Numerical study of electron tunneling through heterostructures

B. Méndez and F. Domínguez-Adame
Departamento de Física de Materiales, Facultad de Físicas, Universidad Complutense, E-28040 Madrid, Spain

(Received 11 February 1993; accepted 3 June 1993)

A numerical scheme based on the discretized form of the one-dimensional Schrödinger equation is presented. Using a transfer matrix method we recursively compute the transmission coefficient for electrons in arbitrary potentials. The computation time and storage are much reduced so that the code may be implemented by most programmable pocket calculators. The numerical method is used to study electron tunneling through single and double heterostructures, and the accuracy of the method is discussed.

I. INTRODUCTION

Tunneling through heterostructures plays a major role in the physics of many electronic devices.[1] In particular, resonant tunneling of electrons through quantum well structures has recently attracted considerable attention because of its possible application to ultrahigh speed devices.[2] Quantum dynamics of electrons in such systems has been widely studied within the framework of the effective mass approximation. From a theoretical point of view, one must solve the scattering problem at the semiconductor heterojunction, starting from a Schrödinger-like equation as an effective mass equation for electrons. The electronic motion parallel to the layers is that of free particles with an effective mass different from the actual electron mass. Perpendicular to the layers, neglecting band bending and space charges, the potential across a single junction is taken as a step potential, the height of the step being the offset of conduction bands in both semiconductors. Therefore, calculations actually reduce to a simple one-dimensional scattering problem.

The methods of solving the quantum mechanics of electron tunneling through heterostructures generally break down into two different approaches: the Wentzel–Kramers–Brillouin (WKB) approximation[3] and the transfer matrix approach.[4] The WKB solutions become valid whenever the potential barrier varies slowly compared with the electron wavelength, that is, the wavelength must be small compared with the distance over which the potential changes appreciably. This assumption is incorrect in devices of technical interest, namely, those with narrow layers, especially at low energy. On the other hand, the transfer matrix method requires the solution of the Schrödinger equation in each layer of the device, assuming that the

effective mass is constant. Two linearly independent solutions in each region and their derivatives are arranged into a 2×2 matrix. By matching the wavefunction and its derivative at each interface one can obtain the transmission coefficient simply by successive multiplication of these matrices. Although students of introductory courses in quantum mechanics can easily understand this method, mathematical difficulties may arise in dealing with realistic electronic devices. This occurs, for example, in the case of resonant tunneling through parabolic quantum wells, where solutions involve confluent hypergeometric functions.[5] Even if one considers the simple case of a single heterostructure and assumes that the potential is constant in each semiconductor layer (square well or barrier potential), analytic solutions are complicated when a bias voltage is applied because Airy functions appear. Hence, it would be desirable to find an alternative method to obtain accurate results without many mathematical tools.

The introduction of computers into elementary physics courses could provide invaluable help in understanding many different phenomena. Of particular interest are those schemes that teach students to use computers to solve the problems in the same way as do working physicists. The aim of this paper is to present a simple numerical method to study electron tunneling through arbitrary barriers, with application to semiconductor devices. The method is based on previous work of the authors,[6] in which a discretization scheme of the Schrödinger equation for general periodic potentials was presented to determine one-dimensional band structures. With minor modifications, this numerical method can be extended to study the electron scattering in general one-dimensional potentials. We feel that the numerical technique we present may be included in courses in solid-state physics as well as in courses in elementary quantum mechanics or nuclear physics (for instance, to study transmission through fission barriers): the mathematics is not too involved whereas the underlying ideas are easily understood with only some basic notions in quantum physics.

II. NUMERICAL METHOD

Let us consider an electron of energy E and effective mass m^*, moving in the x direction, incident on a heterostructure of width L. Regarding the application of results to electronic devices, we refer energies and potentials to the conduction-band edge of the bulk semiconductor. We suppose that the potential is constant outside the heterostructure with value zero at the left and $-V$ at the right when a bias voltage is applied. Also we assume that the effective mass m^* is constant throughout. The magnitude of the effective mass reflects the strength of the crystal potential and varies distinctly from material to material. In the case we shall deal with, however, the band structures of both semiconductors are very similar near the band edges and we ignore the difference in the effective mass. Hence the corresponding Schrödinger equation is written as

$$-\frac{\hbar^2}{2m^*}\frac{d^2\psi(x)}{dx^2}+V(x)\psi(x)=E\psi(x), \tag{1}$$

where the exact form of the potential $V(x)$ is immaterial for the moment and will be specified later.

In order to numerically solve this equation, let us take an integration step $s=L/N$, N being a positive integer, and define the grid points $x_n=ns$, with n an arbitrary integer. The Taylor expansion of the electron wave function at points $x_{n\pm1}$ yields

$$\psi(x_{n\pm1})=\psi(x_n)\pm s\frac{d\psi(x_n)}{dx}+\frac{s^2}{2}\frac{d^2\psi(x_n)}{dx^2}\pm\frac{s^3}{6}\frac{d^3\psi(x_n)}{dx^3}+O(s^4). \tag{2}$$

On adding $\psi(x_{n+1})$ and $\psi(x_{n-1})$ we obtain

$$\frac{d^2\psi(x_n)}{dx^2}=\frac{1}{s^2}[\psi(x_{n+1})+\psi(x_{n-1})-2\psi(x_n)]+O(s^2). \tag{3}$$

Neglecting the last term in Eq. (3), provided that s is small enough, we obtain from Eq. (1) the discretized form of the Schrödinger equation at any point of the grid

$$\psi(x_{n+1})+\psi(x_{n-1})=\alpha_n\psi(x_n), \tag{4}$$

where for simplicity we have defined

$$\alpha_n\equiv2+\frac{2m^*}{\hbar^2}[V(x_n)-E]s^2. \tag{5}$$

For an electron incident from the left, the wave function outside the heterostructure is of the form

$$\psi(x_n)=\begin{cases}e^{iqn}+re^{-iqn} & n\leqslant0\\ te^{iq'n} & n\geqslant N\end{cases}, \tag{6a}$$

where, for small s,

$$q\equiv\sqrt{\frac{2m^*Es^2}{\hbar^2}},\quad q'\equiv\sqrt{\frac{2m^*(E+V)s^2}{\hbar^2}}. \tag{6b}$$

As usual in one-dimensional scattering problems,[7] the wave function at the left is the superposition of the incident wave, whose amplitude is set to unity without loss of generality, and a reflected wave, whereas at the right there is only a transmitted wave. Here r and t denote the reflection and transmission amplitudes, respectively, so that the reflection and transmission coefficients are $\rho=|r|^2$ and $\tau=|t|^2\sin q'/\sin q$, respectively.

Notice that Eq. (4) is similar to a tight-binding Hamiltonian. By this we mean a discrete Hamiltonian that relates the electronic wave function at three consecutive lattice sites, so the corresponding Hamiltonian matrix is tridiagonal (with identical off-diagonal elements in our case). Hence one can find τ by means of the transfer matrix technique, in analogy to usual tight-binding calculations. At this point we should stress the transfer matrix formalism we use is similar, although not identical, to that followed by Tsu and Esaki.[4] These authors solve Eq. (1) analytically for an unbiased heterostructure so the (continuous) wave function are combination of plane waves in every semiconductor layer. The transfer matrix then relates the complex amplitudes of those plane waves at both ends of the heterostructure (r and t in our notation), and it is obtained as the product of several complex matrices. In contrast, the transfer matrix we use relates the value of the wave function at different grid points, as we shall see below. It presents the great advantage that its elements are all real instead of complex, thus requiring no advanced computing techniques (in fact, we have obtained our results with a BASIC code).

Table I. Comparison between the transmission coefficient calculated numerically for different numbers of grid points with exact and WKB results in a single GaAs–$Ga_{1-x}Al_xAs$ heterostructure. The potential barrier is 0.25 eV height, 50 Å width and the energy of the incident electron is 50 meV. All values $\times 10^{-3}$.

$N=10$	$N=100$	$N=500$	Exact	WKB
7.0278	6.7780	6.7756	6.7755	2.6506

Let us return to the discrete Schrödinger equation. Equation (4) may be cast in a matrix form

$$\begin{pmatrix}\psi(x_{n+1})\\ \psi(x_n)\end{pmatrix}=\begin{pmatrix}\alpha_n & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\psi(x_n)\\ \psi(x_{n-1})\end{pmatrix}\equiv P_n\begin{pmatrix}\psi(x_n)\\ \psi(x_{n-1})\end{pmatrix}, \tag{7}$$

and iterating this equation

$$\begin{pmatrix}\psi(x_{N+1})\\ \psi(x_N\end{pmatrix}=P_N\cdots P_0\begin{pmatrix}\psi(x_0)\\ \psi(x_{-1})\end{pmatrix}\equiv T(N)\begin{pmatrix}\psi(x_0)\\ \psi(x_{-1})\end{pmatrix}. \tag{8}$$

$T(N)$ is the transfer matrix of the whole system, and relates the wave function at both edges of the heterostructure. Two important properties will be used below, namely, $T(N)$ is real and $\det[T(N)]=1$. Using Eqs. (6) and (8) one gets after a little algebra

$$\rho=\rho(E)=\frac{|T_{12}e^{-iq}+T_{11}-T_{21}e^{iq'}-T_{22}e^{i(q'-q)}|^2}{|T_{11}e^{-iq}+T_{12}-T_{22}e^{iq'}-T_{21}e^{i(q'-q)}|^2}, \tag{9}$$

where we have dropped the explicit dependence on N of the transfer matrix elements. The denominator of Eq. (9) is obtained from the numerator replacing $T_{11}\leftrightarrow T_{12}$ and $T_{21}\leftrightarrow T_{22}$. Consequently we must only evaluate the numerator (say). Once ρ is explicitly evaluated (recall that T is real), taking into account the unitarity condition $\tau=1-\rho$ and the fact that $\det(T)=1$ we finally arrive at the following expression for the transmission coefficient

$$\tau=\tau(E)=\frac{4\sin q\sin q'}{D(E)}, \tag{10}$$

where $D(E)$ is the denominator of (9)

$$D(E)=T_{11}^2+T_{12}^2+T_{21}^2+T_{22}^2+2(T_{11}T_{12}+T_{21}T_{22})\cos q-2(T_{11}T_{21}+T_{12}T_{22})\cos q'$$
$$-2(T_{11}T_{22}+T_{12}T_{21})\cos q\cos q'+2\sin q\sin q'. \tag{11}$$

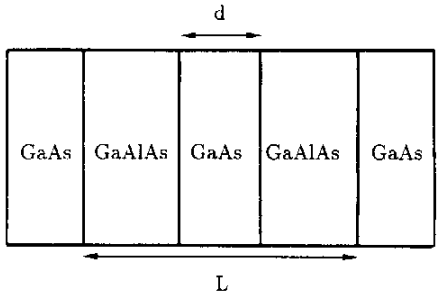

Fig. 1. Schematic representation of a double barrier GaAs–GaAlAs heterostructure.

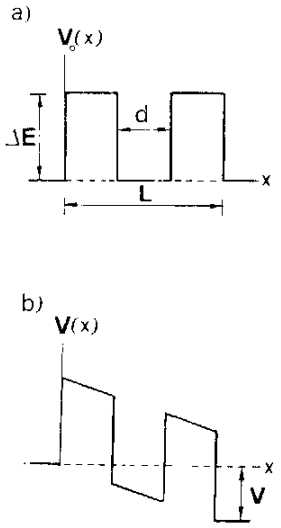

Fig. 2. Potential energy diagram in the double barrier heterostructure: (a) unbiased and (b) under bias conditions.

Hence τ is recursively computed from the matrix elements of $T(N)$. Taking into account the fact that $T(N)=P_NT(N-1)$ and $T(0)=P_0$ we find the following recurrence relations involving only real parameters:

$$\begin{aligned}T_{11}(N)&=\alpha_NT_{11}(N-1)-T_{11}(N-2),\\ T_{12}(N)&=\alpha_NT_{12}(N-1)-T_{12}(N-2),\\ T_{21}(N)&=T_{11}(N-1),\\ T_{22}(N)&=T_{12}(N-1),\quad N=1,2,...\end{aligned} \tag{12}$$

with the initial conditions $T_{ij}(-1)=\delta_{ij}$, $T_{11}(0)=\alpha_1$, $T_{12}(0)=-1$, $T_{21}(0)=1$ and $T_{22}(0)=0$. It is worth mentioning that the recurrence relations (12) involve at most three consecutive values of the transfer matrix elements, so no storage of former values is required. Therefore, the code may be implemented with small storage capabilities.

III. RESULTS AND DISCUSSIONS

In order to facilitate a direct comparison of our numerical procedure with analytical results, we first consider an unbiased ($V=0$) single GaAs–$Ga_{1-x}Al_xAs$ heterostructure of $L=50$ Å width. In such a case $m^*/m=0.067$ (see Ref. 1, p. 42) so that $2m^*/\hbar^2=0.0176$ eV^{-1} $Å^{-2}$. The conduction-band *offset* (the difference between the conduction-band edges in $Ga_{1-x}Al_xAs$ and GaAs) is about $\Delta E=0.25$ eV (positive) for 30% Al content. Therefore, the potential is simply a square barrier of the form

$$V(x)=\begin{cases}\Delta E, & 0<x<L\\ 0, & \text{otherwise}\end{cases}. \tag{13}$$

In this simple structure the transmission coefficient may be calculated analytically. For an incident electron with energy below the barrier height we get

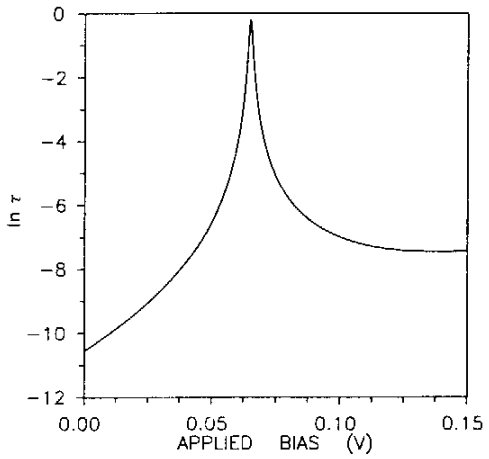

Fig. 3. Transmission coefficient obtained numerically as a function of the applied voltage in a double barrier heterostructure. The potential barrier is 0.25 eV height, $L=3d=150$ Å and the energy of the incident electron is 50 meV.

$$\tau_{\mathrm{ANA}}=\left(1+\frac{(\Delta E)^2\sinh^2[L\sqrt{2m^*(\Delta E-E)/\hbar^2}]}{4E(\Delta E-E)}\right)^{-1}. \tag{14}$$

In addition, the WKB method gives the expression

$$\tau_{\mathrm{WKB}}=\exp[-2L\sqrt{2m^*(\Delta E-E)/\hbar^2}]. \tag{15}$$

We have found in our numerical studies that good accuracy is obtained even for a small number of grid points ($N\sim 100$). Comparison of the transmission coefficient obtained numerically, τ_{NUM}, for different values of N with τ_{ANA} and τ_{WKB} is presented in Table I, when the energy of the incident electron is 50 meV above the GaAs conduction-band edge. Note that for N as low as 100 (s of order of the Bohr radius) one obtains three exact figures, whereas the WKB prediction is rather poor.

Having tested the reliability of the numerical method, let us study a second example. We consider a double barrier GaAs–$Ga_{1-x}Al_xAs$ heterostructure, as depicted in Fig. 1. Resonant tunneling through the double barrier occurs when the energy of the incident electron coincides with that of an unoccupied discrete state in the well (GaAs) between two confining barriers ($Ga_{1-x}Al_xAs$). Calling L the total width of the device and d the width of the well, we assume the potential to be of the form [see Fig. 2(a)]

$$V_0(x)=\begin{cases}\Delta E, & 0<x<(L-d)/2 \text{ or } (L+d)/2<x<L\\ 0, & \text{otherwise}\end{cases}, \tag{16}$$

in the unbiased device. The energy of the incident electron is nearly the Fermi energy, which is constant in the device. Therefore, to obtain a resonant tunneling current one must modify the energy of the discrete states in the well. A simple way to do that is to apply a bias voltage. This voltage causes the occurrence of an electric field in the device. Assuming that this field is constant throughout, the potential is now given by [see Fig. 2(b)]

$$V(x)=\begin{cases}0, & x<0\\ V_0(x)-(x/L)V, & 0<x<L,\\ -V, & x>L\end{cases} \tag{17}$$

where $V_0(x)$ is the potential in the unbiased device and $V>0$ is the applied bias.

We have numerically studied a double barrier heterostructure with $L=3d=150$ Å, $\Delta E=0.25$ eV and $m^*/m=0.067$. Using the numerical method, we have determined $\tau(E)$ for different values of the applied voltage V and $E=50$ meV. The obtained results are shown in Fig. 3 for $N=1000$. We observe the occurrence of a very narrow resonance at $V=65.1$ mV. As mentioned above, the external electric field modifies the quasilevels in the well, and their energies are lowered as the applied voltage is increased. The strong peak in the transmission coefficient then indicates that the Fermi level matches one of these quasilevels. The tunneling current will be high enough only if this matching occurs. For instance, increasing the voltage from 50 to 65.1 meV the transmission coefficient increases about three orders of magnitude. The full width at half maximum of the peak is only $\Gamma=1.5$ mV. This means that the resonant tunneling current is very sensitive to minor changes of the applied voltage. The value of Γ allows us to determine the lifetime of the quasilevel in the well. Assuming that this lifetime is not strongly dependent on the position of the level (so that it remains almost unchanged under minor variations of the applied voltage), it is clear that the width of the resonance peak in the transmission coefficient is approximately equal to the width of the level. Hence the lifetime is found to be $\approx\hbar/\Gamma=10^{-14}$ s. Therefore if initially there are no electrons in the well and electrons are made to tunnel by the applied voltage, only after a time of order of a few $\hbar/\Gamma$ the high transmission at resonance is achieved by multiple reflection. Thus this time sets an upper limit to achievable frequencies in the ac regime. This theoretical limit is very high and shows the capabilities of resonant heterostructures as ultrahigh speed devices. It is clear, however, that in a real device impurity scattering, space charge and many other phenomena reduce this upper limit. In fact, better estimates indicate that such *cutoff* frequency is about 60 GHz.[2] Fortunately this frequency is still high for technical applications!

IV. CONCLUSIONS

In this paper we have presented a numerical scheme to evaluate scattering parameters (transmission coefficient) of an electron in arbitrary one-dimensional potentials. The method is based on the analogy between the discretized form of the Schrödinger equation and tight-binding Hamiltonians. According to this analogy, a transfer matrix technique is used to compute the transmission coefficient for each energy in a recursive way. The method has been applied to several semiconductor heterostructures, and it has been found that good numerical accuracy is obtained even for not very small integration steps. The advantage of this method is that one can include more sophisticated and more realistic potentials, for which analytical solutions are difficult or impossible to obtain. In this sense, the effects of band bending at interfaces, inhomogeneous electric fields and different effective masses in both semiconductors can be very easily included in our numerical approach.

[1]M. Jaros, *Physics and Application of Semiconductor Microstructures* (Clarendon, Oxford, 1989), p. 190.
[2]F. Capasso and R. A. Kiehl, "Resonant tunneling transistor with quantum well base and high-energy injection: A negative differential resistance device," J. Appl. Phys. **58**, 1366–1368 (1985).
[3]J. Heremans, D. L. Partin, P. D. Dresselhaus, and B. Lax, "Tunneling through narrow-gap semiconductor barriers," Appl. Phys. Lett. **48**, 644–646 (1986).
[4]R. Tsu and L. Esaki,"Tunneling in a finite superlattice," Appl. Phys. Lett. **22**, 562–564 (1973).
[5]H. Cruz, A. Hernández-Cabrera, and A. Muñoz, "Resonant tunneling of electrons through parabolic quantum wells: An analytical calculation of the transmission coefficient," Semicond. Sci. Technol. **6**, 218–222 (1991).
[6]B. Méndez, F. Domínguez-Adame, and E. Maciá, "A transfer matrix method for the determination of one-dimensional band structures," J. Phys. A.: Math. Gen. **26**, 171–177 (1993).
[7]F. Schwabl, *Quantum Mechanics* (Springer, New York, 1992), p. 56.