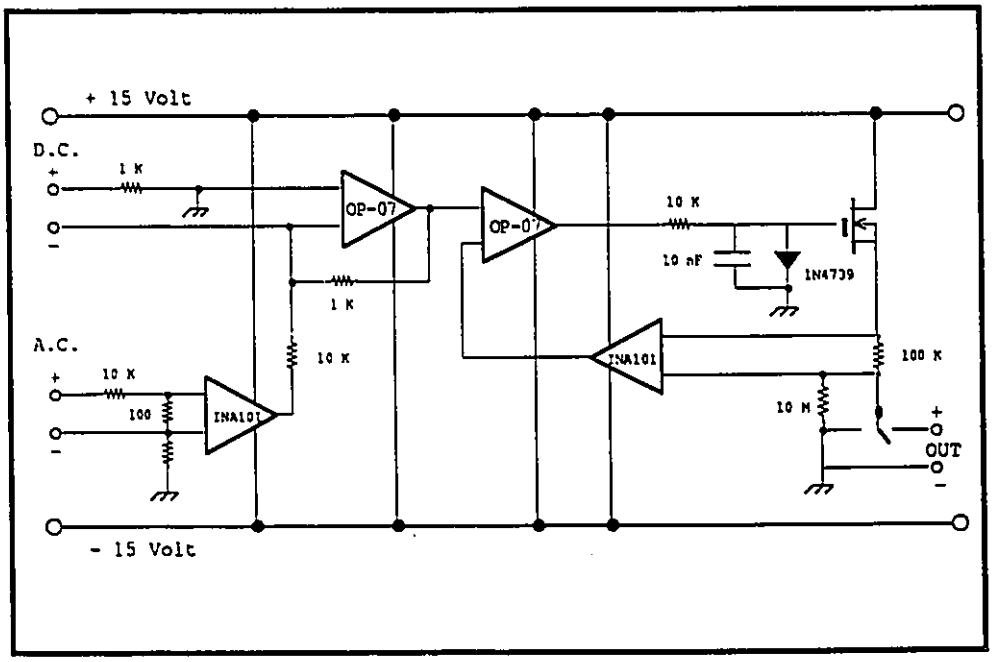


Fig. 5.7b The circuit diagram to the modulation box designed for differential measurements.

current which equalizes the voltages at the inputs. For the DC component of the signal the transfer function is given by the ratio of the two resistors:

$$\frac{I_{out}}{I_{in}} = \frac{V_{out}}{V_{in}} \cdot \frac{1\,\text{k}\Omega}{100\,\text{k}\Omega} = \frac{1}{100}$$

Similarly for the AC component, the resulting AC signal is

$$I_{out}(ac) = \left[\frac{\left(\frac{V_{ac}}{100}\right)}{10}\right] / 100\,\text{k}\Omega = 10^{-8}\,V_{ac}$$

The large reduction of the AC input signal is necessary since the synthesizer provides a minimum output of 1 mV and the measurement requires that the AC modulation be small compared to the DC component for a true differential signal (Thomas and Klein, 1963).

5.4 Data Acquisiton

To facilitate data acquisition, a Hewlett-Packard 9816 Series 200 computer was used as a controller for all the measuring instruments addressed on an IEEE-488 bus interface, and readings were acquired by accessing each meter in succession. Usually each data point consisted of an average of at least ten readings. The data was archived on both dot matrix printer and floppy disk for further numerical analysis purposes. At least 400 points were acquired per full voltage sweep (0.2 – 1.0 V) in order to accurately define the position of the resonance peaks and features in the tunneling current from the differential conductance. As well as performing a direct measurement of the differential conductance, a numerical derivative of the data was often taken by a 6–point sliding quadratic least squares fit method.

VI. Results and Discussion

6.1 Resonant Tunneling

6.1.1 Zero magnetic field

Most of the measurements reported in this chapter were performed at 1.7 K in the top loading dilution refrigerator system. The results described in §6.3.2 were obtained in the temperature range 4.2 K < T < 77 K using a cryostat with a variable temperature insert. The current-voltage characteristic for MBE 519B is shown in Figure 6.1. These data were taken for a mesa diameter of 100 μm, but the features shown are typical of all mesa sizes. The plot shows the tunneling current for bias voltages in both polarities up to 1.0 V. The polarity convention is defined as being positive when the top contact (dot) is at a higher potential relative to the bottom contact (ring). It is seen that the results differ slightly in the two bias directions, with the magnitude of the current under negative bias being larger. Although the double barrier structure was designed to be symmetric, diffusion of dopants into the $Si_{0.78}Ge_{0.22}$ cladding layers is difficult to control and may account for the observed differences. For the data presented in this thesis the structures were biased to yield the lowest current density, which usually corresponded to the negative bias direction.

Two resonances in the tunneling current are clearly seen for both positive and negative bias directions. It is noted that the current is discontinuous and unstable in the negative differential resistance (NDR) regions. This instability is due to oscillations with the measuring circuit, and could be eliminated by the addition of a small shunt resistance in parallel with the sample. The resonances are broad in comparison to those observed for III-V systems (Ben Amor *et al.*, 1989; Eaves *et al.*, 1988), and the peak-to-valley current ratios are approximately 2.8:1 and

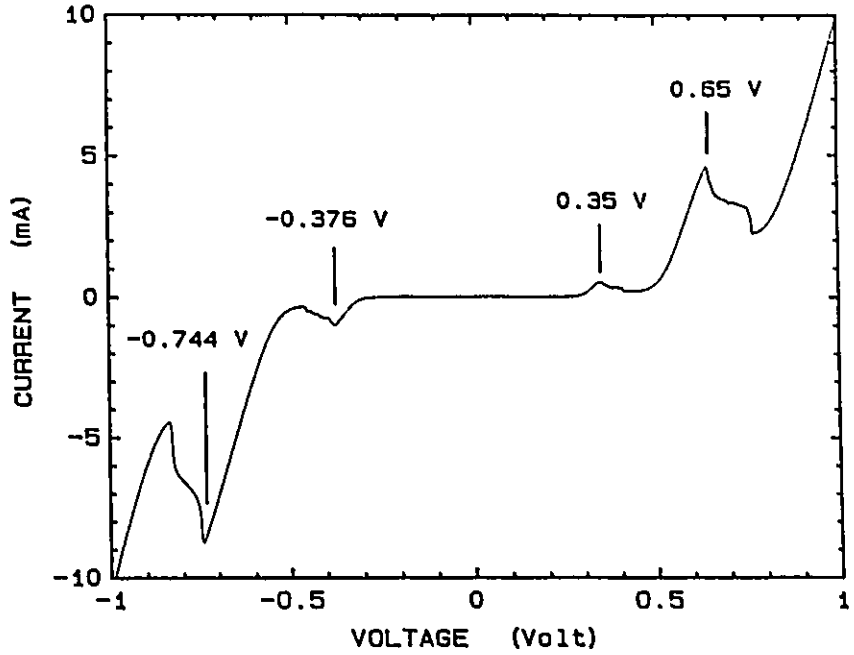


Fig. 6.1 The current-voltage characteristic of MBE 519B (100 μm mesa). The bias positions of the resonance peaks are indicated by the arrows. The NDR regions are unstable due to oscillations with the measuring circuit.

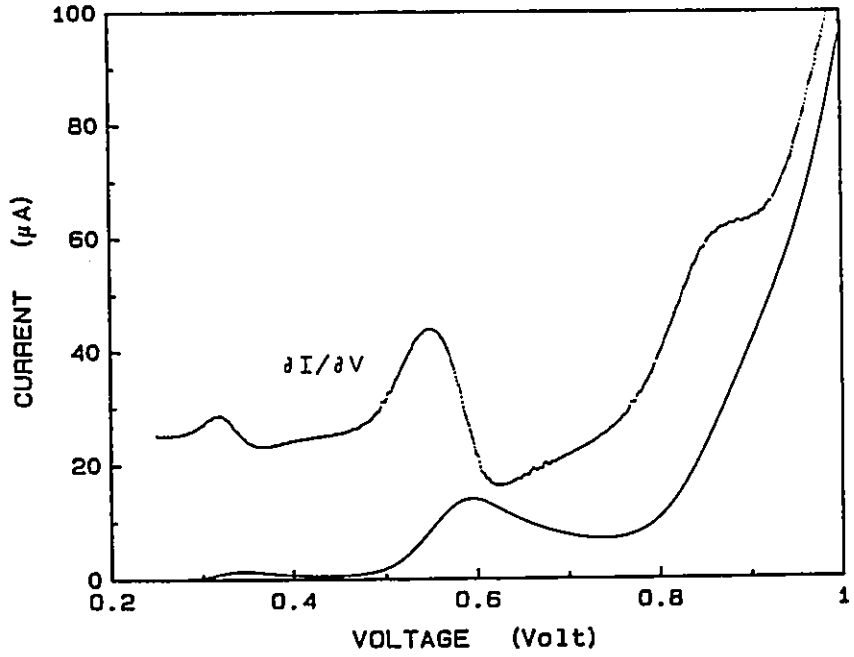


Fig. 6.2 The current-voltage characteristic of MBE 519D (7 μm mesa) along with the differential conductance (dotted line) shown in arbitrary units.

2.4:1 for the lower and higher bias voltage resonances respectively. Although these ratios are small compared to peak-to-valley ratios measured for electron tunneling (in excess of twenty in GaAs/GaAlAs and greater than two hundred in InGaAs/InAlAs), these results are excellent for hole tunneling and confirm the high quality of the MBE growth.

Figure 6.2 shows a typical I-V characteristic of a 7 μm diameter mesa (MBE 519D). It is seen in this plot that the current in the NDR regions is now stable, and a smooth curve is obtained for all bias voltages. The peaks of the resonances occur at bias voltages of 0.344 V and 0.59 V, somewhat lower than the positions measured for the 100 μm sample of Figure 6.1, that is, 0.35 V and 0.65 V. A series resistance of approximately 20 Ω has been assumed (Liu *et al.*, 1989) to take into account the contribution from the metal contacts, the doped p^+ contact electrodes and the two heterojunctions formed at the Si/SiGe interfaces away from the double barrier structure. Since the resistances of the small diameter devices are large compared to the series resistance, the correction to the bias voltage positions of the observed resonances is negligible. The resistance of the larger devices are comparable to the series resistance and the positions of the resonances are corrected to lower the bias potential. If the value of the series resistance is taken to be 10 Ω, the agreement in the bias positions between all mesa sizes is excellent. From this it is concluded that the bias positions of the two observed resonances are 0.344 V and 0.595 V. This series resistance also accounts for the current instability observed in the NDR regions for the larger mesas: the magnitude of the sample and series resistances are of the same order. Conversely, the smaller structures exhibit no oscillation because their resistances are much larger ($\sim 10^3 - 10^6$ Ω).

Additionally, a new feature appears at a higher bias voltage ~ 0.9 V, displayed as a change in the curvature of the I-V characteristic. Although this

structure is not easily resolved in the I-V curve, a numerical derivative of the current with respect to bias voltage makes it more prominent. In the differential conductance, the curvature change appears as a plateau, as shown in the overlay to Figure 6.2. Such a feature is typical of a weak resonance, which contributes a small additional tunneling current to a large background current.

A comparison of the experimental data with results from the subband dispersion calculations given in §4.4 indicate the excellent agreement for the positions of the light hole resonance LH1 at 0.344 V and the heavy hole resonance HH2 at 0.595 V. No resonance was observed at the calculated position of the HH1 level which may be attributed to the very low current density at that bias voltage. At higher bias voltages the effective barrier height will be significantly altered by the applied bias, and the dispersion calculation must be redone to take into account another resonance at approximately 0.9 V.

6.2 Magnetic Field Dependence

6.2.1 Field applied normal to the layers

The effect of a magnetic field on the I-V characteristic curve is shown in Figure 6.3a for several values of the field applied parallel to the tunneling direction ($\mathbf{B}||\mathbf{J}$) or equivalently normal to the plane of the interfaces.

Examining the first resonance (LH1), more easily seen in the inset to Figure 6.3a, it is observed that the bias potential corresponding to the peak current remains essentially the same at all fields below 10 Tesla, and moves slightly towards higher bias above 10 Tesla. The shift of the peak at higher fields varies in direction and amplitude from sample to sample and is believed to be associated with a change in the lineshape of the resonance itself. A suppression of the tunneling current is

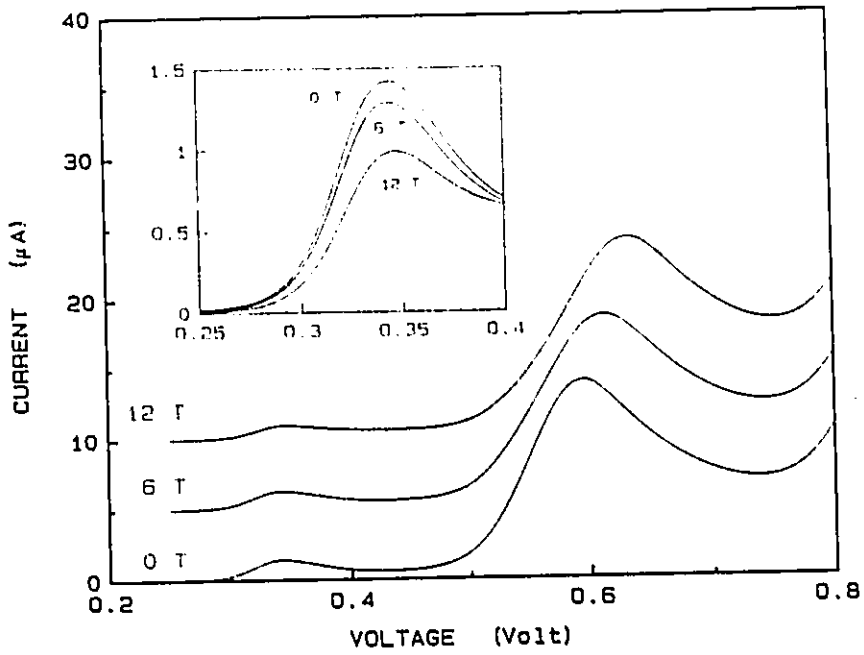


Fig. 6.3a The magnetic field dependence of the tunneling current for MBE 519D (7 μm) for fields B = 0, 6, and 12 Tesla applied normal to the plane of the layers. The inset shows a blow-up of the LH1 resonance

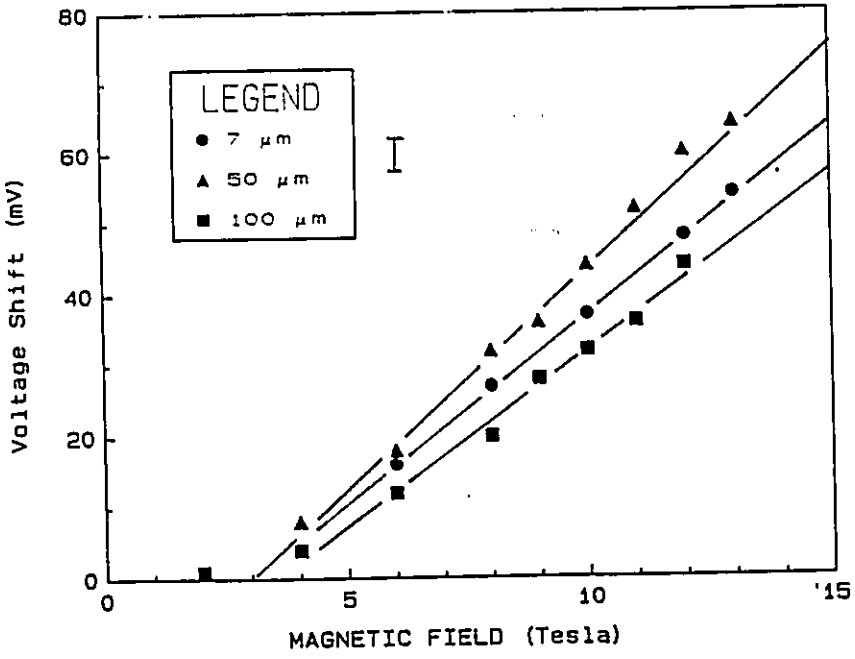


Fig. 6.3b The shift of the HH2 resonance as a function of applied field normal to the plane of the layers for several mesa diameters. The shift is linear above 2 Tesla. The typical error bar is drawn next to the legend.

observed with increasing magnetic field for resonant bias voltage conditions while the magnitude of the current prior to the onset of the resonance remains unchanged for all fields. The net effect is a reduction of the peak-to-valley ratio from 2.4:1 at zero field to 1.2:1 at 13 Tesla.

The behaviour of the second resonance (HH2) in the presence of a magnetic field is significantly different from that of LH1. The bias voltage corresponding to the peak tunneling current is magnetic field dependent shifting upwards with increasing field. The field dependence of the voltage shift is linear as shown in Figure 6.3b for several mesa diameters. The magnitude of the peak tunneling current increases slightly ($\sim 6-7\%$) over a field range of 13 Tesla. However, as for the LH1 resonance, there is a net reduction in the peak-to-valley ratio from 2.0:1 at zero field to 1.6:1 at 13 Tesla due to the rapid increase of the valley current with increasing magnetic field.

In the samples investigated, the observed negative differential resistance regions correspond to resonant tunneling of holes from the heavy hole band in the emitter to three quasi-bound states in the quantum well. The emitter region is treated as a bulk strained $Si_{0.78}Ge_{0.22}$ layer. The strain due to the lattice mismatch between Si and $Si_{0.78}Ge_{0.22}$ induces a splitting of the light and heavy hole bands and only heavy holes exist in the emitter at low temperatures. It is assumed that these heavy holes are confined to a 2D subband formed at the $Si/Si_{0.78}Ge_{0.22}$ interface. As discussed in §4.3, the strain also leads to an anisotropic effective mass, with the in-plane mass equal to 0.16 m_o and the perpendicular mass equal to 0.44 m_o for the heavy hole band in the emitter. The magnetic field dependence suggests that the first resonance corresponds to heavy holes in the emitter tunneling to a light hole band (LH1) in the well with the effective mass in both layers approximately equal to account for the constant resonant energy as a function of magnetic field. The

calculated in-plane dispersion for LH1 is electron-like at $k = 0$ with an effective mass – 0.14 m_0; there is a maximum in the dispersion at $k = 0.02$ Å^{-1} where the local effective mass is 0.3 m_0 (Liu *et al.*, 1989). At non-zero in-plane wavevector, mixing of the light and heavy hole states can occur because of the loss of full rotational symmetry. Tunneling from a heavy hole band in the emitter to a light hole band in the well is taken into account in a one-particle treatment by switching from a heavy hole to a light hole character in the first barrier (*cf.* Buchanan *et al.*, 1990). From the results presented in Figure 6.3a, there is no measurable shift of the LH1 resonance below 10 Tesla; this is consistent with the $k = 0$ in-plane effective mass associated with the LH1 band in the well and the HH band in the emitter, 0.14 m_0 and 0.16 m_0 respectively.

The bias position of the HH2 resonance is highly field dependent, with a shift to higher bias voltage of 60 mV over a range of 13 Tesla. By using Equation (3.23), the shift in bias voltage can be related to the difference in effective masses of the emitter and well states. The size of the shift indicates that the planar effective masses are substantially different, and the direction of the shift towards higher bias implies that the effective mass in the quasi-bound level of the well must be smaller than that in the emitter. This is consistent with resonant tunneling from the heavy hole band in the emitter with $m^* = 0.16\ m_0$ to the heavy hole band HH2 in the well with $m^* = 0.06\ m_0$ as predicted by the in-plane dispersion calculation presented in §4.5.

In §3.5, it was shown that the application of a voltage across the double barrier structure results in a gradient of the potential in each layer. For small bias voltages, the gradient in the well will be small and the corrections to the bound state energies, treated by perturbation theory, will be correspondingly small. For larger voltages, the gradients will be steep and the effective width of the barriers may be altered significantly. At sufficiently high voltages the barriers will appear triangular

and particles will traverse the barrier via Fowler-Nordheim tunneling (Duke, 1969; Mendez, 1987). This change of the barrier profile results in the barriers becoming narrower and transparent: a rapid rise in the tunneling current is then expected. The current-voltage characteristic of the samples investigated displays this monotonic increase in current beyond the HH2 resonance. The feature at ~ 0.9 V is weak and is only resolved in the differential conductance as a plateau in a rapidly increasing background current. While it is difficult to measure the dependence of this resonance with respect to magnetic field, an upper bound of +60 mV over 13 Tesla can be estimated on the shift in bias position. Assuming that 1/3 to 1/5 of the total applied voltage is dropped across the double barrier structure, then the planar mass associated with this quasi-bound level must be significantly different from that of the emitter (of the order of ~0.08 m_o). Calculations of the transmission coefficient of an incoming HH state through a $Si/Si_{1-x}Ge_x$ DBTS (Wessel and Altarelli, 1989) suggests that the observed resonance is due to tunneling from an incoming heavy hole state to an outgoing LH2 state, possibly mixed with a spin-orbit split-off hole (SOH) band state in the well.

6.2.2 Field applied in the plane of the layers

The effect of a strong magnetic field applied perpendicular to the tunneling current ($\mathbf{B} \perp \mathbf{J}$) or in the plane of the interfaces was examined in a range of magnetic fields up to 14.3 Tesla. The results for several representative fields are shown in Figure 6.4 for a 12 μm mesa diameter. The position of the peak current is shifted towards higher bias voltage for both LH1 and HH2 resonances, and the shift of the HH2 peak is greater than that measured with the field applied normal to the interface, increasing 85 mV over 14.3 Tesla. As displayed in the inset to Figure 6.4, a quadratic dependence of the shift on the magnetic field is an acceptable fit to the

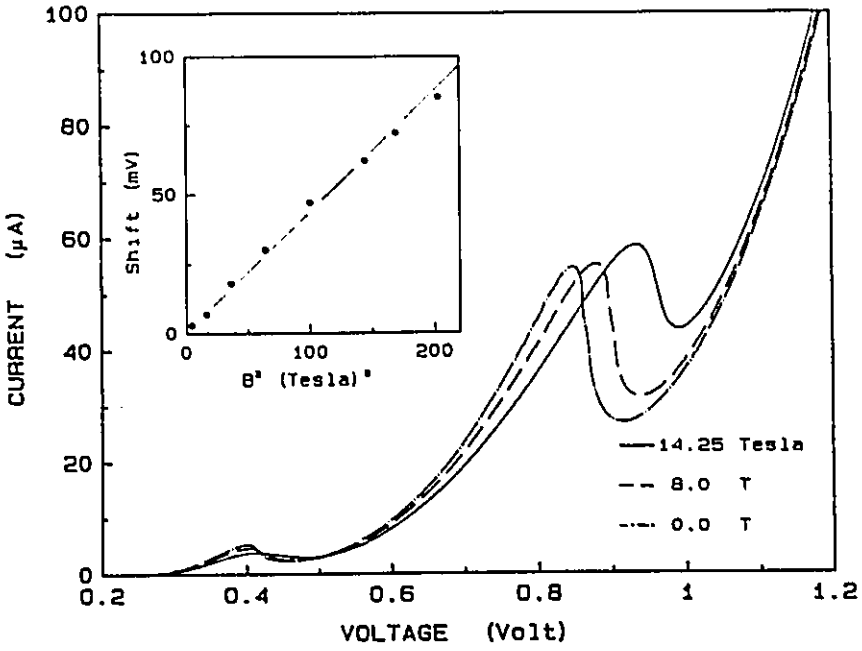


Fig. 6.4 The current-voltage characteristic of MBE 519A (25 μm) in the presence of a planar magnetic field. The two resonances are both shifted towards higher bias voltages. The inset shows the shift of the HH2 resonance fitted to a B^2 dependence.

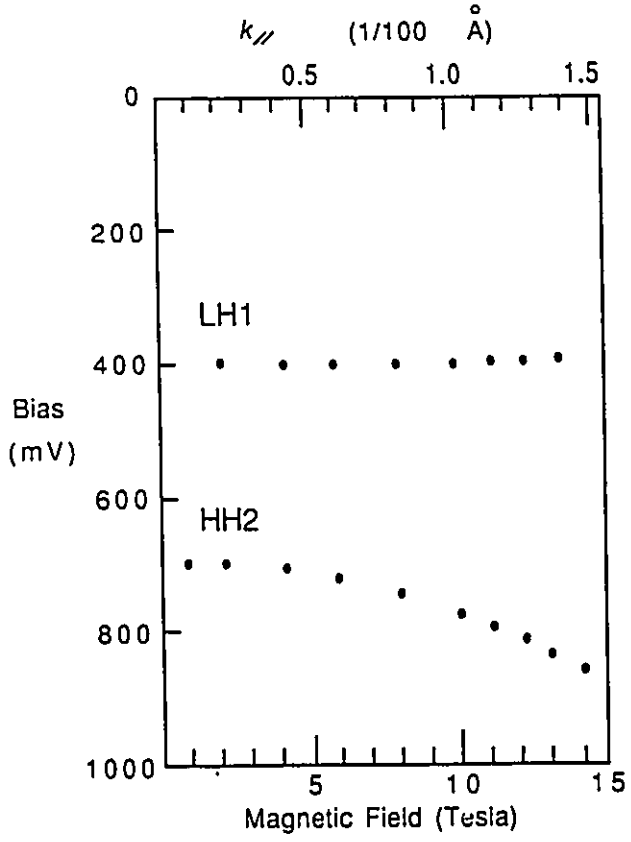


Fig. 6.5
The bias position of the resonance peaks plotted as a function of planar magnetic field. The top scale shows the equivalent planar wavevector, calculated using Equation (3.26). Hole energy is plotted downwards in this diagram.

data. Similar to the results of §6.2.1, there is a slight increase in the magnitude of the peak current of ~ 5% over the entire field range.

The behaviour of the peak of the LH1 resonance is not a simple function of the magnitude of the field: the shift of the peak current is initially towards lower bias, changing direction and reaching a local maximum at approximately 6 – 8 Tesla, and then moving again towards lower bias voltages. The entire range of voltages over which the peak is measured is only 6 mV, which is 15× smaller than the shift of the HH2 resonance.

In Figure 6.5, the bias position of the resonances is plotted as a function of magnetic field (**B**⊥ **J**). The top axis is scaled in equivalent units of $k_{//}$, calculated from Equation (3.26) where the Fang-Howard parameter is taken to be 5 nm^{-1}. In this diagram, the energy scale increases from top to bottom for the purpose of comparison with the calculated dispersion curves of Figure 4.5.

As discussed in §3.7, a magnetic field applied parallel to the interface can be used to determine the dispersion curves of holes tunneling into the quantum well. In this field configuration, the tunneling resonances are observed at higher bias voltage because the momentum wavevector k_y in the emitter is shifted by $2eBz/\hbar$. Holes with different momenta k_y will not be subjected to the same shift in effective potential and this redistribution of energies is manifested by a broadening in the tunneling current near the peak of the resonances. The experimental results are quite consistent with this interpretation. A plot of the shifts of the resonance peaks as a function of B^2 reveals a quadratic shift of the HH2 peak attesting to the parabolicity of the band. There is very little displacement of the LH1 peak with magnetic field, and the small shift to lower biases is consistent with its electron-like dispersion. The magneto-tunneling results reflect the calculated dispersion of the

LH1 and HH2 states in the well shown in Figure 4.5. The variation of the planar magnetic field is equivalent to sweeping through the planar wavevector and can be used as a spectroscopic technique to map the dispersion curves $E(k_{//})$ of the quasi-bound states in the well.

6.3 Non-resonant tunneling

6.3.1 Magnetic field dependence

A detailed examination of the tunneling current was undertaken, which revealed additional structure in the valley current of the LH1 resonance. In Figure 6.6, it is seen that the lineshape of the valley current changes dramatically with magnetic field; an additional feature is readily resolved by ~ 13 Tesla, appearing as a secondary hump in the I-V curve. This feature moves towards higher bias voltage as a function of magnetic field, and extrapolates back to the position of the LH1 resonance at zero field. As this feature is not easily observed in the broad valley of the LH1 resonance, its magnetic field dependence was obtained by measuring the tunneling current as a function of field while maintaining a fixed bias voltage. An oscillatory behaviour of tunneling current is observed as a function of the magnetic field, as displayed in Figure 6.7 for three bias voltages. Although the zero field current value differs for each bias voltage, all curves exhibit an initial decrease with field and a subsequent change in curvature leading to a local minimum in the tunneling current. The current then peaks and decreases again with increasing magnetic field. This data can be analyzed to follow the displacement of the emerging feature in the valley current as a function of field.

In taking a numerical derivative of the current with respect to the voltage, an additional feature centered at ~ 0.55 V is resolved when the magnetic field is

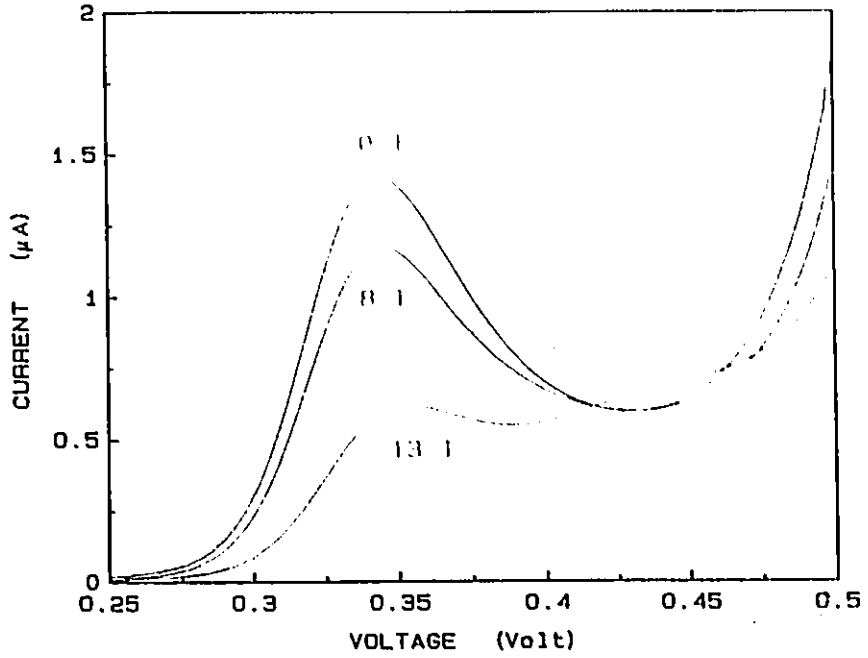


Fig. 6.6 The valley current of the LH1 resonance (MBE 519D, 7 μm) for the magnetic field applied normal to the layers B = 0, 8 and 13 Tesla. Along with the large reduction of the peak amplitude a secondary peak emerges at ~ 0.42 V.

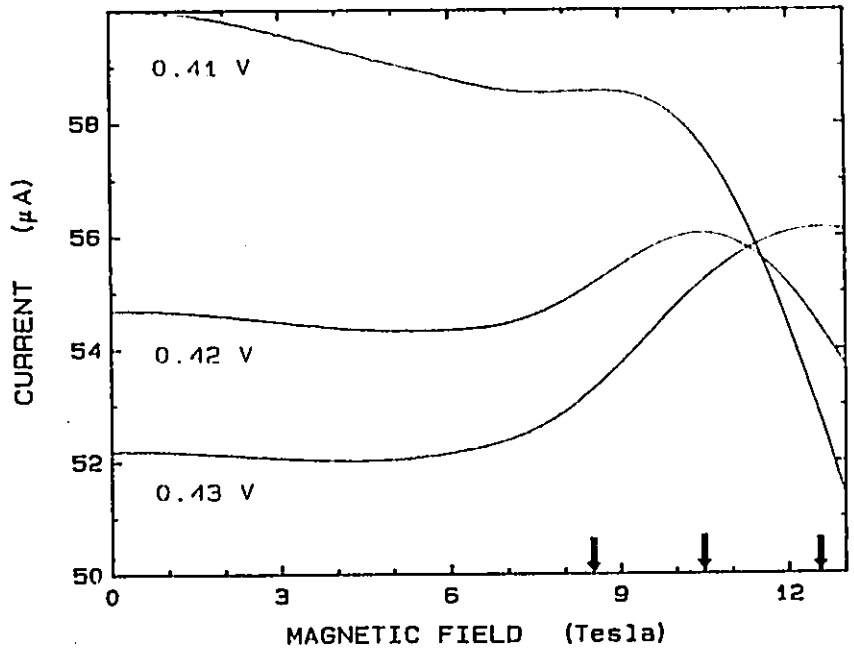


Fig. 6.7 The magnetic field dependence of the valley current for bias voltages of V = 0.41 V, 0.42 V and 0.43 V. The arrows mark the field positions of the local peaks in the current.

increased beyond 6 Tesla. This is displayed in Figure 6.8a where a broad shoulder is seen to emerge in the differential conductance; it is most prominent between 6 – 8 Tesla, appearing as a double hump in the differential conductance. To further investigate this feature, an AC measurement of the differential conductance was made, and the results are shown in Figure 6.8b. An advantage of this method is that noise in the data is greatly reduced from that observed in a numerical derivative without having to filter or smooth the data and thereby lose sensitivity in the measurement. As expected, the shoulder appears in the differential conductance at the same field and voltage values as observed in Figure 6.8a, which confirms the validity of both means of differentiation.

Whereas resonant tunneling requires that the total energy and momentum in the plane of the layers be conserved, non-resonant tunneling can occur via elastic and inelastic scattering events which relax the conservation conditions. As discussed in §3.8, holes may tunnel non-resonantly via elastic scattering from the $m = 0$ emitter level to the first excited Landau level ($n = 1$) in the well or inelastically from the $m = 0$ level in the emitter to Landau levels ($n = 0, 1, 2, ...$) in the well with the emission of an optical phonon of energy $\hbar\omega_L = (n - m)\hbar\omega_c$. These processes can be distinguished by examining the magnetic field dependence of the bias voltages at which they occur. It is useful to plot the bias voltage of both the LH1 and HH2 resonance peaks and their associated features as a function of magnetic field, as illustrated in Figure 6.9 for MBE 519A (50 μm). This is known as a fan diagram and the data fall onto sets of lines which, in order of increasing bias voltage, are identified as: the LH1 resonance which is independent of magnetic field, an associated feature which extrapolates to the LH1 bias position at B = 0 T, a feature at constant energy spacing from LH1 unchanged by the magnetic field and finally the HH2 resonance.

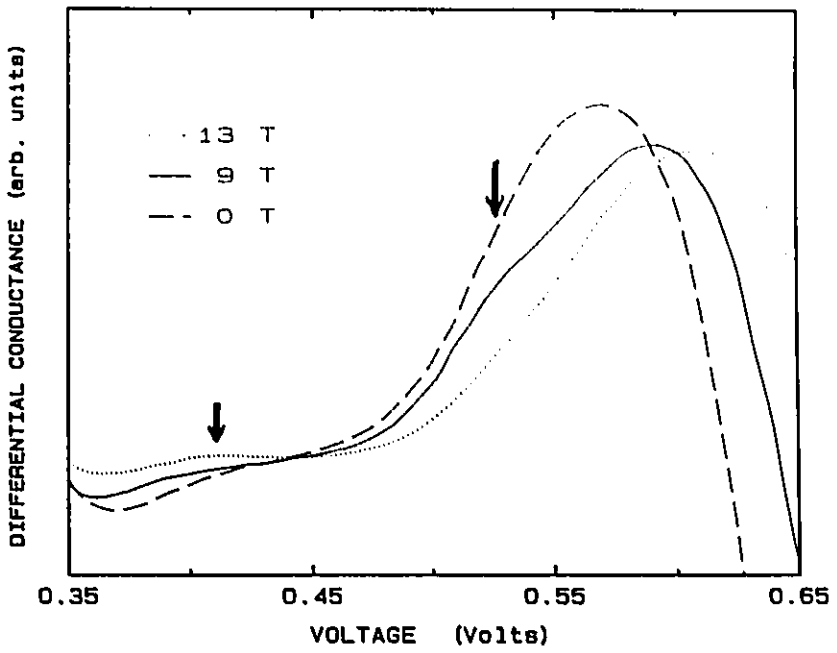


Fig 6.8a The differential conductance (numerical derivative) of MBE 519D (7 μm) for magnetic fields normal to the layers B = 0, 9 and 13 Tesla. Secondary features are observed at ~ 0.42 V and 0.55 V.

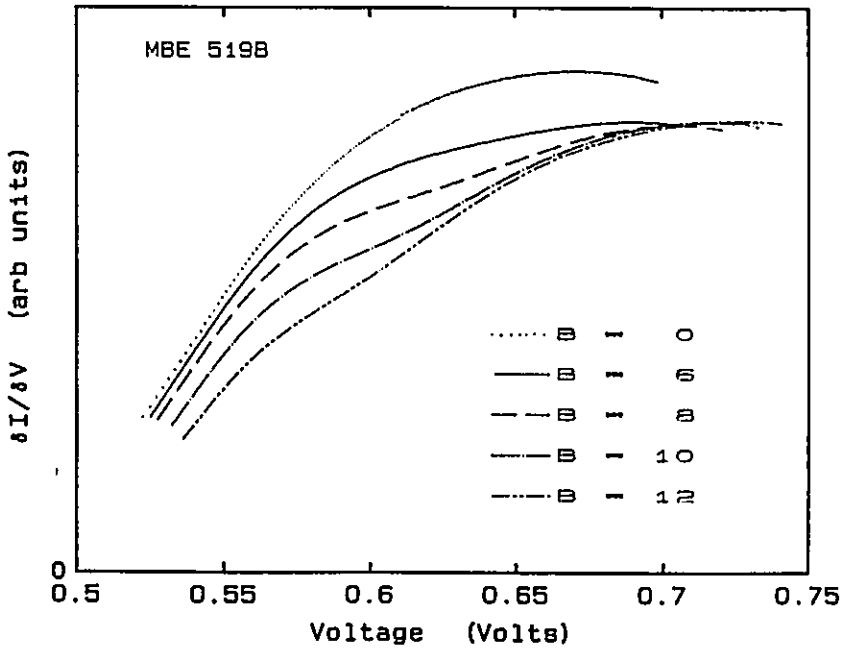


Fig. 6.8b The differential conductance of MBE 519B (100 μm) obtained from an AC modulation technique.

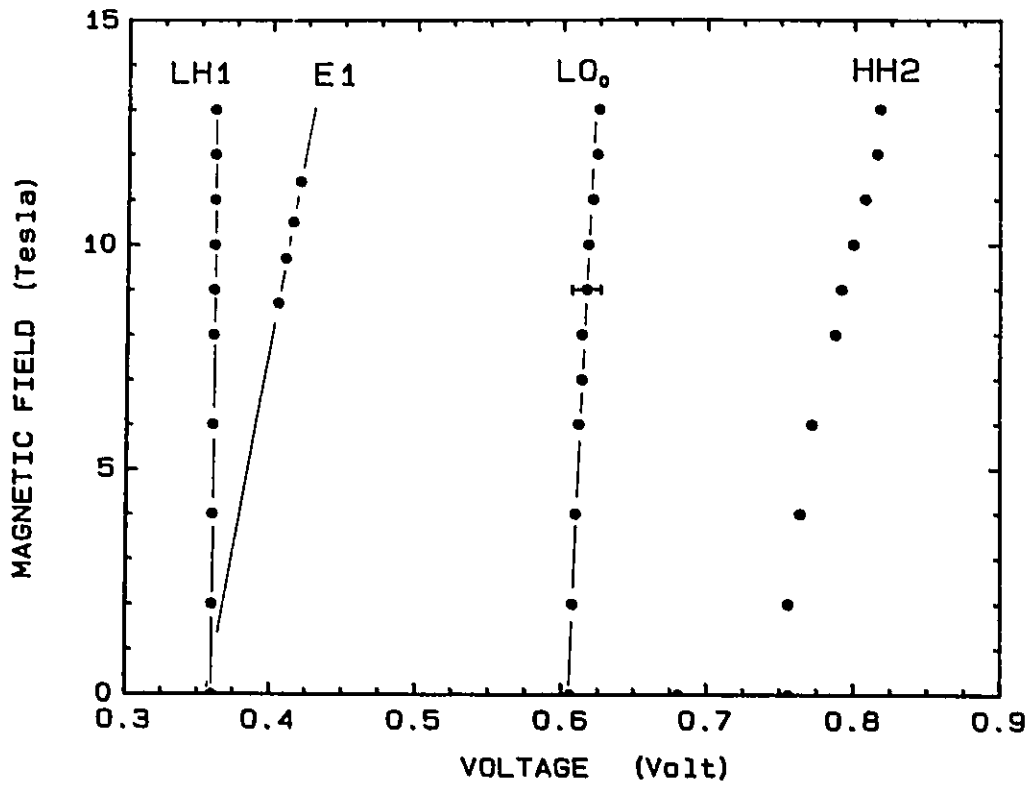


Fig. 6.9 A plot of the bias positions of features in the tunneling current as a function of magnetic field applied normal to the interfaces of the double barrier structure. LH1 the light hole resonance peak, E1 the elastic scattering process, LO_0 a inelastic scattering process for which $n - m = 0$, and HH2 the resonance due to the second heavy hole subband in the well. An error bar is drawn for the LO_0 process which indicates the uncertainty in the determination of this position. The uncertainty associated with the other points is too small to be displayed on this scale.

This plot clearly shows the appearance of two non-resonant tunneling features which vary linearly with field at ~ 0.42 V and ~ 0.60 V. The first feature displays a distinct magnetic field dependence obtained from the I-B curves of Figure 6.7. An extrapolation to zero magnetic field shows that this feature originates from the LH1 resonance. This set of secondary tunneling maxima is attributed to non-resonant tunneling involving an elastic scattering-induced transition from the $m = 0$ Landau level in the emitter to a $n = 1$ Landau level in the quantum well. This transition can take place via scattering from ionized impurities or roughness at the interface. The fact that this feature is more prominent in the smaller diameter devices supports this identification: the effect of the etchant on the edges of the mesa will be more important in devices where the diameter is comparable to the total thickness.

The feature at ~ 0.60V does not exhibit a significant shift in bias position, as indicated by the representative error bar drawn in Figure 6.9. Its position with respect to the LH1 resonance is approximately 220 mV, and remains constant for all magnetic fields. Assuming a drop of ~ 1/3 of the total applied bias voltage for consistency with the energy calculations for LH1, this corresponds to an energy of ~ 62 meV. Referring to the calculated phonon mode energies in Ghanbari and Fasol (1989), the agreement with the longitudinal optical Si-Si phonon mode at 62.5 meV is excellent. Therefore the shoulder in the differential conductance is interpreted as inter-Landau level inelastic scattering of the tunneling particle from the $m = 0$ Landau level in the emitter to the $n = 0$ level in the well with the emission of an LO phonon.

6.3.2 Temperature dependence

Although strongest at low temperatures, the resonant tunneling features are known to persist up to 150 K (Liu *et al.*, 1989). In order to shed some light on the origin of the non-resonant tunneling features, additional measurements were taken over a range of temperatures. Figure 6.10 shows the tunneling current about the LH1 resonance at 4.2 K and 77 K. The increase in temperature results in a broadening of the resonances and a decrease of the peak-to-valley ratios. There is a slight temperature dependence of the position of the resonance peaks, resulting in small (~ 2 mV) shift towards lower bias with increasing temperature.

The non-resonant features are best resolved in a numerical derivative of an AC differential measurement, as shown in Figure 6.11 for a field ($\mathbf{B}||\mathbf{J}$) of 7 Tesla. In this plot of $\partial^2 I/\partial V^2$ the local minima correspond to features in the I-V curve. The broad non-resonant feature observed at ~ 0.6 V decreases in amplitude with increasing temperature and by 77 K has virtually disappeared. The elastic scattering feature which is observed in the I-V curve above 12 Tesla could not be investigated above 4.2 K since fields of that magnitude are beyond the limits of the apparatus. However, a feature in the differential conductance is seen at 0.44 V at all magnetic fields $0 < B < 8$ Tesla; it is strongest at low temperatures and by 77 K has almost disappeared. It is observed at the same bias position irrespective of the magnetic field.

In plotting $\partial I/\partial V$ and $\partial^2 I/\partial V^2$ two new features are now resolved, appearing in the onset of the LH1 resonance (~ 0.25 V) and very close to the peak tunneling current of the HH2 resonance (~ 0.7 V). The feature at lower bias is only observed at temperatures T ~ 20 K, and the amplitude increases with temperature up

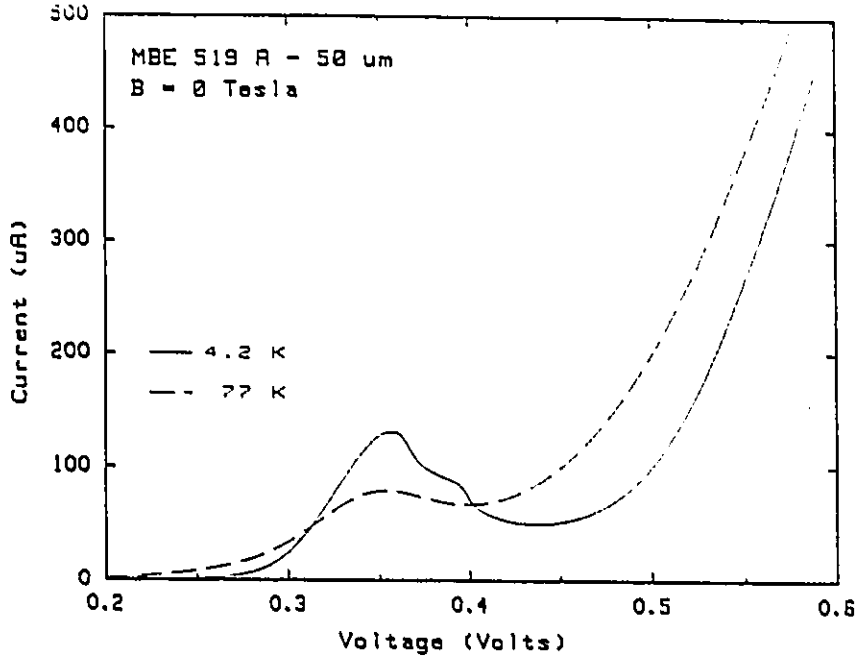


Fig. 6.10 The current-voltage characteristic of MBE 519A (50 μm) for bias voltages about the LH1 resonance, at temperatures T = 4.2 and 77 K.

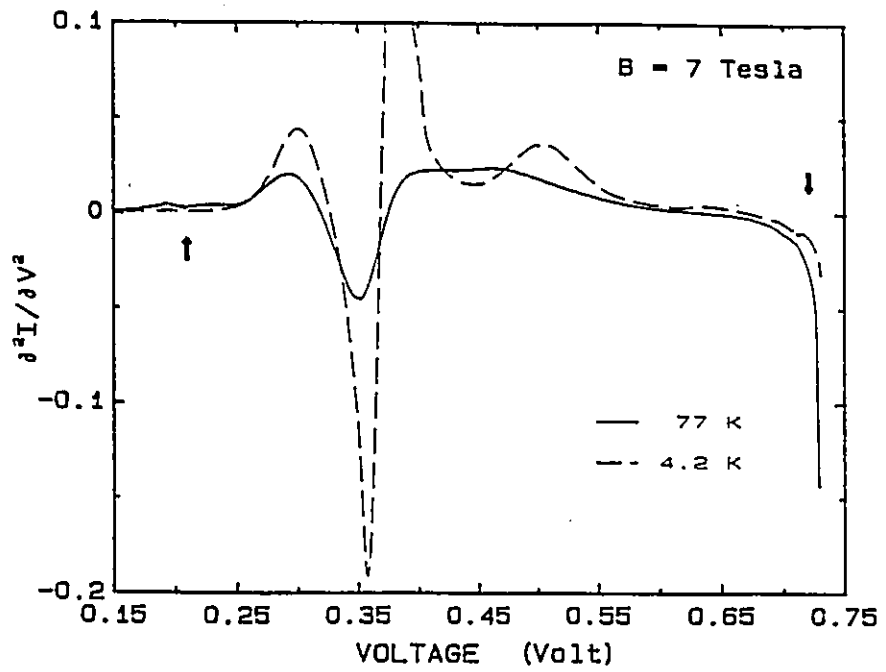


Fig. 6.11 The second derivative of the current with respect to bias voltage for a magnetic field B = 7 Tesla normal to the layers.

to 77 K. No field dependence of either the magnitude or position was observed up to 7 Tesla. Although the feature near the peak of the HH2 resonance is extremely weak at all temperatures and can only be seen in $\partial^2 I/\partial V^2$, it is strongly dependent on magnetic field, shifting ~ 20 – 30mV over a range of 7 Tesla. This shift is therefore comparable in magnitude to that of the HH2 resonance.

The states occupied by the electron gas are described by the Fermi-Dirac distribution. Only those carriers within ~ kT of the Fermi energy participate in electrical transport (*cf.* Equation 3.2). At 4.2 K, kT is approximately 0.5 meV whereas at 77 K this corresponds to an energy of ~ 7 meV. Consequently, as the temperature is increased carriers with a wider range of energies are available for tunneling. This broadening with increasing temperature is indeed observed in the samples investigated. Additionally, the scattering probability increases with temperature, and this accounts for the reduced peak-to-valley ratios observed. Because of thermal broadening of the Landau levels the cyclotron energy is only defined to within kT, the mean thermal energy. This implies that inter-Landau level transitions will be strongest at zero temperature and are expected to be smeared out at sufficiently high temperature. The broad feature centred around ~ 0.6 V attributed to a LO phonon emission process disappears above 40 K. The feature at ~ 0.44 V is suppressed with increasing temperature and by 77 K is almost indiscernible. It does not shift with increasing magnetic field and is therefore believed to be associated with the LH1 resonance. A tentative identification of this non-resonant feature is an inelastic scattering process mediated by phonon emission possibly an optical Ge-Ge phonon mode. This must be confirmed through further experimental studies and theoretical calculations of phonon modes for the Si/SiGe double barrier structures.

The weak feature seen in the onset of the LH1 resonance appears only at higher temperatures. A possible explanation of this feature may be found in examining the distribution of carriers in the emitter. All previous derivations assumed that the holes in emitter were tunneling from 2D states in the accumulation layer. However, there may also be a population of holes occupying 3D states with energies between E_F and the conduction band edge E_c. Resonant tunneling for these carriers occurs at a bias $V = 2E_F/e$ and cut-off for a bias $V = 2E_c/e$, prior to the cut-off for the 2D states. Since the occupancy of the 3D states increases with temperature, their contribution to the tunneling current is expected to increase. A precise quantitative analysis of this effect is not possible from the data, but qualitatively the results support such a mechanism. Experimental evidence for tunneling via 3D band edge states has been reported in n-type structures (in particular Buchanan *et al.*, 1990) and supported by the theoretical work of Chevoir and Vinter (1990). The feature at the onset of the HH2 resonance (~ 0.7 V) is very weak and has been observed only at temperatures above 4.2 K. It has a field dependence similar to that of HH2 and although the lack of experimental data does not allow a convincing identification of this feature, it is worth noting that the presence of Landau levels has been reported at a similar bias position by others (Schuberth *et al.*, 1991).

VII. Conclusions

The study of the magneto-transport properties of holes in $Si/Si_{0.78}Ge_{0.22}/Si$ double barrier structures presented in this thesis is one of the first to explore the resonant and non-resonant tunneling processes of holes in either group III-V or group IV materials. This work was performed in the temperature range 1.7 K to 77 K, and in planar and perpendicular magnetic fields up to 14.3 Tesla. The diameter of the devices ranged from 7 μm to 100 μm. In addition to the resonant tunneling features previously observed in these structures, new evidence was obtained for tunneling via a third quasi-bound state at high bias voltage. Although non-resonant tunneling processes have previously been observed in n-type double barrier structures, the results reported in this thesis constitute the first such observation in p-type systems. Non-resonant features associated with the LH1 resonance were observed in perpendicular fields and correspond to inter-Landau level transitions via elastic scattering mediated by impurities and inelastic scattering with the emission of optical phonons. The interpretation of the data in terms of inelastic scattering processes mediated by optical phonons is supported by theoretical calculations of phonon modes in Si/SiGe superlattices (Ghanbari and Fasol, 1989). A novel spectroscopy of the quasi-bound states in the well was performed by monitoring the shift in energy of the resonant tunneling features as a function of applied planar magnetic field. This yielded the dispersion curves $E(k_{//})$ of the resonant states and confirmed the parabolic and non-parabolic nature of the HH2 and LH1 subbands respectively. The appearance of an additional feature at the onset of the LH1 resonance with increasing temperature was tentatively attributed to tunneling of holes from 3D states at the heavy hole band edge. This interpretation is supported by experimental studies in GaAs (Buchanan *et al.*, 1990) and calculations by Chevoir and Vinter (1990).

It would be informative to extend this work to structures with a wider quantum well to decrease the energy of the third quasi-bound state allowing further characterization of this resonant state. A recent claim by Schuberth *et al.* (1991) that Landau levels can be observed at the onset of the HH2 resonance should be checked in samples with similar growth parameters. The Si/SiGe/Si double barrier structure has proven to be an excellent system to demonstrate the presence of non-resonant tunneling processes and the experimental results are consistent with recent theoretical calculations.

References

Anderson, R. L., I.B.M. J. Res. and Dev. 4, 283 (1960).

Ando, T., Fowler, A. B. and Stern, F., Rev. Mod. Phys. **54**, 437 (1982).

Ashcroft, Neil W., and Mermin, N. David, Solid-State Physics, (Holt, Rinehart and Winston, Philadelphia, 1976).

Bauer, G., F. Kuchar, and Heinrich, H., Ed., Two-dimensional Systems, Heterostructures, and Superlattices, Springer Series in Solid-state Sciences **53** (1984).

Beerens, J., Ph. D. Thesis, Université de Sherbrooke, Sherbrooke, Canada (1986).

Ben Amor, S., Martin, K. P., Rascol, J. J. L., Higgins, R. J., Torabi, K. P. A., Harris, H. M. and Summers, H. J., Appl. Phys. Lett. **53**, 2541 (1988).

Ben Amor, S., Martin, K. P., Rascol, J. J. L, Higgins, R. J., Potter, R. C., Lakhani, A. A. and Hier, H., Appl. Phys. Lett. **54**, 1908 (1989).

Binnig, G., Rohrer, H., Gerber, C. and Weibel, E., App.. Phys. Lett. **40**, 178 (1981).

Buchanan, M., Liu, H. C., Powell, T. G. and Wasilewski, Z. R., J. Appl. Phys. **68**, 4313 (1990).

Chang, L. L., Esaki, L. and Tsu, R., Appl. Phys. Lett. **24**, 593 (1974).

Chang, L. L. and Ploog, K., Ed., Molecular Beam Epitaxy and Heterostructures, NATO ASI Series B, (1985).

Chevoir, F. and Vinter, B., "*Resonant and scattering-assisted magnetotunneling*", presented at the NATO Advanced Research Workshop on Resonant Tunneling in Semiconductors: Physics and Applications (May 1990), to be published by Plenum Press, New York.

Cohen-Tannoudji, C., Diu, B. and Laloë, F., Quantum Mechanics, (John Wiley and Sons, New York, 1977).

Duke, C. B., Tunneling in Solids, (Academic Press, New York, 1969).

Eaves, L., Toombs, G. A., Sheard, F. W., Payling, C. A., Leadbeater, M. L., Alves, E. S., Foster, T. J., Simmonds, P. E., Henini, M., Hughes, O. H., Portal, J. C., Hill, G., and Pate, M. A., Appl. Phys. Lett. **52**, 212 (1988).

Eaves, L., Hayden, R. K., Maude, D. K., Leadbeater, M. L., Valadares, E., Henini, M., Hughes, O. H., Portal, J. C., Cury, L., Hill, G. and Pate, M. A., "*Magneto-tunneling spectroscopy to measure the electron and hole $\varepsilon(k)$ dispersion curves in the quantum well of resonant tunnelling structures*", conference proceedings from The Application of High Magnetic Fields in Two-dimensional Transport, Wurzburg, Jul. 31 to Aug. 4, 1990, to be published in the Springer Series in Solid State Physics, (Spring-Verlag, Berlin).

Fowler, A. B., Fang, F. F., Howard, W. E. and Stiles, J. P., Phys. Rev. Lett. **16**, 901 (1966).

Fowler, R. H., and Nordheim, l, Proc. Roy. Soc. (london) **A119**, 173 (1928).

Franz, W., "*WKB Methods*", in Tunneling Phenomena in Solids, edited by E. Burstein and S. Lundqvist (Plenum Press, New York, 1969).

Ghanbari, R. A. and Fasol, G., Sol. State Commun. **70**, 1025 (1989).

Gregoris, G., Ph. D. Thesis, Institut National des Sciences Appliquées, Toulouse, France (1985).

Houghton, D. C., Lockwood, D. J., Dharma-wardana, M. W. C., Fenton, E. W., Baribeau, J. M., and Denhoff, M. W., J. Crystal. Growth **81**, 434 (1987).

Kane, E. O., "*Basic Concepts of Tunneling*", in Tunneling Phenomena in Solids, edited by E. Burstein and S. Lundqvist (Plenum Press, New York, 1969).

Kleiner, W. H. and Roth, L. M., Phys. Rev. Lett. **2**, 334 (1959).

Liu, H. C., Landheer, D., Buchanan, M., and Houghton, D. C., Appl. Phys, Lett. **52**, 1809 (1988).

Liu, H. C., Landheer, D., Buchanan, M., Houghton, D. C., D'Iorio, M. and Kechang, S, Superlattices and Microstructures **5**, 213 (1989).

Mendez, E. E., Esaki, L. and Wang, W. I., Phys. Rev. B **33**, 2893 (1986).

Mendez, E. E., Calleja, E., Goncalves da Silva, C. E. T., Chang, L. L. and Wang, W. I., Phys. Rev. B **33**, 7368 (1986).

Mendez, E. E., Wang, W. I., Calleja, E. and Goncalves da Silva, C. E. T., Appl. Phys. Lett. **50**, 415 (1987).

Mendez, E. E., "*Physics of Resonant Tunneling in Semiconductors*", in Physics and Applications of Quantum Wells and Superlattices, edited by E. E. Mendez and K. von Klitzing, NATO ASI Series B **170**, 159 (1987).

Nicholas, R.J., "*High Field Magnetotransport*", in Physics and Applications of Quantum Wells and Superlattices, edited by E. E. Mendez and K. von Klitzing, NATO ASI Series B **170**, 217 (1987).

O'Reilly, E. P., and Witchlow, G. P., Phys. Rev. B **34**, 6030 (1986).

O'Reilly, E. P., and G. P. Witchiow, Solid State Commun. **62**, 653 (1987).

Patterson, W. R., and Shewchun, J., Rev. Sci. Instr. **35**, 1704, (1964).

People, R., Bean, J. C., Lang, D. C., Sergent, M., Stormer, H. L., Wecht, K. W., Lynch, R. T. and Baldwin, K., Appl. Phys. Lett. **45**, 1231 (1984).

People, R., IEEE J. Quantum Elect. **22**, 1696 (1986).

Price, P. J. and Stern, F., Surf. Sci. **132**, 577 (1983).

Ricco, B. and Azbel, M. Ya., Phys. Rev. B **29**, 1970 (1984).

Roy, D. K., Quantum Mechanical Tunnelling and its Applications, (World Scientific Publishing Co., Singapore, 1986).

Schuberth, G., Abstreiter, G., Gornik, E., Schaffler, F. and Luy, J. F., Phys. Rev. B **41**, (1990) or (1991).

Stern, F. and Das Sarma, S., Surf. Sci. (1984).

Stiles, P., "*Why Interfaces, Quantum Wells, and Superlattices? Some Comments*", in Interfaces, Quantum Wells, and Superlattices, edited by R. Taylor and E. W. Fenton, NATO ASI Series B **179**, 1 (1987).

Stiles, P., "*The effects of a Quantizing Magnetic Field*", in Two-Dimensional Systems, Heterostructures, and Superlattices, edited by G. Bauer, F. Kuchar, and H. Heinrich, (Springer-Verlag, Berlin, 1984).

Thomas, D. E. and Klein, J. M., Rev. Sci. Instr. **34**, 920 (1963).

van de Walle, C. G. and Martin, R. M., J. Vac. Sci. Technol. B 3, 1256 (1985).

Wessel, R. and Altarelli, M., Phys. Rev. B 40, 12457 (1989).

Zawadzki, W., "*Electric and Magnetic Quantization of Two-Dimensional Systems - Elementary Theory*", in Two-Dimensional Systems, Heterostructures, and Superlattices, edited by G. Bauer, F. Kuchar, and H. Heinrich, (Springer-Verlag, Berlin, 1984).