Nonlinear Discretization Scheme for the Density-Gradient Equations

M.G. Ancona
Electronics Science and Technology Division
Naval Research Laboratory, Washington, DC 20375
ancona@estd.nrl.navy.mil

B.A. Biegel
Computer Sciences Corp.
NASA Ames Research Center, Moffett Field, CA 94035

Abstract-**A nonlinear three-point discretization of the density-gradient equations is presented. The new method, an exponential-fitting scheme, is evaluated using numerical examples involving both quantum confinement and tunneling. The nonlinear discretization is shown to perform far better than the conventional linear version allowing for a substantial easing in the mesh refinement especially in tunneling problems.**

I. INTRODUCTION

Density-gradient (DG) theory enables engineering-oriented analyses of electronic devices in which quantum confinement and tunneling phenomena are significant [1-3]. Recently this approach has moved closer to practicality with its application to confinement problems in multi-dimensions by several groups [4,5]. The density-gradient term that represents the quantum effects in the governing equations is a singular perturbation that gives rise to a very thin boundary layer that exists inside the normal Debye boundary layer of the classical theory [6]. Within a DG boundary layer, e.g., in a tunnel barrier, the carrier densities can vary by orders of magnitude on an Ångstrom scale. To resolve this physics a very fine grid is required and this is especially important if the tunneling currents are to be computed accurately. In previous work [1-3] we have employed a standard linear centered discretization for such problems and found it to be accurate but only when exceedingly fine meshes are used, e.g., with mesh spacings as small as 0.01Å. This situation was manageable for 1-D cases, however, for multi-dimensional applications it could impose a significant burden. Hence our interest in non-linear discretizations.

II. DENSITY-GRADIENT THEORY

DG theory is a macroscopic approximation to quantum mechanics in which the nonlocality of quantum mechanics enters the theory solely via an assumption that the equations of state of the electron and hole gases depend not only on the gas densities as in diffusion-drift theory but also on the *gradients* of their densities [1]. The theory built on this simple foundation, though obviously incomplete, has been found to be surprisingly accurate for describing effects of quantum confinement and tunneling in a number of practical situations. Furthermore, because of its simplicity the theory can more readily and efficiently incorporate other complications of real devices such as multi-dimensionality that are not so easily treated by alternative microscopic theories. Finally, and of most relevance to this paper, since DG theory is a generalization of standard diffusion-drift theory, it fits naturally into the framework of conventional device simulation and many of the well-known ideas and methods of the latter carry over directly.

Considering only electrons and neglecting inertia the governing differential equations of DG theory (in 1-D) are

$$J = \mu_n n \phi_x - D_n n_x , \quad J_x = -R$$

$$\frac{2}{s}[b s_x]_x + \psi - \phi = 0 , \quad s = \sqrt{n} \qquad (1)$$

$$[\varepsilon \psi_x]_x = q(n - N_D)$$

As discussed in Ref. 2 the equations appropriate for describing the transport inside tunnel barriers are slightly more complicated in that the electron gas splits into forward-tunneling and back-tunneling electrons. Nevertheless, the form of the equations remains the same and so this issue need not be discussed here. Similarly, the boundary conditions are exactly as described in Ref. 2 and therefore they need not be treated explicitly here either. The neglect of inertia in the above equations means we have subsumed its effect into "virtual anode" boundary conditions, an issue also discussed in Ref. 2. In the numerical simulations of this paper all of these considerations are included. The full system of equations with equations describing electron and hole transport and electrostatics are solved simultaneously.

III. NON-LINEAR DISCRETIZATION

With respect to the electron current equations, $(1)_2$ with $(1)_1$, a nonlinear discretization is already very well-known in the form of the Scharfetter-Gummel scheme [7,8]. Indeed this was the key discovery that made diffusion-drift analysis the standard approach to multi-dimensional device simula-

0-7803-6279-9/00/$10.00 © 2000 IEEE

tion for more than 30 years. This scheme belongs to the general class of numerical techniques for solving partial differential equations known as exponential-fitting methods [9]. It gains its efficiency by building in *a priori* the exponential dependence of the carrier densities on potential. This improvement in the representation of the solution functions translates into a scheme which can achieve greater accuracy and stability on coarser meshes. Of course, on finer meshes the advantage of the nonlinear method degrades because linearization starts to become a better approximation. For the electron current equation the Scharfetter-Gummel discretization is [7]

$$J_{i+1/2} = -\frac{kT}{q}\frac{(\mu_n F)_{i+1/2}}{h_i}\left\{B\left[\frac{q\Delta\phi_i}{kTF_{i+1/2}}\right]n_{i+1} - B\left[-\frac{q\Delta\phi_i}{kTF_{i+1/2}}\right]n_i\right\} \quad (2)$$

where

$$h_i = x_{i+1} - x_i\,, \quad \Delta\phi_i = \phi_{i+1} - \phi_i\,, \quad B(y) = \frac{y}{e^y - 1}$$

and $F_{i+1/2}$ is a numerical factor that equals one in the Maxwell-Boltzmann case and is density-dependent in the Fermi-Dirac case [10]. The nonlinearity in this scheme enters through the Bernoulli function B(y) which is plotted in Fig. 1. As discussed in Ref. 8 this function can be viewed as stabilizing the scheme by adding in numerical diffusion when $\Delta\phi_i$ is large.

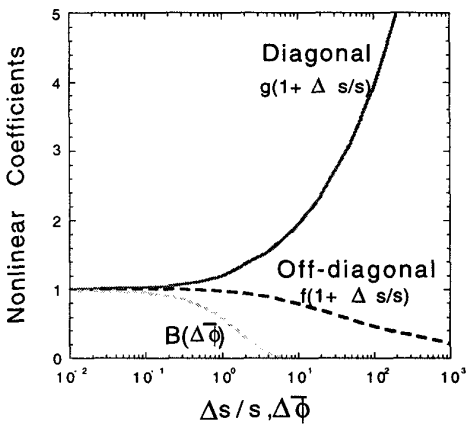

Fig. 1. The functions B(y), f(z) and g(z) versus the normalized differences in ϕ and s between grid points. As Δs increases the DG discretization becomes increasingly diagonally dominant.

The physics of the DG equation $(1)_3$ produces similar exponential character in the solution functions (especially in tunneling situations) and building this behavior into the numerical scheme can be expected to have similar benefits. However, a numerical implementation directly analogous to Scharfetter-Gummel does not appear to be feasible in the DG case. The problem is that in diffusion-drift theory the diffusion is balanced by drift so that, although the carrier density varies exponentially, its associated current density varies only slowly and so can be given a linear discretization. In contrast, in DG theory there is no mathematical equivalent to "drift" so the theory's equivalent "current" also varies exponentially and hence it too must be given an exponential discretization. Therefore to develop an exponential-fitting scheme for the DG equation $(1)_3$ we instead start by assuming that the solution between grid points i−1 and i+1 is of the form:

$$\begin{aligned} s &= Ae^{\alpha x} \quad \text{for} \quad x \in [x_{i-1}, x_i] \\ s &= Be^{\beta x} \quad \text{for} \quad x \in [x_i, x_{i+1}] \end{aligned} \quad (3)$$

This obviously allows for a much better representation of the exponential character of the DG boundary layer. The four parameters A, B, α and β are determined by equating s to its grid point values at i-1, i and i+1. Then integrating the DG equation $(1)_3$ from i−1/2 to i+1/2 and assuming that in this region $\phi - \psi \equiv (\phi - \psi)_i$, we obtain the nonlinear discretization

$$\frac{b_{i+1/2}}{h_i} f\left(\frac{s_{i+1}}{s_i}\right)(s_{i+1} - s_i) - \frac{b_{i-1/2}}{h_{i-1}} f\left(\frac{s_i}{s_{i-1}}\right)(s_i - s_{i-1}) - \frac{1}{2}(\phi_i - \psi_i)\left[h_i g\left(\frac{s_{i+1}}{s_i}\right) + h_{i-1} g\left(\frac{s_{i-1}}{s_i}\right)\right]s_i = 0 \quad (4)$$

where the functions

$$f(z) = \frac{\sqrt{z}}{z - 1}\ln(z) \quad \text{and} \quad g(z) = \frac{\sqrt{z} - 1}{\ln(z)}$$

are plotted in Fig. 1. For small z these functions approach one and the scheme reduces to the linear centered discretization while for larger z the nonlinearity leads to increased diagonal dominance.

One final point concerning these discretizations, whether linear or nonlinear, is that in diffusion-drift theory (i.e., with b = 0) $\phi = \psi$, so (2) inserted into a discretized version of $(1)_2$ is a finite difference equation for n and the discretized version of $(1)_5$ is the equation for ψ. In DG theory instead, $(1)_5$ is again the equation for ψ but since $\phi \neq \psi$, $(1)_2$ with (2) becomes the equation for ϕ and (4), derived from $(1)_3$, is the equation for n.

IV. Simulation Results

As a primary test problem for the nonlinear discretization we study MOS capacitors both with and without tunneling in 1-D. In Fig. 2 the carrier concentration profiles calculated by DG theory for a non-tunneling capacitor biased into inversion (holes) are shown. The decline in the hole density near x = 0 is caused by the quantum repulsion associated with the Si-SiO_2 barrier. In the Figure we compare the carrier concentrations obtained using a coarse grid with 11 points spread non-uniformly over 0.5μm and employing the linear and nonlinear discretizations with an "exact" result computed using a fine grid (145 points). Both the linear and nonlinear discretizations do quite well, however, the nonlinear version does do significantly better especially at lower densities. This is seen more clearly in a plot of the relative

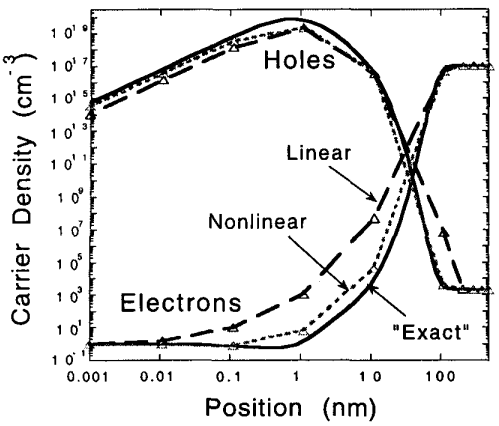

Fig. 2. Comparison of carrier density profiles computed with linear and nonlinear discretizations for a 0.5μm thick inverted semiconductor using 11 non-uniform grid points.

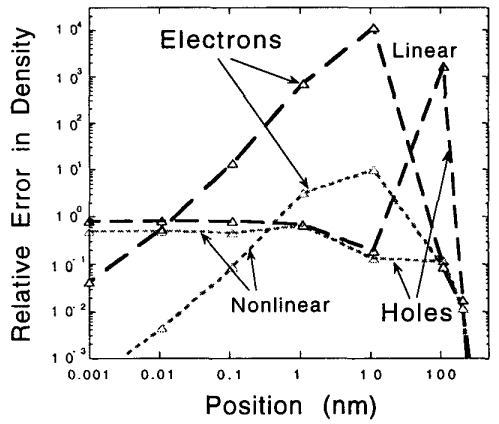

Fig. 3. Relative error in carrier densities computed by DG theory for the device of Fig. 2 using linear and nonlinear discretizations. The errors occur where the carrier densities change most rapidly and are considerably larger for the linear discretization.

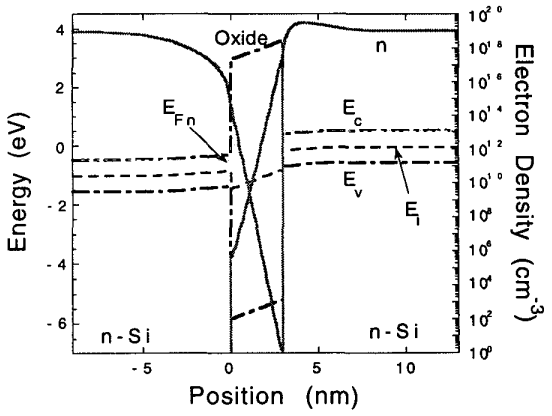

Fig. 4. Band diagram and electron concentration profiles in a silicon (100nm)-oxide (30Å)-silicon (100nm) structure as computed by DG theory on a fine grid.

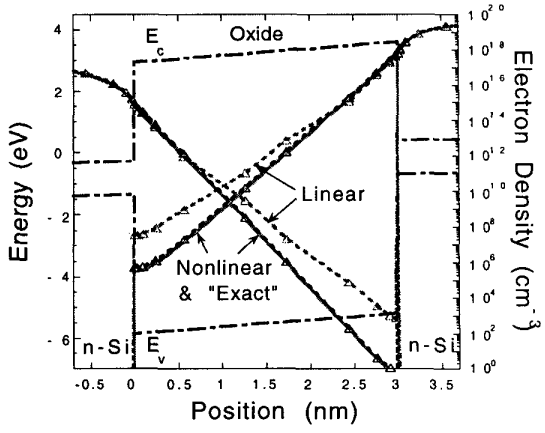

Fig. 5. Forward and backward electron density profiles computed with linear and nonlinear discretizations on a coarse grid (triangles) with 12 grid points in the 3nm barrier and 13 in each 100nm thick semiconductor region.

errors shown in Fig. 3. As expected, the biggest errors occur when the change in density between grid points is largest.

In Figs. 4 and 5 the calculated band diagrams and electron concentrations are plotted for an SOS capacitor with a 30Å gate oxide biased at -1V and with tunneling included. That tunneling is occuring is evident from the exponentially decaying electron concentration profiles (for forward and backward electrons [2]) inside the barrier. The expanded view in Fig. 5 compares the "exact" solution (258 oxide mesh points) for the electron densities with results obtained using a coarse grid with 12 mesh points distributed non-uniformly in the oxide and using both linear and nonlinear discretizations. For this tunneling situation the superiority of the nonlinear discretization is even more pronounced. Most importantly, the current density through the barrier is proportional to the carrier densities at the "downstream" end of the barrier. So for example the current associated with electrons tunneling from right to left (the dominant process under these bias conditions) will be proportional to the left-going electron density at the left end (x = 0) whose "exact" value in this case is $3.3\text{x}10^6\text{cm}^{-3}$. But notice that it is at this point that the largest relative discrepancy between the linear and nonlinear discretizations occurs. In effect, the errors accumulate as one moves across the barrier by virtue of the hyperbolic nature of the equations inside the barrier. And so the calculated current is especially sensitive to discretization error. Because of this and because the I-V characteristic is typically what one is most interested in determining, the error criterion on the current generally dictates the mesh refinement. And with a linear discretization it turns out that to obtain reasonable accuracy in the current can require mesh spacings as small as 0.01Å. (With a uniform grid such a small mesh spacing would be almost unworkable; however, the mesh refinement is most critical near the downstream edge and so, by using a non-uniform grid, one can reduce the computational burden significantly). The nonlinear discretization dramatically eases the grid

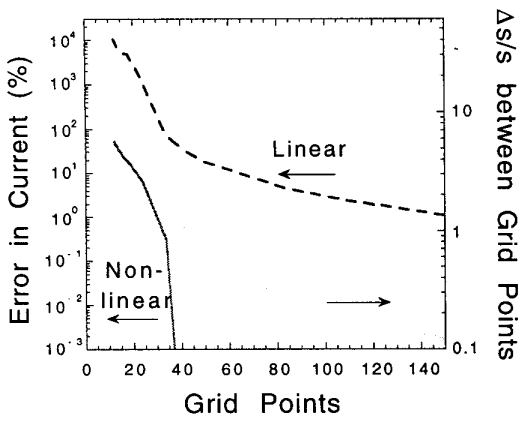

Fig. 6. Relative error in tunneling current densities computed by DG theory for the device of Figs. 3/4 using linear and nonlinear discretizations as a function of the number of barrier grid points. The maximum Δs/s is also shown.

requirements for accuracy in the computed current. To illustrate the benefits of the nonlinear discretization on the calculation of current, in Fig. 6 we plot the errors in the current as computed with both the linear and nonlinear discretizations as functions of the number of (non-uniform) grid points in the barrier. The nonlinear discretization far outperforms the linear version. Also shown in the Figure is the maximum relative difference in s between adjacent grid points showing that the discretization is indeed strongly nonlinear.

It should be recognized that the simulations of Figs. 2-6 test primarily the discretization of the DG equation $(1)_3$ because its resolution requirements are such that a linear discretization of the current equation $(1)_2$ with $(1)_1$ will be adequate. The main reason for continuing to use a nonlinear (Scharfetter-Gummel) discretization for the current equations is that in a typical device in which quantum effects are important, e.g., an EEPROM, large portions of the device actually operate in a classical (diffusion-drift) regime. And for the simulation in these regions one would want to use the relaxed grids made possible by a nonlinear discretization of the diffusion-drift equation.

Finally, we remark that just as with the Scharfetter-Gummel discretization, the primary benefit of the nonlinear discretization of the DG equations is increased efficiency which will be especially pronounced in multi-dimensions. To explore the cost savings involved, efforts are currently underway to carry out 2-D and 3-D device simulations in which the new nonlinear discretization will be used. For this purpose, we are employing the flexible simulator PROPHET developed at Lucent Technologies [4].

V. Summary

In this work we have demonstrated the efficacy of a nonlinear discretization for the numerical solution of the density-gradient equations. The nonlinear discretization is an exponential-fitting scheme that has essentially the same motivation as the conventional Scharfetter-Gummel discretization, although its derivation and details are somewhat different. We have illustrated the new method in both quantum confinement and tunneling situations. The method allows rather coarse grids to be used with DG theory and thereby increases the viability of this theory from a computational standpoint in multi-dimensions.

Acknowledgement: The first author thanks ONR for funding support.

References

1. M.G. Ancona and H.F. Tiersten, *Phys. Rev.* **B35**, 7959 (1987).
2. M.G. Ancona, *Phys. Rev.* **B42**, 1222 (1990).
3. M.G. Ancona, Z. Yu, R.W. Dutton, P.J. Vande Voorde, M. Cao and D. Vook, *Proc. SISPAD*, 235 (1999); a more complete version of this work is to be published in *IEEE Trans. Elect. Dev.* (2000).
4. C.S. Rafferty, B.A. Biegel, Z. Yu, M.G. Ancona, J. Bude and R.W. Dutton, *Proc. SISPAD*, 137 (1998); B.A. Biegel, C.S. Rafferty, M.G. Ancona and Z. Yu, submitted to *IEEE. Trans. Elect. Dev.* (2000).
5. A. Asenov, G. Slavcheva, A.R. Brown, J.H. Davies and S. Saini, *IEDM Tech. Dig.*, 535 (1999).
6. M.G. Ancona, *Proc. Comp. Elect. Wrkshp.*, Champaign-Urbana, IL (1990).
7. S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer-Verlag, New York (1984).
8. An alternative nonlinear discretization of the diffusion-drift equations has been given by J. Kreskovsky, *IEEE Trans. Elect. Dev.* **34**, 1128 (1987).
9. E. Oran and J. Boris, *Numerical Simulation of Reactive Flow*, Elsevier, New York (1987).
10. V.C. Aguilera-Navarro, G.A. Estevez and A, Kostecki, *J. Appl. Phys.* **63**, 2848 (1988).