Chapter 2

Schrödinger-Poisson Model for Quantum Transport

In conventional electron devices, the physical processes contributing to electron transport can be usefully studied in the semi-classical limit by solving the Boltzmann equation for the one-particle distribution function $f(q,p)$ in phase-space

$$\frac{\partial f}{\partial t} = -\frac{\partial \mathcal{E}(q,p)}{\partial p}\frac{\partial f}{\partial q} + \left[\frac{\partial v(q)}{\partial q} + \frac{\partial \mathcal{E}(q,p)}{\partial q}\right]\frac{\partial f}{\partial p} + \left.\frac{\partial f}{\partial t}\right|_{Coll.} \tag{2.1}$$

where $v(q)$ is the self-consistent potential and $\mathcal{E}(q,p)$ is the energy band structure. The collision term takes into account relaxation mechanisms such as electron-phonon scattering. Physical quantities of interest in device analysis such as the particle density and current density can be obtained as averages over the distribution function $f(q,p)$:

$$\begin{aligned} n(q) &= \frac{1}{2\pi}\int dp\, f(q,p) \\ J(q) &= \frac{1}{2\pi}\int dp \frac{\partial \mathcal{E}(q,p)}{\partial p} f(q,p) \end{aligned} \tag{2.2}$$

A major assumption in the semiclassical limit is that the evolution of $f(q,p)$ at any position depends only locally on the potential. This approximation precludes quantum interference effects that arise from the non local action of the potential. Therefore Eq. 2.1 cannot describe phenomenon such as quantum mechanical tunneling through classically forbidden regions. While the semi-classical approach has been very successful in the study of conventional

devices, it fails completely in heterostructure devices that depend on quantum interference effects for their operation.

Quantum interference effects are most easily studied using models based on the effective-mass Schrödinger equation. Electron transport is assumed to be coherent throughout the device, dissipative mechanisms appearing only in the device contacts. In this chapter, a single band Schrödinger model is described and applied to resonant-tunneling diodes.

This study is confined to electron transport in heterostructure devices such as $In_yAl_{1-y}As/In_xGa_{1-x}As$ double barrier diodes. Assuming a simple double barrier potential superimposed on a uniform semiconductor, the validity of the effective-mass approximation, even in the presence of strong barriers, is semi-quantitatively demonstrated using a Wannier function basis constructed by Pedersen et al. (1991) for the Kronig-Penny model. To treat the variation in the bandstructure across a heterojunction, an effective-mass Hamiltonian based on the Weyl correspondence rule, that has not been previously investigated in the study of heterostructures, is proposed. The effective-mass treatment for a heterostructure is facilitated by the concept of generalized Wannier functions. However, since generalized Wannier functions are difficult to construct explicitly for real heterostructures, the significance of finite-difference formulations is discussed. The treatment of the boundary conditions is based on the work of Frensley (1991) and Lent and Kirkner (1990). Finally, the model presented here is extended to include the full , $-X$ conduction band edge in the position representation. The chapter concludes with an illustration of the model.

2.1 Effective-mass Hamiltonians

When the energy band structure or the potential vary significantly over atomic length scales, it is necessary to invoke the Schrödinger equation to describe the

motion of electrons:

$$\left(-\frac{\hbar^2}{2m_o}\nabla^2 - \imath\hbar\frac{\partial}{\partial t}\right)\Psi(z,t) + U(z)\Psi(z,t) = 0 \tag{2.3}$$

where m_o is the free electron mass, $\Psi(z,t)$ is the electron wavefunction, and $U(z)$ is a potential which includes the crystal potential and any external potential. It is appealing to rewrite Eq. 2.3 as

$$\left(\mathcal{H}^o + v\right)\Psi(z,t) = \imath\hbar\frac{\partial}{\partial t}\Psi(z,t) \tag{2.4}$$

where $\mathcal{H}^o$ is the crystal Hamiltonian and v is the external potential. The wavefunctions $\Psi(z,t)$ can be expanded in the basis of Wannier functions as

$$\Psi(z,t) = \sum_{n,R}\zeta_n(R,t)a_n(z-R) \tag{2.5}$$

where a_n is the Wannier function in the n^{th} band localized at the lattice site R and ζ_n the corresponding wave amplitude. Substituting 2.5 in Eq. 2.4, multiplying by $a^*_{n'}(z-R')$ and integrating over the lattice we get

$$\sum_{nR}\int dz\, a^*_{n'}(z-R')(\mathcal{H}^o + v)a_n(z-R) = \imath\hbar\frac{\partial\zeta_{n'}(R',t)}{\partial t} \tag{2.6}$$

Fourier expanding the energy band:

$$\mathcal{E}_n(k) = \sum_R \mathcal{E}_{nR}e^{\imath kR}$$

and using

$$\mathcal{H}^o a_n(z-R) = \sum_{R'}\mathcal{E}_{n,R-R'}a_n(z-R')$$

we get (Ziman, 1972)

$$\sum_{R'}\mathcal{E}_{nR'-R}\zeta_n(R',t) + \sum_{n'R'}\mathcal{V}_{nn'RR'}\zeta'_n(R',t) = \imath\hbar\frac{\partial\zeta_n(R,t)}{\partial t} \tag{2.7}$$

where

$$\mathcal{V}_{nn'RR'} = \int dz\, a^*_{n'}(z-R')v(z)a_n(z-R) \tag{2.8}$$

Alternatively, ζ_n can be treated as a continuous envelope function defined everywhere instead of just at the lattice sites. In that case the following equation results (Ziman, 1972):

$$\left[\mathcal{E}_n(-\imath\nabla) - \imath\hbar\frac{\partial}{\partial t}\right]\zeta_n(z,t)\Big|_{z=R} + \sum_{n',R'}\mathcal{V}_{nn'RR'}\zeta_n(R',t) = 0 \tag{2.9}$$

Then Eq. 2.9 is a Schrödinger equation with the envelope function ζ_n taking the place of a wavefunction. The operator $\mathcal{E}_n$ replaces the crystal Hamiltonian $\mathcal{H}^o$. If the energy dispersion relation $\mathcal{E}_n(k)$ is known, the kinetic energy operator in position representation is obtained by substituting k with $-\imath\nabla$.

If the perturbing potential is not very strong and does not vary rapidly in time, $\mathcal{V}$ is vanishingly small for $n \neq n'$ and a single-band equation results:

$$\sum_{R'}\mathcal{E}_{nR'-R}\zeta_n(R',t) + \sum_{R'}\mathcal{V}_{nnRR'}\zeta_n(R',t) = \imath\hbar\frac{\partial\zeta_n(R,t)}{\partial t} \tag{2.10}$$

For electron transport in devices based on the $In_xAl_{1-x}As/In_yGa_{1-y}As$ heterojunctions, the single-band equation appears to be satisfactory. However, in the study of hole transport or in interband tunneling devices such as those based on the InAs/GaSb/AlSb material system (Collins et al., 1991), interband coupling plays a major role and a multiband treatment of quantum transport is necessary (Ting et al., 1992).

If the potential is also not rapidly varying in space, then

$$\mathcal{V}_{RR'} = v(R)\delta_{RR'} \tag{2.11}$$

and we obtain the single-band effective mass equation

$$\mathcal{H}\zeta = \left[\mathcal{E}(-\imath\nabla) + v(z)\right]\zeta = \imath\hbar\frac{\partial\zeta}{\partial t} \tag{2.12}$$

This is a very practical approximation since the Wannier functions are no longer required.

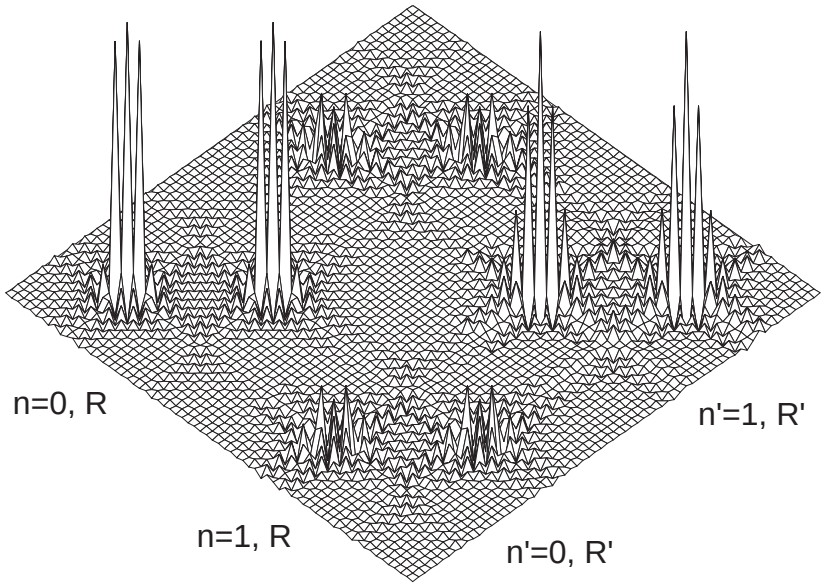

Figure 2.1: Matrix elements $\mathcal{V}_{nn'RR'}$ for a double barrier quantum well potential with 1.7nm wide barriers and a 5nm well.

The validity of the approximations made here can be evaluated using a Kronig-Penney model for which explicit Wannier functions have been constructed by Pedersen et al. (1991). The aim here is only to obtain a feel for the approximations. To this end, the Wannier functions

$$a_n(z) = \frac{(-1)^n}{\sqrt{\Delta}} \cos\left[\frac{(2n+1)\pi z}{2\Delta}\right] \frac{\sin(\pi z/2\Delta)}{\pi z/2\Delta}$$

in the "free-electron" limit are assumed (Ziman, 1972; Pedersen et al., 1991) for a lattice with spacing Δ and Brillouin zone $[-\pi/\Delta, \pi/\Delta]$. The matrix elements $\mathcal{V}_{nn'RR'}$ including only the first two bands are shown in Fig. 2.1. The interband terms are less than 15% in strength compared to the intraband terms. The

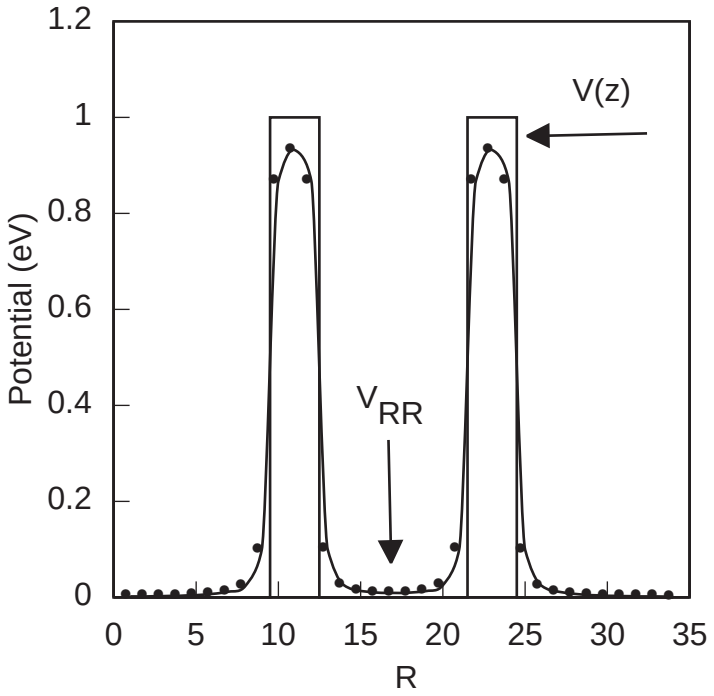

Figure 2.2: Comparison of $\mathcal{V}_{00RR}$ to $v(z)_{z=R\Delta}$ for double barrier quantum well potential with 1.7nm wide barriers and a 5nm well.

approximation that $\mathcal{V}$ is diagonal in the lattice site index becomes poor with increasing band index because the Wannier functions for the higher bands are less localized. The diagonal elements $\mathcal{V}_{RR}$ for n=0 are compared to a double barrier potential in Fig. 2.2. If the Wannier functions are taken to be delta functions ($\delta(z - R\Delta)$), then the matrix elements $\mathcal{V}$ are diagonal in the position representation.

The simplest quantum mechanical treatment of electron motion is based on a single band effective-mass Hamiltonian:

$$\mathcal{H} = -\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} + v_h(z) + v_s(z) \tag{2.13}$$

where v_h is the Hartree potential and is the average effect on the motion of an electron due to the many other electrons and can be obtained from Poisson's equation. The abrupt change in the local band-edge energy across heterojunc-

tions is accounted for by v_s. In the description of devices such as GaAs/AlAs resonant-tunneling diodes where the effective mass in AlAs ($0.15m_o$) is very different from that in GaAs ($0.067m_o$), it is necessary to include the spatial dependence of effective-mass. Regardless of the fact that an energy band is ill-defined in the absence of translational symmetry, $\mathcal{E}_n(-\imath\nabla)$ in Eq. 2.12 can be replaced $\mathcal{E}_n(-\imath\nabla, z)$ (a related notion is the Short-Time Fourier Transform (STFT) (Rioul & Vetterli, 1991) used to represent non-stationary signals). Because $\hat{z}$ and $\hat{k}$ do not commute, the kinetic energy operator corresponding to $\mathcal{E}_n(z, k)$ must be constructed by assuming a correspondence rule. Several forms for the Hamiltonian have been proposed (Zhu & Kroemer, 1983; Morrow & Brownstein, 1984; Einevoll et al., 1990; von Roos, 1983), the most favored being the "minimal Hermitian form,"

$$\mathcal{H} = -\frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m^*(z)}\frac{\partial}{\partial z} + v(z) \tag{2.14}$$

An interesting correspondence rule is the Weyl transform (Groot & Suttorp, 1972; Balescu, 1973; Leaf, 1968). The Weyl prescription for a parabolic energy band in the presence of a spatially varying effective-mass is

$$\mathcal{H} = -\frac{\hbar^2}{8}\left[\frac{1}{m^*(z)}\frac{\partial^2}{\partial x^2} + 2\frac{\partial}{\partial z}\frac{1}{m^*(z)}\frac{\partial}{\partial z} + \frac{\partial^2}{\partial z^2}\frac{1}{m^*(z)}\right] + v(z) \tag{2.15}$$

It should be mentioned that the correspondence rule is only a postulate, and there is no point in trying to prove a particular rule. The results depend only weakly on the correspondence rule, as shown in 2.3, where the conduction curves of a typical resonant-tunneling diode obtained using models based on Eqs. 2.14 and 2.15 are compared. Nevertheless, it is important to consistently use the correspondence rule. For example, as discussed in the next chapter, since the Wigner function is the Weyl transform of the density matrix, Eq. 2.15 should be used in obtaining the Wigner equation of motion.

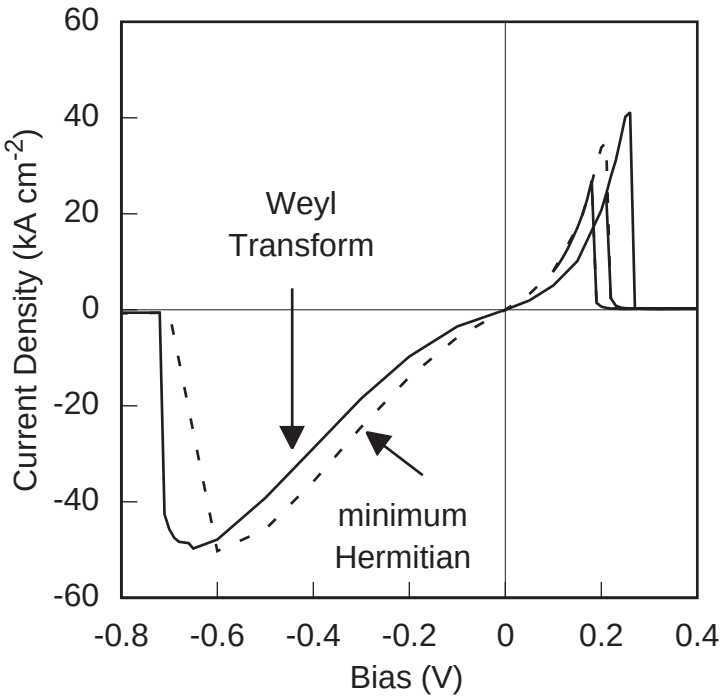

Figure 2.3: Self-consistently calculated conduction curves of a double barrier diode with asymmetric anode and cathode spacers. Under forward bias, the cathode spacer is $25nm$ lightly doped GaAs, and under reverse bias it is $5nm$ lightly doped GaAs. Results obtained using the "minimum Hermitian" Hamiltonian (dashed line) and the Weyl transform Hamiltonian (solid line) are not very different.

2.2 The Schrödinger equation on a discrete lattice

The common approach to a numerical solution of the Schrödinger equation proceeds by dividing the simulation domain into several regions. Using appropriate matching conditions for the envelope function at the interfaces between the different regions, transfer matrices are constructed (Ricco & Azbel, 1989; Pötz, 1989; Ohnishi et al., 1986; Cahay et al., 1987). The matching conditions follow directly from the form of the Hamiltonian. For example, the Hamiltonian 2.14 implies the continuity of ζ and ζ'/m^* across an interface. Given the boundary conditions, the transfer matrices for each mesh interval are cascaded to obtain the solution everywhere. This technique is numerically unstable when there are large regions of the device where the wavefunction is evanescent (Ting et al., 1992). Further, the matching conditions across interfaces are not easy to obtain for more complicated Hamiltonians.

A more suitable approach to modeling semiconductor devices is based on the tight-binding energy band. The tight-binding approach results in a set of linear algebraic equations for the envelope function at the lattice sites. The resulting problem is very similar to that obtained by finite-differencing the continuum Schrödinger equation. The Schrödinger equation on a tight-binding or finite-difference basis (Frensley, 1990), in conjunction with quantum transmitting boundary conditions (Lent & Kirkner, 1990) leads to a stable and more realistic numerical model.

Consider a position-independent tight-binding energy band in which only nearest neighbor interactions are assumed (Feynman et al., 1965)

$$\mathcal{E}(k) = \frac{\hbar^2}{m^*\Delta^2}\left[1 - \cos(k\Delta)\right]$$

Compared to a parabolic energy band that extends to infinity, the tight-binding result is more representative of the nature of energy bands in a semiconductor.

The energy band has the following Fourier expansion

$$\mathcal{E}(k) = \sum_R \mathcal{E}_{nR} e^{\imath k R} = \frac{\hbar^2}{2m^*\Delta^2}\left[-e^{-\imath k\Delta} + 2 - e^{\imath k\Delta}\right]$$

Assuming that Eq. 2.11 holds, in steady state Eq. 2.7 becomes

$$-\frac{\hbar^2}{2m^*}\left(\frac{\zeta_{j-1} - 2\zeta_j + \zeta_{j+1}}{\Delta^2}\right) + v_j\zeta_j = \mathcal{E}\zeta_j \tag{2.16}$$

which is just the second-order finite-difference approximation to Eq. 2.13. The finite-difference formulation to Eq. 2.13 is the exact equation of motion for a nearest-neighbor tight-binding energy band. This relation will be exploited in obtaining the Schrödinger equation when the bandstructure is spatially varying.

To obtain the effective mass equation when the bandstructure is spatially varying requires the evaluation of the Hamiltonian matrix elements of the kinetic energy operator in a basis of generalized Wannier functions (GWF) (Roblin & Muller, 1985; Roblin, 1988). Away from a heterojunction, the GWFs exponentially approach the bulk Wannier functions and the matrix elements can be evaluated easily. Near the heterojunction, however, it is difficult to evaluate the GWFs. Having recognized the connection between the effective-mass equations in the tight-binding and finite-difference bases, it is compelling to finite-difference Eq. 2.15 to obtain the matrix elements. On a uniform spatial mesh in the simulation region, the most consistent finite-difference approximation is as follows:

$$\begin{aligned}
\sum_{j=i-1}^{i+1} \mathcal{H}_{ij}\zeta_j &= \mathcal{E}\zeta_i \\
\mathcal{H}_{ii-1} &= -\frac{1}{4\Delta^2}\left(\frac{1}{m_i^*} + \frac{1}{m_{i-1}^*}\right) \\
\mathcal{H}_{ii} &= \frac{1}{4\Delta^2}\left(\frac{1}{2m_{i-1}^*} + \frac{3}{m_i^*} + \frac{1}{2m_{i+1}^*}\right) + v_i \\
\mathcal{H}_{ii+1} &= -\frac{1}{4\Delta^2}\left(\frac{1}{m_i^*} + \frac{1}{m_{i+1}^*}\right)
\end{aligned} \tag{2.17}$$

to $O(\Delta^2)$. Adjacent to a heterojunction, the diagonal elements obtained using the finite-difference formulation are weighted averages of the bulk diagonal elements. Two lattice constants away from a heterojunction, the diagonal elements are the bulk values. Examining Fig. 2.2, the explicit GWFs will probably yield a smoother transition across the heterojunction. If the "minimal Hermitian" Hamiltonian is assumed, the diagonal elements $\mathcal{H}_{ii}$ become (Frensley, 1990)

$$\mathcal{H}_{ii} = \frac{1}{4\Delta^2}\left(\frac{1}{m^*_{i-1}} + \frac{2}{m^*_i} + \frac{1}{m^*_{i+1}}\right) + v_j$$

If the connection to the finite-difference formulation is ignored, the best that can be done in the absence of explicit GWFs is to evaluate the diagonal matrix elements on either side of the heterojunction using the corresponding bulk Wannier functions. This gives

$$\mathcal{H}_{ii} = \frac{1}{m^*_i \Delta^2} + v_j$$

The non-diagonal elements near the interface remain to be determined. Several approaches have been suggested. For an interface between lattice cites i and $i+1$, Zhu and Kroemer (1983) suggest

$$\mathcal{H}_{ii+1} = -\frac{1}{2\Delta^2}\left(\frac{1}{m^*_i m^*_{i+1}}\right)^{1/2}$$

whereas Bastard (1981) suggests

$$\mathcal{H}_{ii+1} = -\frac{1}{4\Delta^2}\left(\frac{1}{m^*_i} + \frac{1}{m^*_{i+1}}\right)$$

identical to the result of the finite-difference formulation. In AlAs/GaAs heterostructures, the two approaches differ by less than 10%.

2.2.1 Boundary conditions

The set of equations 2.17 can be solved after specifying boundary conditions on ζ. The quantum transmitting boundary method (Lent & Kirkner, 1990)

is adopted here. The general solution to Schrödinger's equation outside the simulation domain $[0, L]$, where the potential is a constant, is

$$\zeta_j = \begin{cases} A_1\gamma_1^{j-1} + B_1\gamma_1^{1-j} & \text{for } j \leq 1 \\ A_N\gamma_N^{N-j} + B_N\gamma_N^{j-N} & \text{for } j \geq N \end{cases} \tag{2.18}$$

where N is the number of mesh nodes in $[0, L]$. A_1 and A_N are the amplitudes of the incoming wave components and B_1 and B_N are those of outgoing wave components at the left and right boundaries respectively. γ is the propagation factor $e^{\imath k\Delta}$. Introducing additional mesh points, $j = 0$ and $j = N + 1$, the solutions at the boundaries are:

$$\begin{aligned} \zeta_0 &= A_1\gamma_1^{-1} + B_1\gamma_1 \\ \zeta_1 &= A_1 + B_1 \\ \zeta_N &= A_N + B_N \\ \zeta_{N+1} &= A_N\gamma_N^{-1} + B_N\gamma_N \end{aligned} \tag{2.19}$$

Eliminating B_1 and B_N we get

$$\begin{aligned} \frac{\zeta_0 - \zeta_1\gamma_1}{\left(\gamma_1^{-1} - \gamma_1\right)} &= A_1 \\ \frac{\zeta_{N+1} - \zeta_N\gamma_N}{\left(\gamma_N^{-1} - \gamma_N\right)} &= A_N \end{aligned} \tag{2.20}$$

These two equations are added to the set 2.17. To simulate electrons injected into the device from the contacts, set $A_1 = 1, A_N = 0$ or $A_1 = 0, A_N = 1$ for plane waves incident from the left or right boundaries respectively. γ_1 and γ_N are obtained by writing Schrödinger's equation at the boundary

$$\begin{aligned} \mathcal{E} &= \mathcal{H}_{10}\zeta_0 + \mathcal{H}_{11}\zeta_1 + \mathcal{H}_{12}\zeta_2 \\ \mathcal{E} &= \mathcal{H}_{NN-1}\zeta_{N-1} + \mathcal{H}_{NN}\zeta_N + \mathcal{H}_{NN+1}\zeta_{NN+1} \end{aligned} \tag{2.21}$$

which, for the left boundary, can be rewritten as

$$\begin{bmatrix} 0 & 1 \\ -\mathcal{H}_{12}^{-1}\mathcal{H}_{10} & -\mathcal{H}_{12}^{-1}(\mathcal{H}_{22} - \mathcal{E}) \end{bmatrix} \begin{bmatrix} \zeta_0 \\ \zeta_1 \end{bmatrix} = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$

In regions of constant potential, the Wannier (or tight-binding) coefficients satisfy the relation $\zeta_{j+1} = \gamma_1 \zeta_j$ and we have the following eigenvalue problem (Ting et al., 1992)

$$\begin{bmatrix} 0 & 1 \\ -\mathcal{H}_{12}^{-1}\mathcal{H}_{10} & -\mathcal{H}_{12}^{-1}(\mathcal{H}_{22} - \mathcal{E}) \end{bmatrix} \begin{bmatrix} \zeta_0 \\ \zeta_1 \end{bmatrix} = \gamma_1 \begin{bmatrix} \zeta_0 \\ \zeta_1 \end{bmatrix} \tag{2.22}$$

Therefore, at the two boundaries, the following quadratic equations must be solved

$$\begin{aligned} \mathcal{E} &= \mathcal{H}_{11} + \mathcal{H}_{12}\left(\gamma_1 + \gamma_1^{-1}\right) \\ \mathcal{E} &= \mathcal{H}_{NN} + \mathcal{H}_{NN+1}\left(\gamma_N + \gamma_N^{-1}\right) \end{aligned} \tag{2.23}$$

The incoming wave components are described by the solutions of the quadratic equations in 2.23 for which $\Im(\gamma) \geq 0$ and $|\gamma| \leq 1$. Equation 2.22 can be extended to treat more than nearest-neighbor interactions.

2.2.2 Resonances

Fundamental to the understanding of resonant-tunneling structures is their transmission spectrum. Peaks in the transmission probability are associated with quasi-bound states in the structure. Also important are the resonance widths which are directly related to the quasibound-state lifetimes which influence the high-frequency characteristics (Brown et al., 1989; Bahder et al., 1987). A study of the resonances in a heterostructure thus reveals some of its fundamental properties. Associated with the Schrödinger equation (augmented with the boundary conditions) is a nonlinear eigenvalue problem (Frensley, 1991):

$$\begin{bmatrix} \alpha_1 & \beta_1 & & & & & \\ \mathcal{H}_{10} & \mathcal{H}_{11} - \mathcal{E} & \mathcal{H}_{12} & & & & \\ & \mathcal{H}_{21} & \mathcal{H}_{22} - \mathcal{E} & \mathcal{H}_{23} & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \mathcal{H}_{NN-1} & \mathcal{H}_{NN} - \mathcal{E} & \mathcal{H}_{NN+1} & \\ & & & & \beta_n & \alpha_n \end{bmatrix} \begin{bmatrix} \zeta_0 \\ \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_N \\ \zeta_{N+1} \end{bmatrix} = 0$$

where α and β depend on energy and are due to the boundary conditions. The eigenspectrum is complex as a consequence of the transmitting boundary conditions. The eigenvalues can be written as $\mathcal{E} = E_n - \imath, _n$, where E_n is the n^{th} quasibound level and 2, $_n$ the corresponding resonance width. The resonance width is related to the quasibound-state lifetimes (or escape time) through the uncertainty principle $\tau = \hbar/, $, i.e., in the absence of any incoming waves, the wave function corresponding to a quasi-bound state decays with time constant $\hbar/, $. For a bound state, the resonance width is zero, the electron being bound by the potential. In the Breit-Wigner formalism, the quasibound states have a Lorentzian line shape $T(\mathcal{E}) = , ^2/[(\mathcal{E} - E_n)^2 + , ^2]$. Therefore a search for the poles of the transmission amplitude has been implemented to isolate the resonances $E - \imath, $. Quadratic interpolation and Brent's method (Press et al., 1992) is used, first to find the peak transmission in real energy E_n and then to isolate the singularity at $E - \imath, $. In Fig. 2.4 the variation of , $_0$ with barrier width is shown for a typical AlAs/GaAs double barrier quantum well structure. The position of the resonance E_0 is insensitive to the variation in barrier width.

2.2.3 Electron density

The electron density in the device is assumed to be comprised of independent contributions due to each contact. It is assumed that the electrons occupy the scattering states incident from each contact with a probability given by the Fermi-Dirac distribution at the respective contacts, and that no transitions between states occur in the interior of the device. The density matrix describing the state of the mixed quantum system is (Frensley, 1990; Cahay et al., 1987)

$$\begin{aligned} \rho(z,z') &= \int_{v_l}^{\infty} \frac{d\mathcal{E}}{2\pi\hbar s_l(\mathcal{E})} g(\mathcal{E}-\mu_l)\zeta^{l\to r}(z,\mathcal{E})\zeta^{*l\to r}(z',\mathcal{E}) \\ &\quad + \int_{v_r}^{\infty} \frac{d\mathcal{E}}{2\pi\hbar s_r(\mathcal{E})} g(\mathcal{E}-\mu_r)\zeta^{r\to l}(z,\mathcal{E})\zeta^{*r\to l}(z',\mathcal{E}) \end{aligned} \qquad (2.24)$$

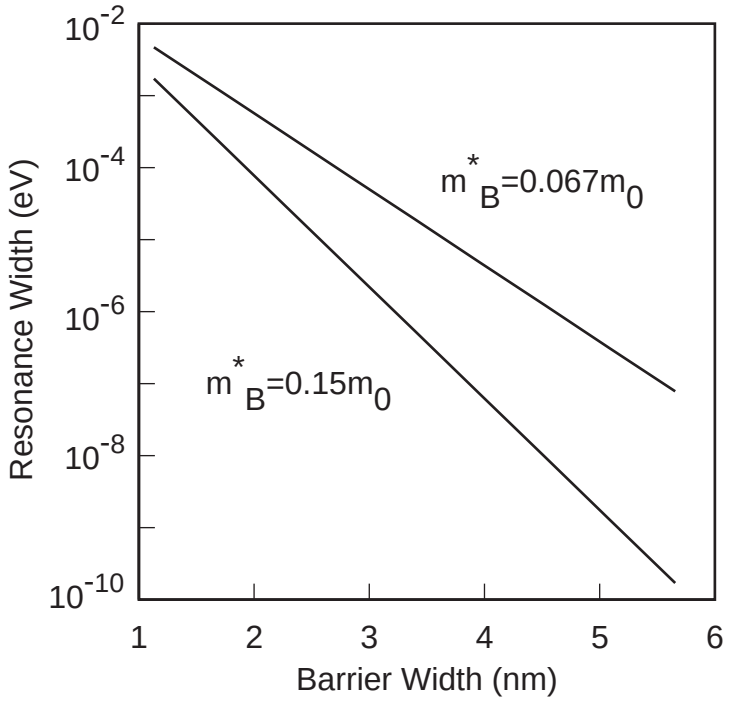

Figure 2.4: Dependence of the resonance width of the lowest resonance on the barrier thickness in a symmetric AlAs/GaAs double barrier structure under flatband and zero applied bias. The GaAs well is 5nm wide. The resonance is much narrower when the higher effective-mass in AlAs is included in the calculation.

where $v_{l,r}$ are the asymptotic potentials to the left and right, and $s_{l,r}$ is the velocity of an electron of energy $\mathcal{E}$ at the respective boundary. g is the Fermi-Dirac distribution function integrated over the transverse momenta and evaluated at the device boundaries:

$$g(\mathcal{E}) = \frac{m^* k_B T}{\pi \hbar^2} \ln\left(1 + e^{-\frac{\mathcal{E}}{k_B T}}\right) \tag{2.25}$$

The chemical potentials in the left and right boundaries $\mu_{l,r}$ are evaluated so that the boundary regions are charge-neutral. The superscript on the wave function ζ indicates whether it is a left or right incident scattering state. The electron density $n(z)$ is the diagonal element of the density matrix $\rho(z,z)$. It is appealing to consider $|\zeta(z)|^2$ as a modified density of states in the device. In a device, the modified density of states is a function of position and can be very different from that implied by the band structure alone.

2.2.4 Current density

The expression for the electron current carried by a solution ζ_j of the effective-mass equation can be obtained by writing a continuity equation for $\zeta^*\zeta$. First we write the time-dependent effective-mass equation and its complex conjugate on the discrete lattice:

$$\begin{aligned} \imath\hbar\frac{\partial \zeta_j}{\partial t} &= \mathcal{H}_{jj-1}\zeta_{j-1} + \mathcal{H}_{jj}\zeta_j + \mathcal{H}_{jj+1}\zeta_{j+1} \\ -\imath\hbar\frac{\partial \zeta_j^*}{\partial t} &= \mathcal{H}_{jj-1}\zeta_{j-1}^* + \mathcal{H}_{jj}\zeta_j^* + \mathcal{H}_{jj+1}\zeta_{j+1}^* \end{aligned} \tag{2.26}$$

Multiplying the first equation by ζ_j^* and the second by $-\zeta_j$ and adding the resulting equations we get

$$\frac{\partial \zeta_j^*\zeta_j}{\partial t} = -\frac{2}{\hbar}\left[\mathcal{H}_{jj+1}\Im(\zeta_j\zeta_{j+1}^*) - \mathcal{H}_{jj-1}\Im(\zeta_{j-1}\zeta_j^*)\right] \tag{2.27}$$

We can now define $J_{j+1/2}$, current density carried by ζ in the interval between nodes j and $j+1$:

$$J_{j+1/2} = \frac{2q\Delta}{\hbar}\mathcal{H}_{jj+1}\Im(\zeta_j\zeta_{j+1}^*) \tag{2.28}$$

The continuity equation for the probability density becomes:

$$\frac{\partial \zeta_j^* \zeta_j}{\partial t} = -\frac{J_{j+1/2} - J_{j-1/2}}{q\Delta} \tag{2.29}$$

The expression for current density is intimately related to the element $\mathcal{H}_{jj+1}$ that couples the lattice-cites j and $j+1$. The expression for current density can be easily verified in the absence of any external potential and effective-mass variations. Under steady state condition we have:

$$\begin{aligned} J_{j+1/2} &= \frac{2q\Delta}{\hbar}\frac{\hbar^2}{2m^*\Delta^2}\Im\left(e^{-\imath kj\Delta}e^{\imath k(j+1)\Delta}\right) \\ &= \frac{\hbar q}{m^*}\frac{\sin(k\Delta)}{\Delta} \end{aligned} \tag{2.30}$$

independent of position. In the limit $\Delta \to 0$, $J_{j+1/2} = \hbar qk/m^*$, an expected result.

The total current density due to contributions from the different scattering states in the system can be obtained as:

$$\begin{aligned} \mathcal{J}_{j+1/2} &= q\int_{v_l}^{\infty}\frac{d\mathcal{E}}{2\pi\hbar s_l(\mathcal{E})}g(\mathcal{E}-\mu_l)J_{j+1/2}^{l\to r}(\mathcal{E}) \\ &\quad - q\int_{v_r}^{\infty}\frac{d\mathcal{E}}{2\pi\hbar s_r(\mathcal{E})}g(\mathcal{E}-\mu_r)J_{j+1/2}^{r\to l}(\mathcal{E}) \end{aligned} \tag{2.31}$$

In steady state, an equivalent expression for current across the device, obtained by evaluating the above expression at the device boundaries, is

$$\mathcal{J} = \frac{q}{2\pi\hbar}\left[\int_{v_l}^{\infty} d\mathcal{E}\, g(\mathcal{E}-\mu_l)T^{l\to r}(\mathcal{E}) - \int_{v_l}^{\infty} d\mathcal{E}\, g(\mathcal{E}-\mu_r)T^{r\to l}(\mathcal{E})\right] \tag{2.32}$$

where

$$\begin{aligned} T^{l\to r} &= \frac{\sin(k_r\Delta)}{\sin(k_l\Delta)}B_N^2 \\ T^{r\to l} &= \frac{\sin(k_l\Delta)}{\sin(k_r\Delta)}B_1^2 \end{aligned} \tag{2.33}$$

are the probabilities of transmission of electron waves incident from the left and right respectively. k_l and k_r are the electron wave-vectors at the left and

right boundaries respectively. B_N and B_1 are obtained by eliminating A_1 and A_N from 2.19. Eq. 2.32 is just the Tsu-Esaki formula for current adapted to the present problem.

2.2.5 Self-consistent potential

The self-consistent potential is obtained by iteratively solving the Schrödinger and Poisson equations until a simultaneous solution is obtained. Näively solving Poisson's equation

$$\frac{d}{dz}\epsilon(z)\frac{d\,v_h^{k+1}(z)}{dz} = q^2\left[N_D^+(z) - n^k(z)\right] \tag{2.34}$$

at the $(k+1)^{st}$ iteration will not guarantee convergence if the initial potential and charge distribution are far from being self-consistent. If global charge neutrality is not satisfied at the k^{th} iteration, unphysically high electric fields can result at the device boundaries. This can lead to divergence of the potential as the iteration proceeds. To stably approach the self-consistent solution, the electron concentration in Poisson's equation at the $(k+1)^{st}$ step must be modified so that large changes $v_h^{k+1} - v_h^k$ are damped. To this end we define a local chemical potential

$$\mu^{k+1}(z) = v_h^k(z) + k_B T \ln\frac{n^k(z)}{N_c} \tag{2.35}$$

where N_c is a scaling factor which appears only temporarily in the calculation. Assuming Boltzmann statistics, we estimate the electron density

$$\begin{aligned} n^{k+1}(z) &= N_c e^{\frac{\mu^{k+1}(z) - v_h^{k+1}(z)}{k_B T}} \\ &= n^k(z) e^{\frac{v_h^k(z) - v_h^{k+1}(z)}{k_B T}} \end{aligned} \tag{2.36}$$

Using 2.36 we get the following non-linear Poisson equation (Venturi, 1989)

$$\frac{d}{dz}\epsilon(z)\frac{d\,v_h^{k+1}(z)}{dz} = q^2\left[N_D^+(z) - n^k(z)e^{\frac{v_h^k(z) - v_h^{k+1}(z)}{k_B T}}\right] \tag{2.37}$$

In an iterative scheme based on Eq. 2.37, changes in the potential between successive iterations are damped out. It is clear that the self-consistent solution satisfies Poisson's equation. Equation 2.37 can be solved by writing the equation for the difference $\delta v^{p+1} = v_h^{k+1,p+1} - v_h^{k+1,p}$, and using Newton iteration

$$\begin{aligned} \frac{d}{dz}\epsilon\frac{d\delta v^{p+1}}{dz} - \frac{q^2}{k_BT}n^k e^{\frac{v_h^k - v_h^{k+1,p}}{k_BT}}\delta v^{p+1} &= q^2\left[N_D^+ - n^k e^{\frac{v_h^k - v_h^{k+1,p}}{k_BT}}\right] \\ &\quad - \frac{d}{dz}\epsilon\frac{dv_h^{k+1,p}}{dz} \end{aligned} \tag{2.38}$$

along with the condition $\delta v = 0$ at the left and right boundaries. Finite-differencing Eq. 2.38 leads to a tri-diagonal system of equations

$$a_{ii-1}\delta v_{i-1}^{p+1} + a_{ii}\delta v_i^{p+1} + a_{ii+1}\delta v_{i+1}^{p+1} = r_i \tag{2.39}$$

where

$$\begin{aligned} r_i &= 2\Delta^2\left[N_D^+ - n_i^k e^{\frac{v_{h,i}^k - v_{h,i}^{k+1,p}}{k_BT}}\right] \\ &\quad - \left[a_{ii-1}v_{h,i-1}^{k+1,p} + a_{ii}v_{h,i}^{k+1,p} + a_{ii+1}v_{h,i}^{k+1,p}\right] \\ a_{ii} &= -\epsilon_{i-1} - 2\epsilon_i - \epsilon_{i+1} - \frac{2\Delta^2}{k_BT}n_i^k e^{\frac{v_{h,i}^k - v_{h,i}^{k+1,p}}{k_BT}} \\ a_{ii-1} &= a_{i-1i} = \epsilon_i + \epsilon_{i-1} \end{aligned} \tag{2.40}$$

The boundary conditions are $\delta v_1 = \delta v_N = 0$ at the first and last node of the spatial mesh.

2.3 $\Gamma - X$ bandstructure

There has been much interest in understanding the effect of the X-valley in the conduction of electrons through AlAs barriers. Clearly, if , $_{GaAs}$ electrons couple significantly to the X_{AlAs} states, the model discussed so far is not valid. Several experimental studies have indicated the importance of , $-X$ transport (Beresford et al., 1989; Bonnefoi et al., 1987; Feldman et al., 1990; Foster

et al., 1989; Mendez et al., 1987; Kyono et al., 1989). Models capable of comprehending , $-X$ transport include the work of Marsh (1987), Ting et al. (1992) and Chiang and Chang (1993).

The model presented here can be extended to include , $-X$ transport by expanding the bandstructure in the , $-X$ direction into its Fourier series

$$\mathcal{E}(k) = \sum_{n=1}^{\infty} \frac{4\hbar^2}{n^2 m_n^* a^2} \left[1 - \cos(\frac{nka}{2})\right] \quad |k| \leq \tfrac{2\pi}{a} \tag{2.41}$$

where $a = 5.6533$Å in GaAs and the Brillouin zone is $|k| \leq 2\pi/a$. The energy band has the following Fourier expansion

$$\mathcal{E}(k) = \sum_{R} \mathcal{E}_R e^{\imath kR} = \sum_{n=1}^{\infty} \frac{2\hbar^2}{n^2 m_n^* a^2} \left[-e^{-\imath nka/2} + 2 - e^{\imath nka/2}\right]$$

First consider a spatially uniform energy band. Using the Fourier components of the band structure in Eq. 2.7, we get the Schrödinger equation

$$-\sum_{n=1}^{\infty} \frac{\hbar^2}{2m_n^*} \left(\frac{\zeta_{j-n} - 2\zeta_j + \zeta_{j+n}}{(na/2)^2}\right) + v_j\zeta_j = \mathcal{E}\zeta_j \tag{2.42}$$

The boundary conditions are obtained by a generalization of the treatment in section 2.2.1 (Ting et al., 1992; Chiang & Chang, 1993).

Alternatively, a simpler scheme is obtained by assuming only nearest-neighbor interactions in the layer adjacent to the boundaries. In a transition region that is at least as long as the interaction range (specified by the number of Fourier components included), with every lattice site away from the boundary, the interaction range can be increased until the desired number of Fourier components is included. Outside this transition region, in the interior of the device, the full , $-X$ bandstructure can be included. To illustrate the approach, the matrix equation for two ($n = 1, 2$) Fourier components is given as

$$\begin{bmatrix} \alpha_1 & \beta_1 & & & & & \\ \mathcal{H}_{10} & \mathcal{H}_{11} & \mathcal{H}_{12} & & & & \\ \mathcal{H}_{20} & \mathcal{H}_{21} & \mathcal{H}_{22} & \mathcal{H}_{23} & \mathcal{H}_{24} & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & \mathcal{H}_{N-1N-3} & \mathcal{H}_{N-1N-2} & \mathcal{H}_{N-1N-1} & \mathcal{H}_{N-1N} - \mathcal{E} & \mathcal{H}_{N-1N+1} & \\ & & & \mathcal{H}_{NN-1} & \mathcal{H}_{NN} - \mathcal{E} & \mathcal{H}_{NN+1} & \\ & & & & \beta_n & \alpha_n & \end{bmatrix} \begin{bmatrix} \zeta_0 \\ \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_{N-1} \\ \zeta_N \\ \zeta_{N+1} \end{bmatrix} = \begin{bmatrix} A_1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ A_N \end{bmatrix}$$

where the matrix elements $\mathcal{H}_{ij}$ are obtained from Eq. 2.42. In this example, nearest-neighbor interaction is assumed in one layer adjacent to the boundary. For the case of spatially varying effective mass, Eq. 2.17 can be generalized to include n Fourier components of the bandstructure:

$$\begin{aligned}
\sum_{j=i-n}^{i+n} \mathcal{H}_{ij}\zeta_j &= \mathcal{E}\zeta_i \\
\mathcal{H}_{ii-m} &= -\frac{1}{m^2a^2}\left(\frac{1}{m^*_{m,i}} + \frac{1}{m^*_{m,i-m}}\right) \\
\mathcal{H}_{ii} &= \sum_{m=1}^{m=n} \frac{1}{m^2a^2}\left(\frac{1}{2m^*_{m,i-m}} + \frac{3}{m^*_{m,i}} + \frac{1}{2m^*_{m,i+m}}\right) + v_i \\
\mathcal{H}_{ii+m} &= -\frac{1}{m^2a^2}\left(\frac{1}{m^*_{m,i}} + \frac{1}{m^*_{m,i+m}}\right)
\end{aligned} \tag{2.43}$$

where $m^*_{m,i}$ refers to the effective-mass in the i^{th} layer corresponding to the m^{th} Fourier component. The current density $J_{j+1/2}$ is now defined as

$$J_{j+1/2} = \frac{qa}{\hbar}\sum_{m=1}^{m=n} \mathcal{H}_{jj+m}\Im(\zeta_j\zeta^*_{j+m}) \tag{2.44}$$

Of course Eq. 2.32 can still be used in the steady state since the bandstructure in the contact regions is determined by nearest-neighbor interactions. The , $-X$ coupling manifests itself by modifying the transmission spectrum. Implementation of a program incorporating the full , $-X$ bandstructure requires the inclusion of the additional matrix elements into the existing code, and calling a band diagonal solver instead of a tridiagonal solver. This is currently underway.

2.4 Resonant-tunneling diodes

Consider a resonant-tunneling diode consisting of two 17Å AlAs barriers sandwiching a 50Å wide quantum well. In a typical diode, on either side of the tunneling structure is a lightly doped spacer. Here we consider a three step spacer with 50Å lightly doped GaAs closest to the barriers, 100Å n-type ($5\times10^{16}\text{cm}^{-3}$)

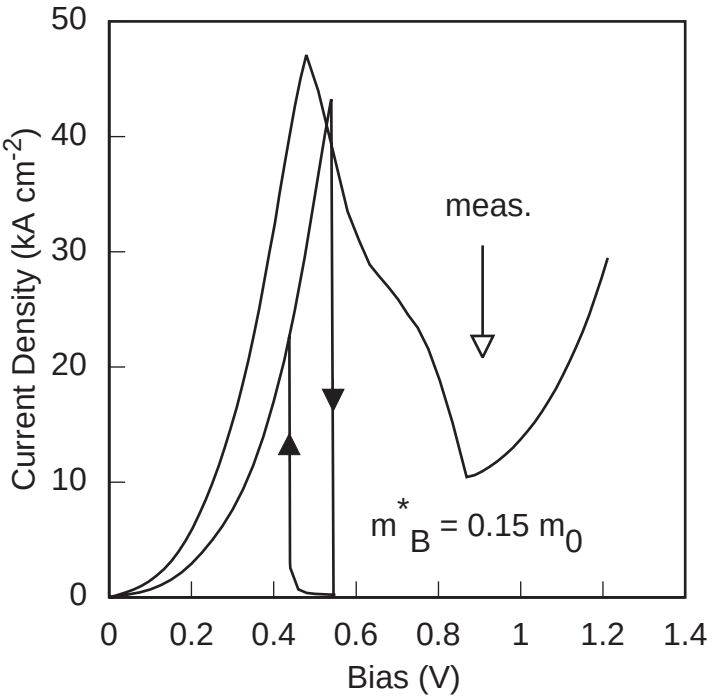

Figure 2.5: Room temperature experimental and calculated $J-V$ curve of a symmetric AlAs/GaAs double barrier diode.

GaAs, 100Å n-type ($6\times10^{17}\text{cm}^{-3}$) GaAs. The contact regions are heavily doped ($4\times10^{18}\text{cm}^{-3}$) GaAs.

The calculated self-consistent conduction ($J-V$) curve is shown in Fig. 2.5. Also shown is the experimental $J-V$ curve at 300K. While the agreement between the peak current densities is good, the valley current is grossly underestimated by the model. This is partly due to the fact that processes such as electron-phonon scattering are ignored in the model, and the nearest-neighbor treatment ignores effects such as , $-X$ tunneling. A more dramatic disagreement with experiment is the significant hysteresis in the calculated $J-V$ curve. That the Schrödinger model predicts hysteresis can be understood by considering the positions of the resonant levels relative to the energy of the virtual cathode. When the resonant level is above the virtual cathode, the resonant transmission is high and the current increases with in-

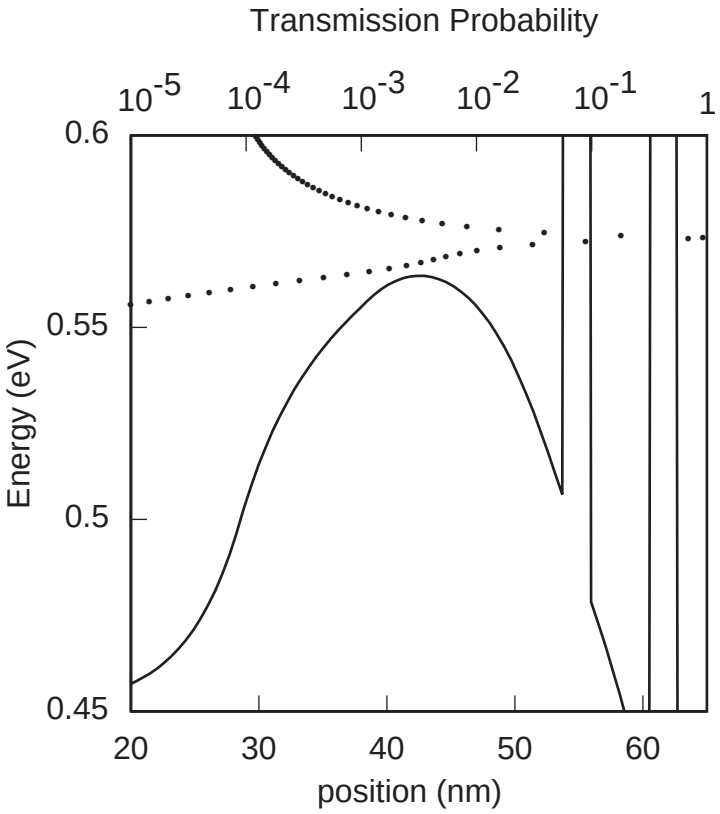

Figure 2.6: Cathode of the resonant-tunneling diode for an applied bias of 0.4508V in the high current state. Shown are the self-consistent potential and the corresponding transmission probability. In the high-current state the lowest resonance is above the virtual cathode potential.

creasing bias. Near the peak, however, two self-consistent solutions exist at a given voltage. In the high current state the resonant level is above the virtual cathode as shown in Fig. 2.6. In the low current state, the lowest resonant level falls below the virtual cathode. Once this happens, the coupling between the triangular well formed in the spacer layer and the main quantum well splits the resonance in two, both with severely degraded peak transmission. This is shown in Fig. 2.7.

The effect of the spacer layer width on the shape of the $J-V$ curve near the peak was illustrated in Fig. 2.3. The spacer layers in the simulated structure

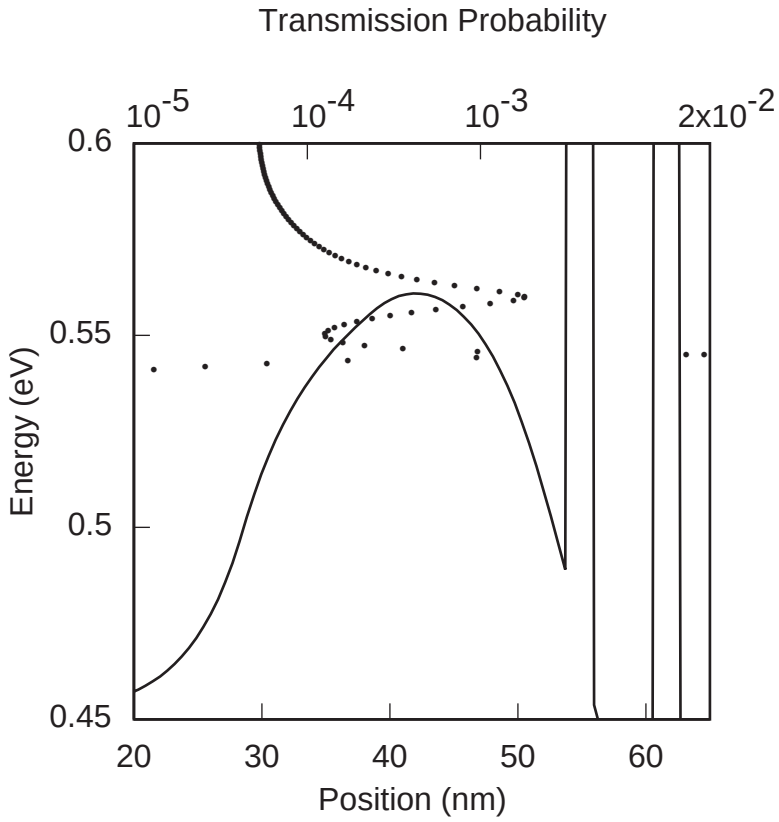

Figure 2.7: Cathode of the resonant-tunneling diode for an applied bias of 0.4508V in the low current state. Shown are the self-consistent potential and the corresponding transmission probability. The lowest resonance levels have fallen below the virtual cathode potential.

are asymmetric with the cathode spacer being thicker (250Å) in forward bias than in the reverse bias when it is only 50Å wide. While the peak is well rounded in reverse bias, for reasons already described, there is considerable hysteresis in forward bias. Although not clear from the figure, due to the higher potential barrier in the thicker spacer, the valley current in forward bias is less than half its value in reverse bias. Experimentally, resonant-tunneling diodes that incorporate about $200 - 300$Å thick, lightly doped spacers have higher peak-to-valley ratios compared to diodes without spacer layers (Huang et al., 1987; Tsao, 1993; Reddy, 1994).

The above discussion is relevant to the design of resonant-tunneling diodes as high power sources. Increasing the width of the anode-side spacer in the resonant tunneling diode, as in the quantum injection transit (QWITT) diode (Kesan et al., 1988), increases the negative differential resistance (NDR) compared to that obtained from a bare resonant tunneling diode. However, if the intrinsic diode does not exhibit NDR, the drift region serves no purpose. It will only increase hysteresis. For application in high power oscillators, it is therefore necessary to design structures that do not exhibit hysteresis. Clearly, the Schrödinger model appears inadequate for the purpose of designing QWITTs. In spite of this, with the hope that the model at least predicted trends, different spacer layer designs were studied to find one for which the model would not predict such a dramatic hysteresis. For example, from Fig. 2.3 it is clear that shortening the spacer layer effectively removes, or reduces the hysteresis. Within the scope of the Schrödinger model, investigating an unusual low-high-low doping profile in the spacer led to an intriguing memory switching phenomenon that has also been experimentally observed. This phenomenon is discussed in Chapter 5.

2.5 Summary

Exploiting the connection of the tight-binding approach to the finite-difference formulation, an effective-mass equation for the case of position-dependent band-structure has been presented. A new Hamiltonian for position dependent parabolic energy bands, based on the Weyl correspondence rule, was finite-differenced to yield the effective-mass equation for a spatially varying tight-binding energy band. It was shown that the form of the current density follows uniquely from the specification of the matrix elements of the Hamiltonian. The expression for the current density reduces to the conventional Tsu-Esaki formula in steady state. The treatment of the boundary conditions followed the work of Frensley (1991) and (Lent & Kirkner, 1990). The formulation has been been extended to include the full , $-X$ conduction band edge in the position representation.

While the effective-mass model is extremely useful in evaluating electronic properties related to resonant states in heterostructures, it precludes the treatment of important effects such as electron-phonon scattering. Electron-phonon interaction plays a very important role in semiclassical devices, and can also be important in quantum devices operating at room-temperature. This is clearly true when the region through which transport is quantum mechanical is small, the rest of the device being essentially semiclassical. An example is the QWITT diode (Kesan et al., 1988). Since the Schrödinger model is very different, in spirit, from traditional transport models such as those based on the Boltzmann equation, it is difficult to analyze such devices. For example, the Schrödinger model grossly underestimates the electron concentration in the drift region of QWITT diodes, thus greatly underestimating the space-charge resistance. This leads to the prediction of unphysically long optimum drift lengths.

Another problem with the Schrödinger equation is the difficulties it

presents in studying time-dependent phenomena. While the evolution of a single state can be simulated (Hellums & Frensley, 1994), the evolution of a system characterized by resonances requires that the initial state of the system must include an infinite basis of wavefunctions. Therefore, approaches based on individual wavefunctions must be abandoned, and a more general quantum kinetic approach, like the Boltzmann transport theory, be taken.

The need is for a model that can treat both quantum interference and electron-phonon scattering. Also the model should enable straightforward time-dependent simulations. Further, it is highly desirable that the model should be similar in spirit (or form) to the Boltzmann transport model. An approach that yields such a transport model is afforded by the Weyl transform and the associated Wigner function. A quantum transport model based on the Wigner function is presented in the next chapter.