JOURNAL OF COMPUTATIONAL PHYSICS 112, 149–161 (1994)

An Accurate Re-formulation of the Wigner Function Method for Quantum Transport Modeling

R. K. MAINS AND G. I. HADDAD

Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, Michigan 48109

Received April 8, 1992; revised October 15, 1992

The Wigner function is a promising method for including such effects as time dependence, self consistency, and inelastic scattering in quantum transport calculations. However, issues regarding the accuracy and consistency of the method need to be resolved. This paper presents a numerical method for determining the Wigner function which is derived from an accurate discretization of the Schrödinger equation. All of the density matrix information is preserved in this method, rather than half, as in previous methods. Results of self-consistent calculations under low bias conditions are presented. Further work must be done on the proper formulation of scattering rates and to determine the device dimensions and k-space discretizations that are required for realistic quantum device calculations.

1. INTRODUCTION

The Wigner function method has been used to include time dependence [1–2], inelastic phonon scattering [3–4], and the self-consistent potential [5–7] in quantum transport modeling. However, questions have been raised about the accuracy of the method, especially when compared with results obtained from an ensemble Schrödinger equation calculation [4]. In practice, consistently good agreement with the Schrödinger equation method has not been obtained, even for the case where inelastic scattering and self-consistency are ignored. It has been suggested that the discrepancy arises from an inherent difference in the assumptions made by these two methods; however, a serious problem remains that cannot be explained by this argument: using only the Wigner function method, the authors have found that consistent results are not obtained as parameters of the simulation are varied. For example, one expects that refining the mesh in k-space would cause the method to converge to the correct answer. In fact, cases have been observed where the method diverges or changes unpredictably. Also, consistent results as physical parameters of the structure are varied are not always obtained; if, for example, the barrier widths of a resonant tunneling diode are increased, the peak current should decrease and the current peak-to-valley ratio should increase. Such behavior is not consistently obtained with Wigner function simulations.

Some work has been carried out to formulate a more accurate approach for solving the Wigner function equations [4, 8, 9]. However, it has been the authors' experience that none of the reported methods has satisfactorily solved the problem.

In this paper, a numerical method is developed for determining the Wigner function, which is based on an accurate discretization of the Schrödinger equation. As will be discussed, problems still remain in the implementation of the method, particularly regarding the proper formulation of scattering rates and the inclusion of sufficient range in the real-space and k-space discretizations.

2. THE PROBLEM WITH CURRENT METHODS

The usual approach is to begin with the analytic form of the equation for the time evolution of the Wigner function $f(x, k)$ [10],

$$\frac{\partial f(x,k)}{\partial t} = -\frac{\hbar k}{m^*}\frac{\partial f(x,k)}{\partial x} + \left(\frac{\partial f(x,k)}{\partial t}\right)_C - \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dk' \times \left\{2\int_0^{\infty} dy \sin([k-k']\,y) \times \left[V\left(x+\frac{y}{2}\right) - V\left(x-\frac{y}{2}\right)\right]\right\} f(x,k'), \quad (1)$$

where the time derivative with subscript "C" denotes the as yet unspecified term due to scattering. The next step in the conventional approach is to discretize the terms in Eq. (1). The spatial derivative term on the right side has been discretized using upwind differencing [2], second-order upwind differencing [9], and explicit Lax–Wendroff differencing [5]. Although some improvement is obtained by using second-order accurate discretizations, it has been the authors' experience that sufficient accuracy is not obtained

0021-9991/94 $6.00

by simply modifying this term. A more serious problem is encountered in discretizing the potential energy term in Eq. (1), which involves two nested, infinite integrals. It would seem natural to discretize the integral over dk' by simply summing over all the values in the k-space discretization. However, it is not clear what the upper limit of the integral over dy should be. In practice, significantly different results may be obtained by using different values for this upper limit.

The analytic expression for the Wigner function in terms of a single wave function is

$$f(x, k) = \frac{1}{\pi} \int_{-\infty}^{\infty} dy\, \psi^*(x+y)\, \psi(x-y)\, e^{2iky}. \tag{2}$$

(Throughout this paper a single wave function will be used for clarity; however, it is understood that f actually consists of a superposition of many terms such as in Eq. (2), each with a statistical weighting.) Equation (2) is typically discretized as

$$f(x_j, k_m) = \frac{1}{N_k\, \Delta k} \sum_{n=-N_{\max}}^{N_{\max}} \psi^*(x_j + n\, \Delta x) \times \psi(x_j - n\, \Delta x)\, e^{i(2\pi mn/N_k)}, \tag{3}$$

where N_k is the number of points in k-space, and $N_{\max}$ is yet to be determined. Note that in arriving at the discretization in Eq. (3), the following relation has been assumed for the steps in real space and k-space,

$$\Delta x\, \Delta k = \pi / N_k, \tag{4}$$

which is required for Fourier completeness. This leads to a k-space discretization of the form:

$$k_m = m\, \Delta k; \qquad m = -N_{\max}, \ldots, N_{\max}. \tag{5}$$

Now, using Eq. (3) as the discrete definition of the Wigner function, the inverse transform is defined as

$$\psi^*(x_j + n'\, \Delta x)\, \psi(x_j - n'\, \Delta x) = \Delta k \sum_m e^{-i(2n'm\pi/N_k)} f(x_j, k_m). \tag{6}$$

In order for Eqs. (3) and (6) to be a valid transform pair, the following restrictions must be placed on the extent of n' in Eq. (6),

$$-\left[\frac{N_k}{2}\right] \leqslant n' \leqslant \left[\frac{N_k}{2}\right], \tag{7}$$

where the square brackets in Eq. (7) denote the greatest integer function. The quantity on the left in Eq. (6) is just the density matrix in the coordinate representation (again, when like terms are summed over a statistical ensemble). However, according to (6), only products of wave functions ψ separated by an even number of space steps Δx are within the space of the inverse transform; the density matrix within the simulated region, however, contains all products of wave functions. Therefore, if only the discrete Wigner function definition according to Eq. (3) is used, half the information contained in the density matrix will be lost. This problem has been previously noted by Frensley [11]. One may wonder why this loss of information necessarily constitutes a problem; cannot one include only half the information contained in the density matrix and still have a useful simulation model? The answer will become clear in the following section; in order to develop a scheme that is consistent with Schrödinger's equation, *all* the information contained in the discrete density matrix must be retained.

3. DEVELOPMENT OF THE NEW FORMULATION

Rather than attempting to discretize Eq. (1), let us start with an accurate discretization of Schrödinger's equation [12]:

$$\frac{d\psi(x + n\, \Delta x)}{dt} = \frac{i\hbar}{2m^*} \frac{\left(\begin{array}{l}\psi(x + (n+1)\, \Delta x) - 2\psi(x + n\, \Delta x) \\ \quad + \psi(x + (n-1)\, \Delta x)\end{array}\right)}{\Delta x^2} - \frac{i}{\hbar} V(x + n\, \Delta x)\, \psi(x + n\, \Delta x). \tag{8}$$

For now, the time derivative in (8) is not discretized. Let us also adopt the definition of $f(x_j, k_m)$, the discrete Wigner function, given by Eq. (3). The parameter N_k must be an odd integer, so that f is symmetric about $k_m = 0$ and includes this value. The parameter $N_{\max}$ in Eq. (3) is taken to be

$$N_{\max} = [N_k/2], \tag{9}$$

where the square brackets denote the greatest integer function (e.g., if $N_k = 61$, then $N_{\max} = 30$). This choice for $N_{\max}$ leads to an equation for $\partial f/\partial t$ that is entirely expressible in terms of defined Wigner function values.

As was previously mentioned, using $f(x_j, k_m)$ only retains half the density matrix values on the simulated region. Also, it is not possible to derive an equation consistent with Eq. (8) for $\partial f/\partial t$ using only f as defined in Eq. (3). To overcome these difficulties, a second Wigner function is defined midway between meshpoints,

$$g(x_j^{\text{mid}}, k_m^g) = \frac{1}{\Delta k_g (N_k+1)} \sum_{\substack{n=-N_k \\ (\Delta n=2)}}^{N_k} \psi^* \left(x_j^{\text{mid}} + \frac{n}{2}\Delta x \right) \times \psi \left(x_j^{\text{mid}} - \frac{n}{2}\Delta x \right) e^{i(2\pi(m/2)(n/2)/(N_k+1))},$$
$$m = -N_k, -N_k+2, \ldots, +N_k,$$
$$\Delta m = 2, \tag{10}$$

where k_m^g denotes k-values associated with the Wigner function g, which are different from the k_m values associated with f, and where x_j^{mid} is located between meshpoints x_j:

$$x_j^{\text{mid}} = x_j + \Delta x/2. \tag{11}$$

Note in Eq. (10) that the indices n and m are incremented by two at each step. Therefore the quantity g is defined in terms of products of wave functions separated by an odd number of mesh points, thereby including the previously missing half of the density matrix. The inverse transform associated with g is

$$\psi^* \left(x_j^{\text{mid}} + n\frac{\Delta x}{2} \right) \psi \left(x_j^{\text{mid}} - n\frac{\Delta x}{2} \right) = \Delta k_g \sum_{\substack{m=-N_k \\ (\Delta m=2)}}^{N_k} g(x_j^{\text{mid}}, k_m^g)\, e^{-i(2\pi(m/2)(n/2)/(N_k+1))},$$
$$n = -N_k, -N_k+2, \ldots, +N_k. \tag{12}$$

Note from Eq. (10) that there are $N_k + 1 = N_{kg}$ values of g at each midpoint; the motivation for this choice will become clearer in the derivation of the update equations. The quantity Δk_g in Eqs. (10) and (12) is also defined so that Fourier completeness is satisfied:

$$\Delta k_g \, \Delta x = \pi / N_{kg}. \tag{13}$$

The derivation of the discrete equation for $\partial f/\partial t$ will now be presented (the scattering term is not included in the derivation; rather it will be added later). Differentiating Eq. (3) yields

$$\frac{\partial f(x_j, k_m)}{\partial t} = \frac{1}{\Delta k N_k} \sum_{n=-[N_k/2]}^{[N_k/2]} e^{i(2\pi mn/N_k)} \times \left\{ \frac{\partial \psi^*(x_j + n\,\Delta x)}{\partial t} \psi(x_j - n\,\Delta x) + \psi^*(x_j + n\,\Delta x) \frac{\partial \psi(x_j - n\,\Delta x)}{\partial t} \right\}. \tag{14}$$

Next substitute the discrete equation (8) and its complex conjugate into Eq. (14) for $\partial\psi/\partial t$ and $\partial\psi^*/\partial t$; the result is

$$\frac{\partial f(x_j, k_m)}{\partial t} = \frac{i\hbar}{2m^*\pi\,\Delta x} \sum_{n=-[N_k/2]}^{[N_k/2]} [\psi^*(x_j + n\,\Delta x) \times \psi(x_j - (n-1)\,\Delta x) + \psi^*(x_j + n\,\Delta x) \times \psi(x_j - (n+1)\,\Delta x) - \psi(x_j - n\,\Delta x) \times \psi^*(x_j + (n+1)\,\Delta x) - \psi(x_j - n\,\Delta x) \times \psi^*(x_j + (n-1)\,\Delta x)]\, e^{i(2\pi mn/N_k)} + \frac{i\,\Delta x}{\pi\hbar} \sum_{n=-[N_k/2]}^{[N_k/2]} [\psi^*(x_j + n\,\Delta x) \times \psi(x_j - n\,\Delta x)\{V(x_j + n\,\Delta x) - V(x_j - n\,\Delta x)\}]\, e^{i(2\pi mn/N_k)}. \tag{15}$$

The next step in the derivation is to replace the products of the wave functions appearing in Eq. (15) by the corresponding Wigner function f or g as defined by Eq. (3) or (10). It is at this point that the difficulty of using only the Wigner function f becomes apparent. Since Eq. (15) contains products of wave functions separated by both odd and even numbers of space steps Δx, it is not possible to express this equation entirely in terms of f-values (it is possible, however, if the step size is doubled to $2\,\Delta x$ in Eq. (8)). For example, the following shows how the first wave function product term may be transformed so that the Wigner function definition may be used:

$$\psi^*(x_j + n\,\Delta x)\, \psi(x_j - (n-1)\,\Delta x) = \psi^*(x_j' + n'\,\Delta x)\, \psi(x_j' - n'\,\Delta x),$$
$$\text{where} \quad x_j' = x_j + \frac{\Delta x}{2}, \qquad n' = n - \frac{1}{2}. \tag{16}$$

It is seen that the x_j' value defined in Eq. (16) is between meshpoints, so that these products may be replaced by g values according to Eq. (10), provided of course that the x_j' and n' values fall within the range of definition of g in Eq. (10). If the mesh is set up so that the first and last meshpoints are f-points and the first g-values on the left are at $x_1 + \Delta x/2$, and on the right at $x_{N_x} - \Delta x/2$, where N_x is the number of meshpoints, then all the x_j' values will be within the range of g provided we are at interior points, i.e., excluding x_1 and x_{N_x}. At the boundaries, different equations must be used for f. For the n' values to be included, from Eq. (12) they must have the property that $2n'$ in each case is an odd integer, within the range $-N_k, \ldots, +N_k$, which is readily verified since n cannot be larger than $[N_k/2]$ (see Eq. (7)).

Therefore, the first four product terms in Eq. (15) may be expressed in terms of g-values as follows, where the

appropriate definition for x_j' and n' must be used for each term:

$$\psi^*(x_j' + n'\,\Delta x)\,\psi(x_j' - n'\,\Delta x) = \Delta k_g \sum_{\substack{m=-N_k \\ (\Delta m=2)}}^{N_k} g(x_j', k_m^g)\, e^{-i(2\pi(m/2)n'/N_{kg})}. \quad (17)$$

Also, the wave function product in the potential energy term of Eq. (15) is directly expressible in terms of f using Eq. (6). Using this equation and incorporating g-values as in Eqs. (16) and (17) leads to the following form for the equation for $\partial f/\partial t$:

$$\begin{aligned}\frac{\partial f(x_j, k_m)}{\partial t} = &-\frac{\hbar}{m^*\,\Delta x^2(N_k+1)} \sum_{\substack{m'=-N_k \\ (\Delta m'=2)}}^{N_k} \sin\left(\frac{\pi m'}{2(N_k+1)}\right) \\ &\times [g(x_j^{\text{mid}}, k_{m'}^g) - g(x_j^{\text{mid}} - \Delta x, k_{m'}^g)] \\ &\times \left[1 + \sum_{n=1}^{[N_k/2]} 2\cos\left(\frac{2\pi mn}{N_k} - \frac{2\pi(m'/2)\,n}{N_k+1}\right)\right] \\ &-\frac{2}{\hbar N_k} \sum_{m'=-[N_k/2]}^{[N_k/2]} f(x_j, k_{m'}) \\ &\times \sum_{n=1}^{[N_k/2]} [V(x_j + n\,\Delta x) - V(x_j - n\,\Delta x)] \\ &\times \sin\left(\frac{2\pi n(m-m')}{N_k}\right). \end{aligned} \quad (18)$$

It is seen from Eq. (18) that the spatial derivative term, $\partial f/\partial x$, in the analytic equation (1), is expressed as space-centered derivatives involving the surrounding values of g. Also, whereas the spatial derivative term occurs for a single k-value in the analytic equation, in Eq. (18) a sum of spatial derivatives over all $k_{m'}^g$ values results; however, the $k_{m'}^g \approx k_m$ term is heavily weighted in Eq. (18), since this case corresponds to $m \approx m'/2$ so that the cosine argument goes to zero. Also, replacing the sine function in (18) by its argument for small m' shows a correspondence with the k factor multiplying the spatial derivative in Eq. (1), at least for small $k_{m'}^g$. The potential energy term in Eq. (18) is a straightforward discretization of the analytic term in (1); however, the ambiguity as to the upper limit of the spatial integration (the sum over n) is now resolved.

It can be seen from the potential energy term in Eq. (18) that, for equations near the boundary, potential energy values are required beyond the domain of simulation. One method of obtaining these values is to assume that the potential is constant in the contacts, given by the last value at the boundaries. A better method is to assume that the potential energy varies linearly in the contacts, with a slope given by the last two points at the boundary. This is a more realistic choice since we expect the electric field in the contacts to be nearly constant, which corresponds to a linear potential variation.

Although some of the terms in Eq. (18) are recognizable from the analytic equation, it is not an intuitive discretization of Eq. (1). However, it is the more accurate result that is consistent with Schrödinger's equation.

If the electron density is given by

$$n(x_j) = \Delta k \sum_{m=-[N_k/2]}^{[N_k/2]} f(x_j, k_m), \quad (19)$$

and if the current density is defined at the midpoints as

$$\begin{aligned} J_n(x_j^{\text{mid}}) = &\frac{q\hbar\,\Delta k N_k}{m^* N_{kg}} \sum_{\substack{m'=-N_k \\ (\Delta m'=2)}}^{N_k} \frac{1}{\Delta x} \\ &\times \sin\left(\frac{\pi m'}{2N_{kg}}\right) g(x_j^{\text{mid}}, k_{m'}^g), \end{aligned} \quad (20)$$

then by summing terms in Eq. (18) over m it is easily shown that a discrete form of the continuity equation is exactly satisfied by this formulation, i.e.,

$$\frac{\partial n}{\partial t} = \frac{-1}{q}\frac{\partial J_n}{\partial x}. \quad (21)$$

If we further define an effective $k_m^{g^*}$,

$$k_m^{g^*} = \frac{1}{\Delta x}\sin\left(\frac{\pi m}{2N_{kg}}\right), \quad (22)$$

and using Δk_g as defined in Eq. (13), the current density may be expressed in a more familier form:

$$J_n(x_j^{\text{mid}}) = \frac{q\hbar}{m^*} \sum_{\substack{m'=-N_k \\ (\Delta m'=2)}}^{N_k} k_{m'}^{g^*}\, g(x_j^{\text{mid}}, k_{m'}^g)\,\Delta k_g. \quad (23)$$

The next step in the derivation is the development of a corresponding equation for the time evolution of g. Proceeding in the same manner, Eq. (10) is differentiated with respect to time, and the resulting $\partial\psi/\partial t$ and $\partial\psi^*/\partial t$ terms are replaced by the discrete equation (8) and its complex conjugate, respectively, which yields

$$\begin{aligned}\frac{\partial g(x_j^{\text{mid}}, k_m^g)}{\partial t} = &-\frac{i\hbar}{2m^*\pi\,\Delta x} \sum_{\substack{n=-N_k \\ (\Delta n=2)}}^{N_k} \left[\psi^*\left(x_j^{\text{mid}} + \frac{n+2}{2}\Delta x\right)\right. \\ &\times \psi\left(x_j^{\text{mid}} - \frac{n}{2}\Delta x\right) + \psi^*\left(x_j^{\text{mid}} + \frac{n-2}{2}\Delta x\right) \\ &\times \psi\left(x_j^{\text{mid}} - \frac{n}{2}\Delta x\right) - \psi^*\left(x_j^{\text{mid}} + \frac{n}{2}\Delta x\right) \end{aligned}$$

$$\times\psi\left(x_j^{\text{mid}}-\frac{n-2}{2}\Delta x\right)-\psi^*\left(x_j^{\text{mid}}+\frac{n}{2}\Delta x\right)$$
$$\times\psi\left(x_j^{\text{mid}}-\frac{n+2}{2}\Delta x\right)\Bigg]e^{i(\pi mn/2(N_k+1))}$$
$$+\frac{i\,\Delta x}{\pi\hbar}\sum_{\substack{n=-N_k\\ \Delta n=2}}^{N_k}\Bigg[\psi^*\left(x_j^{\text{mid}}+\frac{n}{2}\Delta x\right)$$
$$\times\psi\left(x_j^{\text{mid}}-\frac{n}{2}\Delta x\right)\Bigg\{V\left(x_j^{\text{mid}}+\frac{n}{2}\Delta x\right)$$
$$-V\left(x_j^{\text{mid}}-\frac{n}{2}\Delta x\right)\Bigg\}\Bigg]e^{i(\pi mn/2(N_k+1))}. \quad (24)$$

The wave function product terms in the first group on the right side are separated by an even number of meshpoints (since n is odd), so they may be replaced by f-values. In the potential terms, the wave functions are separated by odd numbers of meshpoints so they are expressed in terms of g. For this equation, however, for some values of n the product terms in the first group will be outside the range of the definition of f. The following shows the transformations of the product terms within the first group in preparation for the substitution of f-values, and the value of n for which the substitution is not possible is mentioned for each case:

$$\psi^*\left(x_j^{\text{mid}}+\frac{n+2}{2}\Delta x\right)\psi\left(x_j^{\text{mid}}-\frac{n}{2}\Delta x\right)$$
$$=\psi^*(x_j'+n'\,\Delta x)\,\psi(x_j'-n'\,\Delta x),$$
$$\text{where}\quad x_j'=x_j^{\text{mid}}+\Delta x/2,\qquad n'=(n+1)/2,$$
$$\text{does not work for } n=N_k; \quad (25a)$$

$$\psi^*\left(x_j^{\text{mid}}+\frac{n-2}{2}\Delta x\right)\psi\left(x_j^{\text{mid}}-\frac{n}{2}\Delta x\right)$$
$$=\psi^*(x_j'+n'\,\Delta x)\,\psi(x_j'-n'\,\Delta x),$$
$$\text{where}\quad x_j'=x_j^{\text{mid}}-\Delta x/2,\qquad n'=(n-1)/2,$$
$$\text{does not work for } n=-N_k; \quad (25b)$$

$$\psi^*\left(x_j^{\text{mid}}+\frac{n}{2}\Delta x\right)\psi\left(x_j^{\text{mid}}-\frac{n-2}{2}\Delta x\right)$$
$$=\psi^*(x_j'+n'\,\Delta x)\,\psi(x_j'-n'\,\Delta x),$$
$$\text{where}\quad x_j'=x_j^{\text{mid}}+\Delta x/2,\qquad n'=(n-1)/2,$$
$$\text{does not work for } n=-N_k; \quad (25c)$$

$$\psi^*\left(x_j^{\text{mid}}+\frac{n}{2}\Delta x\right)\psi\left(x_j^{\text{mid}}-\frac{n+2}{2}\Delta x\right)$$
$$=\psi^*(x_j'+n'\,\Delta x)\,\psi(x_j'-n'\,\Delta x),$$
$$\text{where}\quad x_j'=x_j^{\text{mid}}-\Delta x/2,\qquad n'=(n+1)/2,$$
$$\text{does not work for } n=N_k. \quad (25d)$$

To proceed further, Eq. (6) is used to introduce f-values for the terms in Eqs. (25) that can be replaced; this will yield a spatial derivative term involving f-values; however, "boundary terms" will be left over corresponding to the values of n noted in Eqs. (25) (boundary terms in the sense that they are just outside the boundary of the definition for f, not that they necessarily occur on the simulation boundary). Substituting f-values for the terms that can be replaced and using Eq. (12) to replace the wave function products in the potential energy terms of Eq. (24) yields the discrete relation for $\partial g/\partial t$,

$$\frac{\partial g(x_j^{\text{mid}},k_m^g)}{\partial t}=-\frac{\hbar}{m^*\,\Delta x^2 N_k}\sin\left(\frac{\pi m}{2(N_k+1)}\right)$$
$$\times\sum_{m'=-[N_k/2]}^{[N_k/2]}\left[f\left(x_j^{\text{mid}}+\frac{\Delta x}{2},k_{m'}\right)\right.$$
$$\left.-f\left(x_j^{\text{mid}}-\frac{\Delta x}{2},k_{m'}\right)\right]$$
$$\times\left[1+\sum_{\substack{n'=2\\ (\Delta n'=2)}}^{N_k-1}2\cos\left(\frac{\pi mn'}{2(N_k+1)}-\frac{\pi m'n'}{N_k}\right)\right]$$
$$-\frac{2}{\hbar(N_k+1)}\sum_{\substack{m'=-N_k\\ (\Delta m'=2)}}^{N_k}g(x_j^{\text{mid}},k_{m'}^g)$$
$$\times\sum_{\substack{n=1\\ (\Delta n=2)}}^{N_k}\left[V\left(x_j^{\text{mid}}+\frac{n\,\Delta x}{2}\right)\right.$$
$$\left.-V\left(x_j^{\text{mid}}-\frac{n\,\Delta x}{2}\right)\right]\sin\left(\frac{\pi n(m-m')}{2(N_k+1)}\right)$$
$$+\left.\frac{\partial g}{\partial t}\right|_{\text{B.T.}}, \quad (26)$$

where the final term labeled B.T. represents the boundary terms that are not expressible in terms of Wigner function values, either f or g. These boundary terms are, specifically,

$$\left.\frac{\partial g(x_j^{\text{mid}},k_m^g)}{\partial t}\right|_{\text{B.T.}}=\frac{-i\hbar}{2\pi m^*\,\Delta x}$$
$$\times\Bigg\{\Bigg[\psi^*\left(x_j^{\text{mid}}+\frac{N_k+2}{2}\Delta x\right)$$
$$\times\psi\left(x_j^{\text{mid}}-\frac{N_k}{2}\Delta x\right)$$
$$-\psi^*\left(x_j^{\text{mid}}+\frac{N_k}{2}\Delta x\right)$$
$$\times\psi\left(x_j^{\text{mid}}-\frac{N_k+2}{2}\Delta x\right)\Bigg]$$
$$\times e^{i(\pi mN_k/2(N_k+1))}-\text{c.c.}\Bigg\}, \quad (27)$$

where c.c. in (27) denotes the complex conjugate of the previous term.

There are two possible approaches to these boundary terms. One is to simply ignore them and hope that they are insignificant. However, another approach is taken here. For the case where ψ in (27) is an energy eigenstate of the Schrödinger equation, the boundary terms can be reduced to a form that is expressible in terms of defined Wigner function values. To do this, the following property for energy eigenstates is developed. First, multiplying the discretized time-independent Schrödinger equation at point y by $\psi^*(y+n\,\Delta x)$ yields

$$\psi^*(y+n\,\Delta x)\left\{\frac{-\hbar^2}{2m^*}\frac{\psi(y+\Delta x)-2\psi(y)+\psi(y-\Delta x)}{\Delta x^2}+V(y)\,\psi(y)=E\psi(y)\right\}. \tag{28}$$

Similarly, multiplying the complex conjugate of Schrödinger's equation at the point $y+n\,\Delta x$ by $\psi(y)$ and subtracting the result from Eq. (28), and finally setting $y=x-(N_k/2)\,\Delta x$, $n=N_k$, yields

$$\begin{aligned}
&\psi^*\left(x+\left(\frac{N_k}{2}+1\right)\Delta x\right)\psi\left(x-\frac{N_k}{2}\Delta x\right)\\
&\quad-\psi^*\left(x+\frac{N_k}{2}\Delta x\right)\psi\left(x-\left(\frac{N_k}{2}+1\right)\Delta x\right)\\
&=\psi^*\left(x+\frac{N_k}{2}\Delta x\right)\psi\left(x-\left(\frac{N_k}{2}-1\right)\Delta x\right)\\
&\quad-\psi^*\left(x+\left(\frac{N_k}{2}-1\right)\Delta x\right)\psi\left(x-\frac{N_k}{2}\Delta x\right)\\
&\quad+\frac{2m^*\,\Delta x^2}{\hbar^2}\psi\left(x-\frac{N_k}{2}\Delta x\right)\psi^*\left(x+\frac{N_k}{2}\Delta x\right)\\
&\quad\times\left[V\left(x+\frac{N_k}{2}\Delta x\right)-V\left(x-\frac{N_k}{2}\Delta x\right)\right].
\end{aligned} \tag{29}$$

It is seen that the left side of Eq. (29) is just the boundary term of Eq. (27). The right side, however, is still not in a form where the Wigner function definitions can be used. The terms on the right can further be transformed as

$$\begin{aligned}
&\psi^*\left(x+\frac{N_k}{2}\Delta x\right)\psi\left(x-\left(\frac{N_k}{2}-1\right)\Delta x\right)\\
&\quad=\psi^*(x'+n'\,\Delta x)\,\psi(x'-n'\,\Delta x),
\end{aligned}$$

where $x'=x+\Delta x/2$, $n'=(N_k-1)/2$;

$$\begin{aligned}
&\psi^*\left(x+\left(\frac{N_k}{2}-1\right)\Delta x\right)\psi\left(x-\frac{N_k}{2}\Delta x\right)\\
&\quad=\psi^*(x'+n'\,\Delta x)\,\psi(x'-n'\,\Delta x),
\end{aligned} \tag{30}$$

where $x'=x-\Delta x/2$, $n'=(N_k-1)/2$.

The terms in Eq. (30) are now in a form to which Eq. (6) may be applied (it is easily verified that the x' and n' values in (30) are within the range of definition of f). Also, the potential energy wave function product term in Eq. (29) is readily replaced using Eq. (12) for the inverse transform of g. Making these substitutions, the boundary terms in the equation for g may finally be reduced to the form

$$\begin{aligned}
&\left.\frac{\partial g(x,k_m^g)}{\partial t}\right|_{\text{B.T.}}\\
&\quad=\frac{\hbar}{m^*\,\Delta x^2 N_k}\\
&\qquad\times\sum_{m'=-[N_k/2]}^{[N_k/2]}\left[f\left(x+\frac{\Delta x}{2},k_{m'}\right)-f\left(x-\frac{\Delta x}{2},k_{m'}\right)\right]\\
&\qquad\times\sin\left[\frac{\pi m N_k}{2(N_k+1)}-\frac{\pi m'(N_k-1)}{N_k}\right]\\
&\qquad+\frac{2}{\hbar(N_k+1)}\left[V\left(x+\frac{N_k}{2}\Delta x\right)-V\left(x-\frac{N_k}{2}\Delta x\right)\right]\\
&\qquad\times\sum_{\substack{m''=-N_k\\(\Delta m''=2)}}^{N_k}g(x,k_{m''})\sin\left[\frac{\pi(m-m'')\,N_k}{2(N_k+1)}\right],
\end{aligned} \tag{31}$$

where it is understood in (31) that x is between meshpoints. Comparing the potential energy boundary term with the term in Eq. (26), it is seen that it may be taken into account by simply reducing the upper limit of the sum over n from N_k to N_k-2. Including the boundary terms in Eq. (26) then leads to the equation for $\partial g/\partial t$,

$$\begin{aligned}
\frac{\partial g(x_j^{\text{mid}},k_m^g)}{\partial t}=&-\frac{\hbar}{m^*\,\Delta x^2 N_k}\\
&\times\sum_{m'=-[N_k/2]}^{[N_k/2]}\left[f\left(x_j^{\text{mid}}+\frac{\Delta x}{2},k_{m'}\right)\right.\\
&\left.-f\left(x_j^{\text{mid}}-\frac{\Delta x}{2},k_{m'}\right)\right]\left\{\sin\left(\frac{\pi m}{2(N_k+1)}\right)\right.\\
&\times\left[1+\sum_{\substack{n'=2\\(\Delta n'=2)}}^{N_k-1}2\cos\left(\frac{\pi m n'}{2(N_k+1)}-\frac{\pi m' n'}{N_k}\right)\right]\\
&\left.-\sin\left[\frac{\pi m N_k}{2(N_k+1)}-\frac{\pi m'(N_k-1)}{N_k}\right]\right\}\\
&-\frac{2}{\hbar(N_k+1)}\sum_{\substack{m'=-N_k\\(\Delta m'=2)}}^{N_k}g(x_j^{\text{mid}},k_{m'}^g)\\
&\times\sum_{\substack{n=1\\(\Delta n=2)}}^{N_k-2}\left[V\left(x_j^{\text{mid}}+\frac{n\,\Delta x}{2}\right)\right.\\
&\left.-V\left(x_j^{\text{mid}}-\frac{n\,\Delta x}{2}\right)\right]\sin\left(\frac{\pi n(m-m')}{2(N_k+1)}\right).
\end{aligned} \tag{32}$$

It should be remarked that Eq. (32) will be exactly correct under steady state conditions for the case without scattering, since energy eigenstates were assumed to transform the boundary terms; for transient conditions the boundary terms will be only approximately correct. The effect of assuming this form for the boundary terms on transient calculations and calculations including scattering has not yet been determined.

4. BOUNDARY CONDITIONS AND SCATTERING TERMS

As mentioned previously, Eq. (18) is solved at interior meshpoints to determine f. However, at boundary meshpoints (the first and last meshpoints of the simulated region, where f values are defined) this equation cannot be used in its present form. Mathematically, this is so since the surrounding g-values required for the spatial derivatives are not defined on both sides of the meshpoint. Physically, boundary conditions must be set at these points to model the irreversible interaction with the contact regions, as discussed by Frensley [11]: the incident distribution is fixed according to the statistics of the contact region, and outgoing waves should be permitted to exit the simulation region without reflection at the boundaries. This can be accomplished in different ways, with varying degrees of success. One approach is to formulate mathematically absorbing boundary conditions [13, 14]. Another is to introduce imaginary potentials near the boundaries for wave packet simulations, which dissipate the reflected packets [15, 16]; since in the present formulation wave packets are not used and the self-consistent potential is required, this method is not feasible. Another approach, used in the present formulation, is to assume that at the boundaries an equilibrium distribution consistent with the electric field there is established; this can only work provided sufficient scattering is included to drive the distribution to equilibrium at the boundaries.

A further motivation to include scattering is based on the following observation. For the DC problem, if boundary conditions are specified only at the first and last simulation point, the equations so far presented are singular for the case without scattering, so that a well-defined solution does not exist. That this is the case can be mathematically observed; however, it also has a physical basis that is illustrated by the following simple example. Let us assume that at both boundaries, the solution is described by a plane wave traveling to the right of the form,

$$e^{i(k_{m'}x+\pi/4)}, \tag{33}$$

and by a wave of equal amplitude traveling to the left,

$$e^{-i(k_{m'}x+\pi/4)}, \tag{34}$$

where $k_{m'}$ is one of the discrete wavevectors given in Eq. (5). If we assume these waves are uncorrelated, then using Eq. (3) there are two non-zero terms in the Wigner function f:

$$f(x_j, k_m = k_{m'}) = f(x_j, k_m = -k_{m'}) = 1/\Delta k. \tag{35}$$

This result is obtained by applying Eq. (3) to each of the terms in Eqs. (33) and (34) separately. Now, however, if we have a single wavefunction of the following form, where the waves are correlated,

$$\psi = e^{i(k_{m'}x+\pi/4)} + e^{-i(k_{m'}x+\pi/4)}, \tag{36}$$

then we still have the terms in Eq. (35); however, an additional term appears at $k=0$,

$$f(x_j, k_m = 0) = \frac{2}{\Delta k}\cos\left(\frac{2m'j\pi}{N_k}+\frac{\pi}{2}\right). \tag{37}$$

If we chose the simulation x_j boundaries to correspond, for example, with the values of j,

$$j=0, \qquad j=\frac{N_k}{m'}, \tag{38}$$

then the $k=0$ term in Eq. (37) will be zero at the boundaries. Therefore, we have two distinct solutions that have exactly the same boundary conditions, so that the solution is ill-defined if only boundary conditions are specified. Physically, this is so because, in addition to the boundary conditions, the *correlations* of the problem must be specified.

This situation changes if scattering is introduced into the problem. In that case the solution, assuming correlated waves to $x=\pm\infty$, is no longer possible, since scattering processes will eventually destroy the coherence. In practice, it has been found that in solving the equations presented here for the DC case, well-defined solutions exist, provided sufficient scattering is introduced into the problem. (It should be mentioned that for the transient problem, it may not be required to include scattering, since the correlations can be specified in the initial conditions. However for this case, the boundary condition problem is more difficult.)

Based on the forgoing discussion, scattering processes have been included in formulating the boundary conditions for the results presented here. For that case, at equilibrium near a contact boundary the electric field will be constant and the f distribution will be spatially constant, driven to an equilibrium distribution by the scattering processes. To find the distribution, Eq. (18) is solved with scattering included, g is assumed constant (so that the spatial derivative terms involving g disappear), and the potential $V(x)$ (given by

the electric field at the boundary) is assumed to be linear. This solution yields an $f_{eq}(k)$ distribution which is then used to set boundary conditions on f in the simulation.

In order to determine f_{eq}, it is first necessary to specify the form of the scattering rates. For the results presented in this paper, scattering terms of the following form were used in the equations for f:

$$\left.\frac{\partial f(x_j, k_m)}{\partial t}\right|_C = -S^f_{outm} f(x_j, k_m) + \sum_{\substack{l=-[N_k/2] \\ l\neq m}}^{[N_k/2]} S^f_{lm} f(x_j, k_l), \tag{39}$$

where the first term on the right is the scattering rate out of state k_m to all other states, and the second term is the scattering rate from all other states into state k_m. A similar form is used for the g equations; however, for g to have a well-defined distribution at equilibrium, a coupling term for scattering from f states into g states is also included:

$$\left.\frac{\partial g(x_j^{\text{mid}}, k_m^g)}{\partial t}\right|_C = -S^g_{outm} g(x_j^{\text{mid}}, k_m^g) + \sum_{\substack{l=-N_k \\ \Delta l=2; l\neq m}}^{N_k} S^g_{lm} g(x_j^{\text{mid}}, k_l^g) + \sum_{m'=-[N_k/2]}^{[N_k/2]} S^{fg}_{m'm} \frac{f(x_{j+1}, k_{m'}) + f(x_j, k_{m'})}{2}. \tag{40}$$

It is emphasized that Eqs. (39) and (40) are not the only or necessarily the best choices for the forms of the scattering rates. In particular, the coupling term for scattering into g states from f states without a corresponding loss term in the f equations is suspect; further work needs to be done on the formulation of proper scattering rates. It is important that the scattering rates used have certain properties, such as maintaining current continuity and driving the distributions to the equilibrium boundary conditions assumed at the contacts. For the calculations presented here, the scattering rates appearing in Eqs. (39) and (40) are assumed to be characterized by a relaxation time. At equilibrium with zero electric field, the form for the f scattering rates that ensures that Eq. (39) is zero is

$$S^f_{lm} = \frac{f_0(k_m)}{\tau \sum_{i=-[N_k/2]}^{[N_k/2]} f_0(k_i)}, \qquad S^f_{outm} = \sum_{\substack{l=-[N_k/2] \\ l\neq m}}^{[N_k/2]} S^f_{ml}, \tag{41}$$

where τ is a relaxation time related to the material low-field mobility μ as

$$\tau = \mu m^*/q \tag{42}$$

and where f_0 is the zero-field equilibrium distribution given by

$$f_0(k_m) = \frac{m^* k_B T}{2\pi^2\hbar^2} \log[1 + e^{-(\hbar^2 k_m^2/2m^* - E_F)/k_B T}]. \tag{43}$$

(Note that the second relation of Eq. (41) ensures that the sum of terms over m in Eq. (39) will be zero, which guarantees current continuity as described in Eqs. (19)–(23)). For the g scattering terms, the definitions that ensure that Eq. (40) is zero at equilibrium are

$$S^g_{lm} = \frac{g_0(k_m^g)}{D}, \qquad S^g_{outm} = \left(\sum_{\substack{m'=-N_k \\ (\Delta m'=2; m'\neq m)}}^{N_k} S^g_{mm'} + \sum_{i=-[N_k/2]}^{[N_k/2]} f_0(k_i) \right) \Big/ D, \qquad S^{fg}_{m'm} = \frac{g_0(k_m^g)}{D}, \tag{44}$$

where D is given by

$$D = \tau \left(\sum_{i=-[N_k/2]}^{[N_k/2]} f_0(k_i) + \sum_{\substack{m'=-N_k \\ (\Delta m'=2)}}^{N_k} g_0(k_{m'}^g) \right) \tag{45}$$

and where $g_0(k_m^g)$ is obtained from Eq. (43) by replacing the k_m by k_m^g values.

For the case at equilibrium near a contact boundary the electric field will be constant and the f and g distributions will be spatially constant. Retaining only the potential energy term in Eq. (18), expressing the potential energy difference in terms of the electric field E at the boundary, and including the scattering terms of Eq. (39) yields the following set of equations for f_{eq} at steady-state equilibrium:

$$0 = -S^f_{outm} f_{eq}(k_m, E) + \sum_{\substack{m'=-[N_k/2] \\ (m'\neq m)}}^{[N_k/2]} \left[S^f_{m'm} - \frac{4}{\hbar N_k} \times \sum_{n=1}^{[N_k/2]} qEn\,\Delta x \sin\left(\frac{2\pi n(m-m')}{N_k}\right) \right] f_{eq}(k_{m'}, E), \qquad m = -\left[\frac{N_k}{2}\right], \ldots, \left[\frac{N_k}{2}\right]. \tag{46}$$

The above system of equations may be used to determine f_{eq} for a given electric field E. However, summing the equations in (46) over m yields zero, so that the equations are not independent. To uniquely determine f_{eq}, one of the dependent equations in (46) must be replaced by the normalization condition,

$$\Delta k \sum_{m'=-[N_k/2]}^{[N_k/2]} f_{eq}(k_{m'}, E) = N_D, \tag{47}$$

where N_D is the contact doping. Equations (46) and (47) may be solved for the f_{eq} boundary conditions for any electric field E.

A similar set of equations may be solved to determine g_{eq} in the contact regions; this system has a well-defined solution even without a normalization condition, due to the coupling term with f_{eq}.

5. INCLUDING VARIABLE EFFECTIVE MASS

So far in the development of the equations, a constant effective mass m^* has been assumed. To include the effect of variable effective mass, the starting point is the following discretized form of Schrödinger's equation [17], rather than Eq. (8),

$$\frac{\partial\psi(x+n\,\Delta x)}{\partial t}=\frac{i\hbar}{2\,\Delta x^2}\left[\frac{1}{m^*(x+(n+1/2)\,\Delta x)}\times(\psi(x+(n+1)\,\Delta x)-\psi(x+n\,\Delta x))-\frac{1}{m^*(x+(n-1/2)\,\Delta x)}\times(\psi(x+n\,\Delta x)-\psi(x+(n-1)\,\Delta x))\right]-\frac{i}{\hbar}V(x+n\,\Delta x)\,\psi(x+n\,\Delta x), \tag{48}$$

where the variable effective mass m^* is defined on the midpoints (or g-points).

The derivations of the equations for f and g are not given here, but only the results; to derive these equations, a procedure similar to the case for a constant effective mass is followed, but Eq. (48) is used. The counterpart of Eq. (18) for $\partial f/\partial t$ with a variable effective mass is

$$\begin{aligned}\frac{\partial f(x_j,k_m)}{\partial t}=&\frac{-\hbar}{\Delta x^2(N_k+1)}\times\sum_{\substack{m'=-N_k\\(\Delta m'=2)}}^{N_k}\sum_{n=-[N_k/2]}^{[N_k/2]}\times\sin\left(\frac{2\pi mn}{N_k}-\frac{\pi m'n}{N_k+1}+\frac{\pi m'}{2(N_k+1)}\right)\\&\times\left[\frac{g(x_j+\Delta x/2,k^g_{m'})}{m^*(x_j-(n-1/2)\,\Delta x)}-\frac{g(x_j-\Delta x/2,k^g_{m'})}{m^*(x_j+(n-1/2)\,\Delta x)}\right]\\&-\frac{2}{\hbar N_k}\sum_{m'=-[N_k/2]}^{[N_k/2]}f(x_j,k_{m'})\\&\times\sum_{n=1}^{[N_k/2]}[V(x_j+n\,\Delta x)-V(x_j-n\,\Delta x)]\\&\times\sin\left(\frac{2\pi n(m-m')}{N_k}\right)-\frac{\hbar}{\Delta x^2N_k}\\&\times\sum_{m'=-[N_k/2]}^{[N_k/2]}f(x_j,k_{m'})\\&\times\sum_{n=-[N_k/2]}^{[N_k/2]}\sin\left(\frac{2\pi(m'-m)\,n}{N_k}\right)\\&\times\left[\frac{1}{m^*(x_j+(n+1/2)\,\Delta x)}+\frac{1}{m^*(x_j+(n-1/2)\,\Delta x)}\right].\end{aligned} \tag{49}$$

It can be shown that the first term on the right side of Eq. (49) reduces to that of Eq. (18) if m^* is constant. The potential energy term is identical to that of Eq. (18). The last term in Eq. (49) has no counterpart in the constant m^* case; it has a form similar to the potential energy term and couples all the values of f at a single meshpoint. Note that this term vanishes when m^* is constant, since for that case the term is odd in the index n.

The counterpart of Eq. (26) for $\partial g/\partial t$ is

$$\begin{aligned}\frac{\partial g(x_j^{\text{mid}},k^g_m)}{\partial t}=&\frac{-\hbar}{N_k\,\Delta x^2}\times\sum_{m'=-[N_k/2]}^{[N_k/2]}\sum_{\substack{n=-N_k\\(\Delta n=2)}}^{N_k-2}\sin\left(\frac{\pi m'(n+1)}{N_k}-\frac{\pi mn}{2N_{kg}}\right)\\&\times\left[\frac{f(x_j^{\text{mid}}+\Delta x/2,k_{m'})}{m^*(x_j^{\text{mid}}+((n+1)/2)\,\Delta x)}-\frac{f(x_j^{\text{mid}}-\Delta x/2,k_{m'})}{m^*(x_j^{\text{mid}}-((n+1)/2)\,\Delta x)}\right]\\&-\frac{2}{\hbar N_{kg}}\sum_{\substack{m'=-N_k\\(\Delta m'=2)}}^{N_k}g(x_j^{\text{mid}},k^g_{m'})\\&\times\sum_{\substack{n=1\\(\Delta n=2)}}^{N_k}\left[V\left(x_j^{\text{mid}}+\frac{n}{2}\Delta x\right)-V\left(x_j^{\text{mid}}-\frac{n}{2}\Delta x\right)\right]\\&\times\sin\left(\frac{\pi(m-m')}{2N_{kg}}\right)-\frac{\hbar}{\Delta x^2N_{kg}}\sum_{\substack{m'=-N_k\\(\Delta m'=2)}}^{N_k}g(x_j^{\text{mid}},k^g_{m'})\\&\times\sum_{\substack{n=-N_k\\(\Delta n=2)}}^{N_k}\sin\left(\frac{\pi(m-m')\,n}{2N_{kg}}\right)\left[\frac{1}{m^*(x_j^{\text{mid}}+((n+1)/2)\,\Delta x)}+\frac{1}{m^*(x_j^{\text{mid}}+((n-1)/2)\,\Delta x)}\right]+\left.\frac{\partial g(x_j^{\text{mid}},k^g_m)}{\partial t}\right|_{\text{B.T.}},\end{aligned} \tag{50}$$

where, assuming energy eigenstates as before, the boundary terms are given by

$$\left.\frac{\partial g(x_j^{\text{mid}}, k_m^g)}{\partial t}\right|_{\text{B.T.}} = \frac{\hbar}{\Delta x^2 N_k} \times \sum_{m'=-[N_k/2]}^{[N_k/2]} \left[\frac{f(x_j^{\text{mid}} + \Delta x/2, k_{m'})}{m^*(x_j^{\text{mid}} - ((N_k-1)/2)\,\Delta x)} - \frac{f(x_j^{\text{mid}} - \Delta x/2, k_{m'})}{m^*(x_j^{\text{mid}} + ((N_k-1)/2)\,\Delta x)}\right] \times \sin\left(\frac{\pi m N_k}{2N_{kg}} - \frac{\pi m'(N_k-1)}{N_k}\right) + \frac{2}{\hbar N_{kg}}\left[V\left(x_j^{\text{mid}} + \frac{N_k}{2}\Delta x\right) - V\left(x_j^{\text{mid}} - \frac{N_k}{2}\Delta x\right)\right] \sum_{\substack{m'=-N_k \\ (\Delta m'=2)}}^{N_k} g(x_j^{\text{mid}}, k_{m'}^g) \times \sin\left(\frac{\pi(m-m')\,N_k}{2N_{kg}}\right) + \frac{\hbar}{\Delta x^2 N_{kg}} \sum_{\substack{m'=-N_k \\ (\Delta m'=2)}}^{N_k} g(x_j^{\text{mid}}, k_{m'}^g) \times \left[\frac{1}{m^*(x_j^{\text{mid}} + ((N_k+1)/2)\,\Delta x)} + \frac{1}{m^*(x_j^{\text{mid}} + ((N_k-1)/2)\,\Delta x)} - \frac{1}{m^*(x_j^{\text{mid}} - ((N_k-1)/2)\,\Delta x)} - \frac{1}{m^*(x_j^{\text{mid}} - ((N_k+1)/2)\,\Delta x)}\right] \times \sin\left(\frac{\pi(m-m')\,N_k}{2N_{kg}}\right). \tag{51}$$

It can be shown that the first term on the right side of Eq. (50) reduces to that of Eq. (26) if m^* is constant. The potential energy terms are identical for both equations. The third term on the right in Eq. (50) is new and clearly vanishes for the constant m^* case. Comparing the boundary terms in Eq. (51) with Eq. (27), it is seen that the first term on the right is slightly modified for the variable m^* case; the potential energy terms are identical, and a third term on the right of Eq. (51) now appears, which vanishes for the constant m^* case.

It is easily shown that a discrete form of the continuity equation will be satisfied if the current density for the variable m^* case is defined as

$$J_n(x_j^{\text{mid}}) = \frac{q\hbar}{m^*(x_j^{\text{mid}})} \sum_{\substack{m'=-N_k \\ (\Delta m'=2)}}^{N_k} k_{m'}^{g*}\, g(x_j^{\text{mid}}, k_{m'}^g)\, \Delta k_g. \tag{52}$$

6. NUMERICAL METHODS AND RESULTS

The basic problem is to solve Eqs. (18) and (32), subject to boundary conditions for f as determined by solving Eqs. (46) and (47) and by including scattering terms of the form given by Eqs. (39)–(45). Since the boundary conditions obtained from Eq. (46) depend on the electric field at the boundary, it is necessary to perform a self-consistent calculation where Poisson's equation is included at interior meshpoints x_j,

$$\frac{V'_{j+1} - 2V'_j + V'_{j-1}}{\Delta x^2} = \frac{q^2(N_{\text{D}j} - \Delta k \sum_m f(x_j, k_m))}{\varepsilon}, \tag{53}$$

where the second derivative on the left has been discretized and where V'_j represents the potential energy for electrons, but with potential variations due only to material discontinuities not included. At the boundary meshpoints, Dirichlet conditions are set on the potential energy as

$$V'_1 = 0, \qquad V'_{N_x} = -qV_{\text{applied}}, \tag{54}$$

where V_{applied} is the voltage applied across the device with positive reference polarity assumed on the right.

This set of equations constitutes a non-linear problem that must be solved using an iterative technique. These equations may in principle be solved either for the DC or for the transient case. However, for the transient case the equations are solved implicitly so that a large matrix must be solved several times per time step advancement. Considering the computation time required for each matrix solution (about 1.5 h on an Apollo DN4000 workstation for a problem of modest size), it is not presently feasible to carry out a transient case of several hundred time steps. Instead, the DC problem will be addressed here, where on the order of 5–10 matrix solutions are required per DC bias point. (For the time-dependent calculation the matrix structure is the same as presented here, but additional terms from the time derivatives are included in the matrix entries.)

The unknowns are ordered as follows. Starting at the left-hand side, we have a block of f values at x_1, ranging over all values of k_m. This is followed by a block of g values at the x_1^{mid} midpoint between x_1 and x_2, ranging over all values of k_m^g. This structure is repeated as we progress from left to right, with a block of f followed by a block of g at the next midpoint. On the right-hand side, we have the final block of f values. Next the values of the electron potential energy are included as variables, but in order to minimize the number of fills in the resulting matrix these values are ordered from V'_{N_x} to V'_1. Finally, the values of electric field at the boundaries, $E_{0\text{L}}$ and $E_{0\text{R}}$ are included as variables, which are related to the potential energy values as

$$E_{0\text{L}} = \frac{V'_2 - V'_1}{q\,\Delta x}, \tag{55}$$

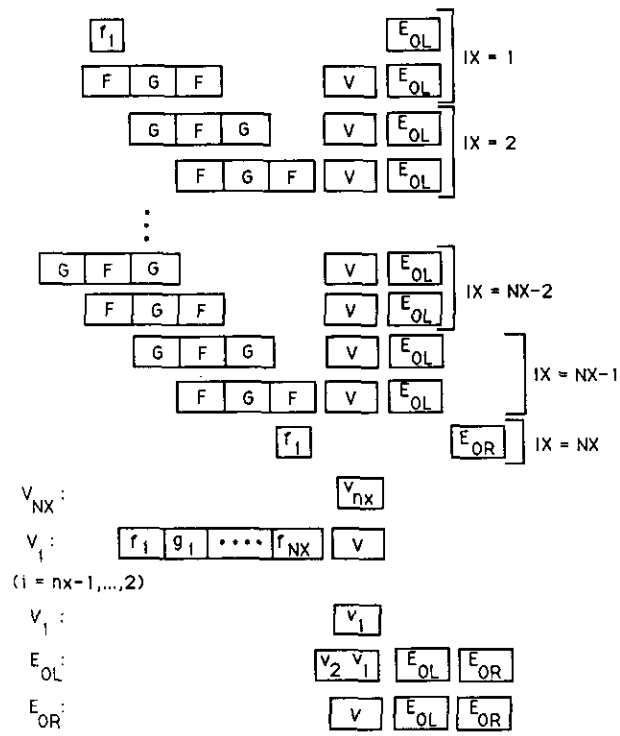

FIG. 1. Structure of the Jacobian matrix entries for solving the self-consistent Wigner function equations.

and

$$E_{0R} = \frac{V'_{N_x} - V'_{N_x-1}}{q\,\Delta x}. \tag{56}$$

Figure 1 shows the resulting matrix structure (for the calculations here an expanded matrix structure was used that contains more entries near the boundaries, since many different boundary treatment methods have been tried in the course of this work). The ordering of the equations (rows) for this matrix exactly corresponds to the ordering of unknowns described earlier. In the figure, the values enclosed in boxes indicate either single or multiple non-zero entries of the matrix. If the enclosed value is indexed, such as f_i or E_{0L}, a single entry is indicated; an enclosed value without an index indicates a block of variables. For example, the enclosed V indicates the block of potential entries $V'_{N_x}, \ldots, V'_1$, and an enclosed F indicates a block of all f-type Wigner function values at a particular meshpoint. The matrix is sparse, so that all entries not indicated with boxes are zero. Some of the entries in Fig. 1 are not directly generated by the equation corresponding to that row but are generated by fills during the Gaussian elimination matrix solution process. Newton's method is used, so that at each iteration a matrix equation of the following form must be solved:

$$\mathbf{J}\,\Delta\mathbf{U} = -\mathbf{EQ}. \tag{57}$$

In Eq. (57), **J** is the Jacobian matrix with the structure of Fig. 1, which is defined as follows: Let eq_i be the equation corresponding to the ith unknown u_i, set up so that eq_i is exactly zero when the equation is solved. Then the (i, j)th entry of J is given by

$$J_{ij} = \frac{\partial\,\mathrm{eq}_i}{u_j}. \tag{58}$$

In Eq. (57), $\Delta\mathbf{U}$ is a vector of the predicted changes for the next iteration of all the variables; and **EQ** is a vector containing the present numerical value of each equation.

For calculation, the unknowns are normalized as follows. Wigner function values are divided by the maximum value of the equilibrium distribution function f_0, potential energy variables are divided by q, and the two electric field variables are multiplied by Δx so they are in units of V.

As an example of the method, a GaAs–$Ga_{0.7}Al_{0.3}As$ structure is simulated with 16.8 Å barriers and a 44.8 Å well. The barrier height is 0.246 eV, and a constant effective mass of $m^* = 0.067m_0$ is used. The donor doping concentration is 10^{24} m^{-3} in the contact regions, and the barriers and well are undoped. The overall device length simulated is 442.4 Å with a space step of $\Delta x = 5.6$ Å, so that $N_x = 80$. For the k-space discretization, $N_k = 61$ is used; according to Eq. (4), $\Delta k = 9.1967 \times 10^7$ m^{-1}. The scattering relaxation time τ is 1.9045×10^{-13} s, corresponding to a low-field mobility of 0.5 m^2/V-s, and the temperature is 300°K.

For this particular problem, the Jacobian matrix of Fig. 1 requires 2,961,765 entries, and since the calculation must be performed in double precision this requires approximately 23.69 Mbytes of numerical storage (in fact, more storage than this is required when the other arrays are included; however the Jacobian matrix is by far the most demanding storage requirement). Since the memory requirements increase faster than linearly with the problem size, it can be seen that the level of complexity of problems that can be treated using this method is limited. (The above memory requirement assumes that full advantage is taken of the sparse nature of the Jacobian matrix, so that only the non-zero matrix entries are processed.)

For the zero-bias case, an initial condition for the DC iterative method was generated by assuming a flat potential profile and solving for an ensemble of energy eigenstates of Schrödinger's equation [13] over the set of k-values used in the f discretization (except that $k = 0$ is not included). Each of these eigenstates produces a contribution to f according to Eq. (3) and to g according to Eq. (10). Starting from this initial condition, five Newton iterations were required to obtain convergence below a tolerance of 10^{-5} (this is determined by monitoring the maximum change of the variables per iteration with respect to the reference values). Figure 2a shows the resulting zero-bias solution for electron concentration and potential energy, where the self-consistent

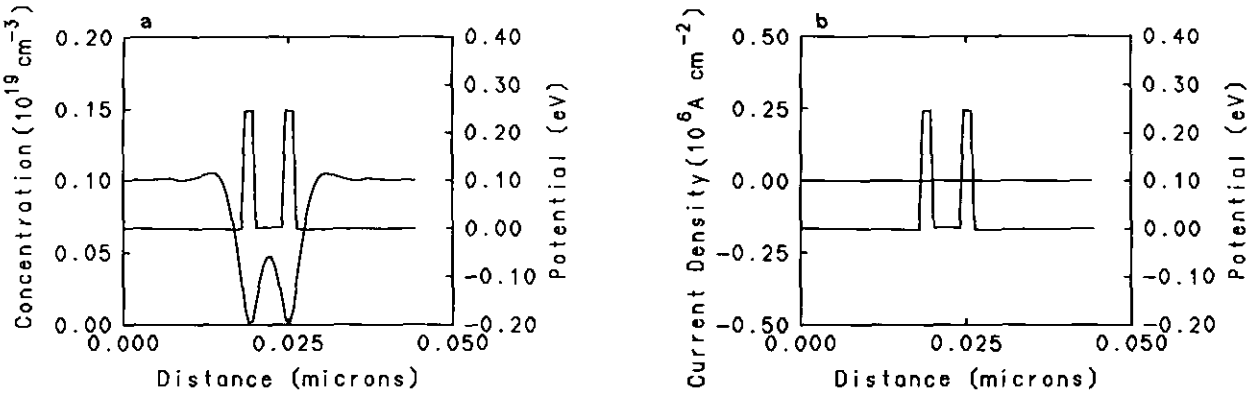

FIG. 2. Zero-bias solution for the GaAs-$Ga_{0.7}Al_{0.3}As$ structure showing potential profile with (a) electron concentration profile and (b) current density profile.

coupling between space charge and potential energy can be observed. Figure 2b shows the zero-bias current density, which is constant (except for some numerical error) as it should be since the discrete equations satisfy the current continuity equation according to Eq. (21), which implies spatially constant current density for the DC case.

Figure 3a shows the electron concentration and Fig. 3b shows the electron current density for a DC solution with an applied voltage of 0.05 V (note that negative current density in the figure implies that electrons are flowing to the right). It can be seen that severe problems are encountered at this bias voltage. The most distressing is that the electron density (calculated using Eq. (19)) is undershooting below zero. Also, oscillations in the electron concentration profile indicate that there are boundary reflection problems. However, the fact that the current density profile of Fig. 3b is constant indicates that the DC equations are being solved correctly.

Because of such problems, it has not yet been possible to obtain a complete I–V (current-voltage) curve for a resonant tunneling device through the negative differential resistance region using this method. Some possible reasons which should be further investigated are as follows. First, the scattering formulation given in Eqs. (39) and (40) is not entirely consistent, since a scattering term from f states into g states is present in the g-equations, without a corresponding loss term in the f-equations. The reasons that this form was adopted for the present calculations are, first, that some coupling of the g scattering terms with f terms is required in order to have a well-defined solution for g at equilibrium, far removed into the contacts. However, if a corresponding loss term is put in Eq. (39) to account for f terms scattering to g states, then the sum over all k_m of the scattering terms is not zero, so that current continuity is lost. It may be that a more generalized definition of current density than that given by Eq. (23) is required to handle this coupling, or perhaps the reader can come up with a different scattering rate formulation which avoids this problem.

Another likely problem that may be causing the difficulties evident in Fig. 3 is that the device may simply not be long enough for the equilibrium boundary conditions set on f to be valid. These boundary conditions assume that the f distribution has relaxed to an equilibrium condition (characterized by the electric field in the respective contact)

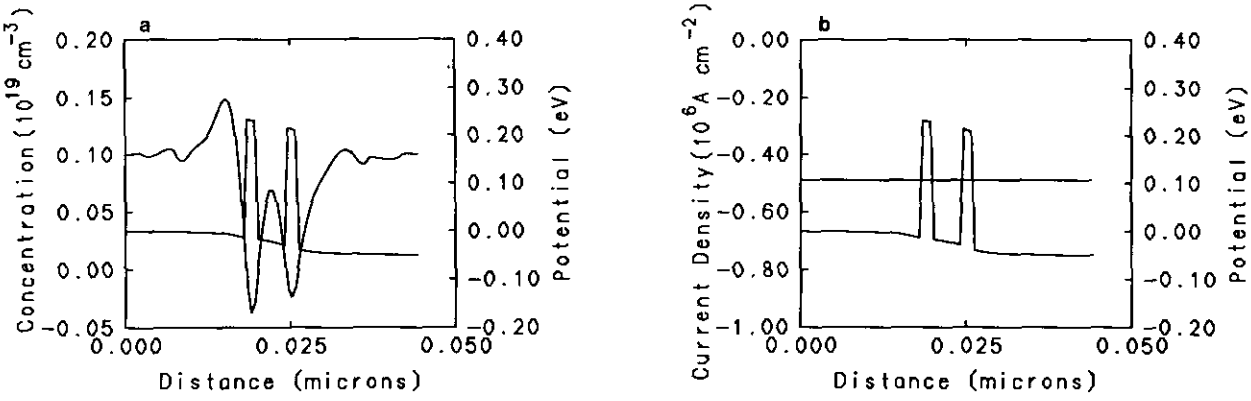

FIG. 3. DC solution with applied voltage of 0.05 V showing potential energy and (a) electron concentration (b) current density profile.

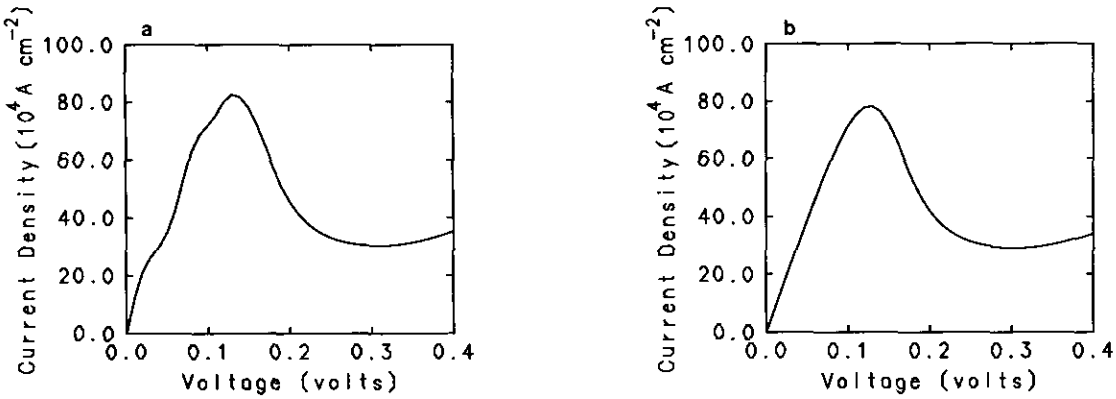

FIG. 4. I–V calculation using the Schrödinger equation (a) for the case with 60 points in k-space and (b) for the case with 200 points in k-space.

by the time the contacts are reached; if this is not the case, reflections will be set up at the boundary. The difficulty at present with testing this hypothesis is that the CPU memory storage and time requirements increase rapidly with the length of the region simulated.

A further potential problem, whose resolution is also limited by available computer resources, is the discretization used in k-space. Figure 4a shows a calculation of the I–V curve for this device using an ensemble of eigenstates from the Schrödinger equation [13], where 60 points were used in k-space (i.e., 60 eigenstates were determined). This k-space discretization corresponds to that used in the Wigner function calculations. It can be seen that the curve is not smooth. To show that the wiggles are an effect of the rather coarse discretization in k-space, Fig. 4b shows the results of a similar calculation but where 200 points in k-space were used (the k-space limits k_{max} and k_{min} were basically the same for the two discretizations). It can be seen that the discretization using 60 points is not sufficient for this problem. The inadequacy may become worse when self-consistency is taken into account, where a "notch" forms on the cathode side of the device which can accommodate (nearly) bound electron states that constitute the accumulation layer on the left-hand side (see Fig. 3).

In spite of these problems, it is believed that the method presented in this paper is a more accurate numerical formulation of the Wigner function equations than has been previously developed. However, further work must be carried out to address these problems, in order to determine the potential of this method for quantum device simulation.

7. CONCLUSION

An accurate formulation of the equations for the Wigner function has been presented, based on a finite-difference discretization of the Schrödinger equation. The resulting matrix structure and the method for solving the equations have been presented. Preliminary results have been shown which indicate that further problems must be resolved before the method can reliably be used. These problems and potential solutions have been discussed.

ACKNOWLEDGMENTS

This work was supported by the U.S. Army Research Office under the URI program, Contract No. DAAL03-87-K-0007. The authors thank Dr. William Frensley for helpful discussions.

REFERENCES

1. W. R. Frensley, *Phys. Rev. Lett.* **57**, 2853 (1986); **60**, 1589(E) (1988).
2. W. R. Frensley, *Phys. Rev. B* **36**, 1570 (1987a); **37**, 10,379(E).
3. W. R. Frensley, *Solid State Electron.* **31**, 739 (1988).
4. R. K. Mains and G. I. Haddad, *J. Appl. Phys.* **64**, 5041 (1988).
5. N. C. Kluksdahl, A. M. Kriman, D. K. Ferry, and C. Ringhofer, *Phys. Rev. B* **39**, 7720 (1989-I).
6. W. R. Frensley, in *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic Press, New York, 1989), p. 231.
7. W. R. Frensley, *Solid State Electron.* **32**, 1235 (1989).
8. K. L. Jensen and F. A. Buot, *J. Appl. Phys.* **65**, 5248 (1989).
9. K. L. Jensen and F. A. Buot, *J. Appl. Phys.* **67**, 2153 (1990).
10. S. R. deGroot and L. G. Suttorp, *Foundations of Electrodynamics* (North-Holland, Amsterdam, 1972).
11. W. R. Frensley, *Rev. Mod. Phys.* **62**, 745 (1990).
12. A. Goldberg, H. M. Schey, and J. L. Schwartz, *Am. J. Phys.* **35**, 177 (1967).
13. R. K. Mains and G. I. Haddad, *J. Appl. Phys.* **64**, 3564 (1988).
14. C. Ringhofer, D. Ferry, and N. Kluksdahl, *Transp. Theory Statist. Phys.* **18**, 331 (1989).
15. D. Neuhauser, M. Baer, R. S. Judson, and D. J. Kouri, *J. Chem. Phys.* **93**, 312 (1990).
16. M. Baer and H. Nakamura, *J. Chem. Phys.* **96**, 6565 (1992).
17. R. K. Mains and G. I. Haddad, *Appl. Phys. Lett.* **55**, 2631 (1989).