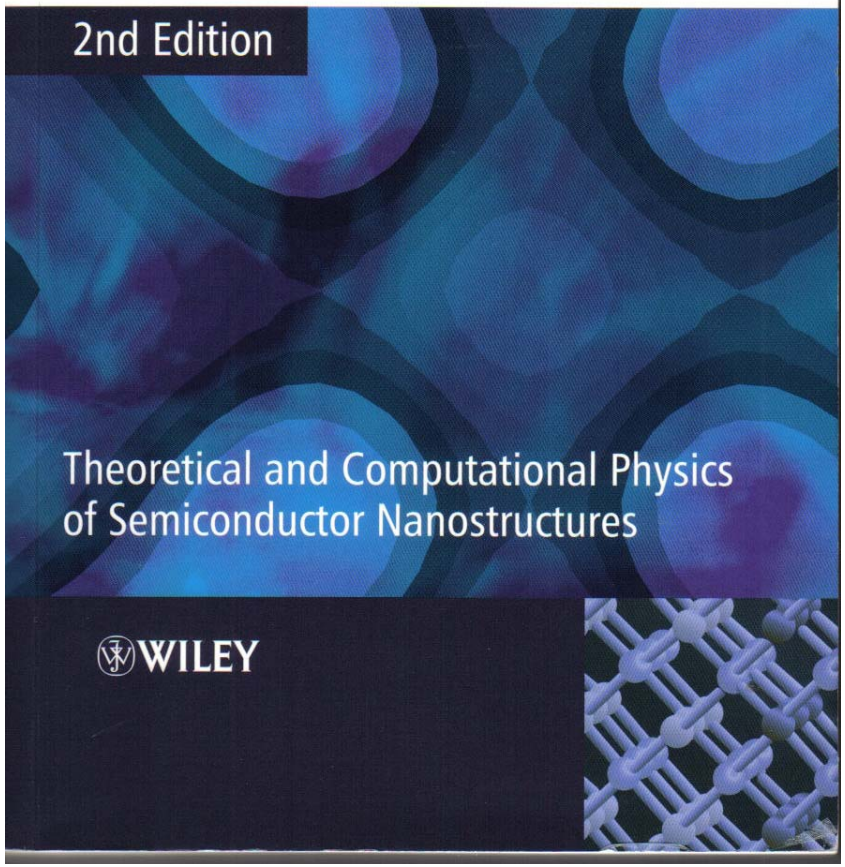
PAUL HARRISON

Quantum Wells, Wires and Dots

2nd Edition



Theoretical and Computational Physics of Semiconductor Nanostructures

WILEY

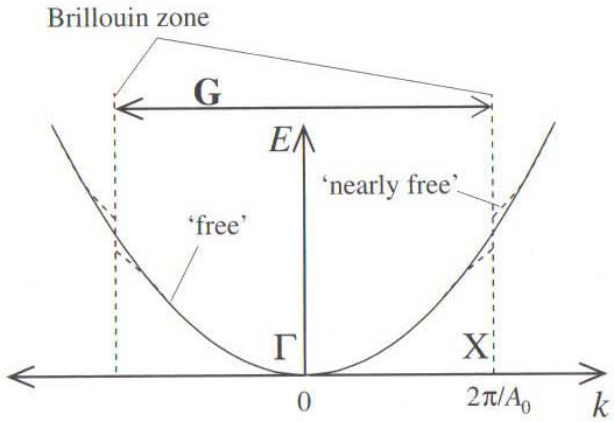


Figure 1.14 Comparison of the free and nearly free electron models

CHAPTER 2

SOLUTIONS TO SCHRÖDINGER'S EQUATION

2.1 THE INFINITE WELL

The infinitely deep one-dimensional potential well is the simplest confinement potential to treat in quantum mechanics. Virtually every introductory level text on quantum mechanics considers this system, but nonetheless it *is* worth visiting again as some of the standard assumptions often glossed over, do have important consequences for one-dimensional confinement potentials in general.

The time-independent Schrödinger equation summarises the wave mechanics analogy to Hamilton's formulation of classical mechanics [22], for *time-independent potentials*. In essence this states that the kinetic and potential energy components sum to the total energy; in wave mechanics, these quantities are the eigenvalues of linear operators, i.e.

$$\mathcal{T}\psi + \mathcal{V}\psi = E\psi \tag{2.1}$$

where the eigenfunction ψ describes the state of the system. Again in analogy with classical mechanics the kinetic energy operator for a particle of *constant* mass is given

by the following:

$$\mathcal{T} = \frac{\mathcal{P}^2}{2m} \tag{2.2}$$

where $\mathcal{P}$ is the usual quantum mechanical linear momentum operator:

$$\mathcal{P} = -i\hbar\nabla = -i\hbar\left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) \tag{2.3}$$

By using this form for the kinetic energy operator $\mathcal{T}$, the Schrödinger equation then becomes:

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi + V(x,y,z)\psi = E\psi \tag{2.4}$$

where the function $V(x,y,z)$ represents the potential energy of the system as a function of the spatial coordinates. Restricting this to the one-dimensional potential of interest here, then the Schrödinger equation for a particle of mass m in a potential well aligned along the z-axis (as in Fig. 2.1) would be:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}\psi(z) + V(z)\psi(z) = E\psi(z) \tag{2.5}$$

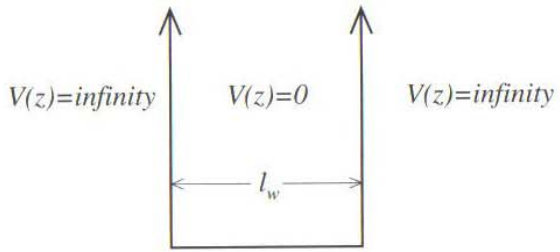


Figure 2.1 The one-dimensional infinite well confining potential

Outside of the well, $V(z) = \infty$, and hence the only possible solution is $\psi(z) = 0$, which in turn implies that all values of the energy E are allowed. Within the potential well, the Schrödinger equation simplifies to:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}\psi(z) = E\psi(z) \tag{2.6}$$

which implies that the solution for ψ is a linear combination of the functions $f(z)$ which when differentiated twice give $-f(z)$. Hence try the solution:

$$\psi(z) = A\sin kz + B\cos kz \tag{2.7}$$

Substituting into equation (2.6) then gives:

$$\frac{\hbar^2k^2}{2m}(A\sin kz + B\cos kz) = E(A\sin kz + B\cos kz) \tag{2.8}$$

$$\therefore \frac{\hbar^2k^2}{2m} = E \tag{2.9}$$

Consideration of the boundary conditions will yield the, as yet unknown, constant k. With this aim, consider again the kinetic energy term for this system, i.e.

$$\mathcal{T} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}\psi(z) \tag{2.10}$$

which can be rewritten as

$$\mathcal{T} = -\frac{\hbar^2}{2m}\frac{\partial}{\partial z}\left(\frac{\partial}{\partial z}\psi(z)\right) \tag{2.11}$$

The mathematical form of this implies that, *as a minimum*, $\psi(z)$ must be continuous. If it is not, then the first derivative will contain poles which must be avoided if the system is to have finite values for the kinetic energy. Given that $\psi(z)$ has already been deduced as zero outside of the well, then $\psi(z)$ within the well must be zero at both edges too.

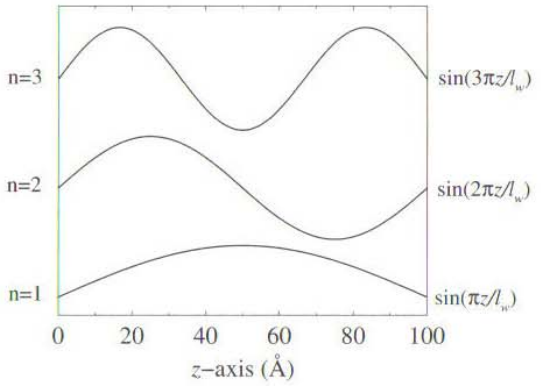


Figure 2.2 Solutions to the one-dimensional infinite well confining potential

If the origin is taken as the left hand edge of the well as in Fig. 2.2, then $\psi(z)$ as defined in equation (2.8) can contain no cosine terms, i.e. $B = 0$, and hence $\psi(z) = A\sin kz$. In addition, for $\psi(0) = \psi(l_w) = 0$:

$$k = \frac{\pi n}{l_w} \tag{2.12}$$

where n is an integer, representing a series of solutions. Substituting into equation (2.9), then the energy of the confined states is given by:

$$E_n = \frac{\hbar^2\pi^2n^2}{2ml_w^2} \tag{2.13}$$

The only remaining unknown is the constant factor A, which is deduced by considering the normalisation of the wave function; as $\psi^*(z)\psi(z)$ represents the probability of finding the particle at a point z, then as the particle must exist somewhere:

$$\int_0^{l_w} \psi^*(z)\psi(z)\ \mathrm{d}z = 1 \tag{2.14}$$

which gives $A = \sqrt{(2/l_w)}$, and therefore

$$\psi_n(z) = \sqrt{\frac{2}{l_w}} \sin\left(\frac{\pi n z}{l_w}\right) \tag{2.15}$$

Under the effective mass and envelope function approximations, the energy of an electron or hole in a hypothetical infinitely deep semiconductor quantum well can be calculated by using the effective mass m^* for the particle mass m of equation (2.13).

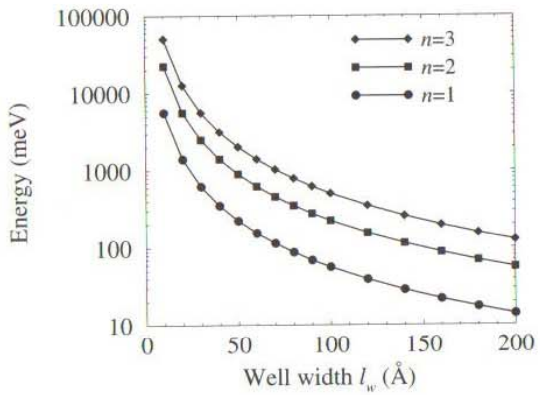


Figure 2.3 First three energy levels versus well width for an electron in a GaAs infinite potential well

Figure 2.3 displays the results of calculations of the lowest three energy states of an electron in a GaAs well of width l_w surrounded by hypothetical infinite barriers (for these and all material parameters see Appendix A). All three states show the same monotonic behaviour, with the energy decreasing as the well width increases.

The sine function solutions derived for this system are completely standard and found extensively in the literature. Although it should be noted that the arguments developed for setting the boundary conditions, i.e. $\psi(z)$ continuous, also implied that the first derivative should be continuous too, although use is never made of this second boundary condition. The limitations of solution imposed by this are avoided by saying

that not only is the potential infinite outside the well, but in addition the Schrödinger equation is not defined in these regions—a slight contradiction with the deduction of the first boundary condition. This point, i.e. that there is still ambiguity in the choice of boundary conditions for commonly accepted solutions, will be revisited later in this chapter.

2.2 IN-PLANE DISPERSION

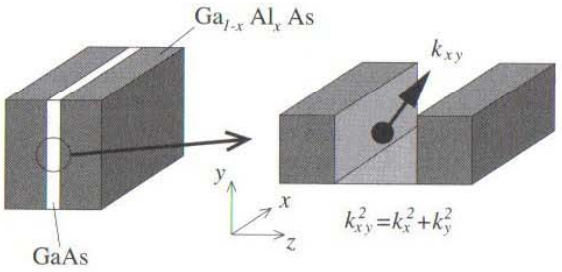


Figure 2.4 A GaAs/$Ga_{1-x}Al_xAs$ layered structure and the in-plane motion of a charge carrier

If the one-dimensional potential $V(z)$ is constructed from alternating thin layers of dissimilar semiconductors, then the particle, whether it be an electron or a hole, can move in the plane of the layers (see Fig 2.4).

In this case, all of the terms of the kinetic energy operator are required, and hence the Schrödinger equation would be as follows:

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi + V(z)\psi = E\psi \tag{2.16}$$

As the potential can be written as a sum of independent functions, i.e. $V = V(x) + V(y) + V(z)$, where it just happens in this case that $V(x) = V(y) = 0$, the eigenfunction of the system be written as a product:

$$\psi(x, y, z) = \psi_x(x)\psi_y(y)\psi_z(z) \tag{2.17}$$

Using this in the above Schrödinger equation then:

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi_x}{\partial x^2}\psi_y\psi_z + \frac{\partial^2\psi_y}{\partial y^2}\psi_x\psi_z + \frac{\partial^2\psi_z}{\partial z^2}\psi_x\psi_y\right) + V(z)\psi_x\psi_y\psi_z = E\psi_x\psi_y\psi_z \tag{2.18}$$

It is then possible to identify three distinct contributions to the total energy E, one from each of the perpendicular x-, y-, and z-axes, i.e. $E = E_x + E_y + E_z$. It is said

that the motions 'de-couple' giving an equation of motion for each of the axes:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi_x}{\partial x^2}\psi_y\psi_z = E_x\psi_x\psi_y\psi_z \tag{2.19}$$

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi_y}{\partial y^2}\psi_x\psi_z = E_y\psi_x\psi_y\psi_z \tag{2.20}$$

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi_z}{\partial z^2}\psi_x\psi_y + V(z)\psi_x\psi_y\psi_z = E_z\psi_x\psi_y\psi_z \tag{2.21}$$

Dividing throughout, then:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi_x}{\partial x^2} = E_x\psi_x \tag{2.22}$$

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi_y}{\partial y^2} = E_y\psi_y \tag{2.23}$$

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi_z}{\partial z^2} + V(z)\psi_z = E_z\psi_z \tag{2.24}$$

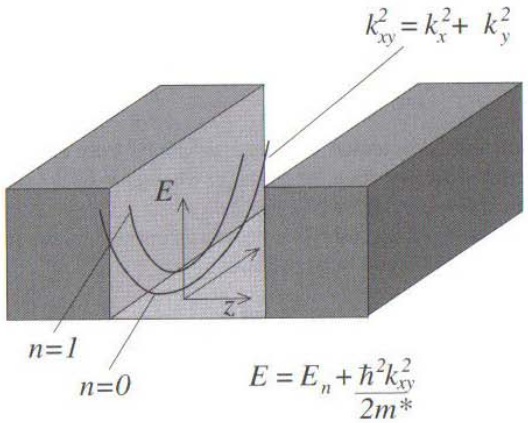


Figure 2.5 Schematic showing the in-plane ($k_{x,y}$) dispersion curves and the subband structure

The last component is identical to the one-dimensional Schrödinger equation for a confining potential $V(z)$ as discussed, for the particular case of an infinite well, in the last section. Consider the first and second components. Again, an eigenfunction f is sought which when differentiated twice returns $-f$; however, in this case it must be remembered that the solution will represent a moving particle. Thus the eigenfunction must reflect a current flow and have complex components, so try the standard travelling wave, $\exp(ik_x x)$. Then:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\exp(ik_x x) = E_x\exp(ik_x x) \tag{2.25}$$

$$\therefore \frac{\hbar^2 k_x^2}{2m} = E_x \tag{2.26}$$

which is clearly just the kinetic energy of a wave travelling along the x-axis. A similar equation follows for the y-axis, and hence the in-plane motion of a particle in a one-dimensional confining potential, but of infinite extent in the x–y plane can be summarised as:

$$\psi_{x,y}(x,y) = \frac{1}{A}\exp[i(k_x x + k_y y)] \quad \text{and} \quad E_{x,y} = \frac{\hbar^2|\mathbf{k}_{x,y}|^2}{2m} \tag{2.27}$$

Therefore, while solutions of Schrödinger's equation along the axis of the one-dimensional potential produce discrete states of energy $E_z = E_n$, in the plane of a semiconductor quantum well there is a continuous range of allowed energies, as illustrated in Fig. 2.5. In bulk materials, such domains are called 'energy bands', while in quantum well systems these energy domains associated with confined levels are referred to as 'subbands'. Therefore the effect of the one-dimensional confining potential is to remove a degree of freedom, thus restricting the momentum of the charge carrier from three-dimensions to two. It is for this reason that the states within quantum well systems are generally referred to as two-dimensional.

Later in this text, quantum wires and dots will be considered which further restrict the motion of carriers in two and three dimensions respectively, thus giving rise to the terms one- and zero-dimensional states.

Summarising then, within a semiconductor quantum well system the total energy of an electron or hole, of mass m^*, with in-plane momentum $k_{x,y}$, is equal to $E_z + E_{x,y}$, which is given by:

$$E = E_n + \frac{\hbar^2|\mathbf{k}_{x,y}|^2}{2m^*} \tag{2.28}$$

2.3 DENSITY OF STATES

Therefore the original confined states within the one-dimensional potential which could each hold two charge carriers of opposite spin, from the Pauli exclusion principle, broaden into subbands, thus allowing a continuous range of carrier momenta. In order to answer the question 'Given a particular number of electrons (or holes) within a subband, what is the distribution of their energy and momenta?', the first point that is required is a knowledge of the density of states, i.e. how many electrons can exist within a range of energies. In order to answer this point for the case of subbands in quantum wells, it is necessary first to understand this property in bulk crystals.

Following the idea behind Bloch's theorem (see reference [1] p. 133) that an eigenstate within a *bulk* semiconductor, which can be written as $\psi = (1/\Omega)\exp(i\mathbf{k}_\bullet\mathbf{r})$, must display periodicity within the lattice, then if the unit cell is of side L:

$$\psi(x,y,z) = \psi(x+L, y+L, z+L) \tag{2.29}$$

$$\therefore \psi(x,y,z) = \frac{1}{\Omega}\exp\{i[k_x(x+L) + k_y(y+L) + k_z(z+L)]\} \tag{2.30}$$

$$\therefore \psi(x,y,z) = \frac{1}{\Omega}\exp\left[i(k_x x + k_y y + k_z z)\right]\exp\left[i(k_x L + k_y L + k_z L)\right] \quad (2.31)$$

Thus for the periodicity condition to be fulfilled, the second exponential term must be identical to 1, which implies that:

$$k_x = \frac{2\pi}{L}n_x \qquad k_y = \frac{2\pi}{L}n_y \qquad k_z = \frac{2\pi}{L}n_z \quad (2.32)$$

where n_x, n_y and n_z are integers. Each set of values of these three integers defines a distinct state, and hence the volume of k-space occupied by one state is $(2\pi/L)^3$. These states fill up with successively larger values of n_x, n_y and n_z, i.e. the lowest energy state has values (000), then permutations of (100), (110), etc., which gradually fill a sphere. At low temperatures, the sphere has a definite boundary between states that are all occupied followed by states that are unoccupied; the momentum of these states is called the *Fermi wave vector* and the equivalent energy is the *Fermi energy*. At higher temperatures, carriers near the edge of the sphere are often scattered to higher energy states, thus 'blurring' the boundary between occupied and unoccupied states. For a more detailed description see, for example, Ashcroft and Mermin [1].

Many of the interesting phenomena associated with semiconductors derive from the properties of electrons near the Fermi energy, as it is these electrons that are able to scatter into nearby states thus changing both their energy and momenta. In order to be able to progress with descriptions of, transport for example (later in this book), it is necessary to be able to describe the density of available states.

The density of states is defined as the number of states per energy per unit volume of *real* space:

$$\rho(E) = \frac{\mathrm{d}N}{\mathrm{d}E} \quad (2.33)$$

In k-space, the total number of states N is equal to the volume of the sphere of radius $\mathbf{k}$, divided by the volume occupied by one state and divided again by the volume of real space, i.e.

$$N = 2\,\frac{4\pi k^3}{3}\,\frac{1}{(2\pi/L)^3}\,\frac{1}{L^3} \quad (2.34)$$

$$\therefore N = 2\,\frac{4\pi k^3}{3(2\pi)^3} \quad (2.35)$$

where the factor 2 has been introduced to allow for double occupancy of each state by the different carrier spins. Returning to the density of states, then:

$$\rho(E) = \frac{\mathrm{d}N}{\mathrm{d}E} = \frac{\mathrm{d}N}{\mathrm{d}k}\frac{\mathrm{d}k}{\mathrm{d}E} \quad (2.36)$$

Now equation (2.35) gives

$$\frac{\mathrm{d}N}{\mathrm{d}k} = 2\,\frac{4\pi k^2}{(2\pi)^3} \quad (2.37)$$

In addition, the parabolic bands of effective mass theory give:

$$E = \frac{\hbar^2 k^2}{2m^*} \qquad \therefore k = \left(\frac{2m^* E}{\hbar^2}\right)^{\frac{1}{2}} \quad (2.38)$$

$$\therefore \frac{\mathrm{d}k}{\mathrm{d}E} = \left(\frac{2m^*}{\hbar^2}\right)^{\frac{1}{2}}\frac{E^{-\frac{1}{2}}}{2} \quad (2.39)$$

Which finally gives the density of states in bulk as:

$$\rho(E) = \frac{1}{2\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{\frac{3}{2}} E^{\frac{1}{2}} \quad (2.40)$$

Thus the density of states within a band, and around a minimum where the energy can be represented as a parabolic function of momentum, is continual and proportional to the square root of the energy.

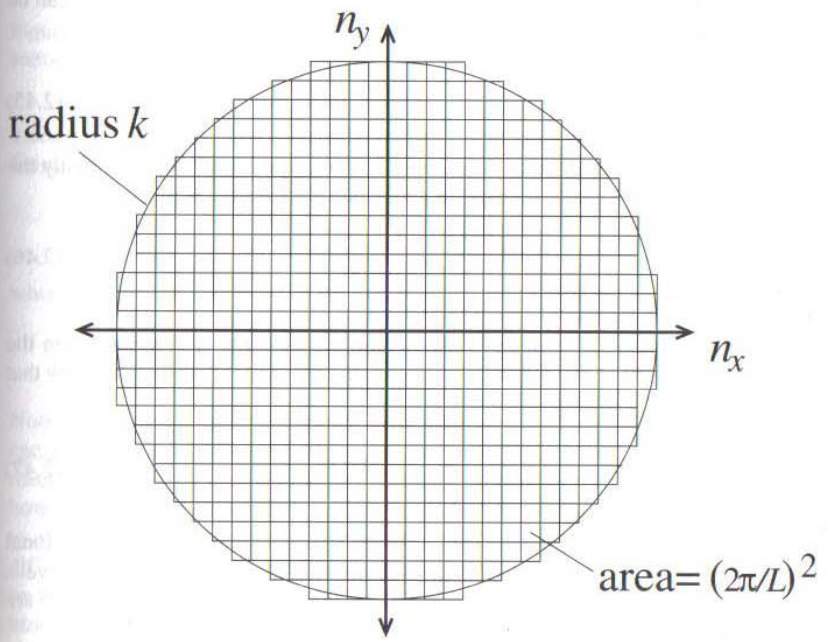


Figure 2.6 Illustration of filling the two-dimensional momenta states in a quantum well

The density of states in quantum well systems follows analogously; however this time, as there are only two degrees of freedom, successive states represented by values of n_x and n_y fill a circle in k-space, as illustrated in Fig. 2.6. Such a situation has

become known as a two-dimensional electron (or hole) gas (2DEG). Hence the total number of states per unit cross-sectional area is given by the spin degeneracy factor, multiplied by the area of the circle of radius k, divided by the area occupied by each state, i.e.

$$N^{2D} = 2\,\pi k^2\,\frac{1}{(2\pi/L)^2}\,\frac{1}{L^2} \tag{2.41}$$

$$\therefore N^{2D} = 2\,\frac{\pi k^2}{(2\pi)^2} \tag{2.42}$$

$$\therefore \frac{\mathrm{d}N^{2D}}{\mathrm{d}k} = \frac{k}{\pi} \tag{2.43}$$

In analogy to the bulk three-dimensional (3D) case, define:

$$\rho^{2D}(E) = \frac{\mathrm{d}N^{2D}}{\mathrm{d}E} = \frac{\mathrm{d}N^{2D}}{\mathrm{d}k}\,\frac{\mathrm{d}k}{\mathrm{d}E} \tag{2.44}$$

As the in-plane dispersion curves are still described by parabolas, then reuse can be made of equation (2.39), as follows:

$$\rho^{2D}(E) = \frac{k}{\pi}\left(\frac{2m^*}{\hbar^2}\right)^{\frac{1}{2}}\frac{E^{-\frac{1}{2}}}{2} \tag{2.45}$$

By substituting for k in terms of the energy E, using equation (2.38) then finally the density of states for a single subband in a quantum well system is given by:

$$\rho^{2D}(E) = \frac{m^*}{\pi\hbar^2} \tag{2.46}$$

in agreement with Bastard [18] p. 12.

If there are many (n) confined states within the quantum well system then the density of states ρ^{2D} at any particular energy is the sum over all subbands *below* that point, which can be written succinctly as:

$$\rho^{2D}(E) = \sum_{i=1}^{n}\frac{m^*}{\pi\hbar^2}\,\Theta(E - E_i) \tag{2.47}$$

where Θ is the unit step function. Fig. 2.7 gives an example of the two-dimensional density of states for a particular quantum well showing the first three confined levels. Note that the steps are of equal height and occur at the subband minima—which are not equally spaced.

2.4 SUBBAND POPULATIONS

The total number of carriers within a subband is given by the integral of the product of the probability of occupation of a state and the density of states. Given that the

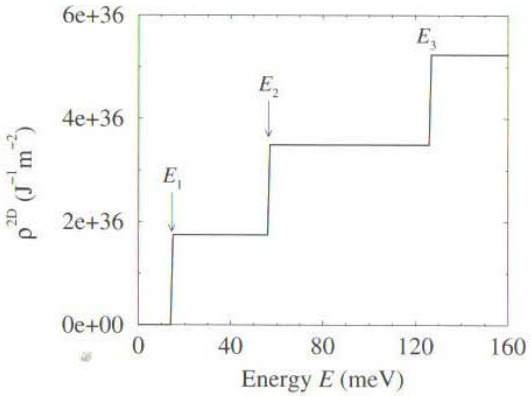


Figure 2.7 The density of states as a function of energy for a 200 Å GaAs quantum well surrounded by infinite barriers

carriers are fermions, then clearly the probability of occupation of a state is given by Fermi–Dirac statistics; hence:

$$N = \int_{\text{subband}} f^{FD}(E)\,\rho(E)\;\mathrm{d}E \tag{2.48}$$

where the integral is over *all* of the energies of a given subband and, of course:

$$f^{FD}(E) = \frac{1}{\exp\left[(E - E_F)/kT\right] + 1} \tag{2.49}$$

Note that E_F is not the Fermi energy in the traditional sense [1]; it is a 'quasi' Fermi energy which describes the carrier population *within* a subband. For systems left to reach equilibrium, the temperature T can be assumed to be the lattice temperature; however this is not always the case. In many quantum well devices which are subject to excitation by electrical or optical means, the 'electron temperature' can be quite different from the lattice temperature, and furthermore the subband population could be non-equilibrium and not able to be described by Fermi–Dirac statistics. For now, however it is sufficient to discuss equilibrium electron populations and assume that the above equations are an adequate description.

Given a particular number of carriers within a quantum well, which can usually be deduced directly from the surrounding doping density, it is often desirable to be able to describe that distribution in terms of the quasi-Fermi energy E_F. With this aim substitute the two-dimensional density of states appropriate to a single subband from equation (2.46) into equation (2.48), then the carrier density, i.e. the number

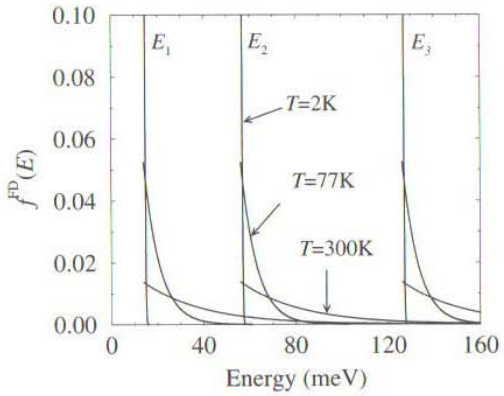


Figure 2.8 Effect of temperature on the distribution functions of the subband populations (all equal to $1\times10^{10}\text{cm}^{-2}$) of the infinite quantum well of Fig. 2.7

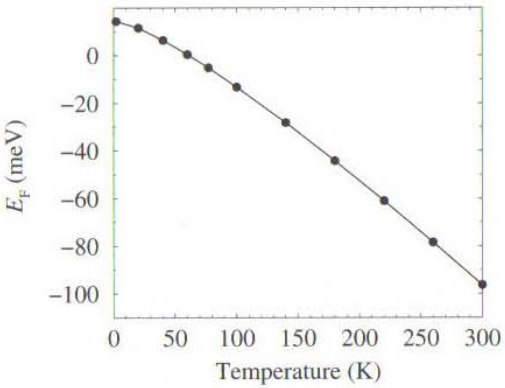


Figure 2.9 Effect of temperature on the quasi-Fermi energy describing the electron distribution of the ground state E_1

per unit area, is given by:

$$N = \int_{E_{\text{min}}}^{E_{\text{max}}} \frac{1}{\exp\left[(E-E_{\text{F}})/kT\right]+1} \frac{m^*}{\pi\hbar^2}\,\mathrm{d}E \tag{2.50}$$

By putting:

$$E' = \frac{E-E_{\text{F}}}{kT}, \quad \text{and then} \quad \mathrm{d}E' = \frac{\mathrm{d}E}{kT} \tag{2.51}$$

equation (2.50) becomes:

$$N = \int_{(E_{\text{min}}-E_{\text{F}})/kT}^{(E_{\text{max}}-E_{\text{F}})/kT} \frac{kT}{\exp E'+1} \frac{m^*}{\pi\hbar^2}\,\mathrm{d}E' \tag{2.52}$$

which is a standard form (see for example, Gradshteyn and Ryzhik [23] equation 2.313.2, p. 112)

$$\int \frac{1}{1+\mathrm{e}^x}\,\mathrm{d}x = x - \ln\left(1+\mathrm{e}^x\right) \tag{2.53}$$

Hence:

$$N = \frac{m^*kT}{\pi\hbar^2}\left[E' - \ln\left(1+\mathrm{e}^{E'}\right)\right]_{(E_{\text{min}}-E_{\text{F}})/kT}^{(E_{\text{max}}-E_{\text{F}})/kT} \tag{2.54}$$

Evaluation then gives:

$$N = \frac{m^*kT}{\pi\hbar^2}\left\{\left[\frac{E_{\text{max}}-E_{\text{F}}}{kT} - \ln\left(1+\mathrm{e}^{(E_{\text{max}}-E_{\text{F}})/kT}\right)\right] - \left[\frac{E_{\text{min}}-E_{\text{F}}}{kT} - \ln\left(1+\mathrm{e}^{(E_{\text{min}}-E_{\text{F}})/kT}\right)\right]\right\} \tag{2.55}$$

The minimum of integration E_{min} is taken as the subband minima and the maximum E_{max} can either be taken as the top of the well, or even $E_{\text{F}}+10kT$, say, with the latter being much more stable at lower temperatures. Given a total carrier density N, the quasi-Fermi energy E_{F} is the only unknown in equation (2.55) and can be found with standard techniques. For an example of such a method, see Section 2.5

Fig. 2.8 gives an example of the distribution functions $f^{\text{FD}}(E)$ for the first three confined levels within a 200 Å GaAs infinite quantum well. As the density of carriers, in this case electrons, have been taken as being equal and of value $1\times10^{10}\text{cm}^{-2}$, then the distribution functions are all identical, but offset along the energy axis by the confinement energies. As mentioned above, at low temperatures the carriers tend to occupy the lowest available states, and hence the transition from states that are all occupied to those that are unoccupied is rapid—as illustrated by the 2 K data for all three subbands. As the temperature increases the distributions broaden and a range of energies exist in which the states are partially filled, as can be seen by the 77 and 300 K data. Physically this broadening occurs due mainly to the increase in electron–phonon scattering as the phonon population increases with temperature (more of this in Chapter 9). Fig. 2.9 displays the Fermi energy E_{F} as a function of temperature T for the ground state of energy $E_1 = 14.031$ meV. At low temperatures, E_{F} is just above the confinement energy, since the electron density is fairly low (1×10^{10} cm^{-2}). As the temperature increases, E_{F} falls quite markedly and below the subband minima. If this seems counterintuitive, it must be remembered that E_{F} is a *quasi*-Fermi energy

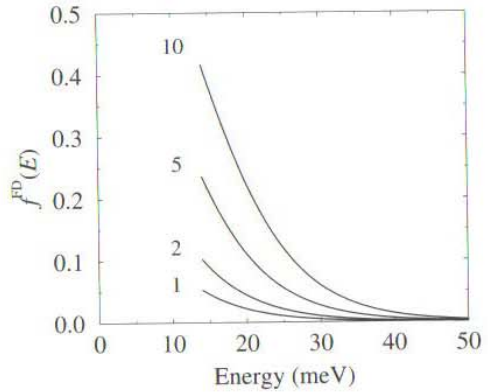


Figure 2.10 Effect of electron density, N=1, 2, 5, 10 ($\times 10^{10}$) cm^{-2}, on the distribution function of the lowest subband of the infinite quantum well of Fig. 2.7

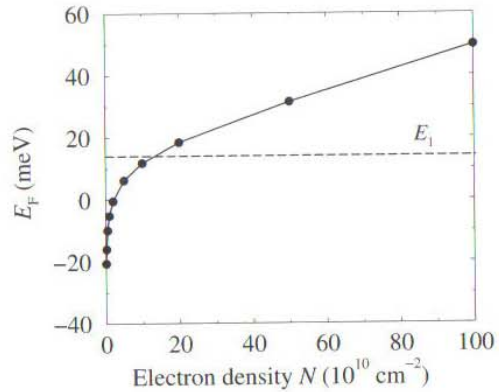


Figure 2.11 Effect of electron density on the quasi Fermi energy describing the distribution of the ground state E_1

whose only physical meaning is to describe the population within a subband—it is not the true Fermi energy of the complete system.

Fig. 2.10 displays the distribution functions for a range of carrier densities, for this same ground state and at a lattice temperature of 77 K. Although not obvious

from the mathematics, $f^{FD}(E)$ at any particular energy E appears to scale with N. The corresponding Fermi energy is illustrated in Fig. 2.11. Clearly, the Fermi energy starts below the subband minima at this mid-range temperature, as discussed above, and as expected increasing numbers of carriers in the subband increases the Fermi energy, i.e. the energy of the state whose probability of occupation is $1/2$.

2.5 FINITE WELL WITH CONSTANT MASS

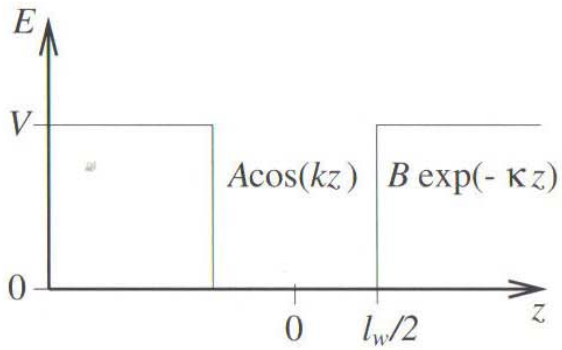


Figure 2.12 Solutions to the finite well potential

While the infinitely deep confining potential has served well as a platform for developing the physics of two-dimensional systems, more relevant to alternating layers of dissimilar semiconductors is the finite quantum well model, which under both the effective mass and envelope function approximations looks like Fig. 2.12. In particular, a layer of GaAs 'sandwiched' between two thick layers of $Ga_{1-x}Al_xAs$ would form a type-I finite quantum well, where the conduction band has the appearance of Fig. 2.12, with the potential energy V representing the discontinuity in the conduction band edge between the materials.

Again taking the simplest starting case of a constant electron mass m^* throughout the dissimilar layers, and neglecting movement within the plane of the layers, then the standard Schrödinger equation can be written for each of the semiconductor layers as follows:

$$-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2}\psi(z) + V\psi(z) = E\psi(z), \quad z \le -\frac{l_w}{2} \tag{2.56}$$

$$-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2}\psi(z) = E\psi(z), \quad -\frac{l_w}{2} \le z \le \frac{l_w}{2} \tag{2.57}$$

$$-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2}\psi(z) + V\psi(z) = E\psi(z), \quad +\frac{l_w}{2} \le z \tag{2.58}$$

Considering solutions to the Schrödinger equation for the central well region, then as in the infinite well case, the general solution will be a sum of sine and cosine terms.

As the potential is symmetric, then the eigenstates will also have a definite symmetry, i.e. they will be either symmetric or antisymmetric. With the origin placed at the centre of the well, the symmetric (even parity) eigenstates will then be in cosine terms, while the antisymmetric (odd parity) states will be as sine waves.

For states confined to the well, the energy E must be less than the barrier height V, thus rearranging the Schrödinger equation for the right hand barrier:

$$\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2}\psi(z) = (V - E)\psi(z) \tag{2.59}$$

Therefore, a function f is sought which when differentiated twice gives $+f$. The exponential function fits this description, therefore consider a sum of growing $\exp(+\kappa z)$ and decaying $\exp(-\kappa z)$ exponentials. In the right-hand barrier, z is positive, and hence as z increases the growing exponential will increase too and without limit. The probability interpretation of the wave function requires that:

$$\int_{\text{all space}} \psi^*(z)\psi(z)\ \mathrm{d}z = 1 \tag{2.60}$$

which further demands that:

$$\psi(z) \to 0 \quad \text{and} \quad \frac{\partial}{\partial z}\psi(z) \to 0, \qquad \text{as} \qquad z \to \pm\infty \tag{2.61}$$

These boundary conditions for states confined in wells will be used again and again and will be referred to as the *standard boundary conditions*. Using this result, the growing exponential components must be rejected and the solutions are for the even parity states, which would follow as:

$$\psi(z) = B\exp(\kappa z), \qquad z \le -\frac{l_w}{2} \tag{2.62}$$

$$\psi(z) = A\cos(kz), \qquad -\frac{l_w}{2} \le z \le \frac{l_w}{2} \tag{2.63}$$

$$\psi(z) = B\exp(-\kappa z), \qquad \frac{l_w}{2} \le z \tag{2.64}$$

Note for later that these wave functions are real, and that the eigenfunctions of this confined system carry no current and hence are referred to as stationary states. Using these trial forms of the wave function in their corresponding Schrödinger equations, gives the, as yet unknown constants:

$$k = \frac{\sqrt{2m^*E}}{\hbar}, \qquad \text{and} \qquad \kappa = \frac{\sqrt{2m^*(V-E)}}{\hbar} \tag{2.65}$$

In order to proceed it is necessary to impose boundary conditions. Recalling the constant mass kinetic energy operator employed in equations (2.56)-(2.58), then in order to avoid infinite kinetic energies:

$$\text{both} \quad \psi(z) \quad \text{and} \quad \frac{\partial}{\partial z}\psi(z) \quad \text{must be continuous}$$

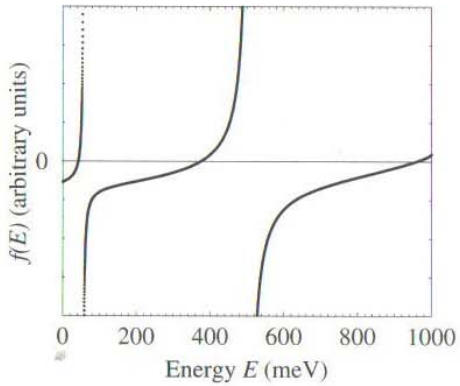


Figure 2.13 Illustration of $f(E)$ as a function of E for the even-parity solutions; where l_w=100 Å, m^*=0.067 m_0 and V=1000 meV

Consider the interface at $z = +l_w/2$; by equating ψ in the well and the barrier:

$$A\cos\left(\frac{kl_w}{2}\right) = B\exp\left(-\frac{\kappa l_w}{2}\right) \tag{2.66}$$

and equating the derivatives gives:

$$-kA\sin\left(\frac{kl_w}{2}\right) = -\kappa B\exp\left(-\frac{\kappa l_w}{2}\right) \tag{2.67}$$

Dividing equation (2.66) by equation (2.67) then gives:

$$-\frac{1}{k}\cot\left(\frac{kl_w}{2}\right) = -\frac{1}{\kappa} \tag{2.68}$$

$$\therefore k\tan\left(\frac{kl_w}{2}\right) - \kappa = 0 \tag{2.69}$$

Odd parity states would require the choice of wave function in the well region as a sine wave, and hence equation (2.63) would become $\psi = A\sin(kz)$; following through the same analysis as above gives the equation to be solved for the odd parity eigenenergies as:

$$k\cot\left(\frac{kl_w}{2}\right) + \kappa = 0 \tag{2.70}$$

Remembering that both k and κ are functions of the energy E, then equations (2.69) and (2.70) are also functions of E only. There are many ways of solving such single-variable equations, and for this particular case the literature often talks of 'graphical

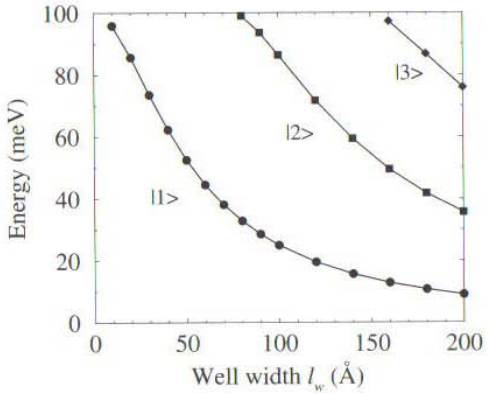


Figure 2.14 Energy levels in a GaAs single quantum well with constant effective mass m^*=0.067m_0 and V=100 meV

methods' [3,7,18]. While it is interesting to view the functional form of the equations, such methods are time consuming and inefficient and wouldn't be employed in the repetitive solution of many quantum wells. Computationally it is much more effective just to treat equations (2.69) and (2.70) with standard techniques, such as Newton–Raphson iteration. In this technique, if $E^{(n)}$ is a first guess to the solution of $f(E) = 0$, then a better estimate is given by:

$$E^{(n+1)} = E^{(n)} - \frac{f(E^{(n)})}{f'(E^{(n)})} \tag{2.71}$$

The new estimate $E^{(n+1)}$ is then used to generate a second approximation to the solution $E^{(n+2)}$, and so on, until the successive estimates converge to a required accuracy.

In order to provide all of the required information to implement the solution, all that remains is to deduce $f'(E)$. For the even parity states:

$$f(E) = k \tan\left(\frac{kl_w}{2}\right) - \kappa \tag{2.72}$$

and so therefore:

$$\frac{\mathrm{d}f}{\mathrm{d}E} = \frac{\mathrm{d}k}{\mathrm{d}E} \tan\left(\frac{kl_w}{2}\right) + k \sec^2\left(\frac{kl_w}{2}\right) \times \frac{l_w}{2}\frac{\mathrm{d}k}{\mathrm{d}E} - \frac{\mathrm{d}\kappa}{\mathrm{d}E} \tag{2.73}$$

where

$$\frac{\mathrm{d}k}{\mathrm{d}E} = \frac{\sqrt{2m^*}}{2\sqrt{E}\hbar}; \qquad \frac{\mathrm{d}\kappa}{\mathrm{d}E} = -\frac{\sqrt{2m^*}}{2\sqrt{V-E}\hbar} \tag{2.74}$$

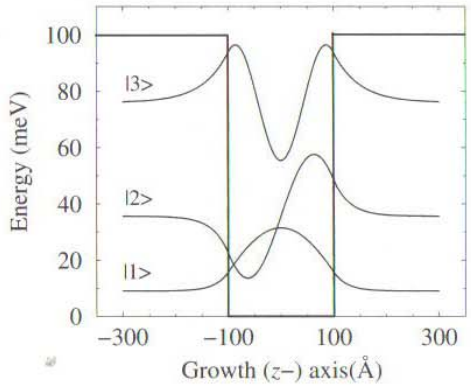


Figure 2.15 Eigenfunctions $\psi(z)$ for the first three energy levels of the 200 Å GaAs well of Fig. 2.14

For the odd-parity states:

$$f(E) = k \cot\left(\frac{kl_w}{2}\right) + \kappa \tag{2.75}$$

$$\therefore \frac{\mathrm{d}f}{\mathrm{d}E} = \frac{\mathrm{d}k}{\mathrm{d}E} \cot\left(\frac{kl_w}{2}\right) - k \csc^2\left(\frac{kl_w}{2}\right) \times \frac{l_w}{2}\frac{\mathrm{d}k}{\mathrm{d}E} + \frac{\mathrm{d}\kappa}{\mathrm{d}E} \tag{2.76}$$

Figure 2.13 illustrates the even parity $f(E)$. In practice, the finite number of discontinuities can lead to solutions being missed if the first guess to the solution $E^{(0)}$ is not close to the true solution. In order to circumvent this computational problem, $f(E)$ is calculated at discrete points along the E axis, separated by an energy thought to be smaller than the minimum separation between adjacent states (generally 1 meV), when $f(E)$ changes sign; then the Newton–Raphson is then implemented to obtain the solution accurately.

Figures 2.14 and 2.15 summarise the application of the method to a GaAs single quantum well, surrounded by barrier of height 100 meV, with the same effective mass. Clearly, as the well width increases, then the energy levels all decrease, with the presence of excited states being also apparent at the larger well widths. The eigenstates are labelled according to their principle quantum number (energy order). The even and odd parities of the states within the well can be seen in Fig. 2.15.

2.6 EFFECTIVE MASS MISMATCH AT HETEROJUNCTIONS

Quantum wells are only fabricated by forming heterojunctions between *different* semiconductors. From an electronic viewpoint, the semiconductors are different because they have different band structures. The difference in perhaps the most fundamental property of a semiconductor, i.e. the band gap, (and its alignment) is accounted for by specifying a band offset, which has been labelled V. Of course, there are many other properties which are also different, such as the dielectric constant, the lattice constant and, what is considered the next most important quantity, the effective mass. It is generally accepted that the calculation of *static* energy levels within quantum wells should account for the variation in the effective mass across the heterojunction.

This problem has been continuously addressed in the literature [17, 24, 25] since the earliest work of Conley *et al.* [26] and BenDaniel and Duke [27], who derived the boundary conditions on solutions of the *envelope functions* as:

$$\text{both} \quad \psi(z) \quad \text{and} \quad \frac{1}{m^*}\frac{\partial}{\partial z}\psi(z) \quad \text{continuous} \tag{2.77}$$

by considering electron transport across a heterojunction. These boundary conditions have become known as the BenDaniel–Duke boundary conditions.

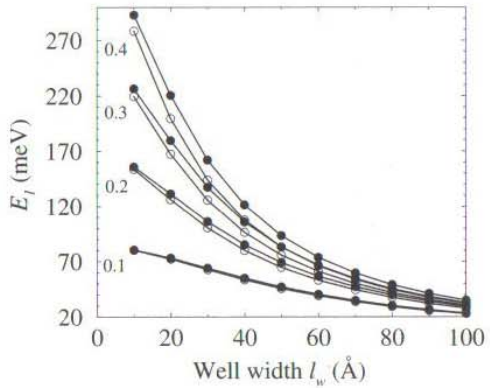


Figure 2.16 Electron ground state energy E_1 as a function of the width l_w of a GaAs well surrounded by $Ga_{1-x}Al_xAs$ barriers, calculated for both the constant mass-model (closed circles) and different barrier masses (open circles) and for a range of barrier alloy concentrations x (= 0.1, 0.2, 0.3, 0.4)

Therefore applying this extension to the finite well of the previous section would require the Schrödinger equation to be specified in each region as follows:

$$-\frac{\hbar^2}{2m_b^*}\frac{\partial^2}{\partial z^2}\psi(z) + V\psi(z) = E\psi(z), \quad z \le -\frac{l_w}{2} \tag{2.78}$$

$$-\frac{\hbar^2}{2m_w^*}\frac{\partial^2}{\partial z^2}\psi(z) = E\psi(z), \quad -\frac{l_w}{2} \le z \le \frac{l_w}{2} \tag{2.79}$$

$$-\frac{\hbar^2}{2m_b^*}\frac{\partial^2}{\partial z^2}\psi(z) + V\psi(z) = E\psi(z), \quad +\frac{l_w}{2} \le z \tag{2.80}$$

with the additional restraint of the matching conditions of equation (2.77).

The solutions follow as previously for the constant-mass case, in equations (2.62), (2.63) and (2.64), but now k and κ contain different effective masses:

$$k = \frac{\sqrt{2m_w^* E}}{\hbar} \qquad \kappa = \frac{\sqrt{2m_b^*(V-E)}}{\hbar} \tag{2.81}$$

The method of solution is almost identical: equating the envelope functions at the interface $z = +l_w/2$

$$A\cos\left(\frac{kl_w}{2}\right) = B\exp\left(-\frac{\kappa l_w}{2}\right) \tag{2.82}$$

and equating $1/m^*$ times the derivative gives

$$-\frac{kA}{m_w^*}\sin\left(\frac{kl_w}{2}\right) = -\frac{\kappa B}{m_b^*}\exp\left(-\frac{\kappa l_w}{2}\right) \tag{2.83}$$

Dividing equation (2.83) by equation (2.82) then:

$$f(E) = \frac{k}{m_w^*}\tan\left(\frac{kl_w}{2}\right) - \frac{\kappa}{m_b^*} = 0 \tag{2.84}$$

and similarly for the odd parity solutions, i.e.

$$f(E) = \frac{k}{m_w^*}\cot\left(\frac{kl_w}{2}\right) + \frac{\kappa}{m_b^*} = 0 \tag{2.85}$$

and obviously, equating the effective masses gives the original constant-mass equations. Again for numerical solution, the derivatives are required, i.e. for even parity:

$$\frac{\mathrm{d}f}{\mathrm{d}E} = \frac{1}{m_w^*}\frac{\mathrm{d}k}{\mathrm{d}E}\tan\left(\frac{kl_w}{2}\right) + \frac{k}{m_w^*}\sec^2\left(\frac{kl_w}{2}\right) \times \frac{l_w}{2}\frac{\mathrm{d}k}{\mathrm{d}E} - \frac{1}{m_b^*}\frac{\mathrm{d}\kappa}{\mathrm{d}E} \tag{2.86}$$

and for odd parity

$$\frac{\mathrm{d}f}{\mathrm{d}E} = \frac{1}{m_w^*}\frac{\mathrm{d}k}{\mathrm{d}E}\cot\left(\frac{kl_w}{2}\right) - \frac{k}{m_w^*}\csc^2\left(\frac{kl_w}{2}\right) \times \frac{l_w}{2}\frac{\mathrm{d}k}{\mathrm{d}E} + \frac{1}{m_b^*}\frac{\mathrm{d}\kappa}{\mathrm{d}E} \tag{2.87}$$

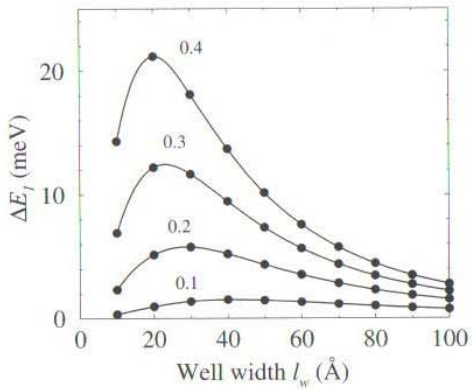


Figure 2.17 Energy difference $\Delta E_1 = E_1(m^*\text{=constant}) - E_1(m^*(z))$ for the structures shown in Fig. 2.16

where

$$\frac{\mathrm{d}k}{\mathrm{d}E} = \frac{\sqrt{2m_w^*}}{2\sqrt{E}\hbar} \qquad \frac{\mathrm{d}\kappa}{\mathrm{d}E} = -\frac{\sqrt{2m_b^*}}{2\sqrt{V-E}\hbar} \tag{2.88}$$

Figure 2.16 compares the electron ground-state energy calculated with a constant GaAs mass, with the energy calculated with a material-dependent mass, for a single GaAs quantum well surrounded by $Ga_{1-x}Al_xAs$ barriers. All of the calculated ground state energies E_1 decrease with increasing well width and increase with increasing Al fraction in the barriers (barrier height $V \propto x$). The effective mass in $Ga_{1-x}Al_xAs$ is greater than in GaAs, hence the variable mass calculations give energies less than the constant-mass model for *all* systems considered here (see later).

Fig. 2.17 displays the calculated ground state energy difference between the two models, $\Delta E_1 = E_1(m^*\text{=constant}) - E_1(m^*(z))$. Clearly, and as would be expected, the larger the difference in the effective masses between the materials, the larger the difference in ground-state energies.

2.7 THE INFINITE BARRIER HEIGHT AND MASS LIMITS

It is interesting to take theoretical models to certain limits as a means of verifying, or otherwise, their behaviour with what might be expected intuitively. Fig. 2.18 illustrates that, in the limit of large barrier heights V, the finite well model recovers the result of the infinite well model, which is what would be hoped for, thus increasing confidence in the derivation.

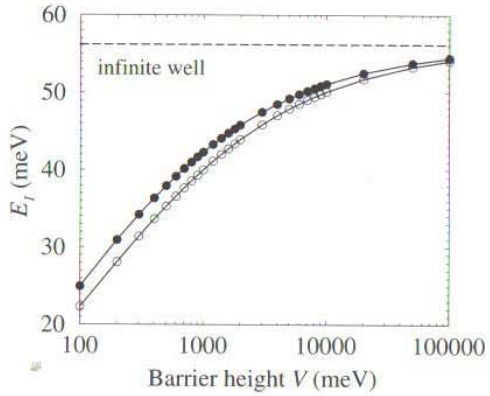


Figure 2.18 Electron ground-state energy E_1 as a function of barrier height V, for a 100 Å GaAs finite well with constant mass (closed circles) and different barrier mass (fixed at mass in $Ga_{0.6}Al_{0.4}As$ (open circles))

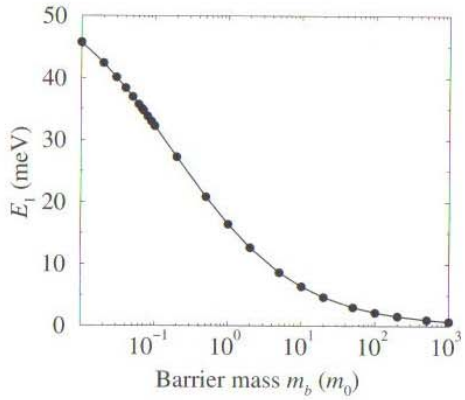


Figure 2.19 Electron ground-state energy E_1 as a function of the mass in the barrier, for a 100 Å GaAs well with a barrier height fixed at that for $Ga_{0.6}Al_{0.4}As$

As has been found in the literature, Fig. 2.19 illustrates the results of allowing the barrier mass to increase without limit, while keeping all other parameters constant. The tendency for the ground-state energy to tend towards zero and the unusual

looking wave functions of Fig. 2.20 have been well documented [18] and are a direct consequence of the second boundary condition, ψ'/m^*.

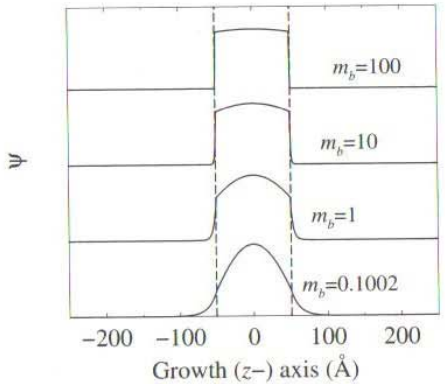


Figure 2.20 Electron ground-state wave functions for several barrier masses, as given, for a 100 Å GaAs well with a barrier height fixed at that for $Ga_{0.6}Al_{0.4}As$

It is worthwhile considering this limit still further. Fig. 2.21 reproduces the results of Fig. 2.19 but for a variety of well widths. Clearly, the ground-state energy tends to zero for all well widths:

$$\lim_{m_b \to \infty} E_1 = 0, \qquad \forall\ l_w \tag{2.89}$$

which at first sight appears to violate Heisenberg's Uncertainty Principle, in that the ground state energy *can* be forced to zero for infinitesimal well widths, and thus the error in the measurement of position and momentum can be made arbitrarily small. This was an argument advanced by Hagston *et al.* [28]; however direct evaluation of the variance in the position and momentum of these states will show that the uncertainty relationships are not violated. At the moment, the techniques for such a calculation have not been covered; hence such discussions will be returned to later in Section 3.15.

2.8 HERMITICITY AND THE KINETIC ENERGY OPERATOR

The changes in the boundary conditions at a heterojunction are only necessary under theoretical models which parameterise physical quantities in terms of a variable mass. This is only encountered when applying the effective mass approximation to semiconductor heterostructures. Theories of semiconductor heterostructures do exist

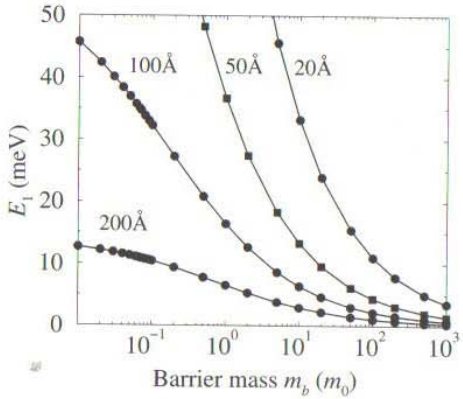


Figure 2.21 Electron ground state energy E_1 as a function of the mass in the barrier, for a variety of GaAs well widths, the barrier height is fixed at that for $Ga_{0.6}Al_{0.4}As$

which do not make this imposition (parameterisation)—within them all of the electron states, within both the conduction and valence band can be derived from the *constant* free electron mass. Therefore the kinetic energy operator remains

$$\mathcal{T} = -\frac{\hbar^2}{2m_0}\nabla^2 \tag{2.90}$$

and hence the boundary conditions are set as continuity in both wave function and derivative—note *wave function*, and not *envelope function*. Such methods are the subject of later chapters, for example, Chapter 11.

However within the auspices of this very successful model of semiconductor band structure, allowance *has* to be made for the additional complication of variable mass. As discussed in Section 2.6, this can be achieved with the original Schrödinger equation applied to each region, taken together with boundary conditions for matching solutions at the heterojunctions between the regions of different effective mass.

However, fundamentally both of these conditions can be incorporated within one Hamiltonian which is valid throughout all space at once. Its construction derives from the cornerstones of quantum mechanics—all physical observables can be represented by linear Hermitian operators [12]. With this aim consider the classical (non-relativistic) kinetic energy:

$$T = \frac{p^2}{2m} \tag{2.91}$$

Quantum mechanically, the linear kinetic energy operator for a variable mass $m^*(z)$ say, is formed by replacing the classical quantities with the corresponding linear

operators, giving:

$$\mathcal{T} = \frac{\mathcal{P}_z^2}{m^*(z)} \tag{2.92}$$

Inspection of equation (2.92) reveals that there is still ambiguity as to the *order* by which the operators will act upon the eigenvector. To resolve this, guidance is sought from a property of Hermitian operators, and that is they are equal to their own Hermitian conjugate (see Schiff [29] p. 151), i.e.

$$\mathcal{T} = \mathcal{T}^\dagger \tag{2.93}$$

In addition the Hermitian conjugate of a product is given by $(\mathcal{ABC})^\dagger = \mathcal{C}^\dagger\mathcal{B}^\dagger\mathcal{A}^\dagger$. Applying this to the effective mass kinetic energy operator then:

$$\mathcal{T} = \mathcal{P}_z \frac{1}{m^*(z)} \mathcal{P}_z \tag{2.94}$$

$$\therefore \mathcal{T} = -\frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m^*(z)}\frac{\partial}{\partial z} \tag{2.95}$$

which would give a Schrödinger equation of the form:

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m^*(z)}\frac{\partial}{\partial z}\psi(z) + V(z)\psi(z) = E\psi(z) \tag{2.96}$$

which in order to avoid differentiating discontinuous functions and producing infinities, clearly demands the boundary conditions of equation 2.77, i.e.,

$$\text{both} \quad \psi(z) \quad \text{and} \quad \frac{1}{m^*}\frac{\partial}{\partial z}\psi(z) \quad \text{are continuous}$$

2.9 ALTERNATIVE KINETIC ENERGY OPERATORS

The Hermiticity requirements above have been used to derive a kinetic energy operator and hence a single Hamiltonian which is valid throughout all space, and reproduces the accepted boundary conditions on the envelope functions of heterostructures. This form is, however, not unique. In particular, any combination of linear momentum operators and effective mass of the form:

$$\mathcal{T} = m^{*\alpha}(z)\mathcal{P}_z m^{*\beta}(z)\mathcal{P}_z m^{*\alpha}(z), \quad \text{where} \quad 2\alpha + \beta = -1 \tag{2.97}$$

will also satisfy these requirements [30,31].

The accepted choice is $\alpha = 0$, $\beta = -1$, but clearly there are infinitely many solutions. Morrow [32] and Hagston *et al.* [28] have independently forwarded $\alpha = -1/2$, β=0, giving

$$\mathcal{T} = \frac{1}{\sqrt{m^*(z)}}\mathcal{P}_z\mathcal{P}_z\frac{1}{\sqrt{m^*(z)}} \tag{2.98}$$

which can be written succinctly as:

$$\mathcal{T} = \mathcal{Q}^\dagger\mathcal{Q}, \quad \text{where} \quad \mathcal{Q} = \mathcal{P}_z\frac{1}{\sqrt{m^*(z)}} \tag{2.99}$$

Transforming the wave function $\psi(z)$ into $\sqrt{m^*(z)}\phi(z)$ and substituting into the Schrödinger equation gives:

$$-\frac{\hbar^2}{2}\frac{1}{\sqrt{m^*(z)}}\frac{\partial^2}{\partial z^2}\frac{1}{\sqrt{m^*(z)}}\sqrt{m^*(z)}\phi(z) + V(z)\sqrt{m^*(z)}\phi(z) = E\sqrt{m^*(z)}\phi(z) \tag{2.100}$$

$$\therefore -\frac{\hbar^2}{2m^*(z)}\frac{\partial^2}{\partial z^2}\phi(z) + V(z)\phi(z) = E\phi(z) \tag{2.101}$$

which implies that $\phi(z)$ is an eigenfunction of this transformed Hamiltonian, but with the same eigenvalue E_n. This new entity, $\phi(z)$, satisfies the boundary conditions:

$$\text{both} \quad \phi(z) \quad \text{and} \quad \frac{\partial}{\partial z}\phi(z) \quad \text{are continuous}$$

Therefore, discontinuous changes in the effective mass $m^*(z)$ at abrupt heterojunctions would give a discontinuity in the *wave function* ψ, a view which has been suggested elsewhere [33].

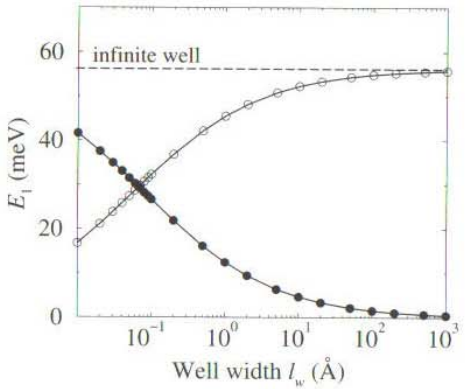


Figure 2.22 Effect of the two different kinetic energy operators on the ground-state energy of a 100 Å GaAs well surrounded by $Ga_{0.8}Al_{0.2}As$ barriers: $\mathcal{T} = \frac{1}{m}\mathcal{P}^2$ (open circles); $\mathcal{T} = \mathcal{P}\frac{1}{m}\mathcal{P}$ (closed circles)

Fig. 2.22 compares the consequences of such an alternative kinetic energy operator on the ground state energy E_1 of a finite quantum well, with that of the accepted

operator. In the limit of large barrier mass, the energy given by the new operator tends towards that of the infinitely deep quantum well. This appears intuitively satisfying, given that the wave function penetration into the barrier, $\exp(-\kappa z)$, has a similar dependence on the barrier mass m_b as on the height V, i.e.

$$\kappa = \frac{\sqrt{2m_b^*(V-E)}}{\hbar} \tag{2.102}$$

Thus mathematically speaking, increasing m_b would appear to produce the same effect as increasing V.

Galbraith and Duggan [25] have compared the results of similar calculations with experimental photoluminescence excitation measurements and have concluded that for their single-quantum-well samples the accepted form of $\mathcal{T} = \mathcal{P}\frac{1}{m}\mathcal{P}$ fits the data better. There is no doubt that this is not the last word on the form of the kinetic energy operator, and the associated boundary conditions, on the effective mass envelope functions which follow. However, for now it is best to conform to the commonly agreed standard and employ $\mathcal{T} = \mathcal{P}\frac{1}{m}\mathcal{P}$ throughout the remainder of this book.

2.10 EXTENSION TO MULTIPLE-WELL SYSTEMS

For a whole variety of reasons semiconductor physicists and electronic device engineers need to design and fabricate heterostructures more complex than the single quantum well [7]. The above techniques of solving Schrödinger's equation in each semiconductor layer separately and deducing the unknown coefficients by implementing boundary conditions can be applied to these multilayered systems. In particular, consider the asymmetric double quantum well of Fig. 2.23.

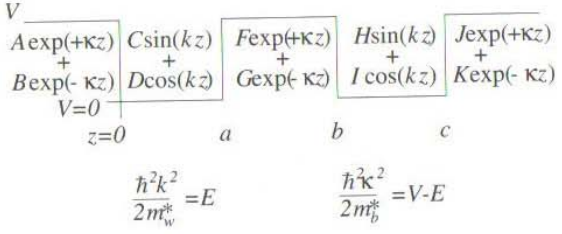


Figure 2.23 Solutions to Schrödinger's equation in a double quantum well

Choosing both wells to be of the same depth, and all of the barriers to be of the same height V, allows the simplification that k and κ are constant throughout the structure; the general solution to the Schrödinger equation in each region are as given in the figure. Proceeding as before, then matching envelope functions and the derivative divided by the mass at the interfaces, then considering $z = 0$ first:

$$A + B = D \tag{2.103}$$

$$\frac{1}{m_b^*}(\kappa A - \kappa B) = \frac{1}{m_w^*}kC \tag{2.104}$$

with $z = a$ then:

$$C\sin ka + D\cos ka = F\exp(+\kappa a) + G\exp(-\kappa a) \tag{2.105}$$

$$\frac{1}{m_w^*}(kC\cos ka - kD\sin ka) = \frac{1}{m_b^*}[\kappa F\exp(+\kappa a) - \kappa G\exp(-\kappa a)] \tag{2.106}$$

$z = b$:

$$F\exp(+\kappa b) + G\exp(-\kappa b) = H\sin kb + I\cos kb \tag{2.107}$$

$$\frac{1}{m_b^*}[\kappa F\exp(+\kappa b) - \kappa G\exp(-\kappa b)] = \frac{1}{m_w^*}(kH\cos kb - kI\sin kb) \tag{2.108}$$

$z = c$:

$$H\sin kc + I\cos kc = J\exp(+\kappa c) + K\exp(-\kappa c) \tag{2.109}$$

$$\frac{1}{m_w^*}(kH\cos kc - kI\sin kc) = \frac{1}{m_b^*}[\kappa J\exp(+\kappa c) - \kappa K\exp(-\kappa c)] \tag{2.110}$$

which can be rewritten more neatly in matrix form as:

$$\begin{pmatrix} 1 & 1 \\ \frac{1}{m_b^*}\kappa & -\frac{1}{m_b^*}\kappa \end{pmatrix}\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{1}{m_w^*}k & 0 \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix} \tag{2.111}$$

$$\begin{pmatrix} \sin ka & \cos ka \\ \frac{1}{m_w^*}k\cos ka & -\frac{1}{m_w^*}k\sin ka \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} \exp(+\kappa a) & \exp(-\kappa a) \\ \frac{1}{m_b^*}\kappa\exp(+\kappa a) & -\frac{1}{m_b^*}\kappa\exp(-\kappa a) \end{pmatrix}\begin{pmatrix} F \\ G \end{pmatrix} \tag{2.112}$$

$$\begin{pmatrix} \exp(+\kappa b) & \exp(-\kappa b) \\ \frac{1}{m_b^*}\kappa\exp(+\kappa b) & -\frac{1}{m_b^*}\kappa\exp(-\kappa b) \end{pmatrix}\begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} \sin kb & \cos kb \\ \frac{1}{m_w^*}k\cos kb & -\frac{1}{m_w^*}k\sin kb \end{pmatrix}\begin{pmatrix} H \\ I \end{pmatrix} \tag{2.113}$$

$$\begin{pmatrix} \sin kc & \cos kc \\ \frac{1}{m_w^*}k\cos kc & -\frac{1}{m_w^*}k\sin kc \end{pmatrix}\begin{pmatrix} H \\ I \end{pmatrix} = \begin{pmatrix} \exp(+\kappa c) & \exp(-\kappa c) \\ \frac{1}{m_b^*}\kappa\exp(+\kappa c) & -\frac{1}{m_b^*}\kappa\exp(-\kappa c) \end{pmatrix}\begin{pmatrix} J \\ K \end{pmatrix} \tag{2.114}$$

Labelling the 2×2 matrix for the left-hand side of the nth interface as $\mathbf{M}_{2n-1}$ and the corresponding matrix for the right hand side of the interface as $\mathbf{M}_{2n}$, n=1, 2, 3, etc., then the above matrix equations would become:

$$\mathbf{M}_1\begin{pmatrix} A \\ B \end{pmatrix} = \mathbf{M}_2\begin{pmatrix} C \\ D \end{pmatrix} \tag{2.115}$$

$$\mathbf{M}_3 \begin{pmatrix} C \\ D \end{pmatrix} = \mathbf{M}_4 \begin{pmatrix} F \\ G \end{pmatrix} \tag{2.116}$$

$$\mathbf{M}_5 \begin{pmatrix} F \\ G \end{pmatrix} = \mathbf{M}_6 \begin{pmatrix} H \\ I \end{pmatrix} \tag{2.117}$$

$$\mathbf{M}_7 \begin{pmatrix} H \\ I \end{pmatrix} = \mathbf{M}_8 \begin{pmatrix} J \\ K \end{pmatrix} \tag{2.118}$$

Now equation (2.115) gives:

$$\begin{pmatrix} A \\ B \end{pmatrix} = \mathbf{M}_1^{-1}\mathbf{M}_2 \begin{pmatrix} C \\ D \end{pmatrix} \tag{2.119}$$

and equation (2.116) gives:

$$\begin{pmatrix} C \\ D \end{pmatrix} = \mathbf{M}_3^{-1}\mathbf{M}_4 \begin{pmatrix} F \\ G \end{pmatrix} \tag{2.120}$$

$$\therefore \begin{pmatrix} A \\ B \end{pmatrix} = \mathbf{M}_1^{-1}\mathbf{M}_2\mathbf{M}_3^{-1}\mathbf{M}_4 \begin{pmatrix} F \\ G \end{pmatrix} \tag{2.121}$$

and eventually:

$$\begin{pmatrix} A \\ B \end{pmatrix} = \mathbf{M}_1^{-1}\mathbf{M}_2\mathbf{M}_3^{-1}\mathbf{M}_4\mathbf{M}_5^{-1}\mathbf{M}_6\mathbf{M}_7^{-1}\mathbf{M}_8 \begin{pmatrix} J \\ K \end{pmatrix} \tag{2.122}$$

The product of the eight 2×2 matrices is still a 2×2 matrix; thus writing:

$$\begin{pmatrix} A \\ B \end{pmatrix} = \mathcal{M} \begin{pmatrix} J \\ K \end{pmatrix} \tag{2.123}$$

then

$$A = \mathcal{M}_{11}J + \mathcal{M}_{12}K \tag{2.124}$$

$$B = \mathcal{M}_{21}J + \mathcal{M}_{22}K \tag{2.125}$$

Again the probability interpretation of the wave function implies that the wave function must tend towards zero into the outer barriers, i.e. the coefficients of the growing exponentials must be zero. *In this case*, with the origin at the 1st interface (see Fig. 2.23), then this implies that $B = 0$ and $J = 0$, and hence the second of the above equations would imply that $\mathcal{M}_{22}$=0. As all of the elements of $\mathcal{M}$ are functions of both k and κ, which are both in turn functions of the energy E, then an energy is sought which satisfies:

$$\mathcal{M}_{22}(E) = 0 \tag{2.126}$$

which can be found by standard numerical procedures, as discussed in Section 2.5. This approach and variations upon it are often referred to as the *Transfer matrix technique*. Once the energy is known, the coefficients A to K follow simply and the envelope wave function can be deduced.

2.11 THE ASYMMETRIC SINGLE QUANTUM WELL

The above system has illustrated a method of solution for a system in which the confinement energy E was always less than the barrier height V. While this is often true, a class of quantum well structures exist in which there is more than one barrier height (or well depth); Fig. 2.24 illustrates one such system.

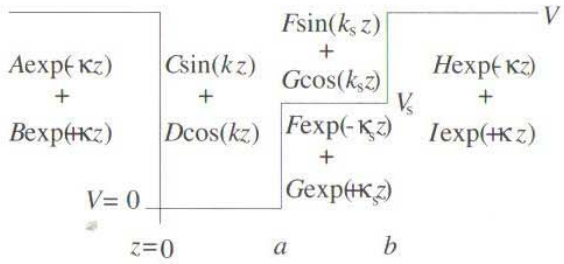


Figure 2.24 Solutions to Schrödinger's equation in a stepped asymmetric quantum well

A confined state could exist which has an energy below or above that of the step height V_s, hence the functional form of the solution to Schrödinger's equation in that region is dependent upon the energy E. This is illustrated schematically in Fig. 2.24. For states above the step, the wave function has the form of a normal well state, but with a k value different from that of the well region. For states of energy below the step potential V_s, the wave function resembles a barrier state. This causes difficulties computationally, as two transfer matrices have to be determined. A better approach is to write the wave function in *every* region (r) as a linear combination of travelling waves, i.e.

$$\psi_r = A' \exp(ik_r z) + B' \exp(-ik_r z) \tag{2.127}$$

where the coefficients A' and B' are now allowed to be complex and the wave vector k is as before:

$$k = \sqrt{\frac{2m^*}{\hbar^2}(E - V_r)} \tag{2.128}$$

Forming the transfer matrix from this standard form for the wave function and applying the boundary condition of decaying exponentials at both ends of the well structure, allows the method of solution as before. For now it is worthwhile confirming that ψ_r can take both the usual well- and barrier-state forms.

Consider regions where $E > V$:

$$\psi_r = A'(\cos k_r z + i \sin k_r z) + B'(\cos k_r z - i \sin k_r z) \tag{2.129}$$

$$\therefore \psi_r = (A' + B')\cos k_r z + i(A' - B')\sin k_r z \tag{2.130}$$

or by collecting real and imaginary components, then:

$$\psi_r = [\Re(A' + B')\cos k_r z - \Im(A' - B')\sin k_r z] \tag{2.131}$$

$$+ i\,[\Im(A' + B')\cos k_r z + \Re(A' - B')\sin k_r z] \tag{2.132}$$

Now eigenstates confined within potential wells are *stationary states*, which means that they have no time dependence and carry no current, i.e. their wave functions are real. Thus it would be expected that the solution would naturally yield such states, and in fact for the imaginary component of ψ_r to be zero, then:

$$\Im(A' + B') = 0 \qquad \text{and} \qquad \Re(A' - B') = 0 \tag{2.133}$$

which are simultaneously satisfied if:

$$B' = A'^* \tag{2.134}$$

This implies that in regions where the energy of the state of interest is greater than the potential, the solution will then give the coefficients as a complex conjugate pair in order to ensure a real wave function.

In regions where $E < V$, then:

$$k = i\sqrt{\frac{2m^*}{\hbar^2}(V_r - E)} = i\kappa \tag{2.135}$$

where κ is the usual decay constant as defined in equation (2.81). Then the general form for a barrier state is generated, i.e.

$$\psi_r = A' \exp(-\kappa z) + B' \exp(+\kappa z) \tag{2.136}$$

As mentioned above, application of the boundary conditions of $\psi \to 0$ as $z \to \pm\infty$, would force the choice that B' was zero and A' was real. In other regions where $E < V$, e.g. the step in Fig. 2.24 or the central barrier of Fig. 2.23, then both A' and B' should naturally arise as real.

2.12 ADDITION OF AN ELECTRIC FIELD

As illustrated in Fig. 2.25, the effect of an electric field along the growth (z-) axis is to add a linear potential, which for an electron of charge $-e$ in the conduction band will be written as $-eFz$. For small electric fields, the effect of this electric field on the confined energy levels within a quantum well can be approximated by first-order perturbation theory.

If the perturbing potential is written as V' then the change in the ground-state energy level is given, for example, by Schiff [29], p. 246:

$$\Delta E^{(1)} = \langle \psi_1 | V' | \psi_1 \rangle \tag{2.137}$$

with the electric field along the growth axis as the perturbation; then this translates to

$$\Delta E^{(1)} = \int_{-\infty}^{+\infty} \psi_1^*(z)(-eFz)\psi_1(z)\ \mathrm{d}z \tag{2.138}$$

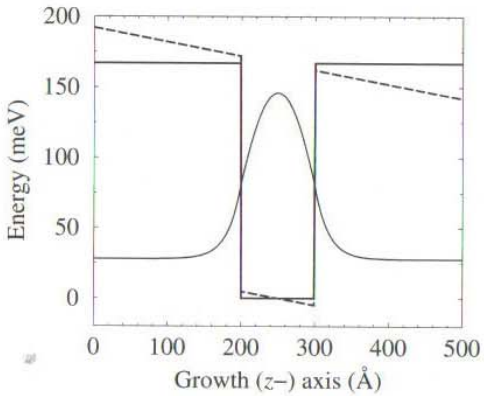


Figure 2.25 The ground state wave function of the zero-field potential (solid line) and the corresponding potential at a field of 10 kVcm^{-1} (dashed line)

The ground state of the *symmetric* quantum well of Fig. 2.25 is an even function, and hence the integrand of equation (2.138) is odd. Therefore, evaluation of the integral gives zero, i.e. to first-order the addition of a *small* electric field has no effect on the *ground-state* energy level. When applying the same logic to the first excited state, although ψ_2 is odd the integrand will still be an odd function, which again results in no change. There is currently much interest in the literature in breaking the symmetry of quantum wells in order to induce a first-order response in the ground state energy to an electric field [34–36].

Larger electric fields would require a more accurate evaluation of the perturbation as can be given by second-order perturbation theory—see Schiff [29], p. 247:

$$\Delta E^{(2)} = \sum_{m=2}^{\infty} \frac{|\langle \psi_m | V' | \psi_1 \rangle|^2}{E_m - E_1} \tag{2.139}$$

which in this case translates to:

$$\Delta E^{(2)} = \sum_{m=2}^{\infty} \frac{|\int_{-\infty}^{+\infty} \psi_m^*(z)(-eFz)\psi_1(z)\ dz|^2}{E_m - E_1} \tag{2.140}$$

where the sum is over *all* excited states, including those with an energy $E_m > V$, i.e. the so called 'continuum' states. As the electric field F does not have a z-dependence, then clearly the change in the energy to second-order $\Delta E^{(2)}$ is proportional to F^2. In addition, a charged particle prefers to move to areas of lower potential, i.e. the electron within the quantum well of Fig. 2.25 moves to the right hand side of the

well, thus lowering its total energy, so indeed $\Delta E^{(2)} \propto -F^2$. This suppression of the confined energy level by an electric field is known as the 'quantum confined Stark effect' [37] and is commonly observed experimentally in heterostructures [35].

While such a perturbative treatise may be adequate for single quantum wells, its accuracy for complicated, perhaps multiple well systems and at high fields, is questionable. In order to improve upon this, a full (non-perturbative) solution to the heterostructure and field potentials is required. Within the envelope function and effective mass approximations, the Schrödinger equation within a particular material (i.e. region of constant mass) and with an applied electric field of strength F, is:

$$-\frac{\hbar^2}{2m^*}\frac{\partial^2\psi}{\partial z^2} + [V(z) - eFz]\,\psi = E\psi \tag{2.141}$$

This new equation does not have the standard solution of linear combinations of trigonometric and exponential functions, and therefore a different approach is required. First, rearrange equation (2.141) to give

$$\frac{\partial^2\psi}{\partial z^2} - \frac{2m^*}{\hbar^2}[V(z) - eFz - E]\,\psi = 0 \tag{2.142}$$

Then by making the substitutions:

$$\alpha = \frac{2m^*}{\hbar^2}[V(z) - E] \tag{2.143}$$

and

$$\beta = \frac{2m^*}{\hbar^2}(-eF) \tag{2.144}$$

the Schrödinger equation becomes:

$$\frac{\partial^2\psi}{\partial z^2} - (\alpha + \beta z)\,\psi = 0 \tag{2.145}$$

Consider the further substitution:

$$z' = \frac{\alpha + \beta z}{\gamma} \tag{2.146}$$

where γ is an, as yet, unknown constant. Then, since z' is a first-order linear function of z:

$$\frac{\partial^2\psi}{\partial z^2} = \frac{\partial^2\psi}{\partial z'^2} \times \left(\frac{\partial z'}{\partial z}\right)^2 \tag{2.147}$$

$$\text{i.e.} \qquad \frac{\partial^2\psi}{\partial z^2} = \frac{\beta^2}{\gamma^2}\frac{\partial^2\psi}{\partial z'^2} \tag{2.148}$$

Substituting into equation (2.145) then gives:

$$\frac{\beta^2}{\gamma^2}\frac{\partial^2\psi}{\partial z'^2} - \gamma z'\psi = 0 \tag{2.149}$$

If γ^3 is set equal to β^2, then the full transformation is:

$$z' = \left(\frac{2m^*}{\hbar^2}\right)^{\frac{1}{3}}\left[\frac{V(z) - E}{(eF)^{\frac{2}{3}}} - (eF)^{\frac{1}{3}}z\right] \tag{2.150}$$

and the Schrödinger equation can be written as:

$$\frac{\partial^2\psi}{\partial z'^2} - z'\psi = 0 \tag{2.151}$$

The reason for such a procedure is that equation (2.151) has a standard solution which is a linear combination of Airy functions (see Abramowitz and Stegun [38], p. 446):

$$\psi(z') = A\,\mathrm{Ai}(z') + B\,\mathrm{Bi}(z') \tag{2.152}$$

The full solution proceeds by matching the wave functions at the heterojunctions according to the BenDaniel–Duke boundary conditions as before.

This is a standard procedure and many examples of its use can be found in the literature [39]. It is included along with the transfer matrix method of the previous sections for completeness. However, it can be appreciated that within these methods the study of a new quantum well system requires quite a large investment in developing new computational solutions. While a general computer code can be written to solve all semiconductor heterostructures, another one has to be written to solve all of these systems again with an electric field. The possibility of adding piezoelectric fields [40], which tilt the bands like an electric field, but with the additional possibility of being in the other direction, would require another computer code, and so on. While we continue in this vain for the remainder of this chapter, in the next chapter a numerical solution is introduced, which although occasionally has a few limitations, is still very powerful because of its *versatility*.

2.13 THE INFINITE SUPERLATTICE

Finite superlattices (or multiple quantum wells) can be treated with the transfer matrix technique, as described above for a double quantum well (a two period finite superlattice). Such approaches yield a discrete step of states due to the confinement along the growth axis of the semiconductor multilayer. However, when many identical quantum wells are stacked within the same semiconductor layer the electrons and holes see a periodic potential, which *can* appear to be infinite in exactly the same way as bulk crystals. When this occurs, the electron and hole wave functions are no longer localised but are of infinite extent and equally likely to be in any of the quantum wells. They are said to occupy 'Bloch states'.

As described in Section 2.11, and following Jaros [5], p. 220, the envelope function within the well region can be written as:

$$\psi_w = A\exp(ik_w z) + B\exp(-ik_w z) \tag{2.153}$$

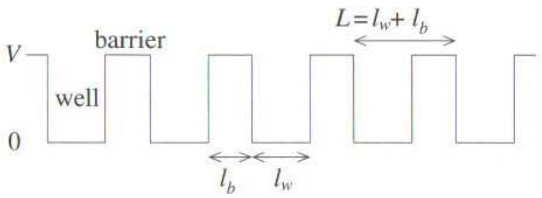


Figure 2.26 Band-edge potential profile of a superlattice of well and barrier widths, l_w and l_b respectively

and in the barrier region as:

$$\psi_b = C\exp(ik_b z) + D\exp(-ik_b z) \tag{2.154}$$

where:

$$k_w = \sqrt{\frac{2m_w^*}{\hbar^2}E} \qquad k_b = \sqrt{\frac{2m_b^*}{\hbar^2}(E-V)} \tag{2.155}$$

and it is assumed (for now) that $E > V$, so that k_b is real. Clearly, and as mentioned above, if the superlattice is infinite then a particle is equally likely to be found in any well, which means that its wave function must be periodic with the lattice, i.e.

$$\psi(z) = \psi(z+L) \tag{2.156}$$

For a travelling-wave state of the form, $\exp(ikz)$, then:

$$\psi(z+L) = \exp[ik(z+L)] = \exp(ikz)\exp(ikL) \tag{2.157}$$

$$\text{i.e.} \qquad \psi(z+L) = \psi(z)\exp(ikL) \tag{2.158}$$

By using this *periodic* form of the first BenDaniel–Duke boundary conditions (ψ continuous) at $z = L$, then:

$$\psi_w(L) = \psi_w(0)\exp(ikL) = \psi_b(L) \tag{2.159}$$

Using the wave functions of equations (2.153) and (2.154), then:

$$(A+B)\exp(ikL) = C\exp(ik_b L) + D\exp(-ik_b L) \tag{2.160}$$

By employing the same periodicity condition in the second of the BenDaniel–Duke boundary conditions (ψ/m), also at $z = L$, then:

$$\frac{ik_w}{m_w^*}(A-B)\exp(ikL) = \frac{ik_b}{m_b^*}\left[C\exp(ik_b L) - D\exp(-ik_b L)\right] \tag{2.161}$$

Then using the BenDaniel–Duke boundary conditions at $z = l_w$:

$$A\exp(ik_w l_w) + B\exp(-ik_w l_w) = C\exp(ik_b l_w) + D\exp(-ik_b l_w) \tag{2.162}$$

and

$$\frac{ik_w}{m_w^*}\left[A\exp(ik_w l_w) - B\exp(-ik_w l_w)\right] = \frac{ik_b}{m_b^*}\left[C\exp(ik_b l_w) - D\exp(-ik_b l_w)\right] \tag{2.163}$$

Equations (2.160), (2.161), (2.162) and (2.163), can be rewritten as:

$$\begin{pmatrix} \exp(ikL) & \exp(ikL) & -\exp(ik_b L) & -\exp(-ik_b L) \\ \frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_b}{m_b^*}\exp(ik_b L) & \frac{k_b}{m_b^*}\exp(-ik_b L) \\ \exp(ik_w l_w) & \exp(-ik_w l_w) & -\exp(ik_b l_w) & -\exp(-ik_b l_w) \\ \frac{k_w}{m_w^*}\exp(ik_w l_w) & -\frac{k_w}{m_w^*}\exp(-ik_w l_w) & -\frac{k_b}{m_b^*}\exp(ik_b l_w) & \frac{k_b}{m_b^*}\exp(-ik_b l_w) \end{pmatrix} \times \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = 0 \tag{2.164}$$

which implies that, for a solution other than the trivial $A = B = C = D = 0$, the determinant of the 4×4 must be zero, i.e.

$$\begin{vmatrix} \exp(ikL) & \exp(ikL) & -\exp(ik_b L) & -\exp(-ik_b L) \\ \frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_b}{m_b^*}\exp(ik_b L) & \frac{k_b}{m_b^*}\exp(-ik_b L) \\ \exp(ik_w l_w) & \exp(-ik_w l_w) & -\exp(ik_b l_w) & -\exp(-ik_b l_w) \\ \frac{k_w}{m_w^*}\exp(ik_w l_w) & -\frac{k_w}{m_w^*}\exp(-ik_w l_w) & -\frac{k_b}{m_b^*}\exp(ik_b l_w) & \frac{k_b}{m_b^*}\exp(-ik_b l_w) \end{vmatrix} = 0 \tag{2.165}$$

Proceeding one step would give:

$$\exp(ikL)\begin{vmatrix} -\frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_b}{m_b^*}\exp(ik_b L) & \frac{k_b}{m_b^*}\exp(-ik_b L) \\ \exp(-ik_w l_w) & -\exp(ik_b l_w) & -\exp(-ik_b l_w) \\ -\frac{k_w}{m_w^*}\exp(-ik_w l_w) & -\frac{k_b}{m_b^*}\exp(ik_b l_w) & \frac{k_b}{m_b^*}\exp(-ik_b l_w) \end{vmatrix}$$

$$-\exp(ikL)\begin{vmatrix} \frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_b}{m_b^*}\exp(ik_b L) & \frac{k_b}{m_b^*}\exp(-ik_b L) \\ \exp(ik_w l_w) & -\exp(ik_b l_w) & -\exp(-ik_b l_w) \\ \frac{k_w}{m_w^*}\exp(ik_w l_w) & -\frac{k_b}{m_b^*}\exp(ik_b l_w) & \frac{k_b}{m_b^*}\exp(-ik_b l_w) \end{vmatrix}$$

$$-\exp(ik_b L)\begin{vmatrix} \frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_w}{m_w^*}\exp(ikL) & \frac{k_b}{m_b^*}\exp(-ik_b L) \\ \exp(ik_w l_w) & \exp(-ik_w l_w) & -\exp(-ik_b l_w) \\ \frac{k_w}{m_w^*}\exp(ik_w l_w) & -\frac{k_w}{m_w^*}\exp(-ik_w l_w) & \frac{k_b}{m_b^*}\exp(-ik_b l_w) \end{vmatrix}$$

$$\exp(-ik_bL)\begin{vmatrix} \frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_w}{m_w^*}\exp(ikL) & -\frac{k_b}{m_b^*}\exp(-ik_bL) \\ \exp(ik_wl_w) & \exp(-ik_wl_w) & -\exp(ik_bl_w) \\ \frac{k_w}{m_w^*}\exp(ik_wl_w) & -\frac{k_w}{m_w^*}\exp(-ik_wl_w) & -\frac{k_b}{m_b^*}\exp(ik_bl_w) \end{vmatrix}$$

which the reader can see becomes very tedious indeed. Fortunately use can be made of the constant mass case of Jaros [5], and generalised to account for $m_w \neq m_b$ as here, giving:

$$\cos(k_wl_w)\cos(k_bl_b) - \sin(k_wl_w)\sin(k_bl_b)\left(\frac{m_b^2k_w^2 + m_w^2k_b^2}{2m_wm_bk_wk_b}\right) = \cos(kL) \tag{2.166}$$

If $E < V$, then following the arguments in Section 2.11, k_b becomes $i\kappa$, where:

$$\kappa = \sqrt{\frac{2m_b^*}{\hbar^2}(V-E)} \tag{2.167}$$

Substituting for k_b into equation (2.166), and using the following identities:

$$\cos(i\kappa l_b) \equiv \cosh(\kappa l_b) \qquad \text{and} \qquad \sin(i\kappa l_b) \equiv i\sinh(\kappa l_b) \tag{2.168}$$

then:

$$\cos(k_wl_w)\cosh(\kappa l_b) - \sin(k_wl_w)\sinh(\kappa l_b)\left(\frac{m_b^2k_w^2 - m_w^2\kappa^2}{2m_wm_bk_w\kappa}\right) = \cos(kL) \tag{2.169}$$

The superlattice dispersion curves, i.e. the energy E of a particle as a function of its wave vector k, are obtained by solving equations (2.166) and (2.169). This is accomplished by using the same methods as in Section 2.5, i.e. equations (2.166) and (2.169) are rewritten in the form $f(E,k) = 0$, and solved for chosen values of k. Again, a Newton–Raphson iteration is efficient; however, in order to avoid having to deduce $f'(E,k)$, a finite difference expansion can be employed:

$$f'(E,k) \approx \frac{f(E+\delta E,k) - f(E-\delta E,k)}{2\delta E} \tag{2.170}$$

Hence the Newton–Raphson iterative equation becomes:

$$E^{(n+1)} = E^{(n)} - \frac{2f(E^{(n)})\,\delta E}{f(E^{(n)}+\delta E) - f(E^{(n)}-\delta E)} \tag{2.171}$$

where δE is smaller than adjacent solutions for E; typically 10^{-3} meV is a good choice. This convenient method of approximating derivatives with nearby values of the function will be returned to in the next chapter and dealt with there in much more detail.

Fig. 2.27 shows the lowest-energy solutions of a 40 Å GaAs/40 Å $Ga_{0.6}Al_{0.4}As$ superlattice, for a range of k values. It can be seen that the energy E is a periodic function of k with period $2\pi/L$. The wave vector k was introduced

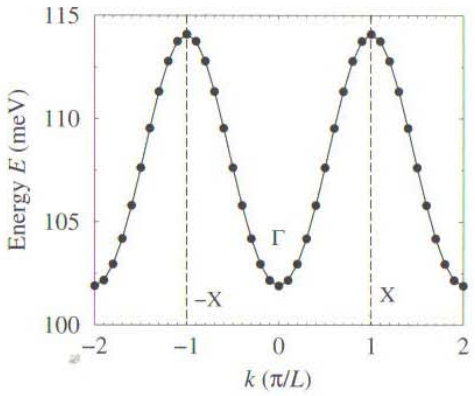


Figure 2.27 Lowest-energy solutions of the Kronig–Penney model of a 40 Å GaAs/40 Å $Ga_{0.6}Al_{0.4}As$ superlattice, for a range of electron momenta along the growth (z-) axis

earlier to describe the wave function as a travelling wave; clearly it can be identified as the momentum of a charge carrier along the growth (z-) axis of an *infinite* superlattice. Using Fig. 2.27, it can be seen that a definite minimum exists when the particle is at rest, i.e. k=0. As the carrier momentum increases, its energy increases too and reaches a maximum at $k = \pm\pi/L$. Thus the carrier within the superlattice occupies a continuous range of energies with a maxima and minima; this domain is the superlattice analogy to the energy bands of a crystal and is referred to as a *miniband*.

In addition the symmetry points of the E–k curves are labelled as in bulk crystals, where k=0 is called Γ and $k = \pi/L$ is referred to as X. This domain in k-space has become known as the *superlattice Brillouin zone* and as in the bulk, is a convenient way to express the relationship between energy and momenta.

Just as single quantum wells can have more than one confined state, superlattices can have more than one miniband. Fig. 2.28 displays the two lowest energy solutions of the longer period 60 Å GaAs/60 Å $Ga_{0.6}Al_{0.4}As$ superlattice; note on this scale that the lowest energy miniband looks almost flat. Clearly, the second miniband has the same periodicity as the first, although it should be noted that its minimum in energy occurs at the edge of the superlattice Brillouin zone and not at the centre. This effect is exploited in the recently developed 'inter-miniband laser' [41–43], which is a very promising high-power source of mid-infrared radiation. This device will be discussed in more detail later, i.e. in the chapter on electron scattering.

The structural dependence of the miniband is illustrated in Fig. 2.29 for a series of superlattices with equal well and barrier widths. The two solid curves display the

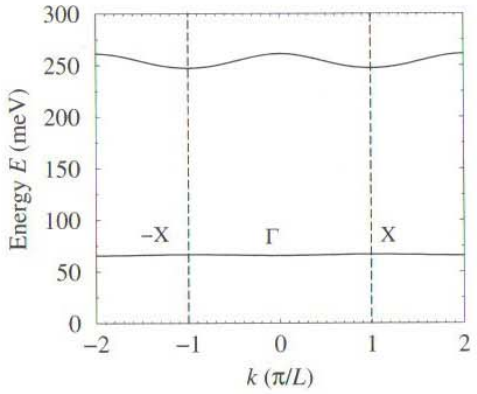


Figure 2.28 Two lowest-energy solutions of an 60 Å GaAs/60 Å $Ga_{0.6}Al_{0.4}As$ superlattice

energy as a function of l_w at Γ and X. It can be seen that these energies fall with increasing well width, as does the miniband width, defined as follows:

$$\text{miniband width} = |E(X) - E(\Gamma)| \tag{2.172}$$

where the modulus bars have been included to account for minibands such as the second, for which $E(X) < E(\Gamma)$. Interestingly, the top and bottom of the miniband straddle the energy of the corresponding single quantum well (given by the dashed line). As the wells become more isolated and further apart, the miniband narrows and the energies tend towards that of the single quantum well. This represents the cross-over from a superlattice to a multiple quantum well—a collection of non-interacting identical wells.

As the miniband width changes with structure, the curvature of the E–k curve changes too. It is clear from Chapter 1 that the effective mass of a particle subject to an energy–momentum dependence of $E(\mathbf{k})$ can be derived from:

$$m^* = \frac{\hbar^2}{\partial^2 E/\partial \mathbf{k}^2} \tag{2.173}$$

Therefore, when an electron or hole is moving along the growth axis of a superlattice, the energy–momentum relationship is such that they appear to have a different effective mass from the bulk. Fig. 2.30 plots this new effective mass at Γ for the range of structures in Fig. 2.29.

It can be seen from the figure that at very short periods the electron's miniband structure produces a new effective mass which is considerably smaller than the bulk value (0.067 m_0). Thus electrons within this miniband will have a much greater

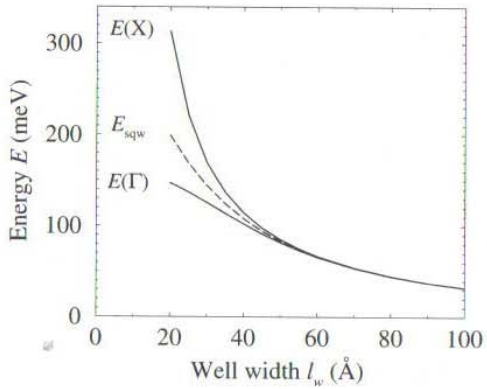


Figure 2.29 Energy at $k = 0\,(\Gamma)$ and $k = \pi/L\,(\mathrm{X})$, showing the lower and upper energy limits of the superlattice minibands, compared with the ground-state energy of the single quantum well of the same width

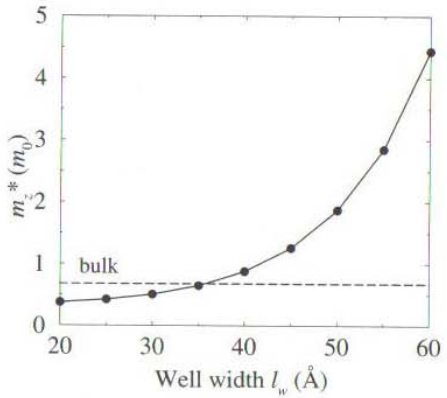


Figure 2.30 The new effective mass for an electron moving in the superlattice miniband

response to an electric field than in bulk, just as from Newton's second law, '$F = ma$', a lower m requires a smaller force in order to produce the same acceleration. In strict semiconductor language, this is quantified in terms of the carrier mobility, which is

given by:

$$\mu = \frac{e\tau_{\text{coll.}}}{m^*} \tag{2.174}$$

where $\tau_{\text{coll.}}$ is the time between collisions, see [1], p. 601, the reciprocal of which is usually referred to as the total scattering rate for all mechanisms, via phonons, ionised impurities, other carriers, etc. Clearly, a reduced effective mass within a miniband gives an increased mobility μ for transport within that miniband, *provided* that all other scattering terms remain the same.

In this series of calculations, attention has been focussed on superlattices with a simple single quantum-well unit cell, and indeed with equal well and barrier widths. If these restrictions are lifted, then there are an almost infinite number of possible superlattices that can be constructed, thus giving the opportunity to engineer, even design, the effective mass to suit the device.

2.14 THE SINGLE BARRIER

Hitherto, only semiconductor layered systems have been considered which form quantum wells, i.e. produce semiconductor layers which trap or confine the electrons along one axis. The 'opposite' of such structures also exist and are known as *barrier structures*. If a layer of a larger-band-gap material, e.g. $Ga_{1-x}Al_xAs$, is sandwiched between layers of a narrower-band-gap material, e.g. GaAs, then a potential barrier *can* result which repels carriers.

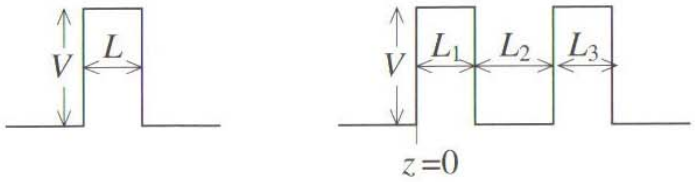


Figure 2.31 Single- and double-barrier structures

Fig. 2.31 illustrates such a single barrier structure. Electrons placed within this layered system simply collect in the GaAs outer (lower energy) regions and there are no quantum-confined energy states. These systems do, however, exhibit quantum behaviour when an electric field is applied perpendicular to the layers (along the growth (z-) axis). Electrons (or holes) arising from doping are accelerated and impinge upon the barrier and even when they have an energy E which is less than the potential energy height V of the barrier, they have a finite probability of passing through the barrier and appearing on the other side. This phenomenon is called *quantum mechanical tunnelling* or often just *tunnelling*. The bizarre nature of quantum mechanical tunnelling can be illustrated with a classical analogy, i.e. one would be pretty disturbed, if when kicking a football against a brick wall, it went straight through! Yet this is exactly what happens with *some* electrons and holes when they meet a potential barrier within a crystal. Tunnelling is an everyday phenomena which occurs in a range of semiconductor devices some of which appear in consumer electronics products, see, for example, Sze [44].

One way of quantifying the proportion of electrons that tunnel through a barrier is in terms of the *transmission coefficient* which is defined as the probability that any single electron impinging on a barrier structure will tunnel and contribute to the current flow through the barrier. Ferry has produced a comprehensive analysis of the transmission coefficient for a single-barrier structure (see [45], p. 60). Suffice here to quote Ferry's result in ([45], equation 3.12), for a constant effective mass across the structure, i.e. the transmission coefficient at an energy E for a barrier of width L and height V is given by:

$$T(E) = \frac{1}{1 + \left(\frac{k^2+\kappa^2}{2k\kappa}\right)^2 \sinh^2(\kappa L)}, \qquad \text{for} \qquad E < V \tag{2.175}$$

where as usual:

$$k = \frac{\sqrt{2m^*E}}{\hbar} \quad \text{and} \quad \kappa = \frac{\sqrt{2m^*(V-E)}}{\hbar} \tag{2.176}$$

For values of the carrier energy E greater than the barrier height V, $\kappa \rightarrow ik'$ (as in Section 2.13), and hence:

$$T(E) = \frac{1}{1 + \left(\frac{k^2-k'^2}{2kk'}\right)^2 \sin^2(k'L)} \qquad \text{for} \qquad E > V \tag{2.177}$$

where

$$k = \frac{\sqrt{2m^*E}}{\hbar} \quad \text{and} \quad k' = \frac{\sqrt{2m^*(E-V)}}{\hbar} \tag{2.178}$$

The mathematics show, that for $E > V$, the transmission coefficient would be expected to oscillate, with resonances when the sine term is zero. These occur when:

$$k'L = n\pi \tag{2.179}$$

which is implied when:

$$E = \frac{(n\hbar\pi)^2}{2m^*L^2} + V \tag{2.180}$$

The squared dependence implies that the resonances when $T = 1$ occur at larger and larger intervals in E, which can be clearly seen in Figs 2.32 and 2.33.

Fig. 2.32 displays the transmission coefficient as a function of the energy E and for a range of barrier widths L. For this range of calculations the barrier height V was fixed at 100 meV, and below this energy the thinner the barrier, then the higher the probability of tunnelling. For $E > V$, the situation is more complex due to the oscillatory nature of T. The trend, however, as highlighted by the curves, is that the

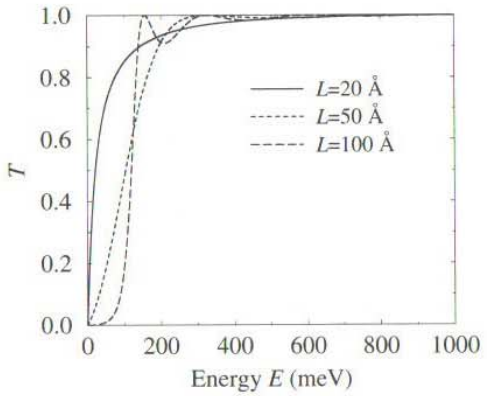


Figure 2.32 Transmission coefficient as a function of the energy through a single barrier for different barrier widths

thicker the barrier, the closer the first resonance ($T = 1$) is to the top of the barrier. This can be understood from equation (2.180):

$$\lim_{L\to\infty} = E^{n=1}_{\text{resonance}} = \lim_{L\to\infty} \frac{(\hbar\pi)^2}{2m^* L^2} + V = V \tag{2.181}$$

Conversely, for a fixed L and a variable V, the first resonance occurs at the same point above the barrier height; this is clearly illustrated in Fig. 2.33.

2.15 THE DOUBLE BARRIER

If two barriers are placed a reasonably small distance apart (in the same crystal, perhaps a few nm) then the system is known as a *double barrier* (see Fig. 2.31), and has quite different transmission properties to the single barrier. Datta [46], p. 33, has deduced the transmission $T(E)$ dependence for the restricted case of symmetric barriers, while Ferry [45], p. 66, has considered asymmetric barriers. In this formalism, allowance will be made for differing barrier widths as well as the discontinuous change in the effective mass between well and barrier materials. Thus, with the aim of deducing the new $T(E)$, consider the solutions to Schrödinger's equation within

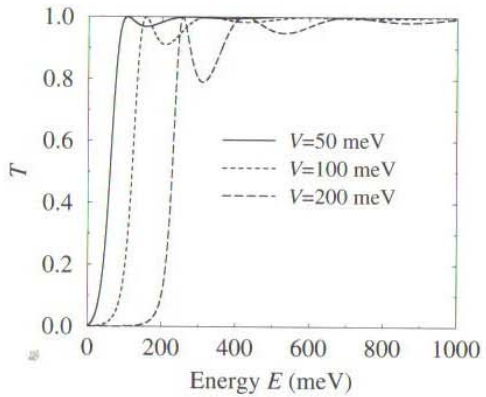


Figure 2.33 Transmission coefficient as a function of the energy through a single 100 Å barrier for different barrier heights

each region for $E < V$:

$$\psi(z) = A\exp(ikz) + B\exp(-ikz), \qquad z < I_1 \tag{2.182}$$
$$\psi(z) = C\exp(\kappa z) + D\exp(-\kappa z), \quad I_1 < z < I_2 \tag{2.183}$$
$$\psi(z) = F\exp(ikz) + G\exp(-ikz), \quad I_2 < z < I_3 \tag{2.184}$$
$$\psi(z) = H\exp(\kappa z) + J\exp(-\kappa z), \quad I_3 < z < I_4 \tag{2.185}$$
$$\psi(z) = K\exp(ikz) + L\exp(-ikz), \quad I_4 < z \tag{2.186}$$

where k and κ have their usual forms as given in equation (2.176) and the positions of the interfaces have been labelled I_1, I_2, I_3, and I_4, respectively. Using the standard BenDaniel–Duke boundary conditions at each interface gives the following

$z = I_1 = 0$:

$$A + B = C + D \tag{2.187}$$

$$\frac{1}{m_w}(ikA - ikB) = \frac{1}{m_b}(\kappa C - \kappa D) \tag{2.188}$$

$z = I_2 = L_1$:

$$C\exp(\kappa I_2) + D\exp(-\kappa I_2) = F\exp(ikI_2) + G\exp(-ikI_2) \tag{2.189}$$

$$\frac{1}{m_b}[\kappa C\exp(\kappa I_2) - \kappa D\exp(-\kappa I_2)] = \frac{1}{m_w}[ikF\exp(ikI_2) - ikG\exp(-ikI_2)] \tag{2.190}$$

$z = I_3 = L_1 + L_2$:

$$F\exp(ikI_3) + G\exp(-ikI_3) = H\exp(\kappa I_3) + J\exp(-\kappa I_3) \tag{2.191}$$

$$\frac{1}{m_w}\left[ikF\exp(ikI_3) - ikG\exp(-ikI_3)\right] = \frac{1}{m_b}\left[\kappa H\exp(\kappa I_3) - \kappa J\exp(-\kappa I_3)\right] \tag{2.192}$$

$z = I_4 = L_1 + L_2 + L_3$:

$$H\exp(\kappa I_4) + J\exp(-\kappa I_4) = K\exp(ikI_4) + L\exp(-ikI_4) \tag{2.193}$$

$$\frac{1}{m_b}\left[\kappa H\exp(\kappa I_4) - \kappa J\exp(-\kappa I_4)\right] = \frac{1}{m_w}\left[ikK\exp(ikI_4) - ikL\exp(-ikI_4)\right] \tag{2.194}$$

The method of solution is the transfer matrix technique as before, writing the above equations in matrix form:

$$\mathbf{M}_1\begin{pmatrix}A\\B\end{pmatrix} = \mathbf{M}_2\begin{pmatrix}C\\D\end{pmatrix} \tag{2.195}$$

$$\mathbf{M}_3\begin{pmatrix}C\\D\end{pmatrix} = \mathbf{M}_4\begin{pmatrix}F\\G\end{pmatrix} \tag{2.196}$$

$$\mathbf{M}_5\begin{pmatrix}F\\G\end{pmatrix} = \mathbf{M}_6\begin{pmatrix}H\\J\end{pmatrix} \tag{2.197}$$

$$\mathbf{M}_7\begin{pmatrix}H\\J\end{pmatrix} = \mathbf{M}_8\begin{pmatrix}K\\L\end{pmatrix} \tag{2.198}$$

Then, as before, the coefficients of the outer regions can be linked by forming the transfer matrix, i.e.

$$\begin{pmatrix}A\\B\end{pmatrix} = \mathbf{M}_1^{-1}\mathbf{M}_2\mathbf{M}_3^{-1}\mathbf{M}_4\mathbf{M}_5^{-1}\mathbf{M}_6\mathbf{M}_7^{-1}\mathbf{M}_8\begin{pmatrix}K\\L\end{pmatrix} \tag{2.199}$$

Clearly, this 2×2 matrix equation still has four unknowns and can't be solved—it is at this point that additional boundary conditions have to be imposed from physical intuition. Whereas before, the standard boundary conditions, i.e. $\psi(z) \to 0$ as $z \to \pm\infty$, were used to solve for the confined states within quantum wells, in these barrier structures these are not appropriate since the travelling waves in the outer layers can have infinite extent. The standard procedure is to assume, quite correctly, that all of the charge carriers approach the double barrier from the same side, as would occur when as part of a biased device, as illustrated schematically in Fig. 2.34. Furthermore, if it is assumed that there are no further heterojunctions to the right of the structure, then no further reflections can occur and the wave function beyond the structure can only have a travelling wave component moving to the right, i.e. the coefficient L must be zero.

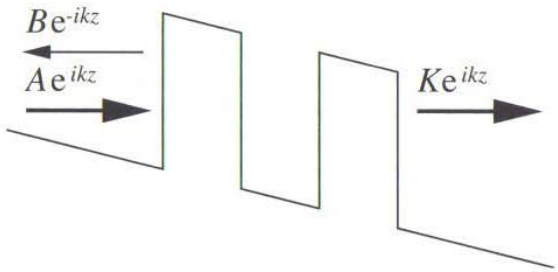


Figure 2.34 The wave function after imposition of the boundary conditions

Therefore, if the 2×2 matrix of equation (2.199) is written as $\mathcal{M}$, then:

$$\begin{pmatrix}A\\B\end{pmatrix} = \mathcal{M}\begin{pmatrix}K\\0\end{pmatrix} \tag{2.200}$$

$$\therefore A = \mathcal{M}_{11}K \tag{2.201}$$

and the ratio of transmitted to incident current, i.e. the transmission coefficient, is simply:

$$T(E) = \frac{K^*K}{A^*A} = \frac{1}{\mathcal{M}_{11}^*\mathcal{M}_{11}} \tag{2.202}$$

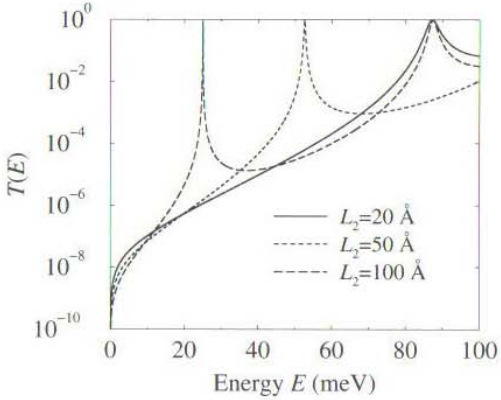


Figure 2.35 Transmission coefficient as a function of the energy through a double barrier of width 100 Å and height 100 meV, separated by a distance L_2

Fig. 2.35 gives an example of the form of $T(E)$ for barriers of height 100 meV and width 100 Å as a function of the distance L_2 between them. The effective masses

were both taken to be equal to the bulk Γ valley electron mass of 0.067 m_0. It can be seen from the figure that the curves contain Dirac δ-functions at certain energies E below the potential barrier height V. This is quite unlike the single barrier case. At these *resonance* energies, the double-barrier system appears transparent and has a transmission coefficient of 1. The wave functions of these states are localised between the barriers and are often referred to as *quasi-bound states* since they resemble the bound states of quantum well structures. However, they are not stationary states in that electrons or holes in such states will eventually scatter into the lower energy states outside of the barriers.

The effect of an increasing barrier height V is shown in Fig. 2.36. It can be seen that, away from a resonance, an increasing barrier height leads, as would be expected, to a decrease in the transmission coefficient T. The classical explanation for this would be, 'it is harder for the electrons to tunnel through higher barriers'. The *resonance energies* increase with increasing barrier height due to confinement effects, and the appearance of the second resonance at higher energies is a reflection on the existence of a second quasi bound state.

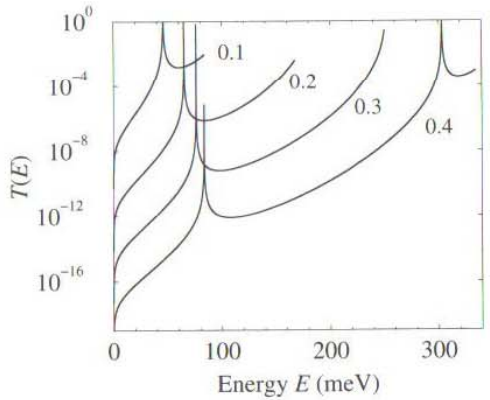


Figure 2.36 Transmission coefficient as a function of the energy through a 100 Å $Ga_{1-x}Al_xAs$ /50 Å GaAs/100 Å $Ga_{1-x}Al_xAs$ double barrier, for x=0.1, 0.2, 0.3, and 0.4

While the transmission coefficient represents a very important parameterisation of the properties of a double-barrier structure, it itself is not a measurable quantity. In fact, the properties of such two-terminal electronic devices are generally inferred (or summarised) from their current–voltage characteristics (I–V curve).

When an electric field is placed across such a double barrier structure, any charge carriers present in the semiconductor, intrinsic or extrinsic, constitute a current which approaches the left-hand barrier. These charge carriers have a distribution of energy

and momenta, often a Fermi-Dirac distribution. Those carriers that are of the same energy as the resonance are able to pass right through the double barrier without hindrance—a phenomenum which has become known as *resonant tunnelling*. As the applied electric field (applied voltage) is increased, the number of carriers with the resonance energy increases and peaks as the Fermi level of the semiconductor to the left of the first barrier is brought into alignment. Therefore the current gradually increases. At higher fields, the current falls away and a period of negative differential resistance ensues [47, 48]. Such a current–voltage characteristic has been exploited in high frequency circuits, thus bringing the *resonant tunnelling diode* to prominence as a very useful electronic device [49].

There are various models of the current–voltage properties of different levels of complexity, the simplest of which would probably be to return to the idea that the current at any particular field would be equal to the number of carriers that tunnelled through the structure. This in turn would be the probability of a particular carrier tunnelling, multiplied by the number of carriers at that energy, i.e.

$$I \propto \int_{\text{band}} T(E) f^{\text{FD}}(E) \rho^{\text{3D}}(E)\ \mathrm{d}E \tag{2.203}$$

As the carriers approaching the barrier structure are in a bulk band, then the integral is over their energies, and the Fermi–Dirac distribution function and density of states have the bulk (3D) forms.

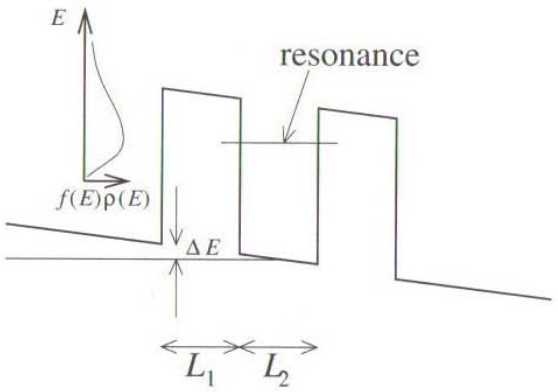


Figure 2.37 Simple model of current flow through a double barrier structure

Figure 2.37 outlines this model. The electric field dependence is introduced via f^{FD} and ρ^{3D}. As the electric field is increased, the bottom of the bulk band is increased in energy relative to the centre of the double-barrier structure by an amount $\Delta E = eF(L_1 + L_2/2)$. Hence, the reference energy (the band minimum) will increase, and by using equation (2.40), the density of states at some energy E, measured from the

band minimum at the centre of the well, would become

$$\rho(E) = \frac{1}{2\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{\frac{3}{2}}(E-\Delta E)^{\frac{1}{2}}\,\Theta(E-\Delta E) \tag{2.204}$$

where the unit step function ensures that the energy of the carriers impinging on the left hand barrier E is greater than ΔE. At the bottom of the band, $E = \Delta E$, and $\rho(E) = 0$. In addition, the Fermi energy of the bulk carriers increases by the same amount, i.e. from equation (2.49):

$$f^{\mathrm{FD}}(E) = \frac{1}{\exp\{[E-(E_{\mathrm{F}}+\Delta E)]/kT\}+1} \tag{2.205}$$

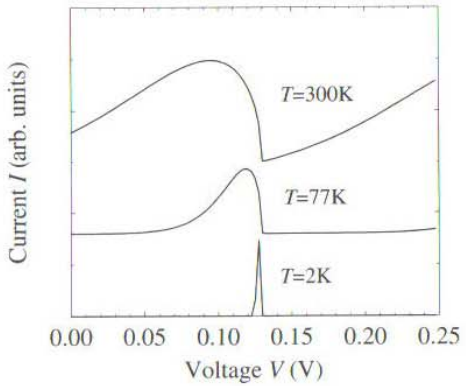


Figure 2.38 Current–voltage curve for the 100 Å $Ga_{0.8}Al_{0.2}As$/50 Å GaAs/100 Å $Ga_{0.8}Al_{0.2}As$ of earlier, obtained at different temperatures

Figure 2.38 illustrates the results obtained by using this model. The current is obtained in arbitrary units and the voltage V has been defined simply as the potential difference across the structure at a particular field, i.e. $V = F(L_1 + L_2 + L_3)$. The I–V curves have been plotted for several temperatures for the $x = 0.2$ structure of Fig. 2.36. The single resonance within this system shows itself as a single peak in the current. Clearly, the current peak broadens as the temperature increases, with this being a direct result of the broadening of the carrier distribution (the number of carriers in a given energy range), given by the product $f^{\mathrm{FD}}(E)\rho(E)$ appearing in the integral for the current, and illustrated schematically in Fig. 2.37. At low temperatures, the carriers occupy a small region of energy space around the band minima. As the field is increased, current only begins to flow when this narrow distribution is brought into

line with the resonance energy, and hence there is a narrow current peak. As the temperature increases, the carrier distribution broadens, therefore there is a greater range of applied voltages that give some degree of alignment of the carriers with the resonance energy. The peak occurs when the peak of the distribution is aligned with the resonance energy. At voltages above this, the number of carriers available for tunnelling decreases, and hence the current also decreases.

The non-zero current at zero field for $T = 300$ K is a consequence of this simplistic model. It arises because even at zero field, there is a finite number of carriers in the broadened distribution which are aligned with the resonance energy and are therefore able to tunnel.

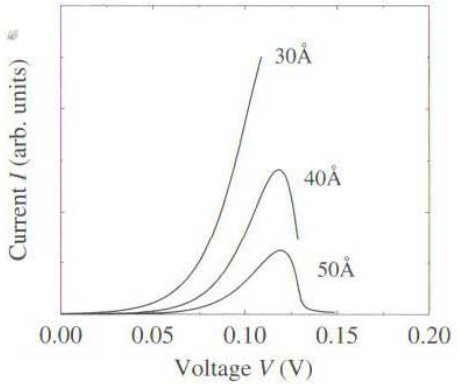


Figure 2.39 Current–voltage curves obtained at 77 K for a variety of barrier widths the structure of Fig. 2.38

Fig. 2.39 displays the results of the same simple model as a function of barrier width for the same structure of Fig. 2.38, but this time at the fixed temperature of 77 K. Again, as noted before, thinner barriers allow the electrons to tunnel more easily and hence give a higher current. An interesting physical point can be inferred from this data. The incomplete nature of the peaks in the 40 and 30 Å data is an indication that a significant fraction of the electrons are passing over the top of the barriers. While the current–voltage model advanced in this section can account for such a situation, the transmission coefficient versus energy data is terminated at the top of the barrier, and so only the fraction of electrons below this are included in the integration for the current in equation (2.203). It is left as an exercise for the reader to generalise the transmission coefficient theory to account for the case when $E > V$, and therefore improve the current–voltage model. This can be achieved by using a similar approach to that of Section 2.13.

The above discussion has been a simple introduction to the modelling of I–V curves for barrier structures, but nonetheless it shows some of the features of real devices. For a much more complete and in depth study see, for example, Mizuta and Tanoue [49].

2.16 EXTENSION TO INCLUDE ELECTRIC FIELD

An obvious improvement to the above model would be to account for the changes in the transmission coefficient as a function of the applied electric field. By using the substitution as before (equation (2.150)), i.e.

$$z' = \left(\frac{2m^*}{\hbar^2}\right)^{\frac{1}{3}} \left[\frac{V(z) - E}{(eF)^{\frac{2}{3}}} - (eF)^{\frac{1}{3}} z\right] \tag{2.206}$$

the solution in each region can then be written as a linear combination of Airy functions, just as for the general electric field case of equation (2.152), i.e.

$$\psi(z) = A\mathrm{Ai}(z') + B\mathrm{Bi}(z'), \qquad z < I_1 \tag{2.207}$$
$$\psi(z) = C\mathrm{Ai}(z') + D\mathrm{Bi}(z'), \quad I_1 < z < I_2 \tag{2.208}$$
$$\psi(z) = F\mathrm{Ai}(z') + G\mathrm{Bi}(z'), \quad I_2 < z < I_3 \tag{2.209}$$
$$\psi(z) = H\mathrm{Ai}(z') + J\mathrm{Bi}(z'), \quad I_3 < z < I_4 \tag{2.210}$$
$$\psi(z) = K\mathrm{Ai}(z') + L\mathrm{Bi}(z'), \quad I_4 < z \tag{2.211}$$

while Airy functions *can* be difficult to work with numerically, an immediate advantage over the zero-field case is that this solution is valid for both $E < V$ and $E > V$, and hence generalisation to this form produces two benefits. The method of solution is analogous to the zero-field case, in that application of the BenDaniel–Duke boundary conditions yields two equations for each interface, which in this case gives eight equations. The unknown coefficients, A and B, are linked to K and L as before by forming the transfer matrix, and are solved by imposition of a boundary condition, which is again a travelling wave in the direction of $+z$ to the right of the barrier structure. It is left to the interested reader to follow through such a derivation. A very general implementation for multiple barrier structures has been reported in the literature by Vatannia and Gildenblat [39].

2.17 MAGNETIC FIELDS AND LANDAU QUANTISATION

If a magnetic field is applied externally to a *non-magnetic* semiconductor heterostructure then the constant effective mass Hamiltonian (familiar from equation (2.5)):

$$\mathcal{H} = -\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} + V(z) \tag{2.212}$$

which can be written:

$$\mathcal{H} = \frac{\mathcal{P}^2}{2m^*} + V(z) \tag{2.213}$$

becomes [50–53]:

$$\mathcal{H} = \frac{1}{2m^*}\left(\mathcal{P} + e\mathbf{A}\right)^2 \mp \frac{1}{2}g^*(z)\mu_B B + V(z) \tag{2.214}$$

where the kinetic energy operator becomes modified by the magnetic field vector potential $\mathbf{A}$, and the second term produces a splitting, known as the 'gyromagnetic spin splitting' between the spin-up ($-$ sign) and spin-down ($+$ sign) electrons. g^* is known as the 'Landau factor' and is really a function of z as it depends on the material, however it is generally assumed to be constant and approximately 2 for conduction band electrons, μ_B is the Bohr magneton and B is the magnitude of the magnetic flux density which is assumed aligned along the growth (z-) axis.

Although the heterostructure potential $V(z)$ remains one-dimensional, the vector potential means that the wave functions are not necessarily one-dimensional so the Schrödinger equation must be written:

$$\left[\frac{1}{2m^*}\left(\mathcal{P} + e\mathbf{A}\right)^2 \mp \frac{1}{2}g^*(z)\mu_B B + V(z)\right]\psi(x,y,z) = E\psi(x,y,z) \tag{2.215}$$

The magnetic field produces a parabolic potential along one of the in-plane axes, the x-axis say, leaving the particle free to move (with a wave vector k_y, say) along the other axis. The standard approach is to employ the Landau gauge $\mathbf{A} = Bx\mathbf{e}_y$, then following the notation of Savić *et al.* [54] the wave function can be written in the separable form $\psi(x,y,z) = \psi_x(x)\psi_y(y)\psi_z(z)$, i.e.

$$\psi(x,y,z) = \frac{1}{\sqrt{L_y}}\psi_j(x - X_{k_y})\mathrm{e}^{ik_y y}\psi(z) \tag{2.216}$$

where L_y is a normalisation constant (the length of the structure along the y-axis, $X_{k_y} = -k_y l_B^2$, where $l_B = \sqrt{\hbar/eB}$ is the Landau length, j is an index over the $\psi_j(x - X_{k_y})$ harmonic oscillator solutions of the parabolic potential and $\psi(z)$ is the usual one-dimensional envelope function of the heterostructure potential without the magnetic field. The harmonic oscillator solutions will be generated numerically in Section 3.5, however they can also be expressed analytically as:

$$\psi_j(x - X_{k_y}) = \frac{1}{\pi^{\frac{1}{4}}\sqrt{2^j j! l_B}}\exp\left(-\frac{(x - X_{k_y})^2}{2l_B^2}\right)H_j\left(\frac{x - X_{k_y}}{l_B}\right) \tag{2.217}$$

where H_j is the jth Hermite polynomial, see Liboff [55] page 200 for a detailed description of the analytical solutions of the harmonic oscillator.

Taking g^* as a constant equal to 2 and the Bohr magneton as 9.274×10^{-24} JT^{-1} then it can be seen that even in a relatively high magnetic field of 10 T, the difference in energy between the spin-up and spin-down electrons is:

$$g^*\mu_B B \approx 2 \times 9.274 \times 10^{-24}\mathrm{JT}^{-1} \times 10\ \mathrm{T} \approx 1\ \mathrm{meV} \tag{2.218}$$

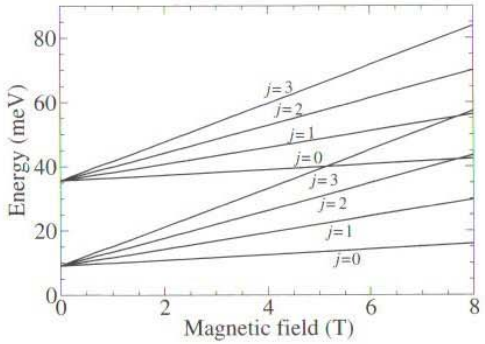


Figure 2.40 The magnetic field induced Landau levels of a 200 Å GaAs quantum well surrounded by $Ga_{1-x}Al_xAs$ barriers of height 100 meV (same structure as Fig. 2.15), with constant effective mass of $0.067m_0$. Note only the first 4 Landau levels associated with each quantum well subband are shown for clarity.

which on the scales of typical quantum well systems and typical carrier densities is relatively small and can generally be ignored. Thus the total energy therefore can be considered to be composed of just two components: the usual energy eigenvalue E_n associated with the nth state of the quantum well without a magnetic field and the harmonic oscillator type energy associated with the in-plane cyclotron motion $(j + \frac{1}{2})\hbar\omega_c$, i.e.

$$E_{n,j} = E_n + \left(j + \frac{1}{2}\right)\hbar\omega_c \tag{2.219}$$

where the 'cyclotron frequency' is the usual one from bulk semiconductors [1,2,56]:

$$\omega_c = \frac{eB}{m^*} \tag{2.220}$$

Fig. 2.40 shows the Landau level splitting induced by the application of an external magnetic field on a 200 Å GaAs quantum well surrounded by $Ga_{1-x}Al_xAs$ barriers, assuming a constant electron effective mass of $0.067m_0$. The figure illustrates the lowest four Landau levels for each of the two lowest energy confined states within the quantum well. A more accurate calculation could include the variation in effective mass of the carrier between the various layers; this would be done by calculating the weighted mean or expectation value of the effective mass according to the probability of finding the carrier in each of the semiconductor layers.

2.18 IN SUMMARY

The consideration of simple layered semiconductor heterostructures with analytical forms for the solutions to Schrödinger's equation has allowed exploration and discovery of the properties of two-dimensional systems. Such models are invaluable and allow a whole range of physical observables of experimental layer structures and electronic devices to be explained. However, implementing these methods computationally *can* be tedious, in that a computer program, written to calculate the energy levels of a single quantum well, has to be rewritten for a double quantum well. In addition, a different program would be needed to solve a triangular well, and furthermore some potential profiles, which can be fabricated, such as diffused quantum wells, have no analytical solution at all. While such continuously varying structures can be expressed with flat step potentials for each monolayer [57], the treatment of an electric field in this manner is questionable.

In the next chapter, a simple, but very general *numerical* solution to Schrödinger's equation will be derived, which will overcome all of these difficulties.