The continuity equations of the drift-diffusion semiconductor device model are hard to discretize because of the severe variations of variables. The Scharfetter - Gummel scheme applies harmonic averages to the conductance function, and usually produces acceptable solutions despite the sharp parameter changes. Some authors attribute the success of this scheme to the relative smoothness of the currents, as compared with carrier concentrations. In this report it is shown that current smoothness cannot be derived from the differential equations, but is related to the specific boundary condition configuration. The success of the Scherfetter - Gummel method is hence easy to justify for nearly 1-D devices, but is harder to justify for some other geometries. A stream function formulation related to that scheme is shown to overcome cases where the continuity equations are ill - conditioned.

Discretization of the Steady State Semiconductor Device Equations

I. Efrat

Research Report YALEU/DCS/RR-745
September 1989

The authors was supported by the Office of Naval Research under Grant N00014-86-J-1906

1 introduction

The drift-diffusion semiconductor equations are usually discretized by the Scharfetter - Gummel scheme. This method applies harmonic averages to the carrier conductance function [2].
Assuming 1-dimensional, constant current, and piecewise linear potential variation, this scheme is exact. Mock [6] shows that the Scharfetter - Gummel discretization is equivalent to a certain finite - element formulation for a stream potential. He attributes the success of this scheme to the experimental fact that currents in semiconductors are usually smoother than carrier concentrations.
Brezzi [3] shows the relaion of this scheme to a hybrid finite element one, where the currents are discretized by constant elements, and carrier concentrations by linear ones. Brezzi further states that "... a solid theoretical argument showing the superiority of the harmonic average is still to be found, in our opinion, except for the obvious 1-D case.." In this report it is shown that in the 2-D case, the equation for the electron stream function is equivalent to the equation for the hole Slotbloom variable, and similarly: the equation for the hole stream function is equivalent to the equation for electron Slotbloom variable.
Smoothness of the stream function can hence be attributed only to the different boundary conditions. The advantage of harmonic averages is consequantly easy to justify for long, narrow devices. It may however not exsist in some other geometrical configurations.
While the Scherfetter - Gummel discretization and the stream function scheme produce the same solution (assuming infinite computational precision), it is shown in section 3 that the latter may be far better numerically conditioned.

2 Model and Discretization

The scaled steady state drift - diffusion semiconductor equations, expressed in Slotbloom Variables are:

$$\lambda^2 \Delta\psi - e^{\psi}u - e^{-\psi}v - C(x) = 0 \tag{1}$$

$$\nabla \cdot e^{\psi} \nabla u = R \tag{2}$$

$$\nabla \cdot e^{-\psi} \nabla v = R \tag{3}$$

Here ψ is scaled electrostatic potential, u and v are exponentials of the pseudo-Fermi potentials, C is the doping concentration, and R is the recombination rate. The electron and hole concentrations n, p are related to u, v by the relations: $n = e^{\psi}u$, $v = e^{-\psi}v$.
For this discussion it will be assumed that the boundary conditions for ψ, u, v are of Dirichlet type on a portion Γ_D of the boundary Γ, and of Neuman type on Γ_N where $\Gamma_N = \Gamma \backslash \Gamma_D$. The Dirichlet portion corresponds to Ohmic boundary conditions, and the Neumann portion to reflecting boundaries. In some cases the carrier concentration variables have internal Neumann boundaries which are not on part of Γ_N. This occurrs at insulator - semiconductor interfaces.
The discretization of (2)(3) may introduce a large truncation error since u, v along with the coefficients e^{ψ}, $e^{-\psi}$ often vary sharply between discretization nodes.

2.1 Scherfetter - Gummel discretization

For any 2 adjacent nodes i, j of a discretization grid, the i, j'th entry in the Jacobian matrix is taken to be:

$$a_{ij}^{(n)} = \frac{c_{ij}}{d_{ij}} d^{\psi_j} B(\psi_j - \psi_i) \tag{4}$$

$$a_{ij}^{(p)} = \frac{c_{ij}}{d_{ij}} d^{-\psi_j} B(\psi_i - \psi_j) \tag{5}$$

where d_{ij} is the distance between the nodes, c_{ij} is a cross - section through which the current is assumed to flow, and $B(y)$ is Bernulli's function:

$$B(y) = \frac{y}{e^y - 1}. \tag{6}$$

This scheme can be derived as an exact solution, assuming constant current flowing parallel to the edge connecting the nodes i, j, and piecewise linear electrostatic potential [5] [2].

2.2 Stream Functions

Assume for simplicity $R \equiv 0$. This assumption will be dropped later. If we denote the electron and hole currents by:

$$J_n = e^{\psi} \nabla u \tag{7}$$

$$J_p = e^{-\psi}\nabla v \tag{8}$$

then (2)(3) can be rewritten as carrier conservation statements:

$$\nabla \cdot J_n = 0 \tag{9}$$

$$\nabla \cdot J_p = 0 \tag{10}$$

A solenoidal vector field can be expressed as a curl of a 'stream' function [7], hence there exist vectors $\vec{\theta}_n$,$\vec{\theta}_p$ such that:

$$J_n = \nabla \times \vec{\theta}_n \tag{11}$$

$$J_p = \nabla \times \vec{\theta}_p \tag{12}$$

In $\mathbf{R}^2$ the vectors $\vec{\theta}_n$, $\vec{\theta}_p$ point to the third direction, and consequently can be regarded as scalar quantities θ_n, θ_p.
Dividing (7),(8) by e^{ψ},$e^{-\psi}$ respectively, and applying the curl operator result in:

$$\nabla \times e^{-\psi} J_n = \nabla \times \nabla u = 0 \tag{13}$$

$$\nabla \times e^{\psi} J_p = \nabla \times \nabla v = 0 \tag{14}$$

Substituting (11), (12) we obtain:

$$\nabla \times e^{-\psi} \nabla \times \vec{\theta}_n = 0 \tag{15}$$

$$\nabla \times e^{\psi} \nabla \times \vec{\theta}_p = 0 \tag{16}$$

In $\mathbf{R}^2$ these reduce to:

$$\nabla \cdot e^{-\psi} \nabla \theta_n = 0 \tag{17}$$

$$\nabla \cdot e^{\psi} \nabla \theta_p = 0 \tag{18}$$

where θ_n,θ_p should be regarded as the scalar magnitude of the corresponding vectors.
Comparison of (2),(3) vs. (17), (18) reveals the following result:

Theorem 1 *In $\mathbb{R}^2$ and without recombination, the equation for u is equivalent to the equation for θ_p, and the equation for v is equivalent to the equation for θ_n.*

At this point, if we are determined to justify the empirical observation that θ_n and θ_p are smoother than u, v then the last resort is the boundary condition configuration.

2.3 Boundary Conditions

The nature of the boundary conditions for θ_n, θ_p is revealed via the following lemma:

Lemma 1 *The equipotential curves of θ_n are tangent to J_n, and normal to the equipotential curves of u. The same statement applies to θ_p, J_p, v.*

Proof: $\nabla\theta_n \perp \nabla u$ since

$$\nabla\theta_n \cdot \nabla u = \nabla\theta_n \cdot e^{-\psi} J_n = \nabla\theta_n \cdot e^{-\psi}\nabla \times \theta_n = 0 \tag{19}$$

The curves $\theta_n = const.$ are normal to $\nabla\theta_n$, hence parallel to ∇u which is J_n/e^{ψ} The curves on which $u = const.$ are normal to ∇u hence also to the curves $\theta_n = const.$ The proof for θ_p, J_p, v is identical. □

Corollary 1 . *θ_n is constant on contiguous segments of Γ_N, and $\partial\theta_n/\partial\nu = 0$ on Γ_D. Similar statements apply to θ_p.*

The resulting problem for θ_n, θ_p is still under determined. For each segment of Γ_N the actual constant value of θ_n, θ_p has to be found. The necessary additional constants are easily derived from the Dirichlet boundary conditions for u, v. Let the Dirichlet boundary be:

$$\Gamma_D = \bigcup_{i=1}^{m} \Gamma_D^i \tag{20}$$

Where each Γ_D^i is connected. Integrating (7), (8) we obtain:

$$u(x_1) = u(x_0) + \int_C J_n(s) e^{-\psi(s)} ds \tag{21}$$

$$v(x_1) = v(x_0) + \int_C J_p(s) e^{\psi(s)} ds \tag{22}$$

where C is any path leading from x_0 to x_1. Choosing arbitrary curves from Γ_D^1 to the $m-1$ other segments, equations of the type (21),(22) give $m-1$ constraints:

$$u_i - u_1 = \int_{C_i} (\nabla \times \vec{\theta}_n) e^{-\psi(s)} ds \tag{23}$$

$$v_i - v_1 = \int_{C_i} (\nabla \times \vec{\theta}_p) e^{\psi(s)} ds \tag{24}$$

There remains an arbitrary constant in the problem, which is physically immaterial. When $m = 1$ then θ_n, θ_p are both constants, and there is no current flowing through the device.
We can now attempt to solve the problem by the following, loosely defined algorithm:
Algorithm 1.

1. Discretize (17),(18).
2. Discretize the m-1 constraints (23),(24)
3. Solve the resulting (under-determined) linear system.
4. Integrate for the values of u, v using (21), (22).

Mock [6] showed that the Scharfetter - Gummel discretization scheme is equivalent to a finite element version of this algorithm, using quadralateral piecewise constant elements for θ_n, θ_p.

2.4 Nonzero Recombination

If $\nabla \cdot J = R \neq 0$ the J is not solenoidal, and cannot be represented as a curl of a vector potential. However if we definethe 2 current components to be:

$$\hat{J}_1(x, y) = J_1(x, y) + \int^x R(t, y) dt \tag{25}$$

$$\hat{J}_2(x, y) = J_2(x, y) \tag{26}$$

then $\hat{J}$ is a solenoidal vector and the preceeding derivation can be applied to it.

3 examples

3.1 A Capacitor

A semiconductor capacitor is a simple 1D device insulated at one end. Assume a p - type silicon piece insulated on the left. A contact attached to the insulator is called the 'gate contact', and a contact at the other end of the device is called the 'source contact'. When a positive bias is applied between gate and source, holes are repelled from the insulator and a depletion layer of unbalanced acceptor ions is formed next to the gate. At a higher voltage there appears a thin layer of electrons near the gate. This layer is called the 'inversion layer'. Being extremely thin and steep, the inversion layer is very hard to resolve by discretization. While being a rather simple device, capacitors are often imbedded in more complex semiconductor devices (particularly MOS devices), and the difficulties mentioned above are inherited by their simulations.

It should be mentioned here that while n, p have a severe boundary layer at the gate, u, v are actually constant for a capacitor. Nevertheless, the conservation equations for either the n, p variables, or for u, v are highly ill-conditioned [1].

The solution of such a capacitor by Pisces [4] requires an extreme mesh refinement at the inversion layer. For gate to source bias which exceeds 0.6 volts, the Jacobian matrices ocurring in the Newton process become too ill conditioned, and Pisces ceases to converge.

In order to apply algorithm 1. to the capacitor we regard it as a long and narrow 2-D device, with carrier Dirichlet boundary conditions at the source, and Neumann boundary conditions along the remaining boundary. While u, v are separated from their boundary data by regions of extreme coefficient variation, θ_n, θ_p are always near their boundary data. The equations for the stream - functions are hence very well conditioned. The integrals (21), (22) may produce large sums, but this integration operation is perfectly well conditioned, and only requires an appropriate quadrature rule.

It should be stressed that despite the similarity between the Scharfetter - Gummel scheme and the stream - function formulation, the latter is very well conditioned in the capacitor case, while the former may be highly ill-conditioned.

3.2 Diodes

np-diodes are characterized by a doping function which is nearly piecewise constant, but varies steeply in a narrow region called 'a junction'. u, v, or n, p normally have corresponding steep gradients (or discontinuities) about the junction. Dirichlet boundary conditioned for the diode are specified at the contacts, and Neumann boundary conditions along the other 2 edges. The carrier variables are always separated from some of their Dirichlet data by the junction. Using the alternative stream function formulation, we observe that the junctions are normal to the Dirichlet boundaries. θ_n, θ_p are 'short-circuited' by the boundary conditions (23), (24), and consequently vary smoothly through the junction. As in the capacitor case, (21), (22) have to be integrated to obtain the carrier concentrations, but using an appropriate quadrature this operation poses no numerical difficulty.

3.3 MOS transistors

3.4 MOS devices

A typical 2-D model for a MOS device is a rectangular silicon piece, part of its upper edge is covered by a thin oxide layer (an insulator). An *npn* type MOS transistor has 2 small heavily doped n regions by the top corners, while the rest of the device is lightly p doped. Source and drain contacts are attached to the n regions, and a gate contact is attached to the insulator. We can distinguish 2 types of substrate boundary conditions (applied at the bottom edge of the device): The substrate may be grounded (Dirichlet boundary condtions) or insulated (Neumann conditions). An insulated substrate leads to a 'floating region', namely, a region without a contact. Floating regions often lead to extreme ill conditioning of the Jacobian matrix [1]

The abrupt interfaces between the p region and the n regions cause sharp variation of the carrier concentrations, as described for the diodes. When a positive voltage is applied to the gate, a depletion layer and an inversion layer form next to it. This thin inversion layer is called 'a channel' since it allows for a controlled amount of current to flow between surce and drain.

In the case of a grounded substrate Γ_D is not much longer then Γ_n, and consequently it is not clear that θ_n, θ_p have more favorable boundary conditions than n, p. In the case of a floating region however, the stream functions have Dirichlet boundary conditions almost around the entire device. Therefore the application of algorithm 1 leads to a well conditioned system, and is numerically superior to the Scherfetter - Gummel discretization.

Notice that even though θ_n has a Dirichlet type boundary at the gate interface, the fact that e^{ψ} is much larger near the inversion layer than in the rest of the semiconductor forms a separation layer, in which the coefficient of (17) is very small. This explains the boundary layer for θ_n which exists in the channel.

4 Conclusion

We have shown that stream potentials for the semiconductor continuity equations obey an adjoint set of boundary conditions. This fact explains the advantage of using harmonic averages for certain geometrical configurations. Ill conditioning of the Jacobian matrices which occurrs in the discretization of capacitors, and floating region devices can be removed by solving the stream function equations directly.

References

[1] U. Ascher, P. A. Markowich, C. Schmeiser, H. Steinruck, and R. Weiss. Conditioning of the steady steady state semiconductor device equations. *Siam J. Appl. Math*, 49(1):165–185, February 1989.

[2] E. M. Buturla and P. E. Cottrell. Two-dimensional finite element analysis of semiconductor steady state transport equations. In *Proc. International Conference "Computer Methods in Nonlinear Mechanics, Austin, Texas*, pages 512–530, 1974.

[3] P. Pietra. F. Brezzi, L.D.Marini. Two-dimensional exponential fitting and applications to semiconductor device equations. Technical Report 597, Instituto di Analisi Numerica del Consiglio Nazionale delle Richerch, 1987.

[4] R. W. Dutton M. R. Pinto, C. S. Rafferty. *PISCES-II User's Manual*, 1984.

[5] P. A. Markowich. *The Stationary Semiconductor Device Eqauations.* Springer-Verlag Wien New York, 1986.

[6] M.S. Mock. Analysis of a discretization algorithm for stationary continuity equations in semiconductor device models ii. *COMPEL 3*, pages 137–149, 1984.

[7] D.E. Rutherford. *Vector Methods Applied to Differential Geometry, Mechanics and Potential Theory.* Oliver and Boyd, 1957.