Computational Methods for Boundary and Interior Layers in Several Dimensions

Edited by
J. J. H. MILLER

3

ON THE SCHARFETTER-GUMMEL DISCRETIZATION FOR DRIFT-DIFFUSION CONTINUITY EQUATIONS

PAUL A. FARRELL AND EUGENE C. GARTLAND, Jr. *

Department of Mathematics and Computer Science, Kent State U.,
314 Merriall Hall, Kent, 44240 Ohio USA;
Department of Mathematics, Southern Methodist U.,
Dallas, 75275 Texas USA

Abstract: Approaches (one-dimensional and two-dimensional or higher) to exponentially fitted methods for continuity equations of semiconductor models are reviewed, all of them variations on the Scharfetter–Gummel scheme. Convergence theory for two and higher dimensions is considered. A test problem is solved using a Scharfetter-Gummel discretization on a rectangular grid, and results are given for numerical simulations on three test problems with varying boundary conditions.

1. Introduction

In this chapter we shall consider a number of approaches which lead to exponentially fitted methods for the continuity equations of semiconductor physics. These are all variations on the well-known Scharfetter-Gummel scheme first proposed in [16]. For simplicity we shall confine ourselves for the most part to derivations on rectangular grids.

The equations describing the electrostatic potential ψ, and the hole and current densities n and p are given by [17]:

$$\epsilon \Delta \psi = q(n - p - N) \qquad \text{(Poisson equation)} \tag{1.1}$$

$$J_n = -q(\mu_n \, n \, \nabla \, \psi - D_n \, \nabla \, n) \qquad \text{(electron current relation)} \tag{1.2}$$

$$J_p = -q(\mu_p \, p \, \nabla \, \psi + D_p \, \nabla \, p) \qquad \text{(hole current relation)} \tag{1.3}$$

$$\nabla \cdot \, J_n = qR(n,p) \qquad \text{(electron} - \text{continuity equation)} \tag{1.4}$$

$$\nabla \cdot \, J_p = -qR(n,p) \qquad \text{(hole} - \text{continuity equation)} \tag{1.5}$$

* Research supported in part by NSF grant DMS-8806733

$$\nabla \cdot J_p = -qR(n,p) \qquad \text{(hole - continuity equation)} \tag{1.5}$$

for $(x,y) \in \Omega \subset \mathbb{R}^2$ (where Ω is bounded, convex domain representing the device geometry) subject to Dirichlet boundary conditions on Γ_c (Ohmic contacts) and homogeneous Neumann boundary conditions for ψ, n, p on Γ_i (insulating boundaries) with $\partial\Omega = \Gamma_c + \Gamma_i$. Here ϵ is the dielectric constant, $N = N_D - N_A$, where N_D and N_A are densities of donor and acceptor ions respectively, J_n and J_p are the electron and hole currents, q is the charge on the electron, μ_n and μ_p are the electron and hole mobilities, which are non-negative slowly varying functions of $\nabla\psi$ and N, and R is the net recombination. D_n and D_p are the electron and hole diffusion coefficients for which we shall assume the Einstein relations $D_n = \mu_n U_T$, $D_p = \mu_p U_T$, where U_T is the thermal voltage. In addition, we shall assume that μ_n and μ_p and hence D_n and D_p are constants.

We shall next exhibit more explicitly the singularly perturbed nature of these equations. To do this we transform them following the approach used in [10]. Let ℓ be the characteristic length of the device under investigation (for example the length of a diode). Define scaled versions of the variables by

$$\psi_s = \frac{\psi}{U_T}, \quad n_s = \frac{n}{\bar{N}}, \quad p_s = \frac{p}{\bar{N}}, \quad J_{n_s} = \frac{\ell J_n}{D_n q \bar{N}}, \quad J_{p_s} = \frac{\ell J_p}{D_p q \bar{N}}, \quad D = \frac{N}{\bar{N}}, \tag{1.6}$$

where $\bar{N} := \max_{(x,y)\in\Omega\cup\partial\Omega} |N(x,y)|$ and

$$x_s = \frac{x}{\ell}, \quad y_s = \frac{y}{\ell}. \tag{1.7}$$

Then, on dropping the subscript s again, (1.1)-(1.5) become

$$\lambda^2 \Delta\psi = n - p - D(x,y,\lambda) \tag{1.8}$$

$$J_n = -n \nabla \psi + \nabla n \tag{1.9}$$

$$J_p = -(p \nabla \psi + \nabla p) \tag{1.10}$$

$$\nabla \cdot J_n = \beta_n S(n,p,\gamma\lambda) \tag{1.11}$$

$$\nabla \cdot J_p = -\beta_p S(n,p,\gamma\lambda) \tag{1.12}$$

with

$$\lambda^2 = \left(\frac{\lambda_D}{\ell}\right)^2 = \frac{\epsilon U_T}{\ell^2 q \bar{N}}, \quad \gamma^2 = \frac{\ell q n_i}{\epsilon U_T}, \quad \beta_n = \frac{\ell^2}{D_n \tau_n}, \quad \beta_p = \frac{\ell^2}{D_p \tau_p}$$

Here λ_D is the minimal Debeye length. It is worthwhile to note at this point the relative magnitude of the terms appearing in (1.8)-(1.12). For a silicon device, with $\ell = 2.5 \times 10^{-3}$, at room temperature $T = 300K$ we have

$$\lambda^2 \approx 0.4 \times 10^{-6}, \quad \gamma^2 \approx \beta_n \approx \beta_p \approx 0.25.$$

The singularly perturbed nature of the Poisson equation (1.8) is immediately obvious and becomes more pronounced as the maximal doping $\bar{N}$ increases.

It is normal to postulate jump discontinuities or exponential layers in the doping profile D. Under these conditions the potential ψ also exhibits an exponential layer of width $0(\lambda|\ln\lambda|)$. The exponential decay is perpendicular to the junction and at a rate proportional to $1/\lambda$. The electron and hole densities n and p exhibit similar behavior. The normal components of the currents J_n and J_p do not however exhibit such layer behavior.

We remark that a number of alternative formulations may be used for (1.8)-(1.12). We shall restrict ourselves to considering those for the hole current p. It is clear that (1.10) and (1.12) can be combined to give :

$$\nabla \cdot (p\nabla\psi + \nabla p) = \beta_p S =: f. \tag{1.13}$$

This may also be rewritten in the Slotboom variable $v = e^{\psi}p$ to give

$$\nabla \cdot (e^{-\psi}v) = f. \tag{1.14}$$

We shall encounter all these forms in the context of the derivation of numerical schemes in later sections.

Before proceeding to consideration of difference schemes for the continuity equations, we shall first outline how these may be derived by a conservation law argument. This will provide motivation for some of the later finite difference schemes.

We shall first assume that the average drift velocity of holes v_p is given by $v_p = \mu_p E$ where $E = -\nabla\psi$ is the Electric field. Now consider a region K with boundary ∂K. The total charge in K due to holes is given by $\int_K qp$, and the rate of flow of electric charge due to holes across ∂K is given by $\int_{\partial K} J_p \cdot \hat{n}$, where J_p is the hole current density. Thus rate of change of charge in K is the sum of the flow across the boundary and the rate of loss of charge due to recombination of electron hole pairs, $-\int_K qR$, that is

$$\frac{d}{dt}\int_K qp = -\int_{\partial K} J_p \cdot \hat{n} - \int_K qR.$$

Using the Divergence Theorem we have

$$\frac{d}{dt}\int_K qp = -\int_K \nabla \cdot J_p - q\int_K R.$$

The steady state version of this is

$$\int_K \nabla \cdot J_p = -q\int_K R \tag{1.15}$$

which is just the integral form of equation (1.5).

2. One Dimensional Analysis

We now consider the solution of the hole continuity equation

$$\nabla \cdot (p\nabla\psi + \nabla p) = \beta_p S =: f. \tag{2.1}$$

For clarity we initially consider only the one dimensional case.

It is well known in the literature (*cf.* [16]) that straightforward application of a standard difference scheme leads to numerical instability. To see this, consider the discretization of the one dimensional analogue of (2.1)

$$\frac{d}{dx}(p\frac{d\psi}{dx} + \frac{dp}{dx}) = f(x) \tag{2.2}$$

using a centered difference approximation for the derivatives.

Writing this in matrix form it is clear that the matrix of the scheme is not an M-matrix unless h is small compared to the scale of variation of ψ. Hence inverse positivity and stability are not guaranteed. In practice the scheme does exhibit instability, giving rise to "wiggles" (*cf.* [8]), unless the mesh spacing h is chosen unrealistically fine. The problem here is violation of the well-known cell Reynolds number restriction which requires that:

$$\max_i |\psi_{i+1} - \psi_i| \leq 2.$$

It was to remedy this instability that Scharfetter and Gummel proposed their modified difference scheme in [16]. We first outline that derivation and then examine alternative approaches including those based on exact discretization of (2.2) using quadratures.

2.1. Scharfetter-Gummel Derivation

Consider the one dimensional analogue of (1.10) and (1.11), that is

$$\frac{dJ_p}{dx} = f \tag{2.3}$$

$$J_p = p\frac{d\psi}{dx} + \frac{dp}{dx}. \tag{2.4}$$

Now discretize (2.3) by the standard scheme

$$\frac{J_{p,i+1/2} - J_{p,i-1/2}}{(h_i + h_{i+1})/2} = f(x_i), \tag{2.5}$$

where $h_i = x_i - x_{i-1}$. Rather than discretizing (2.4) using a standard scheme, which would give rise to numerical instability, assume that J_p is constant and ψ is linear between mesh points and solve (2.4) explicitly as a differential equation for p. Note that the assumption that ψ is linear means that $\frac{d\psi}{dx}$ is constant and (2.4) is thus a constant coefficient differential equation. Considering the interval $[x_i, x_{i+1}]$ we obtain

$$\frac{dp(x)}{dx} + \frac{\psi_{i+1} - \psi_i}{h_{i+1}} p(x) = J_{p,i+1/2}, \tag{2.6}$$

subject to

$$p(x_i) = p_i, \;\; p(x_{i+1}) = p_{i+1}.$$

Applying an integrating factor $e^{(\psi_{i+1}-\psi_i)(x-x_i)/h_{i+1}}$ and integrating from x_i to x_{i+1}, if $\psi_{i+1} \neq \psi_i$, we obtain

$$\begin{aligned} e^{\psi_{i+1}-\psi_i} p_{i+1} - p_i &= J_{p,i+1/2} \int_{x_i}^{x_{i+1}} e^{\frac{\psi_{i+1}-\psi_i}{h_{i+1}}(x-x_i)} dx \\ &= \frac{h_{i+1}}{\psi_{i+1} - \psi_i} (e^{\psi_{i+1}-\psi_i} - 1) J_{p,i+1/2}. \end{aligned}$$

Thus

$$J_{p,i+1/2} = \frac{\psi_{i+1} - \psi_i}{h_{i+1}} \left[\frac{p_{i+1}}{1 - e^{-(\psi_{i+1}-\psi_i)}} + \frac{p_i}{1 - e^{\psi_{i+1}-\psi_i}} \right]. \tag{2.7}$$

On the other hand, if $\psi_{i+1} = \psi_i$, we obtain

$$J_{p,i+1/2} = \frac{p_{i+1} - p_i}{h_{i+1}}. \tag{2.8}$$

We may combine (2.7) and (2.8) by writing :

$$J_{p,i+1/2} - J_{p,i-1/2} = \frac{1}{h_{i+1}} \left[B(\psi_i - \psi_{i+1}) p_{i+1} - B(\psi_{i+1} - \psi_i) p_i \right], \tag{2.9}$$

where

$$B(z) := \frac{z}{e^z - 1}, \qquad B(0) := 1, \tag{2.10}$$

is the Bernoulli generating function.

The discretization given by (2.5) and (2.9) is what has become known as the Scharfetter-Gummel discretization. We remark that there were two fundamental assumptions here,

i) J_p is constant between mesh points.

ii) ψ is linear between mesh points.

2.2. Box Method

An alternative approach to the derivation of difference schemes arises from the scaled version of (1.13) namely

$$\int_R \frac{d\ J_p}{dx} dx = \int_R f(x) dx.$$

Applying this on each interval $[x_{i-1/2},\ x_{i+1/2}]$ gives

$$J_{p,i+1/2} - J_{p,i-1/2} = \int_{x_{i-1/2}}^{x_{i+1/2}} f(x) dx. \tag{2.11}$$

Assuming J_p is piecewise constant, we can use an integrating factor, similarly to the previous case, to obtain

$$J_{p,i+1/2} = \frac{e^{\psi_{i+1}} p_{i+1} - e^{\psi_i} p_i}{\int_{x_i}^{x_{i+1}} e^{\psi(t)} dt}.$$

Now assuming $\psi(x)$ is piecewise linear or approximating it by its linear interpolant, we again obtain (2.9). Thus this derivation gives rise to a scheme which differs only in the handling of the right hand side term $f(x)$. The two become identical if we approximate the integral in (2.11) by

$$\int_{x_{i-1/2}}^{x_{i+1/2}} f(x) dx = \frac{(h_i + h_{i+1})}{2} f(x_i).$$

2.3. Exact Discretization

An exact solution of (2.2) can be written in integral form. To derive this we rewrite the equation in the Slotboom variable $v = pe^{\psi}$. In this variable (2.2) becomes formally self-adjoint:

$$Lv := \frac{d}{dx}(e^{-\psi} \frac{dv}{dx}) = f. \tag{2.12}$$

We may write the solution of (2.12) in terms of local Green's functions, that is

$$v(x) = v_{i-1} r(x) + v_{i+1} s(x) + \int_{x_{i-1}}^{x_{i+1}} g_i(x,y) f(y) dy, \quad x_{i-1} < x < x_{i+1} \tag{2.13}$$

where $r(x)$ and $s(x)$ satisfy the homogeneous equations

$$\begin{aligned} Lr = Ls &= 0 \\ r(x_{i-1}) = 1, r(x_{i+1}) &= 0, \\ s(x_{i-1}) = 0, s(x_{i+1}) &= 1. \end{aligned}$$

Solving these explicitly in terms of integrals we get

$$r(x) = \frac{\int_x^{x_{i+1}} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_{i+1}} e^{\psi(t)}dt}, \quad s(x) = \frac{\int_{x_{i-1}}^{x} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_{i+1}} e^{\psi(t)}dt}. \tag{2.14}$$

and $g_i(x,y)$ is given by

$$g_i(x,y) = -\frac{1}{\int_{x_{i-1}}^{x_{i+1}} e^{\psi(t)}dt} \begin{cases} \int_{x_{i-1}}^{y} e^{\psi(t)}dt \int_x^{x_{i+1}} e^{\psi(t)}dt, x_{i-1} < y < x < x_{i+1}, \\ \int_{x_{i-1}}^{x} e^{\psi(t)}dt \int_y^{x_{i+1}} e^{\psi(t)}dt, x_{i-1} < x < y < x_{i+1}. \end{cases} \tag{2.15}$$

Substituting (2.14) and (2.15) in (2.13) and using it to evaluate $v(x_i)$ we get

$$v_i = v_{i-1}\frac{\int_{x_i}^{x_{i+1}} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_{i+1}} e^{\psi(t)}dt} + v_{i+1}\frac{\int_{x_{i-1}}^{x_i} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_{i+1}} e^{\psi(t)}dt} -$$

$$-\frac{1}{\int_{x_{i-1}}^{x_{i+1}} e^{\psi(t)}dt}\left\{\int_{x_i}^{x_{i+1}} e^{\psi(t)}dt \int_{x_{i-1}}^{x_i}\int_{x_{i-1}}^{y} e^{\psi(t)}dt f(y)dy + \right.$$

$$\left. + \int_{x_{i-1}}^{x_i} e^{\psi(t)}dt \int_{x_i}^{x_{i+1}}\int_{y}^{x_{i+1}} e^{\psi(t)}dt f(y)dy\right\}.$$

Converting back to the hole density function p and rewriting in three point difference form we get

$$\frac{1}{\int_{x_{i-1}}^{x_i} e^{\psi(t)}dt} e^{\psi(x_{i-1})}u_{i-1} - \left(\frac{1}{\int_{x_{i-1}}^{x_i} e^{\psi(t)}dt} + \frac{1}{\int_{x_i}^{x_{i+1}} e^{\psi(t)}dt}\right) e^{\psi(x_i)}u_i +$$

$$+ \frac{1}{\int_{x_i}^{x_{i+1}} e^{\psi(t)}dt} e^{\psi(x_{i+1})}u_{i+1} =$$

$$= \int_{x_{i-1}}^{x_i} \frac{\int_{x_{i-1}}^{x} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_i} e^{\psi(t)}} f(x)dx + \int_{x_i}^{x_{i+1}} \frac{\int_x^{x_{i+1}} e^{\psi(t)}dt}{\int_{x_i}^{x_{i+1}} e^{\psi(t)}dt} f(x)dx. \tag{2.16}$$

Remark: This same exact discretization can alternatively be obtained by integrating by parts the local relation:

$$(Lv, \phi_i) = (f, \phi_i),$$

where ϕ_i is the L-spline generalized "hat" function :

$$\phi_i(x) = \begin{cases} \dfrac{\int_{x_{i-1}}^{x} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_i} e^{\psi(t)}}, x_{i-1} \le x \le x_i, \\ \dfrac{\int_{x}^{x_{i+1}} e^{\psi(t)}dt}{\int_{x_i}^{x_{i+1}} e^{\psi(t)dt}}, x_i \le x \le x_{i+1}. \end{cases}$$

This is essentially the approach taken by Miller in [13].

2.4. Approximation of Integrals by Quadrature Rules

It is, in general, impossible to evaluate the integrals in (2.16) exactly. However, we can obtain a family of three point difference schemes by approximating the integrals.

In fact, if we again assume that $\psi(x)$ is piecewise linear we obtain a scheme whose left hand side is identical to the Scharfetter-Gummel and Box schemes. It remains to choose a quadrature rule to evaluate the right hand side in (2.16). If this is chosen to be a trapezoidal rule we obtain :

$$\int_{x_{i-1}}^{x_i} \frac{\int_{x_{i-1}}^{x} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_i} e^{\psi(t)}dt} f(x)dx \approx \frac{hi}{2} \left\{ \frac{\int_{x_{i-1}}^{x_{i-1}} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_i} e^{\psi(t)}dt} f_{i-1} + \frac{\int_{x_{i-1}}^{x_i} e^{\psi(t)}dt}{\int_{x_{i-1}}^{x_i} e^{\psi(t)}dt} f_i \right\}$$
$$= \frac{hi}{2} f_i.$$

Similarly

$$\int_{x_i}^{x_{i+1}} \frac{\int_{x}^{x_{i+1}} e^{\psi(t)}dt}{\int_{x_i}^{x_{i+1}} e^{\psi(t)}dt} f(x)dx \approx \frac{h_{i+1}}{2} f_i,$$

which makes the right hand side identical to that in the Scharfetter-Gummel Scheme also.

We remark that it is clearly possible to obtain further schemes by using higher order quadrature rules to evaluate the right hand side. These would be similar to those derived in [12], using Marchuk identities. In principle it is also possible to assume that $\psi(x)$ is merely quadratic or cubic rather than linear. The schemes thus obtained would however involve more complicated functions than exponentials, which would render them impractical.

We now consider a number of finite element methods for the derivation of schemes for the solution of (2.2). We shall confine ourselves here to the

one dimensional analogues of schemes on rectangles. In particular we shall not consider methods dealt with elsewhere in this monograph.

We remark that as in the finite difference case the derivation of successful finite element schemes will rely on certain assumptions to the effect that the current J_p is constant on each element and the potential ψ is piecewise linear.

2.5. Hybrid Methods - Brezzi, Marini, Pietra

We first consider a hybrid method due to Brezzi, Marini and Pietra [3]. This is obtained by writing (2.2) as a system of first order differential equations. These are in fact the one dimensional analogue of the original equations (1.10), (1.12) rewritten in the Slotboom variables $v = pe^{\psi}$, that is

$$\frac{dv}{dx} = e^{\psi} J_p \tag{2.17}$$

$$\frac{dJ_p}{dx} = f. \tag{2.18}$$

These may be subject to Dirichlet boundary conditions

$$v(0) = e^{\psi(0)}A, \qquad v(1) = e^{\psi(1)}B \tag{2.19}$$

or mixed boundary conditions

$$v(0) = e^{\psi(0)}A, \qquad J_p(1) = 0. \tag{2.20}$$

For convenience we shall confine ourselves to the former. The latter may be incorporated by the standard finite element approaches. We first derive a weak formulation of (2.17)(2.18). We define

$$H_E^1(0,1) = \{y \in H^1(0,1), \quad y(0) = e^{\psi(0)}A, \quad y(1) = e^{\psi(1)}B\}$$
$$H_0^1(0,1) = \{y \in H^1(0,1), \quad y(0) = y(1) = 0\}.$$

Now multiplying by "test functions" and integrating we get

$$\int_0^1 \frac{dv}{dx}\,\tau\,dx = \int_0^1 e^{\psi} J_p\,\tau\,dx, \quad \forall\,\tau \in L^2(0,1)$$
$$\int_0^1 \frac{dJ_p}{dx}\,\chi\,dx = \int_0^1 f\,\chi\,dx, \quad \forall\,\chi \in H_0^1(0,1).$$

On integrating the latter by parts we obtain

$$\begin{aligned}\int_0^1 \frac{dJ_p}{dx}\chi\,dx v &= J_p\,\chi|_0^1 - \int_0^1 J_p \frac{d\chi}{dx}dx \\ &= J_p(1)\chi(1) - J_p(0)\chi(0) - \int_0^1 J_p\frac{d\chi}{dx}dx \\ &= -\int_0^1 J_p\frac{d\chi}{dx}dx\end{aligned}$$

since $\chi(0) = \chi(1) = 0$. Hence the weak formulation becomes:

Find $v \in H^1_E$ and $J_p \in L^2(0,1)$ such that

$$\int_0^1 \frac{dv}{dx}\,\tau\,dx = \int_0^1 e^{\psi} J_p\,\tau\,dx, \quad \forall\,\tau \in L^2(0,1) \tag{2.21}$$

$$-\int_0^1 J_p \frac{d\chi}{dx}dx = \int_0^1 f\,\chi\,dx, \quad \forall\,\chi \in H^1_0(0,1). \tag{2.22}$$

We now proceed to show that the finite element approximation of (2.21) (2.22) proposed in [3] leads to a scheme which is similar to the Scharfetter Gummel scheme.

We now seek a piecewise linear approximation v^h to v and a piecewise constant approximation J_p^h to J_p. Define $I_i = (x_i, x_{i+1})$, $I = \cup_{i=0}^{k-1} \bar{I}_i$,

$$V^h = \{y:\ y \in H^1_E(0,1),\ y|_{I_i} \in P_1\}\ ,\ V_0^h = \{y:\ y \in H^1_0(0,1),\ y|_{I_i} \in P_1\},$$

$$W^h = \{y:\ \ y \in L^2(0,1),\ \ y|_{I_i} \in P_0\},$$

where P_0 and P_1 are the spaces of polynomials of degree 0 and 1 respectively. Then the hybrid finite element approximation is:

Find $v^h \in V^h$, $J_p^h \in W^h$ such that

$$\int_0^1 \frac{dv^h}{dx}\,\tau\,dx = \int_0^1 e^{\psi} J_p^h\,\tau\,dx, \quad \forall\,\tau \in W^h, \tag{2.23}$$

$$\int_0^1 J_p^h \frac{d\chi}{dx}dx = \int_0^1 f\,\chi\,dx, \quad \forall\,\chi \in V_0^h. \tag{2.24}$$

Let us now examine the discretization arising from this method. Let the constant $J^h_{i+1/2} = J_p^h|_{I_i}$ and let $v_i^h = v^h(x_i)$, and further let $\chi_i(x)$ be the usual "hat" function basis of V_0^h, that is $\chi_i(x_j) = \delta_{ij}$. Now $v^h \in V^h$, hence $\frac{dv^h}{dx}|_{I_i} = \frac{v^h_{i+1} - v^h_i}{h_{i+1}}$ and thus (2.23) gives

$$v^h_{i+1} - v^h_i = J^h_{i+1/2} \int_{x_i}^{x_{i+1}} e^{\psi}\,dx. \tag{2.25}$$

Similarly, considering the contribution of χ_i from (2.24) we get,

$$\begin{aligned}
\int_0^1 f\chi_i dx &= \int_0^1 J_p^h \frac{d\chi_i}{dx}dx \\
&= -J^h_{i-1/2}\int_{x_{i-1}}^{x_i} \frac{d\chi_i}{dx}dx - J^h_{i+1/2}\int_{x_i}^{x_{i+1}} \frac{d\chi_i}{dx}dx \\
&= -J^h_{i-1/2}(\chi_i(x_i) - \chi_i(x_{i-1})) - J^h_{i+1/2}(\chi_i(x_{i+1}) - \chi_i(x_i)) \\
&= -J^h_{i-1/2} + J^h_{i+1/2},
\end{aligned} \tag{2.26}$$

by the definition of χ_i.

Combining (2.25) and (2.26) we obtain

$$\frac{v_{i+1}^h - v_i^h}{\int_{x_i}^{x_{i+1}} e^{\psi} dx} - \frac{v_i^h - v_{i-1}^h}{\int_{x_{i-1}}^{x_i} e^{\psi} dx} = \int_0^1 f\, \chi_i\, dx$$

or

$$\frac{1}{\int_{x_{i-1}}^{x_i} e^{\psi} dx} v_{i-1}^h - \left(\frac{1}{\int_{x_{i-1}}^{x_i} e^{\psi} dx} + \frac{1}{\int_{x_i}^{x_{i+1}} e^{\psi} dx} \right) v_i^h + \frac{1}{\int_{x_i}^{x_{i+1}} e^{\psi} dx} v_{i+1}^h = \int_0^1 f\, \chi_i\, dx, \tag{2.27}$$

which if we assume that ψ is piecewise linear again gives us the same discretization matrix as Scharfetter-Gummel. Again note the key role the assumption that the current J_p was piecewise constant played in the derivation.

2.6. Finite Element Method — Zlamal

We now consider the finite element method derived in Zlamal [19]. Again the assumption is made that J_p is constant and ψ linear on each element.

The hole continuity equation is formulated in the standard Galerkin manner as a variational form in the Slotboom variables

$$\int_0^1 \frac{d}{dx}(e^{-\psi}\frac{dv}{dx})\chi dx = \int_0^1 f\chi\, dx, \qquad \forall \quad \chi \in H_0^1, \tag{2.28}$$

and is rewritten as:

$$\pi(\psi; v, \chi) := -\int_0^1 e^{-\psi} \frac{dv}{dx}\frac{d\chi}{dx} dx = \int_0^1 f\chi\, dx.$$

Now assuming $J_p^h = e^{-\psi}\frac{dv}{dx}$ is piecewise constant and $J_p^h|_{I_i} = J_{i+1/2}, i = 0, \cdots, k$, we define an approximate form π_h by:

$$\begin{aligned} \pi_h(\psi; v, \chi) &= \sum_{i=0}^{k-1} J_{i+1/2} \int_{x_i}^{x_{i+1}} \frac{d\chi}{dx} dx \\ &= \sum_{i=0}^{k-1} J_{i+1/2}(\chi(x_{i+1}) - \chi(x_i)). \end{aligned} \tag{2.29}$$

Zlamal then proceeds to determine an appropriate value for $J_{i+1/2}$, which is consistent with the equation $J_p^h = e^{-\psi}\frac{dv}{dx}$. To do this consider interval I_i where

$$\frac{dv}{dx} = e^{\psi} J_{i+1/2}$$

and integrate

$$\int_{x_i}^{x_{i+1}} \frac{dv}{dx} dx = \int_{x_i}^{x_{i+1}} e^{\psi} J_{i+1/2}\, dx$$

or, writing v_i for $v(x_i)$ etc.,

$$v_{i+1} - v_i = J_{i+1/2} \int_{x_i}^{x_{i+1}} e^{\psi} dx, \tag{2.30}$$

which is identical to (2.25).

Thus, assuming that χ_i are the normal piecewise linear basis elements, that is, $\chi_i(x_j) = \delta_{ij}$, (2.28),(2.29) and (2.30) will lead to a discrete system of the form (2.27). Again Zlamal makes the assumption that ψ is replaced by a piecewise linear interpolant and thus we again get the Scharfetter-Gummel discretization matrix.

In conclusion therefore we remark that all the schemes considered so far are variations on the Scharfetter-Gummel scheme in that, under appropriate assumptions, they all give rise to a discretization matrix which is identical to that of Scharfetter-Gummel. They differ only in their treatment of the source term f and the extent to which they can be generalized to higher dimensions.

2.7. Relationship with Classical Exponentially Fitted Methods

All the derivations considered so far have employed the physics of the problem directly. Let us now briefly consider the formal application of exponentially fitted methods (*cf.* [4]) directly to (2.2). From an asymptotic analysis [10], it is clear that near an interface, at $x = x^*$,

$$|\psi^{(k)}(x;\lambda)| \leq C\{1 + \lambda^{-k} e^{-\frac{|x-x^*|}{\lambda}}\}, \tag{2.31}$$

and that $\frac{d\psi(x)}{dx}$ is $O(1)$ elsewhere. Thus multiplying by the scaling factor λ we get an equation in singularly perturbed form :

$$\frac{d}{dx}(\lambda \frac{dp(x)}{dx} + (\lambda \frac{d\psi}{dx}) p(x)) = \lambda f(x).$$

We now write $b(x) := \lambda \frac{d\psi}{dx}$, to give

$$\frac{d}{dx}(\lambda \frac{dp(x)}{dx} + b(x)p(x)) = \lambda f(x), \tag{2.32}$$

and note that the resulting equation is, in practice, a mildly semi-linear singularly perturbed equation in conservation form. Further $\|b(x)\|_\infty \leq C$, where C is independent of λ.

Let us now consider a standard exponentially fitted scheme for this equation [5, page 107]:

$$\delta \left(\lambda \, \sigma_i \, \delta p_i^h \; + \; b_i^h \, \mu \, p_i^h \right) \;\; = \;\; \lambda \, f_i^h, \tag{2.33}$$

where

$$\delta z_i = \frac{z_{i+1/2} - z_{i-1/2}}{x_{i+1/2} - x_{i-1/2}} \quad , \quad \mu z_i = \frac{z_{i+1/2} + z_{i-1/2}}{2} \quad ,$$

$$\sigma_{i+1/2} = \frac{b_{i+1/2}^h \, h_{i+1}}{2\lambda} \coth \frac{b_{i+1/2}^h \, h_{i+1}}{2\lambda}.$$

We remark that this scheme preserves the conservation form of (2.32). The proofs of uniform convergence in [5] do not, of course, apply in this case, since they require the coefficients to be smooth – in particular $b(x) \in C^2(\Omega)$.

We will now show that this scheme, with an appropriate choice of $b_{i+1/2}^h$, is in fact the Scharfetter-Gummel scheme. First note that (2.33) may be written as :

$$\frac{K_{i+1/2} - K_{i-1/2}}{(h_i + h_{i+1})/2} = \lambda f_i^h, \tag{2.34}$$

where

$$K_{i+1/2} = \lambda \, \sigma_{i+1/2} \, \delta p_{i+1/2}^h \; + \; b_{i+1/2}^h \, \mu \, p_{i+1/2}^h. \tag{2.35}$$

It is clear that (2.34) is in the same form as (2.5) and can be made equivalent by an appropriate choice of f_i^h. It suffices to show therefore that (2.35) is equivalent to (2.9) for an appropriate choice of $b_{i+1/2}^h$. Let us choose

$$b(x_{i+1/2}) = \lambda \frac{d\psi(x)}{dx}|_{x_{i+1/2}} \approx \lambda \frac{\psi_{i+1} - \psi_i}{h_{i+1}} =: b_{i+1/2}^h.$$

Notice that this is exact if $\psi(x)$ is piecewise linear. Now, since

$$\frac{z}{2} \coth \frac{z}{2} = \frac{z}{e^z - 1} + \frac{z}{2} = B(z) + \frac{z}{2} = B(-z) - \frac{z}{2},$$

we have

$$K_{i+1/2} = \frac{\lambda}{h_{i+1}} \left[\left(B(\psi_{i+1} - \psi_i) + \frac{\psi_{i+1} - \psi_i}{2} \right) (p_{i+1}^h - p_i^h) + \right.$$

$$\left. + (\psi_{i+1} - \psi_i) \left(\frac{p_{i+1}^h + p_i^h}{2} \right) \right] =$$

$$= \lambda [B(\psi_i - \psi_{i+1}) p_{i+1}^h - B(\psi_{i+1} - \psi_i) p_i^h] / h_{i+1} \quad = \quad \lambda J_{p,i+1/2},$$

which shows that (2.33) is again the Scharfetter-Gummel scheme.

2.8. Uniform Convergence Results

It is only in the one-dimensional case that researchers have been able to prove satisfactory uniform convergence results for discretizations of Scharfetter-Gummel type. In [10] and [9,§5.3] the error estimate

$$\max_{i=1,\ldots,n} \left(|p_i^h - p(x_i;\lambda)| + |J_{p,i-1/2}^h - J_p(x_{i-1/2};\lambda)| \right) \leq K(h + \lambda |\ln \lambda|)$$

is established (with a constant K that does not depend on λ) for the scaled one dimensional equation (2.2) subject to either the Dirichlet or mixed boundary conditions corresponding to (2.19) and (2.20). Here λ is the singular-perturbation parameter from (1.8) and the scaled potential ψ is $O(1)$ with a finite number of internal layers $x = x^*$ at which it satisfies (2.31).

The analysis uses explicit representations for the solutions of the difference equations and singular-perturbation analysis to justify estimates like those on ψ above. It is indicated that the analysis can be refined to the point of indicating that truly $O(h)$ convergence is obtained "away from" the layers of ψ, but that the scheme cannot be expected to uniformly resolve the concentration and current near points like $x = x^*$. As far as we can tell, this has always been believed to be the case.

It has only recently been established, by Gartland [6], that the one dimensional Scharfetter-Gummel discretization is truly uniformly first-order accurate under assumptions that encompass the situation analyzed by Markowich, Ringhofer, Selberherr, and Lentini cited above. In fact, there is a natural way to extend the computed discrete approximations p_i^h and $J_{p,i+1/2}^h$ to a global approximation, for which is established the bound

$$\max\{\|p - p^h\|_\infty, \|J_p - J_p^h\|_\infty\} \leq Ch,$$

where the constant C is independent of λ. Here h is the maximum spacing of a not-necessarily-uniform mesh (with *no* other restrictions on it).

The only assumptions required of ψ and f are that the norms $\|\psi\|_\infty$, $\|\psi'\|_1$, $\|f\|_\infty$, and $\|f'\|_1$ be bounded independent of λ; this includes the (realistic) cases where ψ has a finite number of internal layers of the form (2.31). The global extension of p^h is accomplished via a local exponential interpolation formula; whereas J_p^h is extended to be piecewise constant. The analysis is broken down into two parts: the first, a sensitivity analysis of the solution operator with respect to replacing ψ by its piecewise-linear interpolant, and the second, a standard consistency-and-stability argument to bound the "quadrature" errors incurred in discretizing the latter problem (with its piecewise-linear potential). The details are in [6]. This gives a theoretical quantification (at least in the one dimensional case) of the "robustness" of this widely used discretization technique.

3. Two Dimensional Generalizations of Scharfetter-Gummel

We shall now proceed to examine a number of generalizations of the Scharfetter-Gummel method to two and higher dimensions. We restrict ourselves to those methods defined on rectangular meshes.

It is clear that to derive two dimensional analogues of the Scharfetter-Gummel scheme we shall have to make assumptions, similar to those made in the one dimensional case, concerning the behaviour of the hole and electron currents J_n and J_p and the electrostatic potential ψ.

If we make no such assumptions we can of course write a solution of the continuity equations in exact form as in the one dimensional case. However, for the hole density p this would be of the form

$$p(x,y) = \int_{\partial\Omega} k(x,y;\varsigma,\eta)p(\varsigma,\eta)d\sigma + \int\int_{\Omega} g(x,y;\varsigma,\eta)f(\varsigma,\eta)d\varsigma d\eta.$$

We do not however have explicit forms for the kernels k and g. It is clear that the value of the solution at the center of a rectangular cell (for example) depends on the values on the *entire* perimeter; therefore exact 9-point schemes cannot, in general, exist.

3.1. Box Method

The simplest approach to generalization is to use a Box method (*cf.* [18]) to derive a finite difference scheme. This was essentially the approach used in [2,1]. This has some physical justification moreover, since it corresponds closely to the (integral) conservation law form of the problem given by (1.15).

To be more precise, substituting (1.10), we obtain

$$\int_R \nabla\cdot(\nabla p + p\nabla\psi) = \int_R f\ .$$

Now applying the Divergence Theorem we obtain

$$\int_{\partial R} (\nabla p + p\nabla\psi)\cdot\hat{n} = \int_R f\ . \tag{3.1}$$

To derive the difference scheme we now apply (3.1) to a box

$$R = [x_{i-1/2}, x_{i+1/2}]\times[y_{j-1/2}, y_{j+1/2}].$$

Firstly, approximate $\int_R f$ by $f_{i,j}*$ area (R) that is

$$\int_R f\ dr \equiv f_{i,j}\ \bar{h}_i\bar{k}_j$$

where

$$\bar{h}_i = \frac{h_i + h_{i+1}}{2}, \quad \bar{k}_j = \frac{k_j + k_{j+1}}{2}.$$

Now

$$\int_{\partial R} (\nabla p + p\nabla\psi)\cdot\hat{n} = \int_{\Gamma_1} (p_x + p\psi_x) + \int_{\Gamma_2} (p_y + p\psi_y) + \int_{\Gamma_3} (-p_x - p\psi_x) + \\ + \int_{\Gamma_4} (-p_y - p\psi_y),$$

where $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ denote the "east", "north", "west" and "south" edges of ∂R. Now approximate the integral along each edge by an average value of $J_p = p_x + p\psi_x$ along the edge times the length of the edge, for example,

$$\bar{J}_p \times \text{ length } (\Gamma_1) = \bar{J}_p \bar{k}_j.$$

It remains to choose an appropriate value for $\bar{J}_p$.

We must once again make a simplifying assumption — in this case that $J_p = p_x + p\psi_x$ is constant along the mesh lines connecting (x_i, y_j) to the adjacent nodes. But this is precisely the same assumption as in the one dimensional Scharfetter-Gummel derivation and solving the equation, for example on Γ_1,

$$p_x + p\psi_x = J, \qquad x_i < x < x_{i+1}$$
$$p(x_i) = p_i, \qquad p(x_{i+1}) = p_{i+1}$$

again yields

$$J = \frac{e^{\psi_{i+1,j}} p_{i+1,j} - e^{\psi_{i,j}} p_{i,j}}{\int_{x_i}^{x_{i+1}} e^{\psi(x,y_j)} dx}.$$

To evaluate the integral in the denominator, we must either assume ψ is piecewise linear or approximate it by its linear interpolant at (x_i, y_j) and (x_{i+1}, y_j), in which case we obtain

$$\int_{x_i}^{x_{i+1}} e^{\psi(x,y_j)} dx \equiv \begin{cases} h_{i+1} \frac{e^{\psi_{i+1,j}} - e^{\psi_{i,j}}}{\psi_{i+1,j} - \psi_{i,j}}, & \text{if } \psi_{i+1,j} \neq \psi_{i,j} \\ h_{i+1} e^{\psi_{i,j}}, & \text{if } \psi_{i+1,j} = \psi_{i,j}. \end{cases}$$

Thus we get a scheme which is a natural generalization of Scharfetter-Gummel:

$$\frac{B(\psi_{i,j} - \psi_{i+1,j})}{h_{i+1}} p_{i+1,j} + \frac{B(\psi_{i,j} - \psi_{i,j+1})}{k_{i+1}} p_{i,j+1} + \frac{B(\psi_{i-1,j} - \psi_{i,j})}{h_i} p_{i-1,j} + \\ + \frac{B(\psi_{i,j-1} - \psi_{i,j})}{k_i} p_{i,j-1} - \left[\frac{B(\psi_{i+1,j} - \psi_{i,j})}{h_{i+1}} + \frac{B(\psi_{i,j+1} - \psi_{i,j})}{k_{i+1}} + \right. \\ \left. + \frac{B(\psi_{i-1,j} - \psi_{i,j})}{h_i} + \frac{B(\psi_{i,j-1} - \psi_{i,j})}{k_i} \right] p_{i,j} = \\ = \bar{h}_i \bar{k}_j \, f_{i,j}$$

where $B(z)$ is the Bernoulli generating function defined in (2.10).

3.2. Finite Element Derivations

We shall now consider some finite element derivations. We shall mainly confine our attention to derivations on rectangular regions.

3.2.1. Zlamal's Method

In [19], Zlamal considers solution of the semiconductor equations in two and three dimensions. In the former case the methods use piecewise bi-linear polynomials on four node isoparametric elements.

We consider the solution of (1.13) subject to a mixture of Dirichlet and Neumann boundary conditions on $\partial\Omega = \Gamma^1 \cup \Gamma^2$ that is

$$p|_{\Gamma^1} = p^*, \qquad \frac{\partial p}{\partial n}|_{\Gamma^2} = 0. \tag{3.3}$$

Let us define

$$H_0^1 = \{v \in H^1(\Omega), \qquad v|_{\Gamma^1} = 0\}.$$

Then, once again using the Slotboom variables and integrating by parts we obtain the following variational formulation :

$$\pi(\psi; v, \chi) = \int_\Omega f\chi, \qquad \forall\, \chi \in H_0^1, \tag{3.4}$$

where

$$\pi(\psi; v, \chi) := \int_\Omega e^{-\psi} \nabla v \cdot \nabla \chi = \int_\Omega J_p \cdot \nabla \chi. \tag{3.5}$$

Zlamal considers a partition into arbitrary triangles or quadrilaterals and, in order to proceed further, must map these onto a reference element, which is either a triangle or square. In order to simplify the notation and extract the essential details, we confine ourselves to consideration of a cover R_h of $\bar{\Omega}$ by rectangles K. Let $h_K = \text{diam}\,(K)$, $h = \max\limits_{K \in R_h} h_K$, $\bar{\Omega}_h = \bigcup\limits_{K \in R_h} K$ and $\Gamma_h = \partial\Omega_h = \Gamma_h^1 \cup \Gamma_h^2$.

We shall seek an approximation P to p in the space

$$V^h = \{v \in C^0(\bar{\Omega}_h), \quad v|_K \text{ bilinear } \}.$$

We shall use the test space

$$V_0^h = \{v \in V^h, \quad V|_{\Gamma_h^1} = 0\}.$$

In order to construct a discrete analogue of $\pi(\psi; v, \chi)$ we require that the approximation J_p^h to $J_p = e^{-\psi}\nabla v$ be constant on each element K. Thus

$$\nabla v = e^{\psi} J_p^h. \tag{3.6}$$

Consider an element $K_{i,j} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$, and let $J_p^h|_{K_{i,j}} = J$. Then, on $K_{i,j}$, (3.6) becomes

$$\nabla v = e^{\psi} J$$

or

$$\begin{pmatrix} \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \end{pmatrix} = e^{\psi} \begin{pmatrix} J_1 \\ J_2 \end{pmatrix}. \tag{3.7}$$

Now integrating the first component of (3.7) from x_i to x_{i+1}, along $y = y_j$, we obtain

$$v(x_{i+1}, y_j) - v(x_i, y_j) = J_1 \int_{x_i}^{x_{i+1}} e^{\psi(x, y_j)} dx$$

Thus

$$J_1 = \frac{v(x_{i+1}, y_j) - v(x_i, y_j)}{\int_{x_i}^{x_{i+1}} e^{\psi(x, y_j)} dx}. \tag{3.8}$$

Similarly, integrating from y_j to y_{j+1}, along $x = x_i$, we get

$$J_2 = \frac{v(x_i, y_{j+1}) - v(x_i, y_j)}{\int_{y_j}^{y_{j+1}} e^{\psi(x_i, y)} dy} \tag{3.9}$$

Now consider the approximation $v^h \in V^h$ to v and assume further that ψ is also in V^h. Note that this is consistent with ψ being a piecewise bilinear solution to the Poisson equation determined previously or simultaneously. Then, substituting v^h for v and using the bilinearity of ψ and v^h, we obtain

$$\begin{aligned} J_1 &= \frac{v^h_{i+1,j} - v^h_{i,j}}{h_{i+1}} \frac{\psi_{i+1,j} - \psi_{i,j}}{e^{\psi_{i+1,j}} - e^{\psi_{i,j}}} \\ &= \frac{\partial v^h}{\partial x} e^{-\psi_{i,j}} B(\psi_{i+1,j} - \psi_{i,j}), \end{aligned} \tag{3.10}$$

where $B(z)$ is defined in (2.10). Similarly,

$$J_2 = \frac{\partial v^h}{\partial y} e^{-\psi_{i,j}} B(\psi_{i,j+1} - \psi_{i,j}). \tag{3.11}$$

Now, combining (3.7), (3.10) and (3.11) we obtain

$$\begin{pmatrix} J_1 \\ J_2 \end{pmatrix} = e^{-\psi_{i,j}} \begin{pmatrix} B(\psi_{i+1,j} - \psi_{i,j}) & 0 \\ 0 & B(\psi_{i,j+1} - \psi_{i,j}) \end{pmatrix} \nabla v^h,$$

or

$$J = e^{-\psi_K} B^K \nabla v^h \tag{3.12}$$

where $B^K = \mathrm{diag}(B(\psi_{i+1,j} - \psi_{i,j}), B(\psi_{i,j+1} - \psi_{i,j}))$ and ψ_K is the corresponding value of ψ at the south-west corner of K, in this case $\psi_{i,j}$. In order to define the discrete analogue of $\pi(\psi; v, \chi)$, let us first define $\chi_{i,j}(x, y)$, for $x_i, y_j \in \Omega_h \setminus \Gamma_h^1$, to be the usual two dimensional "hat" function basis element for V_0^h, which satisfies $\chi_{i,j}(x_\ell, y_m) = \delta_{i\ell}\delta_{jm}$. Thus the support for $\chi_{i,j}(x, y)$, which we shall denote by $S_{i,j}$, is just the four elements containing the vertex (x_i, y_j). Now, for convenience, assuming $\Omega_h = \Omega$,

$$\pi(\psi; v, \chi_{i,j}) = \sum_K \int_K J_p \cdot \nabla \chi_{i,j}.$$

Since the only non-zero contributions to this come from $S_{i,j}$, the support of $\chi_{i,j}(x, y)$, we define the discrete analogue of $\pi(\psi; v, \chi)$ by

$$\begin{aligned} \pi_h(\psi; v, \chi_{i,j}) &= \sum_{K \in S_{i,j}} \int_K J_p \cdot \nabla \chi_{i,j} \\ &= \sum_{K \in S_{i,j}} e^{-\psi_K} \int_K B^K \nabla v^h \cdot \nabla \chi_{i,j}. \end{aligned}$$

Now, let q be the number of nodes in $\Omega_h \setminus \Gamma_h^1$, then there exists a basis of V_0^h consisting of q functions of the form $\chi_{i,j}$. For convenience let us denote these as $\chi_k, k = 1, \cdots, q$. Then for any element χ of V_0^h , $\chi = \sum_{k=1}^q \chi^k \chi_k(x, y)$ and we define

$$\pi_h(\psi; v, \chi) = \sum_{k=1}^q \chi^k \pi_h(\psi; v, \chi_k). \tag{3.13}$$

We remark that the form π_h is symmetric with respect to both v^h and $\chi_{i,j}$. As shown in section 2, it reduces to the Scharfetter-Gummel discretization in one dimension. Zlamal shows that the solution of the discrete problem exists, under certain consistency assumptions. A maximum principle is also proved for rectangles, parallelograms with diagonals making angles not greater than $\pi/2$ and acute triangles.

The procedure for obtaining the approximation to J, given by (3.10) and (3.11), may alternatively be considered as replacing e^ψ by an inverse average type approximation along element sides. Further analysis of such schemes is given in [11].

3.2.2. Hybrid Methods – Brezzi, Marini, Pietra

We shall consider a class of methods proposed in [3]. Previous methods such as those discussed above use some harmonic average on suitably chosen lines, for example boundaries of the elements. In order to evolve a method which is less one dimensional in nature and more amenable of analysis, Brezzi, Marini and Pietra propose discretizing the continuity equation by a mixed or hybrid method, which effectively uses a harmonic average of e^{ψ} over the whole element. This permits a more standard finite element analysis of the scheme.

The first method they propose is a hybrid method on triangles and rectangles. The derivation and analysis on triangles is significantly simpler in this case and we shall consider it first. As above we assume that there is a mixture of Dirichlet and Neumann boundary conditions satisfying (3.3). We define

$$V_E = \{y \in C^0(\bar{\Omega}), \quad y|_{\Gamma_0} = p*\} \quad , \quad V_0 = \{y \in C^0(\bar{\Omega}), \quad y|_{\Gamma_0} = 0\},$$

$$W = \{y : \ y \in [L^2(\Omega)]^2\},$$

Again we rewrite (1.10) and (1.12) in the Slotboom variables $v = pe^{\psi}$ and derive a weak formulation:

Find $v \in V_E$ and $J_p \in W$ such that

$$\int_\Omega \nabla v \cdot \tau = \int_\Omega e^{\psi} J_p \cdot \tau \ , \quad \forall\, \tau \in W \tag{3.14}$$

$$-\int_\Omega J_p \cdot \nabla \chi = \int_\Omega f\, \chi \ , \quad \forall\, \chi \in V_0 \tag{3.15}$$

We now seek a piecewise linear approximation v^h to v and a piecewise constant approximation J_p^h to J_p. Let $\mathcal{T}_h$ be a triangulation of Ω and define the spaces

$$V_E^h = \{y : \ y \in V_E, \ y|_T \in P_1 \ , \ \forall\, T \in \mathcal{T}_h\},$$

$$V_0^h = \{y : \ y \in V_0, \ y|_T \in P_1 \ , \ \forall\, T \in \mathcal{T}_h\},$$

$$W^h = \{y : \ y \in W, \ y|_T \in P_0 \ , \ \forall\, T \in \mathcal{T}_h\},$$

where P_0 and P_1 are again the spaces of polynomials of total degree 0 and 1 respectively. The hybrid finite element approximation is then given by:

Find $v^h \in V^h$, $J_p^h \in W^h$ such that

$$\int_\Omega \nabla v^h \cdot \tau = \int_\Omega e^{\psi} J_p^h \cdot \tau \ , \quad \forall\, \tau \in W^h, \tag{3.16}$$

$$\int_\Omega J_p^h \cdot \nabla \chi = \int_\Omega f\, \chi \ , \quad \forall\, \chi \in V_0^h. \tag{3.17}$$

In practice, one normally assumes that ψ is piecewise linear so the computations in (3.16)(3.17) can be performed.

We now define a piecewise constant function $\bar{\psi}$, satisfying

$$e^{\bar{\psi}}|_T = \int_T e^{\psi}/|T| \quad , \quad \forall\, T \in \mathcal{T}_h. \tag{3.18}$$

Substituting (3.18) in (3.16) and considering $\tau_{i,1}$ and $\tau_{i,2}$ defined by

$$\tau_{i,1}|_{T_i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \tau_{i,1}|_{\Omega \setminus T_i} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\tau_{i,2}|_{T_i} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad , \quad \tau_{i,2}|_{\Omega \setminus T_i} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

we obtain, for $k = 1, 2$,

$$\sum_{T_j \in \mathcal{T}_h} \int_{T_j} \nabla v^h \cdot \tau_{i,k} = \sum_{T_j \in \mathcal{T}_h} \int_{T_j} e^{\psi} J_p^h \cdot \tau_{i,k}.$$

Now, using $v^h|_{T_j} \in P_1$ and $J_p^h|_{T_j} \in P_0$, we can rewrite this as

$$\sum_{T_j \in \mathcal{T}_h} \nabla v^h \cdot \tau_{i,k} \int_{T_j} 1 \; = \sum_{T_j \in \mathcal{T}_h} J_p^h \cdot \tau_{i,k} \int_{T_j} e^{\psi}$$

or, recalling the definition of $\tau_{i,1}$ and $\tau_{i,2}$,

$$\nabla v^h \; |T_i| \;\; = \;\; J_p^h \int_{T_i} e^{\psi},$$

which is just

$$\nabla v^h \;\; = \;\; J_p^h e^{\bar{\psi}}. \tag{3.19}$$

Substituting (3.19) into (3.17) gives a new *conforming discretization* of (1.13), which uses the *harmonic average* $e^{\bar{\psi}}$ for e^{ψ} rather than the usual one:

$$\int_\Omega e^{\bar{\psi}} \nabla v^h \cdot \nabla \chi = \int_\Omega f \; \chi \; , \quad \forall \; \chi \in V_0^h. \tag{3.20}$$

In practice, since ψ can be large, $e^{\bar{\psi}}$ can give rise to overflow problems on a computer. Therefore it is standard practice to revert to the hole density p, obtaining the scheme:

$$\int_\Omega e^{\bar{\psi}} \nabla (e^{\psi} p^h)^I \cdot \nabla \chi = \int_\Omega f \; \chi \; , \quad \forall \; \chi \in V_0^h. \tag{3.21}$$

where z^I denotes the piecewise linear interpolant to z at the vertices of the triangulation $\mathcal{T}_h$. It is shown, in [3], that the discretization matrix of this scheme is an M-matrix provided that all the triangles are of weakly acute type.

The equivalent derivation for rectangles or squares does not lead to a single conforming discretization similar to (3.21). Instead we must eliminate J_p^h from the system of equations by static condensation. The resulting matrix will however be a symmetric positive-definite M-matrix.

Brezzi, Marini and Pietra also consider Mixed-Methods on triangles and squares. These differ from the previous approaches, particularly in that they seek piecewise linear approximations to the hole current J_p and only piecewise constant approximations to the $v = pe^{\psi}$. To be more precise, in the case of a triangularization $\mathcal{T}_h$ of Ω define

$$RT(T) = \{\mathbf{z} = (z_1, z_2) : \ z_1 = a + bx, \ z_2 = c + by, \ \ a, b, c \in I\!R\}$$

and

$$W^h = \{\mathbf{z} : \ \mathbf{z} \in [L^2(\Omega)]^2, \nabla \cdot \mathbf{z} \in L^2(\Omega), \\ \mathbf{z} \cdot \hat{n} = 0 \ \text{ on } \ \Gamma_1, \ \mathbf{z}|_T \in RT(T) \ , \ \forall \ T \in \mathcal{T}_h\},$$

$$V^h = \{z : \ \ z \in L^2(\Omega), , \ \ z|_T \in P_0(T) \ , \ \forall \ T \in \mathcal{T}_h\}.$$

The mixed finite element approximation is then given by:

Find $v^h \in V^h$, $J_p^h \in W^h$ such that

$$\int_\Omega e^{\psi} J_p^h \cdot \tau = \int_\Omega v^h \nabla \cdot \tau + \int_{\Gamma_0} \chi \tau \cdot \hat{n} \ , \quad \forall \ \tau \in W^h, \tag{3.22}$$

$$\int_\Omega \nabla \cdot J_p^h \phi = \int_\Omega f \ \phi \ , \quad \forall \ \phi \in V_0^h. \tag{3.23}$$

The resulting discretization matrix is symmetric, but is neither positive-definite nor an M-matrix. Further it is not possible to eliminate J_p^h directly by static condensation. It is however possible to produce an "extended" discrete problem, which after application of element-wise static condensation twice, leads to a symmetric positive-definite M-matrix, provided that the triangulation $\mathcal{T}_h$ is of weakly acute type. A similar derivation can be performed on squares.

3.3 Convergence Theory in Two Dimensions

The convergence theory for two and higher dimensions poses significantly greater difficulty than in the one dimensional case. The authors are unaware of any uniform convergence results in the literature comparable to those discussed previously for the one dimensional case. In fact most classical convergence results are

unduly pessimistic in the sense that, in practice, the Scharfetter-Gummel scheme and its variations approximate the true solution under much less stringent assumptions on regularity of the solution and fineness of the mesh.

The most satisfactory analyses appear to be those using the finite-element formulation. However classical finite-element proofs do not tend to yield *uniform* convergence results. This is consistent with experience elsewhere in the analysis of numerical schemes for singularly perturbed problems. For example, Brezzi, Marini and Pietra [3], employ standard techniques but establish error bounds only for the current approximation in their hybrid and mixed methods. These are in a weighted L^2-norm given by

$$a(J,\tau) := \int_\Omega J \cdot \tau e^{\psi} \quad , \quad \forall\, J,\tau \in [L^2(\Omega)]^2 \,,$$
$$\|J\|_a := a(J,J)^{1/2}.$$

They bound $\|J_p - J_p^h\|_a$ in terms of certain other projections of J_p onto W^h. These are somewhat technical to describe, and the results are still deemed "unsatisfactory" by the authors in [3].

In [14,15], Mock analyses a class of finite-element methods on triangles and quadrilaterals, which includes Scharfetter-Gummel type schemes. The aim of the analysis is to show that, under realistic assumptions on the physical entities, the schemes are convergent. In particular, he avoids assumptions on the variation of p, $v = pe^{\psi}$, and e^{ψ} between mesh points. He assumes only that J_p and some functions of the right-hand side f are approximated accurately on the mesh. The method involves writing (2.1) in the mixed form in the Slotboom variables and using a special formulation involving stream functions and a function space of "pipes". This is necessary, since v varies rapidly near interfaces and cannot be approximated accurately, on a coarse mesh, from a space of piecewise polynomials. Mock concludes that only variations in J_p and f have to be resolved by the mesh in order to obtain reasonable discretization errors.

In [11], Markowich and Zlamal consider a generalization of the method in [19] using piecewise linear trial and test functions on a triangular mesh of acute type satisfying a minimum angle requirement. The assumptions made are that Ω has a polygonal boundary $\partial\Omega = \Gamma_1 + \Gamma_2$, with Dirichlet conditions on Γ_1 of the form $p|_{\Gamma_1} = p^*$, where $p^* \in H^1(\Omega) \cap C(\Gamma_1)$, and $\frac{\partial p}{\partial n}|_{\Gamma^2} = 0$. In addition, it is assumed that $f \in H^2(\Omega)$ and that

$$-\infty < \underline{\psi} \leq \psi \leq \bar{\psi} < \infty.$$

Under these assumptions they derive a bound on the nodal error in the H^1-norm for the error in approximating the hole density p. To be more precise, they bound

the difference $p^I - p^h$ between the piecewise linear interpolant p^I of p and the discrete solution p^h :

$$\|p^I - p^h\|_{H^1(\Omega)} \leq \frac{K}{e^{\underline{\psi}}} \left(h \| \left| \frac{\partial J_p}{\partial x} \right| \|_\infty + h^2 \|f\|_{H^2(\Omega)} \right),$$

where $|\frac{\partial J_p}{\partial x}|$ is the Jacobian of J_p, $h = \max_{T \in \mathcal{T}_h} \operatorname{diam}(T)$, and K depends only on the minimum angle of $\mathcal{T}_h$, on Ω and on Γ_1. This does not translate however to a similar bound on $\|p - p^h\|_{H^1(\Omega)}$, since the inverse average finite-element method used only depends on the values of ψ along the element edges. Nevertheless this is a uniform nodal approximation result of a form familiar from the singular perturbation literature and moreover it is shown under reasonable assumptions on J_p. Markowich and Zlamal also derive an L^2-type estimate on the approximation of the current J_p. They remark that in this method the current on a triangle $T \in \mathcal{T}_h$ is not uniquely defined. This holds, since J^h is element-wise constant and any of the three values at the vertices $J^h_{T,i}, i = 1, 2, 3$ can be regarded as the appropriate approximation, in the sense that:

$$\left(\sum_{T \in \mathcal{T}_h} \int_T | J_p - J^h_{T,i}|^2 \right)^{1/2} \leq \frac{\bar{\psi}}{\underline{\psi}} \left(h \| \left| \frac{\partial J_p}{\partial x} \right| \|_\infty + h^2 \|f\|_{H^2(\Omega)} \right).$$

They also show that the error in the outflow current at a contact is bounded in a similar manner.

The only convergence analysis of the full nonlinear system (1.8)-(1.12), that we are aware of, is given by Jerome and Kerkhoven in [7]. It uses a calculus (developed by Krasnosel'skii and co-workers) for fixed points of nonlinear operators and their discretizations. The main thrust is to verify a nonlinear approximation theory, in the energy norm, with convergence rates keyed to those in the linear theory.

3.4 Numerical Results

We shall conclude by presenting some numerical results and observations. We shall consider the same test problem as in [3], that is:

$$\begin{aligned} Lp := \nabla \cdot (\nabla p + p \nabla \psi) &= f, \quad \text{in } \Omega \\ p &= g, \quad \text{on } \Gamma_1 \\ (\nabla p + p \nabla \psi) \cdot \hat{n} &= h, \quad \text{on } \Gamma_2 \end{aligned}$$

where the domain Ω is the unit square and f, g, and h are given functions. Further assume that $\Gamma_1 \cap \Gamma_2 = \phi$ and $\overline{\Gamma}_1 \cup \overline{\Gamma}_2 = \partial\Omega$. Here Γ_1 represent the ohmic contacts (sources, drains, etc.), and, with $h \equiv 0$, Γ_2 consists of the

insulation contacts. For the specific example in the illustrations we assume that the boundary conditions on Γ_1 are given by :

$$p(x,y) = \begin{cases} .3, & x = 0, \quad y \leq .25 \\ .3, & x \leq .25, \quad y = 0 \\ 0, & x = 1, \quad y \geq .75 \\ 0, & x \geq .75, \quad y = 1 \end{cases},$$

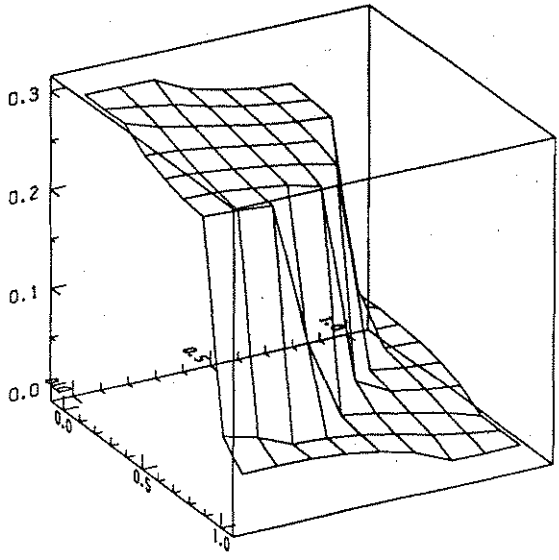

Fig. 1: Scharfetter-Gummel, $h = 1/8$, $k = 1$

We shall, in addition, assume that the electrostatic potential ψ has been calculated previously and is given by $\psi(x,y) = 10^{-k}\psi_0(x,y)$, where $\psi_0(x,y)$

is given by

$$\psi_0(x,y) = \begin{cases} 0, & 0 \le r \le .8 \\ 2(r-.8), & .8 \le r \le .9 \\ .2, & .9 \le r \end{cases},$$

where $r := \sqrt{x^2+y^2}$, and $k = 0, 1, 2$. It is clear that we may alter the steepness of the gradient of ψ in the layers by adjusting k.

We shall now solve this problem using a Scharfetter-Gummel discretization on a rectangular grid. The results are illustrated in figures 1, 2 and 3.

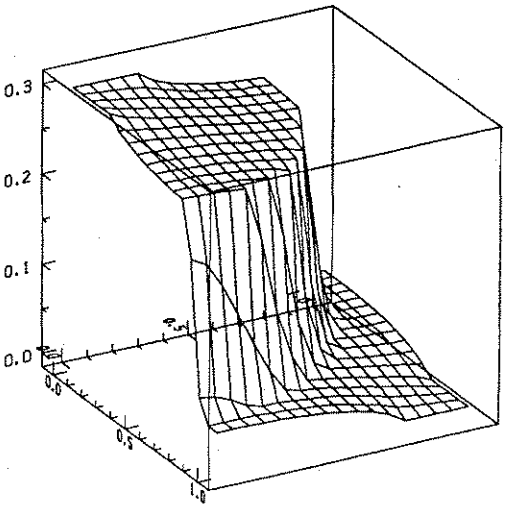

Fig. 2: Scharfetter-Gummel, $h = 1/16$, $k = 1$

Since Scharfetter-Gummel is an exponentially fitted method we expect that it will provide qualitatively correct results even on relatively coarse grids. This proves to be the case. It should be noted that it is not necessary to refine the mesh in the neighbourhood of the layers. This is, as remarked in [10], in contrast to the solution of the Poisson equation for ψ, where refinement of the mesh is required in addition to exponential fitting.

We remark that the Scharfetter-Gummel scheme is formally second-order accurate in the $L^2(\Omega)$ norm. An examination of the errors for various numerical test problems reveals that this order is indeed achieved *provided* the true solution has sufficient regularity—it is required that u is in $H^2(\Omega)$. However for problems with "realistic" boundary conditions, a lesser rate of convergence is observed. Table 1 below contains the results of numerical simulations on three test problems. Problem 1 is precisely the problem described above, whose approximate solution profiles are pictured in Figures 1, 2, and 3. Problem 2 is a very similar problem; it differs by having Dirichlet boundary conditions all around, however these are constructed to have a "square-root" singularity identical to that of Problem 1. Problem 3 has "smooth" Dirichlet boundary conditions.

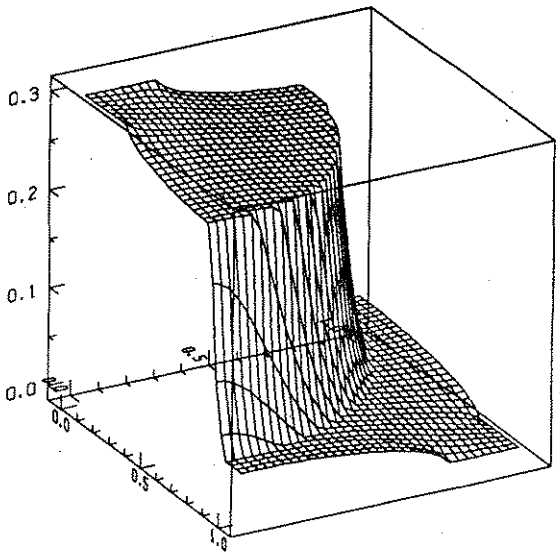

Fig. 3: Scharfetter-Gummel, $h = 1/32$, $k = 1$

The true solutions for these problems are not known; the errors were estimated using extrapolations on three consecutive mesh refinements. The third

problem indicates the expected second-order convergence for smooth solutions. The exact solutions of the first two problems are only in $H^{3/2}(\Omega)$, and the exhibited convergence rates are $O(h)$ and $O(h^{3/2})$ respectively. Both problems suffer degradation due to the lack of regularity of u, and in addition, the order of the approximation associated with Problem 1 is lower than that suggested by standard approximation theory. To our knowledge, this is the first numerical illustration of this particular phenomenon.

	Problem 1		Problem 2		Problem 3	
h	$\|e^h\|_{h,2}$	p	$\|e^h\|_{h,2}$	p	$\|e^h\|_{h,2}$	p
1/4	.10(-1)	.90	.43(-2)	1.55	.20(-1)	2.17
1/8	.54(-2)	.99	.15(-2)	1.54	.45(-2)	2.06
1/16	.27(-2)		.51(-3)		.11(-2)	

Table 1: Errors in the discrete L^2 norm

Acknowledgments

The authors wish to express their gratitude to Professor John J.H. Miller and Professor Joseph W. Jerome respectively, for their help and encouragement in cultivating an interest in and an understanding of the numerical solution of the semiconductor equations. We also wish to express our thanks to Reinhardt R. Zeller for performing the numerical simulations presented above.

References

[1] **R.E. Bank, D.J. Rose** *Some error estimates for the box method,* SIAM J. Numer. Anal. 24(4):777–787, August 1987.

[2] **R.E. Bank, D.J. Rose, W. Fichtner** *Numerical methods for semiconductor device simulation,* SIAM J. Sci. Stat. Comput. 4(3):416–435, September 1983.

[3] **F. Brezzi, L.D. Marini, P. Pietra** *Méthodes d'éléments finis mixtes et schéma de Scharfetter-Gummel,* C. R. Acad. Sci. Paris Sér. I Math. 305(13):599–604, 1987.

[4] **F. Brezzi, L.D. Marini, P. Pietra** *Two-dimensional exponential fitting and applications to drift-diffusion models,* SIAM J. Numer. Anal. 26(6):1342–1355, December 1989.

[5] **P.A. Farrell** *Sufficient conditions for uniform convergence of a difference scheme for a singularly perturbed problem in conservation form,* in Proc. BAIL III (J.J.H. Miller ed.) 203–208, Boole Press, Dublin, 1984.

[6] **C.J. Fitzsimons** *A novel method for solving the continuity equations,* COMPEL 6(2):71–76, 1987.

[7] **K. Hess, J.-P. Leburton, U. Ravaioli** *Computational Electronics. Semiconductor Transport and Device Simulation,* Kluwer Academic Publishers, Boston/Dordrecht/London, 1991.

[8] **J.W. Jerome, T. Kerkhoven** *A finite element approximation theory for the drift diffusion semiconductor model*, University of Illinois Report no. UIUCDCS-R-89-1542 from the Department of Computer Science, 1304 W. Springfield Avenue, Urbana, IL 61801 (to appear in SIAM J. Appl. Math. in 1991).

[9] **P.A. Markowich** *The Stationary Semiconductor Device Equations*, Springer-Verlag, Wien, 1986

[10] **P.A. Markowich, C.A. Ringhofer, S. Selberherr** *A singular perturbation approach for the analysis of the fundamental semiconductor equations*, MRC Technical Summary Report #2482, Mathematics Research Center, University of Wisconsin-Madison, 610 Walnut Street, Madison, Wisconsin 53706, February 1983.

[11] **P.A. Markowich, C.A. Ringhofer, S. Selberherr, M. Lentini** *A singular perturbation approach for the analysis of the fundamental semiconductor equations*, IEEE Trans. Electron Devices ED-30(9):1165–1180, 1983.

[12] **P.A. Markowich, M. Zlámal** *Inverse-average-type finite element discretizations of selfadjoint second-order elliptic problems*, Math. Comp. 51(184):431–449, 1988.

[13] **J.J.H. Miller** *On the inclusion of the recombination term in discretizations of the semiconductor device equations*, Math. Comput. Siconductors, Bell Syst. Tech. J. 29:560–607, 1950.

[14] **R.S. Varga** *Matrix Iterative Analysis* Prentice Hall, New York, 1962.

[15] **M. Zlámal** *Finite element solution of the fundamental equations of semiconductor devices. I. Math. Comp. 46(173):27–43, 1986.*