

Discretization techniques for the efficient solution of the eigenvalue problem in heterostructures

Elias Kougianos^{1,*}, Saraju P Mohanty²

¹ *Dept of Engineering Technology, University of North Texas, Denton, TX 76203, USA. (Phone: +1-940-891-6708, Fax: +1-940-565-2666, Email: eliask@unt.edu)*

² *Dept of Computer Science and Engineering, University of North Texas, Denton, TX 76203, USA. (Phone: +1-940-565-3276, Fax: +1-940-565-2799, Email: smohanty@cse.unt.edu)*

SUMMARY

A systematic development of efficient discretization schemes for the numerical evaluation of the eigenvalues of the Single Band Effective Mass Equation (SBEME) that describes the motion of electrons in an ideal periodic crystal is presented. The approach presented makes use of the translational symmetry of the crystal lattice and utilizes the quantum mechanical properties of the momentum operator as the generator of spatial translation. Boundary conditions satisfied at the heterointerfaces are explicitly incorporated in the discretization procedure and the effects of this approach in overall accuracy are evaluated by studying a prototype quantum well heterostructure. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: Effective Mass Equation, Discretization, Eigenvalue Problem, Heterostructures

1. Introduction

There are numerous classes of electronic and optoelectronic devices that rely on the formation of structures from dissimilar semiconducting, insulating and metallic materials. Common examples are Resonant Tunneling Diodes (RTDs), Heterojunction Bipolar Transistors (HBTs), High-Electron Mobility Transistors (HEMTs), Quantum Wires, and Quantum Dots. In all of these devices, the spatial configuration of the constituent materials frequently translates to the formation of quantum wells of various shapes with rectangular and triangular wells being the most common.

Any type of analysis that attempts to describe the macroscopic behavior of these devices needs to take the presence and properties of these quantum wells into account. Even though several analytical approximation techniques have been developed in the past, oftentimes it is desirable to be able to simulate these devices via computer, which necessitates the development of efficient numerical techniques for the solution of the resulting equations.

*Correspondence to: Dept of Engineering Technology, University of North Texas, Denton, TX 76203, USA. (Phone: +1-940-891-6708, Fax: +1-940-565-2666, Email: eliask@unt.edu)

In the case of quantum wells formed in heterostructures, the starting point for the formulation and solution of these equations is often the so-called Single-Band Effective (SBEM) Equation [1], which describes an the equation of motion satisfied by the envelope function of individual electrons.

Our objective in this paper is to exhaustively investigate an operator based method via which the final equation that describes these electronic states is discretized and numerically solved. The result of the numerical procedure is the eigenvalues of the Hamiltonian in the SBEME which provide the energies of electrons bound in the quantum well.

Even though the following discussion is one-dimensional, the devices we consider (like the majority of devices on which we collect data) are three-dimensional. The nature of a heterostructure does not permit a complete separation of the longitudinal and transverse components in the describing equation. The effects of the transverse direction are incorporated via the concept of transverse electron energy, a procedure we have outlined in detail in Ref. [2].

We present the majority of this discussion within the framework of Dirac's bra-ket notation. A short summary of the notation conventions followed throughout the text is given in Appendix A.

2. Derivation of the SBEM Equation

In this section we review the Single-Band Effective Mass theory for bound and propagating states in an ideal periodic crystal.

2.1. Basic Considerations

In any physical system that must be treated in a quantum mechanically consistent way, three major parts in the development of the desired solution are identified:

- 1 The system itself and its surrounding environment must be properly identified and demarcated, which may not always be a trivial consideration.
- 2 Once the system is identified, the Hamiltonian operator \hat{H} for that particular system must be expressed as a function of the operators that correspond to the dynamical variables of the system. The Hamiltonian along with the boundary conditions imposed on the system, which in general will depend on the outcome of the analysis in part 1, define the appropriate Hilbert space that contains the states of the system.
- 3 The last, and typically most difficult stage, is the diagonalization of the Hamiltonian, i.e. the determination of a set of vectors that comprise a basis of the Hilbert space that contains the states of the system, in which basis the Hamiltonian has a diagonal matrix representation.

Once the Hamiltonian has been identified and diagonalized, any macroscopic property of the system can be calculated, in principle, from the wave function of the system. This wave function is a linear combination of the basis vectors determined in part 3 and depends entirely upon the initial conditions imposed on the system.

The device is considered to be one-dimensional, extending along the x -direction, while its lateral dimensions are considered to be very large, so as to exclude the appearance of any two-dimensional effects. In essence, for all practical calculations, we will consider the cross-sectional area A to approach infinity. It should be noted here that the device is not *truly* one-dimensional but rather three-dimensional, as any real device is, with an infinitely large cross-section thus enabling us to concentrate along the dimension (x) where changes occur.

Within the single-electron, single-band effective mass approximation, the SBEM Hamiltonian \hat{H} (ignoring spin effects) is:

$$\hat{H} = \frac{1}{2} \left(\hat{P} \frac{1}{\hat{M}^*(\hat{R})} \hat{P} \right) + \hat{E}_c(\hat{R}), \quad (1)$$

where \hat{P} and \hat{R} are the momentum and position operators, respectively, \hat{M}^* is the spatially varying effective mass operator, $\hat{E}_c(\hat{R})$ is the edge of the conduction band, and $\hat{U}_E(\hat{R})$ is the potential energy due to externally applied biases. $\hat{E}_c(\hat{R})$ is possibly modulated by an externally applied field, and given by: $\hat{E}_c(\hat{R}) = \hat{E}_{co}(\hat{R}) + \hat{U}_E(\hat{R})$. $\hat{E}_{co}(\hat{R})$ is the zero-bias conduction edge and is spatially varying for heterostructures and compositionally and doping varying semiconductor layers. $\hat{E}_{co}(\hat{R})$ is determined solely by the symmetry properties of the structure and is available for all semiconductors of interest.

Our choice for the Hamiltonian conforms with the so-called Ben Daniel - Duke Hamiltonian [3]. Several forms for the Hamiltonian have been suggested [4, 5, 6] but the Ben Daniel - Duke form is commonly used due to its simplicity and manifestly Hermitian properties [7].

Within the framework of the SBEM approximation, an electron in the crystal is described by an ‘‘envelope function’’ $|\Phi\rangle$ which modulates the Bloch functions for the conduction band and satisfies the Schrödinger-like SBEM equation [1]:

$$i\hbar \frac{\partial}{\partial t} |\Phi\rangle = \hat{H} |\Phi\rangle = \left[\frac{1}{2} \left(\hat{P} \frac{1}{\hat{M}^*(\hat{R})} \hat{P} \right) + \hat{E}_c(\hat{R}) \right] |\Phi\rangle. \quad (2)$$

Our main task, which is the primary contribution of this paper, is to discretize and determine numerically the eigenvalues of this Hamiltonian to facilitate computer-based simulation of heterostructure devices.

2.2. Matrix Form of the SBEM Equation

The form of the Single Band Effective Mass equation presented in Eq. (2) is three-dimensional. We will now decouple it into two components: one longitudinal, along the x direction and one transverse, along the yz plane. We expect all physically interesting quantities to vary along the longitudinal direction but we will see that the effect of the transverse properties of the device cannot be simply ignored.

We define the inverse effective mass operator as:

$$\hat{Q}(\hat{R}) = \frac{1}{\hat{M}^*(\hat{R})}. \quad (3)$$

Eq. (2) becomes now:

$$i\hbar \frac{\partial}{\partial t} |\Phi\rangle = \left[\frac{1}{2} \left(\hat{P} \cdot \hat{Q}(\hat{R}) \cdot \hat{P} \right) + \hat{E}_{co}(\hat{R}) + \hat{U}_E(\hat{R}) \right] |\Phi\rangle, \quad (4)$$

where the contributions to the conduction band edge $\hat{E}_c(\hat{R})$ due to an externally applied potential $\hat{U}_E(\hat{R})$ have been explicitly noted.

The momentum \hat{P} and position \hat{R} operators can be easily split into a direct sum of three orthogonal operators, each acting in a subspace of the complete Hilbert space and each being orthogonal to the other two. Thus we can write:

$$\hat{P} = \hat{P}_x \oplus \hat{P}_y \oplus \hat{P}_z, \quad (5)$$

$$\hat{R} = \hat{X} \oplus \hat{Y} \oplus \hat{Z}. \quad (6)$$

The inverse mass operator $\hat{Q}(\hat{\mathbf{R}})$, the zero-bias conduction band edge operator $\hat{E}_{co}(\hat{\mathbf{R}})$ and the external potential operator $\hat{U}_E(\hat{\mathbf{R}})$ are all functions of the position operator \hat{X} alone and do not depend on the other two position operators, \hat{Y} and \hat{Z} . We can thus write:

$$\hat{Q}(\hat{\mathbf{R}}) = \hat{Q}(\hat{X}) = \hat{Q}_x(\hat{X}) \oplus \hat{Q}_t(\hat{X}), \quad (7)$$

$$\hat{E}_{co}(\hat{\mathbf{R}}) = \hat{E}_{co}(\hat{X}), \quad (8)$$

$$\hat{U}_E(\hat{\mathbf{R}}) = \hat{U}_E(\hat{X}). \quad (9)$$

where \hat{Q}_x and \hat{Q}_t are the longitudinal and transverse components of the inverse effective mass. Eq. (4) can now be expressed as:

$$i\hbar \frac{\partial}{\partial t} |\Phi\rangle = \left[\frac{1}{2} (\hat{P}_x \oplus \hat{P}_y \oplus \hat{P}_z) \cdot \hat{Q}(\hat{X}) \cdot (\hat{P}_x \oplus \hat{P}_y \oplus \hat{P}_z) + \hat{E}_{co}(\hat{X}) + \hat{U}_E(\hat{X}) \right] |\Phi\rangle. \quad (10)$$

The presence of the operator $\hat{Q}(\hat{X})$ “sandwiched” between the orthogonal components of $\hat{\mathbf{P}}$ prevents us from fully decoupling Eq. (4). Using the well-known commutation relations obeyed by the momentum and position operators, as well as their functions [8]:

$$[\hat{P}_\alpha, \hat{P}_\beta] = 0 \quad \text{for any } \alpha, \beta, \quad (11)$$

$$[\hat{X}, \hat{Y}] = [\hat{Y}, \hat{Z}] = [\hat{Z}, \hat{X}] = 0, \quad (12)$$

$$[\hat{Q}(\hat{X}), \hat{P}_\alpha] = \begin{cases} 0 & \text{if } \alpha = y \text{ or } z, \\ i\hbar \frac{d}{d\hat{X}} \hat{Q}_x(\hat{X}) & \text{if } \alpha = x, \end{cases} \quad (13)$$

we can rewrite Eq. (4) as:

$$i\hbar \frac{\partial}{\partial t} |\Phi\rangle = \left\{ \left[\frac{1}{2} \hat{P}_x \cdot \hat{Q}_x(\hat{X}) \cdot \hat{P}_x + \hat{E}_{co}(\hat{X}) + \hat{U}_E(\hat{X}) \right] + \left[\frac{1}{2} \hat{Q}_t(\hat{X}) (\hat{P}_y^2 + \hat{P}_z^2) \right] \right\} |\Phi\rangle. \quad (14)$$

As mentioned earlier, this equation is non-separable because of the presence of the $\hat{Q}_t(\hat{X})$ factor in the second bracket. Since we will be studying steady-state problems, we will eliminate time derivatives using the substitution:

$$|\Phi\rangle = |\Psi\rangle e^{-iEt/\hbar}. \quad (15)$$

This yields the time-independent SBEM equation for stationary states:

$$\left\{ \left[\frac{1}{2} \hat{P}_x \cdot \hat{Q}_x(\hat{X}) \cdot \hat{P}_x + \hat{E}_{co}(\hat{X}) + \hat{U}_E(\hat{X}) \right] + \left[\frac{1}{2} \hat{Q}_t(\hat{X}) (\hat{P}_y^2 + \hat{P}_z^2) \right] \right\} |\Psi\rangle = E |\Psi\rangle. \quad (16)$$

We will now reason as in the Born-Oppenheimer approximation to decompose the electron envelope function as the tensor product of two envelope functions, one along the longitudinal direction and one along the transverse:

$$|\Psi\rangle \sim |\psi_x\rangle \otimes |\psi_{yz}(x)\rangle. \quad (17)$$

Because of the imperfect decoupling of the Hamiltonian in Eq. (16), this decomposition is itself imperfect. However, if we consider the transverse envelope function to change adiabatically (which

essentially means that $|\psi_x\rangle$ follows any changes in longitudinal effective mass $\hat{Q}_x(\hat{X})$ rapidly while $|\psi_{yz}(x)\rangle$ adapts to the changes in the transverse effective mass $\hat{Q}_t(\hat{X})$ much more slowly) we can “decouple” Eq. (16):

$$E_x(x)|\psi_x\rangle = \left[\frac{1}{2}\hat{P}_x \cdot \hat{Q}_x(\hat{X}) \cdot \hat{P}_x + \hat{E}_{co}(\hat{X}) + \hat{U}_E(\hat{X}) \right] |\psi_x\rangle, \quad (18)$$

$$E_{yz}(x)|\psi_{yz}(x)\rangle = \left[\frac{1}{2}\hat{Q}_t(\hat{X})(\hat{P}_y^2 + \hat{P}_z^2) \right] |\psi_{yz}(x)\rangle, \quad (19)$$

where $E_x(x) + E_{yz}(x) = E = \text{const.}$

These are *not* eigenvalue problems because neither $E_x(x)$ nor $E_{yz}(x)$ are constants. We can convert them to eigenvalue problems within the framework of the adiabatic approximation if we know the value of $E_{yz}(x) = \varepsilon_t$ at the injecting contact, i.e. the transverse energy with which electrons are injected at each contact.

Eq. (19) can be immediately recognized as a field-free wave equation. Its solution kets are characterized by the wave vectors k_y and k_z : $|\psi_{yz}(x)\rangle = |k_y, k_z\rangle$ where:

$$\langle yz | k_y, k_z \rangle = \frac{1}{\sqrt{A}} e^{i(k_y y + k_z z)}, \quad (20)$$

with A an arbitrary transverse normalizing area, and

$$\frac{\hbar^2}{2m_t^*(x)} (k_y^2 + k_z^2) = E_{yz}(x), \quad (21)$$

where $m_t^*(x)$ is the transverse effective mass. Since at the injecting contact:

$$\frac{\hbar^2}{2m_{ic}^*} (k_y^2 + k_z^2) = \varepsilon_t, \quad (22)$$

where m_{ic}^* is the transverse effective mass at the contact, we have:

$$E_{yz}(x) = \frac{m_{ic}^*}{m_t^*(x)} \varepsilon_t, \quad (23)$$

and

$$E_x(x) = E - E_{yz}(x) = E - \frac{m_{ic}^*}{m_t^*(x)} \varepsilon_t. \quad (24)$$

Substituting Eq. (24) back in Eq. (18), we have:

$$\left[\frac{1}{2}\hat{P}_x \cdot \hat{Q}_x(\hat{X}) \cdot \hat{P}_x + \hat{E}_{co}(\hat{X}) + \hat{U}_E(\hat{X}) + m_{ic}^* \varepsilon_t \hat{Q}_t(\hat{X}) \right] |\psi_x\rangle = E |\psi_x\rangle. \quad (25)$$

Now, this *is* an eigenvalue problem since $E = \text{const.}$

This form also shows the difference between a “true” one-dimensional problem and a one-dimensionalized three-dimensional problem. The additional term $m_{ic}^* \varepsilon_t \hat{Q}_t(\hat{X})$ does not appear in the former case.

For simplicity we reincorporate the effect of the external potential in the term for the overall conduction band edge $\hat{E}_c(\hat{X})$:

$$\left[\frac{1}{2}\hat{P}_x \cdot \hat{Q}_x(\hat{X}) \cdot \hat{P}_x + \hat{E}_c(\hat{X}) + m_{ic}^* \varepsilon_t \hat{Q}_t(\hat{X}) \right] |\psi_x\rangle = E |\psi_x\rangle. \quad (26)$$

Eq. (26) is just the SBEM equation, decoupled according to an adiabatic approximation. From now on, we need only concentrate on the x direction. The effect of the transverse direction (yz -plane) is incorporated in the transverse energy at the injecting contact, ε_t . For the moment we will assume that this energy is known and has a definite value.

Eq. (26) admits two kinds of solutions: one kind with continuous spectrum, characterized by values of E that satisfy the condition $E > E_c(ic)$, where $E_c(ic)$ is the value of the conduction band edge at the injecting contact and another kind with discrete spectrum. Discrete solutions might or might not exist depending on the particular form of the confining potential and whether or not it possesses local minima. The continuous solutions are characterized by the continuous wave vector $k_x: |\psi_x^{k_x}\rangle$, while the discrete solutions are characterized by distinct indices $i: |\psi_x^i\rangle$. The generic notation $|\psi_x\rangle$ will refer to solutions of Eq. (26) that could be either discrete or continuous.

Before considering the various kinds of solutions in detail, it will be instructive to rewrite Eq. (26) in a matrix form, thus allowing us to use any basis of the Hilbert space we find convenient. Almost exclusively in the remainder of this work, this basis will be that of the eigenkets of the position operator \hat{X} , $|x\rangle$. The ensuing formulation however remains independent of our basis choice and Eq. (26) itself is basis-independent.

We will assume that the eigenkets $|x\rangle$ form a complete and orthonormal basis:

$$\langle x|x'\rangle = \delta(x - x'), \quad (27)$$

$$\int |x\rangle\langle x| dx = \hat{1}, \quad (28)$$

where $\hat{1}$ is the unit operator. Multiplying Eq. (26) from the left with the bra $\langle x|$ and inserting unit operators as needed, the eigenvalue problem becomes:

$$\begin{aligned} E\langle x|\psi_x\rangle &= \frac{1}{2} \iiint \langle x|\hat{P}_x|x'\rangle \langle x'|\hat{Q}_x(\hat{X})|x''\rangle \langle x''|\hat{P}_x|x'''\rangle \langle x'''|\psi_x\rangle dx' dx'' dx''' \\ &+ \int \langle x|\hat{E}_c(\hat{X})|x'\rangle \langle x'|\psi_x\rangle dx' \\ &+ m_{ic}^* \varepsilon_t \int \langle x|\hat{Q}_t(\hat{X})|x'\rangle \langle x'|\psi_x\rangle dx', \end{aligned} \quad (29)$$

or, in terms of matrix elements:

$$\begin{aligned} E\psi_x(x) &= \frac{1}{2} \int \int \int p_x(x, x') q_x(x', x'') p_x(x'', x''') \psi_x(x''') dx' dx'' dx''' \\ &+ \int e_c(x, x') \psi_x(x') dx' \\ &+ m_{ic}^* \varepsilon_t \int q_t(x, x') \psi_x(x') dx'. \end{aligned} \quad (30)$$

These matrix elements are:

$$p_x(x, x') = \langle x|\hat{P}_x|x'\rangle \quad (31)$$

$$q_x(x, x') = \langle x|\hat{Q}_x|x'\rangle \quad (32)$$

$$q_t(x, x') = \langle x|\hat{Q}_t|x'\rangle \quad (33)$$

$$e_c(x, x') = \langle x|\hat{E}_c|x'\rangle \quad (34)$$

$$\psi_x(x) = \langle x|\psi_x\rangle, \quad (35)$$

Since \hat{Q}_x , \hat{Q}_t and \hat{E}_c are functions of \hat{X} alone, their matrix elements in the $|x\rangle$ representation are seen immediately to be:

$$q_x(x, x') = Q_x(x) \delta(x - x') \quad (36)$$

$$q_t(x, x') = Q_t(x) \delta(x - x') \quad (37)$$

$$e_c(x, x') = E_c(x) \delta(x - x'), \quad (38)$$

while the matrix elements of \hat{P}_x are [8]:

$$p_x(x, x') = \frac{\hbar}{i} \frac{d}{dx} \delta(x - x'). \quad (39)$$

The more familiar form of Eq. (26):

$$-\frac{\hbar^2}{2} \frac{d}{dx} \left(Q_x(x) \frac{d\psi_x}{dx} \right) + [E_c(x) + m_{ic}^* \varepsilon_t Q_t(x)] \psi_x(x) = E \psi_x(x), \quad (40)$$

is immediately recovered by using Eq. (37) - Eq. (39).

Finally we note the matrix form of Eq. (26):

$$E \Psi_x = \left[\frac{1}{2} P_x \cdot Q_x \cdot P_x + E_c + m_{ic}^* \varepsilon_t Q_t \right] \Psi_x. \quad (41)$$

The matrices that appear in this equation are infinite dimensional. We will proceed to reduce their dimensionality in the next section.

3. Bound States

In this section we present a discretization scheme that will allow us to examine the eigenvalue problem for bound states to a high level of accuracy and detail.

3.1. The Discretization of the SBEM Equation

In order to solve the eigenvalue problem for the discrete spectrum, we must convert the infinite dimensional matrix Eq. (41) to a finite dimensional one, suitable for solution by computer. This is achieved by converting the continuous set of eigenkets $|x\rangle$ to a discrete set. Out of all the possible kets $|x\rangle$, we select a finite number of them, say N , and attempt to use them as a basis of the Hilbert space. To qualify as a basis, this set of vectors must satisfy the equivalent of Eq. (27) - Eq. (28):

$$\langle x_i | x_j \rangle = \delta_{ij}, \quad (42)$$

$$\sum_{i=1}^N |x_i\rangle \langle x_i| = \hat{1}. \quad (43)$$

where δ_{ij} is the Kronecker Delta, defined by:

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \quad (44)$$

In general, these relations will *not* be satisfied due to the truncation of the basis. This is the inherent error source of such discretization schemes. These relations however are very useful in developing the finite dimensional representation of eigenvalue problems, if we recognize that there is an unavoidable source of error in them. Since we will not be explicitly concerned with an error analysis in this work, we will use them as if they were exact. Certain limits on the accuracy of a particular discretization scheme (DS) we will develop will however be established.

Using these relations, the discrete space equivalent of Eq. (29) becomes:

$$\begin{aligned} E\langle x_i|\psi_x\rangle &= \frac{1}{2} \sum_j \sum_k \sum_l \langle x_i|\hat{P}_x|x_j\rangle \langle x_j|\hat{Q}_x(\hat{X})|x_k\rangle \langle x_k|\hat{P}_x|x_l\rangle \langle x_l|\psi_x\rangle \\ &+ \sum_j \langle x_i|\hat{E}_c(\hat{X})|x_j\rangle \langle x_j|\psi_x\rangle \\ &+ m_{ic}^* \varepsilon_t \sum_j \langle x_i|\hat{Q}_t(\hat{X})|x_j\rangle \langle x_j|\psi_x\rangle. \end{aligned} \quad (45)$$

This is identical with Eq. (41), only in this case all the matrices are finite dimensional. Matrices \hat{P}_x , \hat{Q}_x , \hat{Q}_t and \hat{E}_c are of dimension $N \times N$ while Ψ_x is an $N \times 1$ vector. N is the number of grid points we use or, equivalently, the number of discrete basis vectors. In correspondence with Eq. (31) - Eq. (35):

$$p_x(i, j) = \langle x_i|\hat{P}_x|x_j\rangle, \quad (46)$$

$$q_x(i, j) = \langle x_i|\hat{Q}_x(\hat{X})|x_j\rangle, \quad (47)$$

$$q_t(i, j) = \langle x_i|\hat{Q}_t(\hat{X})|x_j\rangle, \quad (48)$$

$$e_c(i, j) = \langle x_i|\hat{E}_c(\hat{X})|x_j\rangle, \quad (49)$$

$$\psi(i) = \langle x_i|\psi_x\rangle. \quad (50)$$

Most of these are straightforward due to the fact that they are plain functions of x :

$$q_x(i, j) = Q_x(x_i)\delta_{ij}, \quad (51)$$

$$q_t(i, j) = Q_t(x_i)\delta_{ij}, \quad (52)$$

$$e_c(i, j) = E_c(x_i)\delta_{ij}, \quad (53)$$

$$\psi(i) = \psi_x(x_i). \quad (54)$$

The evaluation of $p_x(i, j)$ is however impossible to achieve exactly in a truncated basis. The physical reason behind this is that the action of \hat{P}_x in the $|x\rangle$ representation is equivalent to that of differentiation with respect to x . It is impossible to calculate exactly the derivative of a function at a given point with knowledge of the function only at a finite number of points with the possible exception of very simple functions.

We will present now the discretization scheme that will allow us to evaluate $p_x(i, j)$ with as much accuracy as we desire, up to the maximum accuracy allowed by the particular truncated basis at our disposal.

We employ a simple but powerful linear interpolation scheme to express the action of any power of the momentum operator on the $|x_i\rangle$ kets. Let us consider the ket $|x_j\rangle$. The action of the M -th power of \hat{P}_x on $|x_j\rangle$ is expressed as (see Fig. 1):

$$\hat{P}_x^M |x_j\rangle = \sum_{m=-K}^{m=L} \alpha_{j+m} |x_{j+m}\rangle, \quad (55)$$



Figure 1. Node labeling for the 1-D discretization scheme.

where the undetermined coefficients α are the weights of the interpolation scheme. Thus, we use the values not only at the j -th point, but the values at K points to the left of j and L points to the right of j , as well. We will call this DS a (K, L) DS. The choice of K and L is entirely arbitrary but must be guided by computational intensity considerations. The greater the value of (K, L) , the more accurate the results but the computational penalty might be heavy. A study of these considerations for a prototype potential will be presented in Section IV.

Our task is now to calculate the $(K + L + 1)$ coefficients α_{j+m} . To this purpose, we will take advantage of the properties of the linear momentum operator as the generator of the Lie group of infinitesimal translations [9], namely:

$$\begin{aligned}
 |x_{j+m}\rangle &= \exp\left[-\frac{i}{\hbar}(x_{j+m} - x_j)\hat{P}_x\right]|x_j\rangle \\
 &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n (x_{j+m} - x_j)^n \hat{P}_x^n |x_j\rangle \\
 &= \sum_{n=0}^{K+L} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n (x_{j+m} - x_j)^n \hat{P}_x^n |x_j\rangle \\
 &\quad + O\left((x_{j+m} - x_j)^{K+L}\right).
 \end{aligned} \tag{56}$$

Substituting Eq. (56) in Eq. (55) and projecting on the bra $\langle x_i|$, we have:

$$\begin{aligned}
 \langle x_i|\hat{P}_x^M|x_j\rangle &\simeq \\
 \sum_{n=0}^{K+L} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \sum_{m=-K}^{m=L} \alpha_{j+m} (x_{j+m} - x_j)^n \langle x_i|\hat{P}_x^n|x_j\rangle.
 \end{aligned} \tag{57}$$

This expression is accurate to $O\left((x_{j+m} - x_j)^{K+L}\right)$. Equating similar terms, we obtain the following

system of $(K + L + 1)$ equations:

$$\sum_{m=-K}^{m=L} \alpha_{j+m} = 0, \quad (58)$$

$$\sum_{m=-K}^{m=L} \alpha_{j+m} (x_{j+m} - x_j) = 0, \quad (59)$$

$$\dots$$

$$\sum_{m=-K}^{m=L} \alpha_{j+m} (x_{j+m} - x_j)^M = M! \left(-\frac{\hbar}{i}\right)^M, \quad (60)$$

$$\dots$$

$$\sum_{m=-K}^{m=L} \alpha_{j+m} (x_{j+m} - x_j)^{K+L} = 0. \quad (61)$$

This can be expressed in matrix form as:

$$\begin{pmatrix} 1 & \dots & 1 \\ (x_{j-K} - x_j) & \dots & (x_{j+L} - x_j) \\ \dots & \dots & \dots \\ (x_{j-K} - x_j)^M & \dots & (x_{j+L} - x_j)^M \\ \dots & \dots & \dots \\ (x_{j-K} - x_j)^{K+L} & \dots & (x_{j+L} - x_j)^{K+L} \end{pmatrix} \cdot \begin{pmatrix} \alpha_{j-K} \\ \alpha_{j-K+1} \\ \dots \\ \alpha_{j-K+M} \\ \dots \\ \alpha_{j+L} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ M! \left(-\frac{\hbar}{i}\right)^M \\ \dots \\ 0 \end{pmatrix}. \quad (62)$$

This type of system is of the well-known Vandermonde variety for the solution of which very efficient and robust algorithms exist [10]. It is worthwhile to note that the solution of this system of equations, consisting of a set of $(K + L + 1)$ numbers α_{j+m} actually represents a whole row (the j -th row) of the matrix representation of \hat{P}_x^M . This is easy to see from the following matrix element:

$$\langle x_i | \hat{P}_x^M | x_j \rangle = \sum_{m=-K}^{m=L} \alpha_{j+m} \langle x_i | x_{j+m} \rangle. \quad (63)$$

Taking into account the orthogonality properties of the basis $|x_i\rangle$, Eq. (42), we see that:

$$\langle x_i | \hat{P}_x^M | x_j \rangle = \begin{cases} 0 & \text{if } i < j - K \text{ or if } i > j + L, \\ \alpha_{j+m} & \text{if } j - K \leq i \leq j + L. \end{cases} \quad (64)$$

Thus it is sufficient to solve at most N systems of dimension $(K + L + 1)$ in order to obtain the N rows of matrix P_x . Once P_x is known, direct substitution of all the available matrices in Eq. (41) yields a matrix eigenvalue problem. Direct solution provides the set of bound state energies E_i and the corresponding eigenvectors $|\psi_x^i\rangle$. Specific examples of the application of this method will be presented in Section IV. The accuracy of this (K, L) discretization scheme is $O(h^{K+L})$, where $h = \max\{|x_{j-K} - x_j|, \dots, |x_{j+L} - x_j|\}$.

This method of discretizing equations from a continuous Hilbert space to a discrete functional space is a generalization of the well-known method of polynomial interpolation [11] and, in particular, the approach used by Lyness and Moler [12] and Ballester and Pereyra [13]. Because it is driven by the

physics of the particular problem at hand, as evidenced by the use of the appropriate operator, it is more general and flexible, and applicable to a wide variety of physical problems. For example, the above arguments can be directly used to discretize the radial Schrödinger's equation if we replace the linear momentum operator with the angular momentum operator.

Although the functional method of Refs. [12] and [13] applies to differential operators only, the generalization presented here can be used for any operator regardless of form. In fact, if the Hamiltonian under study is a function of operators which are a subset of the generators of a known Lie group, the previous arguments can be extended to more than one dimension. This situation will occur whenever the symmetry group of the Hamiltonian is a Lie group, a situation frequently encountered in the quantum theory of solids.

3.2. Treatment of Discontinuities

The presence of interfacial areas where the effective mass and the conduction band edge change rapidly, complicates significantly the discretization of the SBEM equation at points where these changes occur. The common approach in the literature has been to ignore the particular location of the interface and consider points only to the left and to the right of the discontinuity where there is no ambiguity as to the precise values of material parameters and envelope functions [14]. Between these well defined points, the material is assumed to change in an unspecified but smooth manner thus eliminating any problems related to uncertain specification of material properties. This particular point of view makes good physical sense when considering what happens at an atomic level at the interface, but is not compatible with the assumptions underlying much of the theory: namely that the interfaces are *abrupt* and not smooth.

It has been pointed out in Ref. [15] that a simple averaging procedure of the potential at abrupt discontinuities significantly improves the accuracy of calculated eigenvalues. In this section we will generalize these arguments and examine the consequences of this analysis.

Instead of simply accounting for the potential discontinuities as in Ref. [15], we must also take into account the abrupt change in effective mass. Referring to the discrete equivalent of Eq. (40), we replace the term $[E_c(i, i) + m_{ic}^* \varepsilon_t Q_t(i, i)] \psi_x(i)$ at the interface position x_i with an average. This average must extend by K points to the left and by L points to the right, otherwise the overall accuracy of the DS scheme will be reduced. Thus we have:

$$\begin{aligned} & \langle [E_c(i, i) + m_{ic}^* \varepsilon_t Q_t(i, i)] \psi(i) \rangle = \\ & \frac{1}{x_{i+L} - x_{i-K}} \int_{x_{i-K}}^{x_{i+L}} [E_c(x) + m_{ic}^* \varepsilon_t Q_t(x)] \psi(x) dx. \end{aligned} \quad (65)$$

Taking into account the fact that multiplication by the matrix P_x corresponds to differentiation with respect to x within a multiplicative factor, and hence multiplication by P_x^{-1} corresponds to integration with respect to x , this equation takes the following matrix form:

$$\begin{aligned} & \langle [E_c(i, i) + m_{ic}^* \varepsilon_t Q_t(i, i)] \psi(i) \rangle = \\ & \left[\frac{1}{x_{i+L} - x_{i-K}} \left(-\frac{\hbar}{i} \right) P_x^{-1} (E_c + m_{ic}^* \varepsilon_t Q_t) \Psi_x \right]_i. \end{aligned} \quad (66)$$

It is thus sufficient to replace the i -th row of the matrix:

$$\left[\frac{1}{2} P_x \cdot Q_x \cdot P_x + E_c + m_{ic}^* \varepsilon_t Q_t \right], \quad (67)$$

with the i -th row of the matrix:

$$\left[\frac{1}{2} \mathbf{P}_x \cdot \mathbf{Q}_x \cdot \mathbf{P}_x + \frac{1}{x_{i+L} - x_{i-K}} \begin{pmatrix} -\hbar \\ i \end{pmatrix} \mathbf{P}_x^{-1} (\mathbf{E}_c + m_{ic}^* \varepsilon_t \mathbf{Q}_t) \right], \quad (68)$$

where the matrix elements of \mathbf{Q}_x at the interfacial node are linearly interpolated from those at neighboring nodes.

This adjustment will not only assure proper handling of the interfaces but will also maintain the overall accuracy of the discretization scheme.

3.3. Modifications for III-V Materials

The basic equations presented in this section require minor modifications to account for second-order effects peculiar to III-V materials. Most prominent among them are the conduction band non-parabolicity, and multiple conduction band valleys ($\Gamma - L - X$). For heavily doped devices, bandgap narrowing must also be accounted for.

In GaAs, it has been shown that non-parabolicity of the Γ valley modifies the effective density of states [16]. However, the effect of a non-parabolic k^4 term in the energy band relation $E(\mathbf{k})$ is to produce high-order derivatives ($\nabla^2 \nabla^2$) in the effective mass theorem which complicate the discretization of Schrödinger's equation. The net effect is densely populated matrices with increased bandwidth and the solution of the eigenvalue problem becomes much more complex and time-consuming with the corresponding gain in accuracy of minimal value, particularly under bias conditions. For these reasons, non-parabolicity effects can be usually ignored. With the help of the discretization schemes developed earlier in this section, the resultant matrices can be calculated even when high order derivatives are involved since differentiation has been reduced to simple matrix multiplication. Thus, non-parabolicity can be taken into account if the need arises.

In materials like AlGaAs, the contribution of all three conduction band valleys ($\Gamma - L - X$) to electronic conduction might be included. This entails solving the Schrödinger-like SBEM equation separately for each band and adding the electron concentrations to obtain the overall concentration. As expected, this addition does increase execution time for the eigenproblem by a factor of 3. In the remainder of this work, we used only one conduction band statistics.

4. Bound States of a Prototype Potential Well

Rectangular potential wells are ideally suited for the study of bound states, mainly because they incorporate all the pertinent features of real devices but in addition their eigenproblems can be solved via simple transcendental equations to a high degree of accuracy. Thus the solution can be considered "quasi-analytical" and is quite useful for comparison with results obtained via other numerical techniques.

In particular, we will consider the rectangular well formed by the conduction band edge of a thin layer of GaAs positioned between two thick layers of $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$. Electrons confined in this well will be bound if their energy corresponds to the eigenvalues of the Hamiltonian operator for this potential profile. The parameters we will use for this prototype potential well will be chosen to be identical with those of Tan *et al.* (see footnote [14]), namely conduction band discontinuity of 0.23eV and electron effective mass $m^* = 0.067$ and $m^* = 0.092$ in the GaAs and $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ layers, respectively.

The geometry of the prototype well is represented in Fig. 2. The length of the well is L_W while the overall length of the device is L_x . We choose $L_x = 1000\text{\AA}$ while $L_W = 56\text{\AA}$. The restriction that

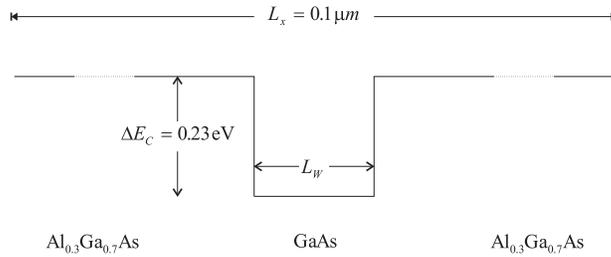


Figure 2. Prototype potential well geometry.

$L_w \ll L_x$ guarantees the validity of the boundary conditions ($\psi(0) = \psi(L_x) = 0$) at the ends of the device. Essentially this requirement is tantamount to $L_x \rightarrow \infty$.

4.1. Uniform vs. Non-uniform Discretization Grid

Careful consideration must be given to the selection of the grid points on which the SBEM equation will be discretized. Before any particular choice of points can be made, the fundamental form of the grid must be decided, i.e. whether the grid should be of uniform or variable spacing. Uniform grids are used extensively because of the ease with which they can be generated as well as the simple form of the resulting equations. It is considerably easier to generate the grid once and for all and not be forced to adjust its spacing throughout the simulation process. It has been however noted (see the paper by Tan *et al.* cited in footnote [14]) that the use of a non-uniform grid significantly improves the accuracy of the calculated eigenvalues.

Intuitively, it is to be expected that a grid that adjusts to the variations of the underlying potential will give more accurate results. There are however more compelling reasons to choose non-uniform grids. In the class of devices we will study, the quantum nature of the devices is manifested in relatively small areas. For example, in a narrow base HBT, the base region is of the order of $50 - 150 \text{ \AA}$. The quantum well formed at the heterointerface of the emitter-base junction is of comparable dimensions. The overall length of the device is however of the order of $2 - 3 \mu m$. A uniform grid of spacing 20 \AA (deemed the absolute maximum to even acknowledge the existence of the base region), would require at least 1000 points to cover the whole device. Beyond the waste of grid points in areas where the potential remains fairly constant, the resultant matrices are of extremely large dimensions, consuming considerable CPU time for their manipulation. Not only it takes orders of magnitude longer to solve these problems, but the results are not as accurate as those obtained via the use of an optimized, much smaller non-uniform grid.

In all subsequent discussions, it will be assumed that the grid is non-uniform and adjusted to the variations of the conduction band edge. For the well of Fig. 2, the potential profile remains constant and no adaptive grid re-allocation technique is required. No space charge effects will be taken into account (i.e. no self-consistency with Poisson's equation will be required.) This will allow us to concentrate on the effects of the discretization alone, unburdened by secondary phenomena. Of the total of N grid points, one third will be allocated for the well itself and one third each for the flat areas of the potential. This strategy guarantees that the physically interesting area (the well) is allocated a large amount of points while the non-essential areas receive a proportionally much smaller number of points.

Finally, in order to treat this as a true one-dimensional problem, it is assumed that electrons are

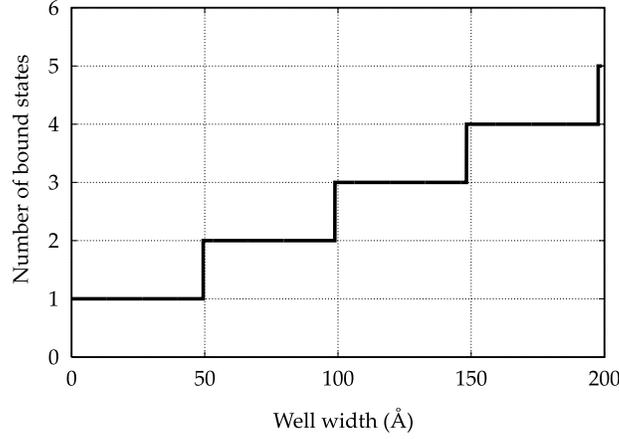


Figure 3. Number of bound states as a function of well width.

injected in the device with zero transverse energy ($\varepsilon_t = 0$).

4.2. Quasi-Analytical Results

This type of well has been studied in Ref. [17]. The number \mathfrak{N} of bound states the well admits is dependent upon the material parameters and the length and depth of the well via the relationship:

$$\mathfrak{N} = 1 + \text{Int} \left[\sqrt{\frac{2m_A^* L_W^2 |V_p|}{\pi^2 \hbar^2}} \right], \quad (69)$$

where m_A^* is the electron effective mass in GaAs and V_p is the depth of the well, 0.23 eV in our case. $\text{Int}(x)$ is the integer part of x . A plot of the number of bound states as a function of the well width L_W for constant $V_p = 0.23$ eV is shown in Fig. 3. This plot also relays information about the strength of the bound state. For example, for our case of $L_W = 56$ Å, we see that the well admits two bound states. Since L_W is very near the transition edge from one to two bound states (at approximately 50 Å), we expect the ground state to be strongly bound but the second state will be weakly bound. The exact values of the energies are given by the solution of the following transcendental equations:

$$\frac{k_w}{m_A^*} \tan \left(k_w \frac{L_W}{2} \right) = \frac{k_b}{m_B^*}, \quad (70)$$

$$\frac{k_w}{m_A^*} \cot \left(k_w \frac{L_W}{2} \right) = -\frac{k_b}{m_B^*}, \quad (71)$$

where m_B^* is the electron effective mass in the AlGaAs layer and the wave vectors k_w and k_b are given by:

$$k_w = \sqrt{\frac{2m_A^* \varepsilon}{\hbar^2}}, \quad (72)$$

$$k_b = \sqrt{\frac{2m_B^* (|V_p| - \varepsilon)}{\hbar^2}}. \quad (73)$$

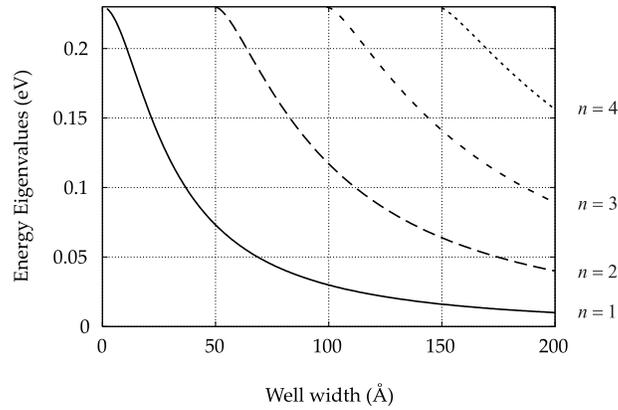


Figure 4. Variation of the first 4 energy eigenvalues (absolute value) of rectangular well for varying widths.

Eq. (70) determines the eigenenergies ε corresponding to states of even symmetry, while Eq. (71) determines those of odd symmetry. These equations were solved numerically to find the first 4 eigenvalues for wells of varying widths. The solutions are shown in Fig. 4 and are accurate to 10^{-12} hence they can be considered as exact.

4.3. Computational Issues and Simulation Results

There are several parameters that must be taken into consideration when deciding on a proper strategy for solving the bound state problem in a heterostructure. One of the fundamental issues is grid spacing. Although the arguments for the use of a non-uniform grid are compelling, from the numerical point of view a uniform grid offers many advantages, the main being the symmetry of the resulting matrix.

Eq. (41) is the defining equation of the eigenproblem. If the spacing of the grid is uniform, this matrix is symmetric. If the spacing is non-uniform, the matrix is non-symmetric. In general the symmetric eigenvalue problem is considerably easier and computationally faster to solve than the general problem [18]. Taking into account these factors, it is appropriate to address the question of using a uniform grid for speed considerations. Given the existence of high-speed and high-quality linear algebra routines that can solve the general eigenvalue problem in times comparable to those of the symmetric problem, any speed gain from the use of a uniform grid is rapidly counterbalanced by the accuracy and device geometry considerations in favor of the non-uniform grid. The linear algebra routines publicly available in the LAPACK [19] package were used throughout this work.

Apart from the grid uniformity issue, grid size is of paramount importance. As expected, a denser grid will provide more accuracy but at what speed penalty? Although this issue is problem-dependent, useful guidelines can be obtained from the solution of the prototype quantum well eigenproblem with different grid sizes and different discretization schemes. The results are shown in Fig. 5 where the overall CPU time to solve the eigenproblem is plotted as a function of grid size. These times were obtained on a 2.4 GHz Intel Pentium IV CPU, running in single user mode.

The standard discretization scheme used in most common approaches, wherein the first derivative is

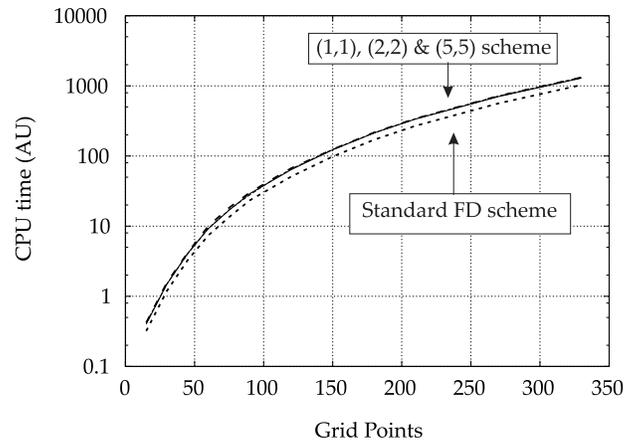


Figure 5. CPU time (in arbitrary units) as a function of grid size and discretization scheme. The numbers in parentheses (K, L) refer to the order of the discretization scheme.

approximated with a two-point central finite difference approximation:

$$\frac{d}{dx}\psi(i) = \frac{\psi(x(i+1)) - \psi(x(i-1))}{x(i+1) - x(i-1)} \quad (74)$$

performs the fastest for all grid sizes. The reason is that the discretization scheme is hard coded in the software since the approximating formulas are already known. In contrast, in our discretization scheme, the approximating formulas must be solved at run-time. This involves the solution of N Vandermonde systems of dimension $(K + L + 1)$ each. Since this involves $O((K + L + 1)^2 N)$ operations and K, L are typically small, the penalty involved is not severe. As seen in Fig. 5, for (K, L) one is likely to encounter in practice, CPU time remains approximately constant and very competitive with the standard DS. Certainly (K, L) larger than $(5, 5)$ are not recommended due to the ill-conditioned nature of high-dimensional Vandermonde problems [11] (but also see the discussion in Ref. [10]).

Before addressing the issue of accuracy, the question of possible improvement, if any, by the detailed treatment of the interfaces as outlined in Sec. 3.2 must be considered. The effect of this approach on CPU time as a function of grid size is shown in Fig. 6.

The results for the standard DS and the $(1, 1)$ DS from Fig. 5 are repeated for comparison. The treatment of the discontinuities, as described in Section 3.2, requires the calculation of the inverse matrix P_x^{-1} , a time-consuming operation. As seen from Fig. 6, the increase in CPU time is acceptable for grids of size $< 100 - 150$ points but increases rapidly for larger grid sizes.

The main motivation for adopting sophisticated DS, namely accuracy, was evaluated for the ground state energy of the well for different schemes and the results are shown in Fig. 7. The error presented in Fig. 7 is obtained by comparing the calculated eigenvalues for each DS with the quasi-analytical results shown in Fig. 4.

Although the standard and $(1, 1)$ schemes, which are entirely equivalent apart from the method of calculating the DS coefficients (in the former case they are hard-coded while in the latter they are calculated during run-time, but the final result is the same), perform satisfactorily, the $(2, 2)$ and $(5, 5)$ schemes offer slight improvement in accuracy. The clearly superior DS though is $(1, 1)$ with handling of

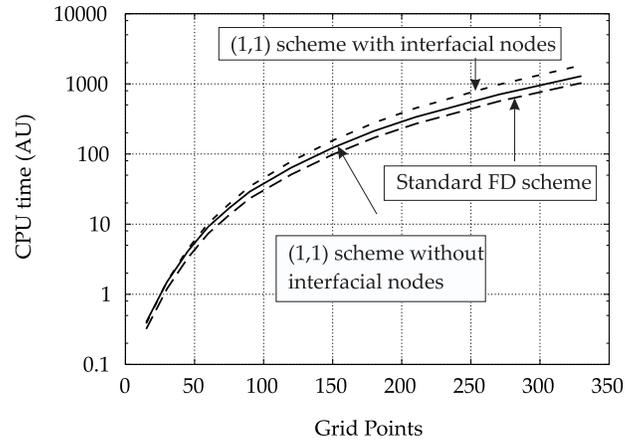


Figure 6. CPU time (in arbitrary units) as a function of grid size, discretization scheme and interface handling method.

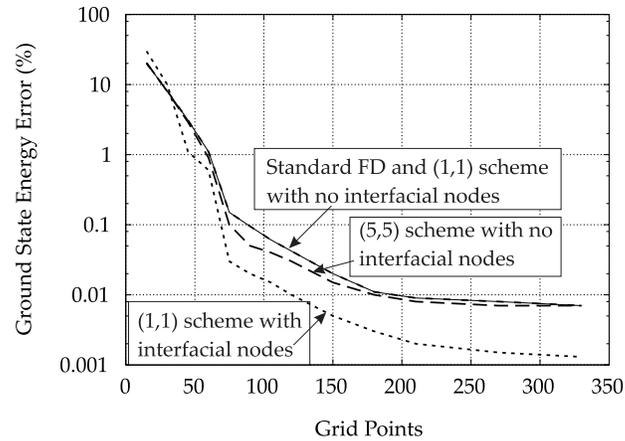


Figure 7. Ground state energy error as a function of grid size and discretization scheme.

discontinuities as described in Section 3.2. In the 100 - 150 grid size region, improvement in accuracy of the order of 3-4x is achieved as compared to standard and (1, 1) DS.

If we take into account the information conveyed in Figs. 5-7, it becomes apparent that a grid size of 100 - 150 points in conjunction with a (1, 1) DS with interface treatment offers superior performance with slight speed penalty. Should higher accuracy be deemed necessary, an increase in grid size should accommodate all requirements. In all cases however, it is evident from the results of Fig. 7 that the interfaces must be taken into account.

One final observation in favor of general (K, L) schemes as opposed to standard DS's. If we plot the distribution of the eigenvalues on the complex plane for a 105 point grid and for various DS's, the series of plots shown in Figs. 8-10 arises. Since the matrices involved are non-Hermitian, their eigenvalues

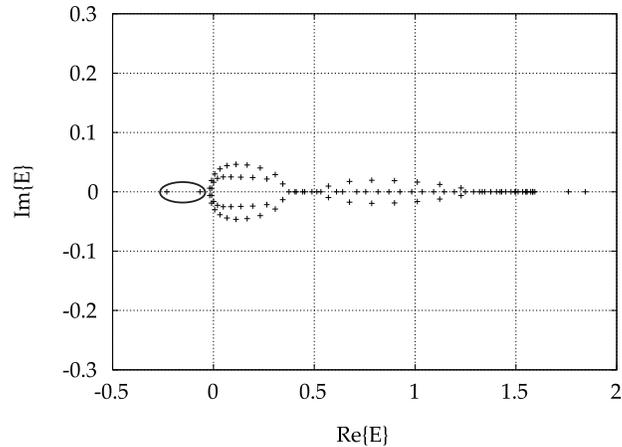


Figure 8. Spectral distribution of eigenvalues for standard and (1,1) DS. The eigenvalues of interest are circled.

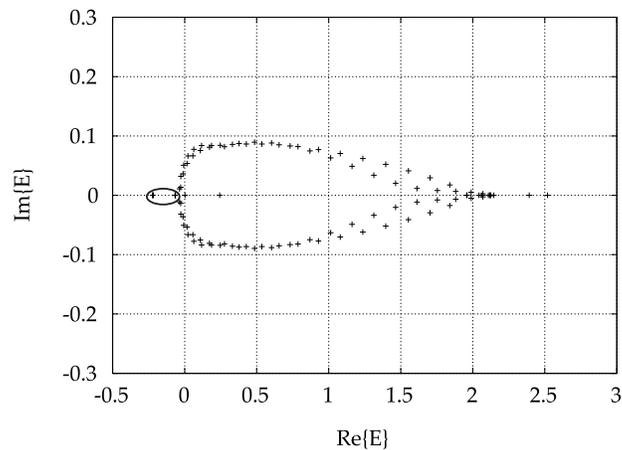


Figure 9. Spectral distribution of eigenvalues for (2,2) DS. The eigenvalues of interest are circled.

are, in general, complex. Out of the multitude of eigenvalues (105 in this case) we are interested in only a few (two). The fact that they must be negative (the energy origin coincides with the top of the conduction band) and greater than -0.23eV helps isolate them. A comparison of their spectral distribution though, shows that they are arranged in different ways on the complex plane. As the order of the DS increases, the complex eigenvalues tend to “spread” away from the origin, a fact that makes their isolation and subsequent rejection easier. In addition, those values with zero imaginary part and positive real part, which correspond to decaying states, tend to migrate from the real axis toward the complex plane. In situations where the energy reference level does not provide sign information about the real bound states, this behavior would be desirable.

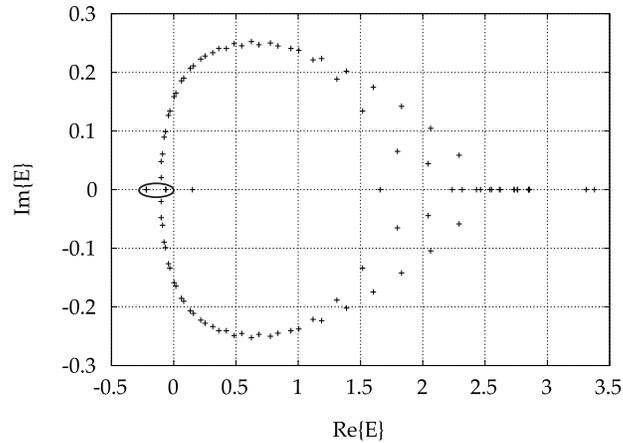


Figure 10. Spectral distribution of eigenvalues for (5,5) DS. The eigenvalues of interest are circled.

5. Conclusion

We presented an operator based family of discretization schemes for the SBEME in a heterostructure. This approach take advantage of the quantum mechanical operator forming the generator of the Lie group of infinitesimal translations, and can be extended to other cases (a rotationally invariant Hamiltonian, for example.)

A prototype potential well was studied and the effects of various discretization schemes on eigenvalue accuracy and execution time were examined. As a consequence of this study, the use of a (1,1) DS with discontinuity treatment is recommended as an optimal trade-off in accuracy vs. CPU time.

APPENDIX

Throughout this work, the formulation of quantum mechanics in Hilbert space [8] and Dirac [20] notation is used, unless reference to a specific representation is required.

Operators acting in Hilbert space are denoted with bold, accented letters. Thus the quantum mechanical momentum operator is typeset as $\hat{\mathbf{P}}$. In situations where the aforementioned operator can be decomposed in two or more orthogonal operators, each acting in its own subspace (orthogonal to all the other subspaces) of the initial Hilbert space, then the component operators are denoted with normal accented letters. Thus the components of $\hat{\mathbf{P}}$ are typeset as \hat{P}_x , \hat{P}_y and \hat{P}_z , with $\hat{\mathbf{P}} = \hat{P}_x \oplus \hat{P}_y \oplus \hat{P}_z$, where \oplus stands for the direct sum of operators.

The matrices representing operators in a given basis are denoted with bold or normal sanserif uppercase letters. So the matrix representing $\hat{\mathbf{P}}$ is typeset as \mathbf{P} and that of \hat{P}_x as P_x . Individual matrix elements are denoted with small roman letters, appropriately indexed. The matrix elements of $\hat{\mathbf{P}}$, for example, are typeset as $p(x, x')$ or $p(i, j)$ depending on whether the basis is continuous or discrete, respectively. In the latter case, the notation p_{ij} is also used.

Functional representations of operators are be denoted by the same letter as the operator, but in a

roman, unaccented typeface. Thus the function representing \hat{P}_x in the $|x\rangle$ representation is typeset as $P_x(x)$.

Finally, quantum mechanical states are denoted in uppercase bold Greek letters as in $|\Phi\rangle$ while their tensor components are denoted with the corresponding lowercase Greek letter as in $|\Psi\rangle = |\psi_x\rangle \otimes |\psi_{yz}\rangle$, where \otimes stands for the tensor product of vectors. The projections of states in the $|x\rangle$ representation (wave functions,) are typeset with the familiar $\psi_x(x) = \langle x|\psi_x\rangle$.

REFERENCES

1. S. Datta, *Quantum Phenomena* (Addison-Wesley, Reading, MA, 1989), pp. 186–192.
2. E. Kougiannos and S. P. Mohanty, “The effect of transverse energy on electronic bound states in heterostructure quantum wells,” *Semicond. Sci. Technol.* **21**, 1472–1744 (2006).
3. D. J. BenDaniel and C. B. Duke, “Space-Charge effects on electron tunneling,” *Phys. Rev.* **2**, 683–692 (1966).
4. R. A. Morrow and K. R. Brownstein, “Model effective mass Hamiltonians for abrupt heterojunctions and the associated wave function matching conditions,” *Phys. Rev. B* **30**, 678–680 (1984).
5. R. A. Morrow, “Effective mass Hamiltonians for abrupt heterojunctions in three dimensions,” *Phys. Rev. B* **36**, 4836–4840 (1987).
6. R. A. Morrow, “Establishment of an effective mass Hamiltonian for abrupt heterojunctions,” *Phys. Rev. B* **35**, 8074–8079 (1987).
7. J.-M. Lévy-Leblond, “Elementary quantum models with position-dependent mass,” *Eur. J. Phys.* **13**, 215–218 (1992).
8. C. Cohen-Tannoudji, B. Diu and F. Laloë, *Quantum Mechanics - Vol. 1* (John Wiley & Sons, New York, 1977), Chap. II.
9. S. Hassani, *Mathematical Physics* (Springer-Verlag, New York, 1999), pp. 815–831
10. Å. Björck and V. Pereyra, “Solution of Vandermonde systems of equations,” *Math. Comp.* **24**, 893–903 (1970).
11. W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1992), 2nd ed., pp. 120–121.
12. J. N. Lyness and C. B. Moller, “Van Der Monde systems and numerical differentiation,” *Num. Math.* **8**, 458–464 (1966).
13. C. Ballester and V. Pereyra, “On the construction of discrete approximations to linear differential equations,” *Math. Comp.* **21**, 297–302 (1967).
14. I-H. Tan, L. Snider, L. D. Chang, and E. L. Hu, “A self-consistent solution of Schrödinger-Poisson equations using a nonuniform mesh,” *J. Appl. Phys.* **68**, 4071–4076 (1990).
15. G. W. Parker, “Discretization of the Schrödinger equation at a discontinuity in the potential with applications to the calculation of transmission and reflection coefficients and eigenvalues,” *Am. J. Phys.* **61**, 756–759 (1993).
16. J. S. Blakemore, “Semiconducting and other Major Properties of Gallium Arsenide,” *J. Appl. Phys.* **53**, R123–R181 (1982).
17. G. Bastard and J. A. Brum, “Electronic States in semiconductor heterostructures,” *IEEE J. Quantum Electron.* **QE-22**, 1652–1644 (1986).
18. G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins, Baltimore, 1989), 2nd ed., Chap. 8.
19. E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov and D. Sorensen, *LAPACK User’s Guide* (SIAM, Philadelphia, 1992).
20. P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1969), 4th ed., Chap. 1.