DRIFT-DIFFUSION SYSTEMS: VARIATIONAL PRINCIPLES AND FIXED POINT MAPS FOR STEADY STATE SEMICONDUCTOR MODELS

Joseph W. Jerome *
Department of Mathematics
Northwestern University
Evanston, IL 60208

Abstract

The mathematical semiconductor device model, consisting of the potential equation and the current continuity subsystem for the carriers, is studied from the standpoint of its decoupling fixed point map and the numerical approximate fixed point map. Variational principles will be discussed for this process and for discretizations achieved by use of generalized splines. By the choice of trial space, these capture the upwinding associated with Scharfetter-Gummel methods. An approximation calculus will be introduced in conjunction with the numerical fixed point map.

1 Introduction.

The standard Van Roosbroeck model for the flow of electrons and holes in a semiconductor device, incorporating diffusion and electric field induced drift, has been much studied since its introduction ([7]). Among the more significant computational ideas introduced in the study of this model were the iterative, decoupling technique, introduced by Gummel ([1]) and the exponential upwinding, introduced by Scharfetter and Gummel ([6]), for the discretization of the current continuity subsystem. Of interest also is the mathematical understanding achieved during this period (cf. [4]).

At the heart of both mathematical analysis and the development of effective computational procedures, is the construction of the system fixed point map. Fixed points can be identified with solutions of the coupled system of Partial Differential Equations (PDEs) in terms of the system dependent variables. The system is not a gradient system, i., e., it does not arise as the Euler system

*The author is supported by the National Science Foundation under grant DMS-8721742

associated with a global convex minimization principle of least "energy". Such variational principles can be derived by partial or total decoupling, however. Since the quasi-Fermi levels appear to be the optimal choice of dependent variables from the standpoint of computing, the bulk of our remarks concerning variational principles at the PDE level will focus on this situation.

A critical issue is the extent to which lagging should be employed in the recombination term. This is taken up in Section 2. The next major topic is taken up in Section 3. Here, we present a novel interpretation of the Scharfetter-Gummel scheme in terms of appropriate local basis functions. Such functions may be identified with generalized splines, and are defined to have the property that the associated flux is piecewise constant. The natural variables here are the Slotboom variables. Since these are not the ultimate computing variables of interest, this should be viewed as intermediate discretization analysis. However, this is consistent with the currently accepted view that upwinding should be applied, as a preliminary step, to the Slotboom variables; then, further linearization should occur with respect to the quasi-Fermi levels, via Newton's method, for example.

The final topic treated is an introduction to a nonlinear approximation theory, which we shall refer to as the Krasnosel'skii calculus. This is taken up in Section 4.

2 The Fixed Point Map

Denote the dimensionless electrostatic potential by u and the quasi-Fermi levels by v and w, respectively. Thus, in intrinsic concentration units, $n = \exp(u - v)$ and $p = \exp(w - u)$ are the expressions for the electron and hole densities. When the Gummel map is defined in terms of v and w, the first (fractional) step is the determination of u by use of the potential equation, where ϵ and k_1 denote the dielectric and doping, respectively:

$$F_1(u, v, w) \equiv -\nabla\cdot(\epsilon\nabla u) + \exp(u - v) - \exp(w - u) - k_1 = 0. \tag{1}$$

It remains to compute the new values, written as v^* and w^*. Denote by $R = R(u, v^*, w^*)$ the Shockley-Read-Hall/Auger recombination term. The variables, v^* and w^*, occur several places in the representation for R. The complete elimination of recombination lag, or, equivalently, the retention of v^* and w^* at each instance of their appearance in R, is one extreme possibility. Such total coupling at first appears desirable, since it achieves quasi-Fermi level dependence, at the PDE level, only upon the intermediately computed electric field,

u. The functional dependence upon the electric field need not be unique or continuous, however, in this case. The other extreme is lag which completely decouples the current continuity subsystem. This amounts to replacement of w^* by w in the electron current equation, and the corresponding operation in the hole current equation. In this case, each decoupled equation satisfies a minimization principle. As an intermediate position, one is led to introduce certain types of partial coupling. In order to fix the notation, let R_{v^*} and R_{w^*} represent the recombination term as it appears in the v^*-equation and w^*-equation, respectively, after insertion of lagging.

- The condition for well-posedness of the current continuity subsystem involves a comparison between the quadratic forms determined by these recombination functionals and a certain fundamental eigenvalue for the diffusion operator.

The details are furnished in [4].

As introduced in [3], one way of viewing the coupled current continuity subsystem is as a so-called obstacle problem. In this analogy, the analytical formalism is that of a variational inequality, and the extreme boundary values, described by the maximum principles, are seen as enforced obstacles. Although the variational inequality serves as a mathematical device, rather than the final formulation, which takes the form of the usual system of equations,

$$F_2(u, v^*, w^*) \equiv -\nabla \cdot J_n - R_{v^*} = 0, \tag{2}$$

$$F_3(u, v^*, w^*) \equiv -\nabla \cdot J_p + R_{w^*} = 0, \tag{3}$$

it serves the essential purpose of defining appropriate maximum principles, thereby providing the underlying stability for the system. The Gummel map, T, may then be written as

$$T : [v, w] \to [v^*, w^*].$$

The maximum principles serve to define the domain of T.

3 Piecewise Constant Flux

We consider the one-dimensional version of the electron current equation, and assume, for simplicity, zero recombination and constant mobility. We employ the notation, $J_n = J$, and require that the discretization scheme be exact for J a piecewise constant flux, whose discontinuities coincide with selected grid

points. If the exactness requirement is interpreted in terms of approximation theory, we are seeking an approximation of n of the form

$$n_h = \sum_i \alpha_i M_i \tag{4}$$

where $\{\alpha_i\}$ is a set of nodal values of n determined by a specified numerical method, and where $\{M_i\}$ is a nodal basis of local support functions of piecewise constant flux. More precisely, given a grid of the interval $G = [0, 1]$, of the form $x_i = ih,\ i = 0, \cdots N$, for $Nh = 1$, M_i is associated with the ith grid point, and is specified by the following requirements for $i \neq 0$ and $i \neq N$ with obvious adjustments for the endpoints:

1. $M_i(x_i) = 1, \quad \text{support } M_i = [x_{i-1}, x_{i+1}]$.
2. M_i is continuous.
3. On each subinterval determined by the grid, M_i has constant flux.

We shall refer to discretizations of the form (4) as being of the class of Scharfetter-Gummel type. The functions M_i are generalizations of the chapeau functions, and are examples of the generalized B-splines introduced in [2]. It is quite easy to give explicit formulas for the B-spline functions, M_i. For example, when $i = 1$,

$$M_1(x) = \begin{cases} \exp[u(x) - u(x_1)] \int_{x_0}^{x} e^{-u(s)}\, ds / \int_{x_0}^{x_1} e^{-u(s)}\, ds, & x_0 \leq x \leq x_1, \\ \exp[u(x) - u(x_1)] \int_{x}^{x_2} e^{-u(s)}\, ds / \int_{x_1}^{x_2} e^{-u(s)}\, ds, & x_1 \leq x \leq x_2. \end{cases}$$

The general formula is obtained via the identifications $0 \to i-1$, $1 \to i$, and $2 \to i+1$. The definition is completed by the support requirement in the second part of item one above.

It is also quite straightforward to compute the piecewise constant flux J. On the subinterval, (x_i, x_{i+1}), the flux has components from both M_i and M_{i+1}. The total flux can then be assembled from the following result. Denoting by $J^-_{M_i}$ the flux due to M_i on (x_{i-1}, x_i), and by $J^+_{M_i}$ the flux due to M_i on (x_i, x_{i+1}), we have

$$\begin{aligned} J^-_{Mi} &= e^{-u(x_i)} / \int_{x_{i-1}}^{x_i} e^{-u(s)}\, ds, \\ J^+_{M_i} &= -e^{-u(x_i)} / \int_{x_i}^{x_{i+1}} e^{-u(s)}\, ds. \end{aligned}$$

In order to evaluate the integrals appearing in the flux representations, it has been common to employ the piecewise linear interpolant of u. When this is

done, direct flux evaluation gives the following representation for the assembled flux on (x_i, x_{i+1}), with $\alpha_i = n_i$:

$$J = \frac{1}{h}[B(\triangle u)n_{i+1} - B(-\triangle u)n_i)], \tag{5}$$

where we have adopted the conventions $\triangle u = u(x_{i+1}) - u(x_i)$ and $B(z) = \frac{z}{\exp(z)-1}$. The latter function is known as the Bernoulli function. Note that an exponential fitting method of this type resolves the currents adequately, even if the mesh allows for substantial variation in the function u. By writing the nodal density values in the form, $n_i = \exp(u_i)\nu_i$, and recasting the vector unknown in terms of the Slotboom vector, ν, we obtain the flux representation in the familiar form involving the hyperbolic sin function.

Although the form given in (5) is well-known, it is not widely understood that it holds for the class (4), irrespective of the numerical method used to characterize the nodal values. The original method of Scharfetter and Gummel was to define these values by the box method, or, in current parlance, a finite volume method. An approximation procedure which may well be superior to the classical procedure just cited is a simple Ritz procedure, based upon a variational principle, wherein (5) is minimized in the integral mean square sense, subject to the boundary conditions. This is known to be second order in h.

4 The Krasnosel'skii Calculus

The generic problem is the local approximation of fixed points x_0 of a smooth mapping T of an open subset of a Banach space E into itself. The admissible approximations are contained in subspaces E_n of E and a suitability hypothesis on the class $\{E_n\}$ is contained in

- If P_n denotes the projection of E onto E_n, then

$$P_n x_0 \to x_0 \text{ as } n \to \infty. \tag{6}$$

The goal of the theory constructed in [5] is to identify hypotheses upon a given family of numerical fixed point approximate maps T_n, with fixed points x_n drawn from E_n, such that the estimate

$$c_1\|R_n x_0\| \leq \|x_n - P_n x_0\| \leq c_2\|R_n x_0\|, \tag{7}$$

holds, for $R_n = P_n T - T_n P_n$. The estimate (7) specifies the truncation error, $\|R_n(x_0)\|$, as the convergence rate governing the numerical scheme. The overall convergence of x_n to x_0 depends upon the rate in (6) as well as in (7).

The authors of [5] identify two fundamental properties which guarantee (7). Although the terms are not used in [5], we shall identify the properties by names familiar to numerical analysts.

1. Consistency: The derivative map of T_n is uniformly continuous in a neighborhood of $P_n x_0$.

2. Stability: The inverse maps of $I - T_n'$ are uniformly bounded in the same neighborhood.

The reader who consults the reference [5] will notice that the hypotheses employed there actually imply (1) and (2) above, which in turn allows the application of a fundamental approximation lemma, derived via a mean value calculus. The form of T_n used in applications decouples as does T, but the individual components are discretized.

References

[1] H. K. Gummel. A self-consistent iterative scheme for one-dimensional steady-state transistor calculations. *IEEE Trans. Electron Devices*, ED-11:455–465, 1964.

[2] Joseph W. Jerome. On uniform approximation by certain generalized spline functions. *J. Approximation Theory*, 7:143–154, 1973.

[3] Joseph W. Jerome. Consistency of semiconductor modeling: An existence/stability analysis for the stationary Van Roosbroeck system. *SIAM J. Appl. Math.*, 45:565–590, 1985.

[4] Joseph W. Jerome. *Analysis of Charge Transport.* Springer, 1996.

[5] M. A. Krasnosel'skii, G. M. Vainikko, P. P. Zabreiko, Ya. B. Rutitskii, and V. Ya. Stetsenko. *Approximate Solution of Operator Equations.* Wolters-Noordhoff, 1972.

[6] H. L. Scharfetter and H. K. Gummel. Large signal analysis of a silicon read diode oscillator. *IEEE Trans. Electron Devices*, ED-16:64–77, 1969.

[7] W. R. Van Roosbroeck. Theory of flow of electrons and holes in germanium and other semiconductors. *Bell System Technical J.*, 29:560–607, 1950.