DISCRETIZATION METHODS FOR THE SEMICONDUCTOR EQUATIONS

S.Selberherr[1)] and Ch.Ringhofer[2)]

1) Abteilung Physikalische Elektronik
Institut für Allgemeine Elektrotechnik und Elektronik
TU Wien, Gußhausstraße 27, A-1040 Wien, AUSTRIA

2) Mathematics Research Center
University of Wisconsin, Madison, WI 53706, USA

Abstract: The basic mathematical results on the elliptic boundary value problem which corresponds to the equations involved in the numerical simulation of semiconductor devices are reviewed. Particularly, smoothness and structure of the solutions of the fundamental semiconductor equations are discussed. The singular perturbation approach to the numerical solution of the semiconductor equations is presented. The implications of the results obtained with the singular perturbation approach on the application of Finite Difference methods and Finite Element methods are discussed. Criteria for an optimal mesh generation strategy are given. An example shows the power of these concepts combined with modern numerical methods in comparison to classical approaches.

1. INTRODUCTION

The characteristic feature of early modeling was the separation of the interiour of the device into different regions, the treatment of which could be simplified by various assumptions like special doping profiles, complete depletion and quasineutrality. These separately treated regions were simply connected to produce the overall solution. If analytic results are intended, any other approach is prohibitive. Fully numerical modeling based on partial differential equations /36/ which describe all different regions of semiconductor devices in one unified manner was first suggested by Gummel /18/ for the one dimensional bipolar transistor. This approach was further developed and applied to pn-junction theory by De Mari /9/, /10/ and to IMPATT diodes by Scharfetter and Gummel /32/. A two dimensional numerical analysis of a semiconductor device was carried out the first time by Kennedy and O'Brien /22/ who investigated the junction field effect transistor. Since then two dimensional modeling has been applied to fairly all important semiconductor devices. There are so many papers of excellent repute that it would be unfair to cite only a few. The time dependence has been investigated by e.g. /24/, /30/ and models for three space dimensions have been announced by e.g. /7/, /39/.

2. THE FUNDAMENTAL SEMICONDUCTOR EQUATIONS

The most familiar model of carrier transport in a semiconductor device has been proposed by Van Roosbroeck /36/. It consists of Poisson's equation (2.1), the current continuity equations for electrons (2.2) and holes (2.3) and the current relations for electrons (2.4) and holes (2.5)

$$\mathrm{div}\ \varepsilon\ \mathrm{grad}\ \psi = -q\cdot(\ p - n + C\ ) \tag{2.1}$$

$$\mathrm{div}\ \vec{J}_n = -q\cdot R \tag{2.2}$$

$$\mathrm{div}\ \vec{J}_p = q\cdot R \tag{2.3}$$

$$\vec{J}_n = -q\cdot(\ \mu_n\cdot n\cdot \mathrm{grad}\ \psi - D_n\cdot \mathrm{grad}\ n\ ) \tag{2.4}$$

$$\vec{J}_p = -q\cdot(\ \mu_p\cdot p\cdot \mathrm{grad}\ \psi + D_p\cdot \mathrm{grad}\ p\ ) \tag{2.5}$$

However, it is of prime importance to note that the equations (2.4) and (2.5) do not characterize effects which are caused by degenerate semiconductors. /25/, /35/, /37/ discuss some modifications of the current relations, which partially take into account the consequences introduced by degenerate semiconductors (e.g. invalidity of Boltzmann's statistics, bandgap narrowing). Just as further examples (2.4) and (2.5) do not describe velocity overshoot phenomena /17/, /34/; and certainly no effects which are due to ballistic transport /16/, the existence of which is still questionable /19/, are included.

3. ANALYTICAL INVESTIGATIONS ABOUT THE SEMICONDUCTOR EQUATIONS

In this chapter we present some of the existing analytical results for the fundamental semiconductor equations. Particularly, we are interested in the possible boundary conditions, dependent variables and an appropriate scaling approach. We shall discuss the structure of solutions to the semiconductor equations, because these results are of importance in both the theoretical and practical context, since - as we shall see in the next chapter - the knowledge of the structure and smoothness properties of solutions is indeed essential for the development of a numerical solution method.

3.1 Domain and Boundary Conditions

Most of the existing programs which solve the semiconductor equations are restricted to a rectangular device geometry. This is not essential as far as the analysis of the equations is concerned. In this chapter we shall assume that the equations (2.1)-(2.5) are posed in a domain D of $\mathbb{R}^n$ (n=1,2,3) with a piecewise smooth boundary ∂D. Equations (2.1)-(2.5) are subject to a mixed set of Dirichlet and Neumann boundary conditions. That means ∂D consists of three parts $\partial D=\partial D_1 \cup \partial D_2 \cup \partial D_3$. ∂D_1 denotes the part of the boundary where the device is surrounded by insulating material. There one assumes the boundary conditions:

$$\partial\psi/\partial\vec{n}_{\perp} = \partial n/\partial\vec{n}_{\perp} = \partial p/\partial\vec{n}_{\perp} = 0 \tag{3.1}$$

Here $\vec{n}_\perp$ denotes the unit normal vector on ∂D which exists anywhere except at a finite number of points (arbitrarily defined corners of the simulation geometry). ∂D_2 denotes the part of the boundary corresponding to the ohmic contacts. There ψ, n and p are prescribed. The boundary conditions can be derived from the applied bias ψ_D and the assumptions of thermal equilibrium and vanishing space charge:

$$\psi = \psi_D + \psi_{built-in}, \quad n \cdot p = n_i^2, \quad n - p - C = 0 \tag{3.2}$$

The last two conditions in (3.2) can be rewritten as:

$$n = (\sqrt{C^2 + 4 \cdot n_i^2} + C)/2$$
$$p = (\sqrt{C^2 + 4 \cdot n_i^2} - C)/2 \tag{3.3}$$

Modeling MOS devices one has also to account for controlled insulator-semiconductor interfaces. ∂D_3 denotes the part of the boundary which corresponds to such an interface. There we have the interface conditions:

$$\vec{J}_n \cdot \vec{n}_\perp = \vec{J}_p \cdot \vec{n}_\perp = 0$$
$$\varepsilon_{sem} \cdot \partial \psi / \partial \vec{n}_\perp \big|_{sem} = \varepsilon_{ins} \cdot \partial \psi / \partial \vec{n}_\perp \big|_{ins} \tag{3.4}$$

Again $\vec{n}_\perp$ denotes the normal vector on ∂D. ε_{sem} and ε_{ins} denote the permittivity constants for the semiconductor and the insulator respectively. $\partial \psi / \partial \vec{n}_\perp |_{sem}$ and $\partial \psi / \partial \vec{n}_\perp |_{ins}$ denote the onesided limits of the derivatives perpendicular to the interface approaching the interface. Within the insulator the Laplace equation: div grad ψ = 0 holds.

3.2 Dependent Variables

For analytical purposes it is often useful to use other variables than n and p to describe the system (2.1)-(2.5). Two other sets of variables which are frequently employed are $(\psi, \varphi_n, \varphi_p)$ and (ψ, u, v) which relate to the set (ψ, n, p) by:

$$n = n_i \cdot e^{(\psi - \varphi_n)/U_t}, \quad p = n_i \cdot e^{(\varphi_p - \psi)/U_t} \tag{3.5}$$

$$n = n_i \cdot e^{\psi/U_t} \cdot u, \quad p = n_i \cdot e^{-\psi/U_t} \cdot v \tag{3.6}$$

(3.5) can be physically interpreted as the application of Boltzmann statistics. However (3.5) also can be regarded as a purely mathematical change of variables so that the question of the validity of the Boltzmann statistics does not need to be considered. The use of $(\psi, \varphi_n, \varphi_p)$ a priori excludes negative carrier densities n and p, which may be present as undesired nonphysical solutions of (2.1)-(2.5) if we use (ψ, n, p) or (ψ, u, v) as dependent variables. As we will see later in this chapter the advantage of the set (ψ, u, v) is that the continuity equations (2.2), (2.3) and current relations (2.4), (2.5)

become self-adjoint. However, owing to the enormous range of the values of u and v, the sets (ψ, n, p) or $(\psi, \varphi_n, \varphi_p)$ have to be prefered for actual computations. We personally favour the set (ψ, n, p).

3.3 Scaling

Since the dependent variables in the system (2.1)-(2.5) are of different order of magnitude and show a strongly different behaviour in regions with small and large space charge the first step towards a structural analysis of (2.1)-(2.5) has to be an appropriate scaling. A standard way of scaling (2.1)-(2.5) has been given by De Mari /9/. There ψ is scaled by the thermal voltage U_t, n and p are scaled by n_i (similar to Mock /29/) and the independent variables are scaled such that all multiplying constants in Poisson's equation become unity. Although physically reasonable this approach has the disadvantage that n and p in general are still several orders of magnitude larger than ψ. A scaling which reduces ψ, n and p to the same order of magnitude has been given by Vasiliev'a and Butuzov /38/. This approach makes the system (2.1)-(2.5) accessible to an asymptotic analysis which is given together with applications in /26/, /27/ and /28/. There n and p are scaled by the maximum absolute value of the net doping C and the independent variables are scaled by the characteristic length of the device. More precisely the following scaling factors are employed.

quantity	symbol	value	
$\vec{x}$	l	$\max(\vec{x}-\vec{y})$, $\vec{x},\vec{y}$ in D	
ψ	U_t	$k \cdot T/q$	(3.7)
n,p	α	$\max\|C\|$	

After scaling the equations become:

$$\lambda^2 \cdot \text{div grad } \psi = n - p - C \tag{3.8}$$

$$\text{div} (\text{ grad } n - n \cdot \text{grad } \psi) = -R$$

$$\text{div} (\text{ grad } p + p \cdot \text{grad } \psi) = -R$$

Here, for simplicity only, μ_n and μ_p have been assumed to be constant. It should be noted that the following analysis also holds if the usual smooth dependence of μ_n and μ_p on n, p and grad ψ e.g./33/ is assumed. Since the independent variable x has been scaled, equations (3.8) are now posed on a domain D^S with maximal diameter equal to one. The small constant λ^2 multiplying the Laplacian in (3.8) is the minimal Debye length of the device:

$$\lambda^2 = \frac{\varepsilon \cdot Ut}{l^2 \cdot q \cdot \alpha} \tag{3.9}$$

l and α are defined in (3.7). Thus for high doping ($\alpha >> 1$) λ^2 will be small. For instance for a silicon device with characteristic length $25 \mu m$ and $\alpha = 10^{20} cm^{-3}$ we compute for λ^2 at approximate room temperature T=300K: $\lambda^2 = 4.10^{-10}$.

R denotes the scaled generation/recombination rate, which is in general a (not necessarily mildly) nonlinear function of n,p and

$\text{grad}\psi$. Thus different models of R may influence analysis results quite drastically. This is obviously to be expected as in many operating conditions the device behaviour depends strongly on the net generation/recombination R.

3.4 The Singular Perturbation Approach

(3.8) represents a singularly perturbed elliptic system with perturbation parameter λ. The advantage of this interpretation is that we can now obtain information about the structure of solutions of (3.8) by using asymptotic expansions: In the subdomains of D^S where the solutions behave smoothly we expand them into power series of the form:

$$w(\vec{x},\lambda) = \sum_{i=0}^{\infty} \tilde{w}_i(\vec{x}) \cdot \lambda^i, \quad w=(\psi,n,p)^T \tag{3.10}$$

which implies a smooth dependence on λ. C - the scaled doping - is smooth in these subdomains and exhibits a sharp transition across the pn-junctions in the device. For the case of an abrupt junction this behaviour is represented by a discontinuity across an n-1 dimensional manifold $\Gamma:(\vec{x}=\vec{x}(s)$, s of $\mathbb{R}^{n-1})$ in the device. Thus Γ is a point in 1 dimension, a curve in 2 dimensions and a surface in 3 dimensions. Of course one curve or surface has to be used for each junction. Since the procedure is the same for each of the junctions it is demonstrated only for one junction. In the case of an exponentially graded doping profile C consists of two parts:

$$C = \tilde{C} + \hat{C} \tag{3.11}$$

where $\tilde{C}$ and $\hat{C}$ are discontinous, $\tilde{C}$ is piecewise smooth and $\hat{C}$ is exponentially decaying to zero away from Γ. In the vicinity of Γ the expansion (3.10) is not valid and has to be supplemented by a "layer" term according to the singular perturbation analysis:

$$w(\vec{x},\lambda) = \sum_{i=0}^{\infty} [\tilde{w}_i(\vec{x}) + \hat{w}_i(s,t/\lambda)] \cdot \lambda^i, \quad w=(\psi,n,p)^T \tag{3.12}$$

Here the following coordinate transformation has been employed: For a point in the vicinity of Γ s denotes the parameter value at the nearest point on Γ and t denotes its distance perpendicular to Γ. Thus the solution of the semiconductor equations exhibits internal layers at pn-junctions.

The $\tilde{w}_i$ and $\hat{w}_i$ in (3.12) can now be determined separately and the structure of the solution is given by its partition into the smooth part $\sum \tilde{w}_i \cdot \lambda^i$ and its rapidly varying part $\sum \hat{w}_i \cdot \lambda^i$. $\tilde{w}_o$ has to satisfy the reduced equations:

$$0 = \tilde{n}_o - \tilde{p}_o - \tilde{C} \tag{3.13}$$

$$\text{div}\,(\text{grad}\,\tilde{n}_o - \tilde{n}_o \cdot \text{grad}\tilde{\psi}_o) = -\tilde{R} \tag{3.14}$$

$$\text{div}\,(\text{grad}\,\tilde{p}_o + \tilde{p}_o \cdot \text{grad}\tilde{\psi}_o) = -\tilde{R} \tag{3.15}$$

For the sake of simplicity but without loss of generality the mobilities μ_n and μ_p have been assumed to be constant. (3.13)-(3.15) is subject to the boundary conditions (3.1)-(3.4). Of course the condition of vanishing space charge is redundant with (3.13). Since $C^{\sim}$ is discontinous at Γ and (3.13)-(3.15) represents a second order system of two equations four "interface conditions" have to be imposed at Γ. They are of the form:

$$n_o^{\sim}\cdot e^{-\psi_o^{\sim}}|_{\vec{x}=\vec{x}-} = n_o^{\sim}\cdot e^{-\psi_o^{\sim}}|_{\vec{x}=\vec{x}+} \tag{3.16}$$

$$p_o^{\sim}\cdot e^{\psi_o^{\sim}}|_{\vec{x}=\vec{x}-} = p_o^{\sim}\cdot e^{\psi_o^{\sim}}|_{\vec{x}=\vec{x}+} \tag{3.17}$$

$$\vec{J}_{n_o}^{\sim}\cdot\vec{n}_{\perp}|_{\vec{x}=\vec{x}-} = \vec{J}_{n_o}^{\sim}\cdot\vec{n}_{\perp}|_{\vec{x}=\vec{x}+} \tag{3.18}$$

$$\vec{J}_{p_o}^{\sim}\cdot\vec{n}_{\perp}|_{\vec{x}=\vec{x}-} = \vec{J}_{p_o}^{\sim}\cdot\vec{n}_{\perp}|_{\vec{x}=\vec{x}+} \tag{3.19}$$

where $w|_{\vec{x}-}$ and $w|_{\vec{x}+}$ denote the onesided limits of w as $\vec{x}$ tends to Γ from each side. $\vec{n}_{\perp}$ denotes the unit normal vector on Γ. $\vec{J}_{n_o}^{\sim}$ and $\vec{J}_{p_o}^{\sim}$ are the zeroth order terms of the smooth parts of the (scaled) electron and hole current densities.

(3.13)-(3.15) together with (3.16)-(3.19) and the boundary conditions (3.1)-(3.4) define the reduced problem whose solution is an $O(\lambda)$ approximation to the full solution away from Γ. As we will see in the next chapter the reduced problem is a useful tool for the development and analysis of numerical methods, since it (especially the conditions (3.16)-(3.19)) has to be solved implicitly by any discretisation method which requires a reasonable number of grid points.

The equations for the rapidly varying parts $w_i^{\wedge}$ reduce to ordinary differential equations. That means that only derivatives with respect to the "fast" variable t/λ occur. Since the rate of decay of $w_i^{\wedge}$ depends heavily on ψ the width of the layer grows with the applied voltage; a fact which is absolutely well known by device physicists, but which becomes nicely apparent by the singular perturbation approach.

4. NUMERICAL SOLUTION OF THE SEMICONDUCTOR EQUATIONS

In this chapter we discuss some of the problems occuring in the numerical solution of the semiconductor equations and the analysis of existing numerical methods. From the viewpoint of numerical analysis there are essentially three major topics to be considered. The first one is the type of discretisation to be used. There exist programs for both Finite Element and Finite Difference discretisations of the system (2.1)-(2.5). As outlined in the previous chapter the solution exhibits a smooth behaviour in some subregions of the domain whereas in others it varies rapidly. Thus a nonuniform mesh is mandatory and adaptive mesh refinement is desirable. So the second topic is the question how to set up the mesh refinement algorithm i.e. which quantities have to be used to control the mesh. Each type of discretisation will lead to a large sparse system of nonlinear equations and so the solution of this system is the third topic. However, this topic is not dealt with herein; extensive treatment can be found in the literature e.g. /3/, /11/, /18/, /33/.

For the sake of simplicity in nomenclature we shall only consider the two-dimensional case in this chapter. However, all results given in the following can be generalized to three dimensions in a straightforward manner. So, the equations are posed in a domain D of $\mathbb{R}^2$ and $\vec{x} = (x,y)^T$ denotes the independent variable.

4.1 Discretisation Schemes

Using Finite Elements or Finite Differences one has to take into account that Poisson's equation (2.1) is of a different type than the continuity equations. Poisson's equation - in the scaling of Markowich /26/ using the variables (ψ,u,v)

$$\lambda^2 \cdot \text{div grad } \psi = e^{\psi} \cdot u - e^{-\psi} \cdot v - C \tag{4.1}$$

is a singularly perturbed elliptic problem whose right hand side has a positive derivative with respect to ψ. Thus it is of a standard form (as discussed in e.g. /14/) except for the discontinuous or exponentially graded term C. Equations of that type are generally well behaved and it suffices to apply a usual discretisation scheme. In the case of Finite Differences using the index convention given in Fig. 1 equation (4.1) is discretized by:

$$\lambda^2 \cdot (\text{div grad } \psi)_{ij} = n_{ij} - p_{ij} - C(x_i, y_j) \tag{4.2}$$

$$E^x_{i+1/2,j} = (\psi_{i+1,j} - \psi_{i,j})/h_i$$

$$E^y_{i,j+1/2} = (\psi_{i,j+1} - \psi_{i,j})/k_j \tag{4.3}$$

$$h_i = x_{i+1} - x_i, \quad k_j = y_{j+1} - y_j$$

$$(\text{div grad } \psi)_{i,j} = 2 \cdot (E^x_{i+1/2,j} - E^x_{i-1/2,j})/(h_i + h_{i-1}) +$$

$$+ 2 \cdot (E^y_{i,j+1/2} - E^y_{i,j-1/2})/(k_j + k_{j-1}) \tag{4.4}$$

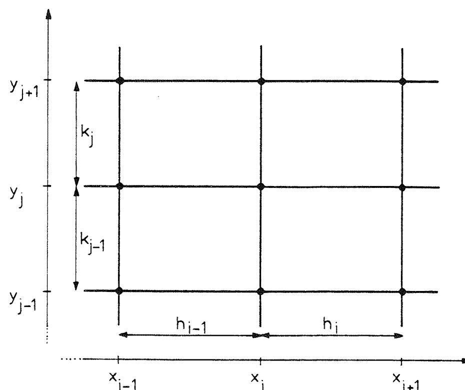

Fig.1 The index convention used

If one of the neighbouring gridpoints (x_{i+1}, y_j), (x_{i-1}, y_j), (x_i, y_{j+1}), (x_i, y_{j-1}) does not exist - as possible in a terminating line approach /1/, /2/ or in the Finite Boxes approach /15/ - (4.4)

has to be modified. In the case of Finite Elements classical shape functions can be used (i.e. linear shape functions for triangular elements, bilinear shape functions for rectangular elements).

It turns out that the discretisation of the continuity equations is more crucial than the discretisation of Poissons's equation. The usual error analysis of discretisation methods provides an error estimate of the form:

$$\max |w_h - w| <= c \cdot H \qquad (4.5)$$

w_h denotes the numerical approximation to $w(x,y)=(\psi,n,p)^T$. H denotes the maximal gridspacing. The constant c will in general depend on the higher order derivatives of w. The singular perturbation analysis /27/ shows that derivatives of $\hat{\psi}$, $\hat{n}$ and $\hat{p}$ in (3.12) are of magnitude $O(\lambda^{-3})$ - $O(\lambda^{-4})$ locally near the junction (λ is defined in (3.9)). /27/ shows also that, even if a nonuniform mesh is used, the amount of gridpoints required to equidistribute the error term in (4.5) can be proportional to λ^{-2} which is of course prohibitive. Therefore a discretisation scheme is needed where the constant c in (4.5) does not depend on the higher derivatives of the rapidly varying terms $\hat{\psi}$, $\hat{n}$ and $\hat{p}$. For the case of Finite Differences such a scheme was given by Scharfetter and Gummel /32/. They approximate:

$$\vec{J}_n = \text{grad } n - n \cdot \text{grad } \psi \qquad (4.6)$$

$$\text{div } \vec{J}_n = \partial J_n^x / \partial x + \partial J_n^y / \partial y = R \qquad (4.7)$$

by:

$$J^x_{n\,i+1/2,j} = \gamma((\psi_{i+1,j} - \psi_{i,j})/2) \cdot (n_{i+1,j} - n_{i,j})/h_i - $$
$$- (n_{i,j} + n_{i+1,j})/2 \cdot (\psi_{i+1,j} - \psi_{i,j})/h_i$$
$$(4.8)$$
$$J^y_{n\,i,j+1/2} = \gamma((\psi_{i,j+1} - \psi_{i,j})/2) \cdot (n_{i,j+1} - n_{i,j})/k_j - $$
$$- (n_{i,j} + n_{i,j+1})/2 \cdot (\psi_{i,j+1} - \psi_{i,j})/k_j$$

$$\gamma(s) = s \cdot \coth(s)$$

$$2 \cdot (J^x_{n\,i+1/2,j} - J^x_{n\,i-1/2,j})/(h_i + h_{i-1}) +$$
$$+ 2 \cdot (J^y_{n\,i,j+1/2} - J^y_{n\,i,j-1/2})/(k_j + k_{j-1}) = R_{i,j} \qquad (4.9)$$

The continuity equation for holes is discretized analogously. Scharfetter and Gummel gave a physical reasoning for the derivation of their scheme. Markowich et al. /27/ proved that in one dimension the Scharfetter-Gummel scheme is uniformly convergent. That means that the error constant c in (4.5) does not depend on the derivatives of $\hat{\psi}$, $\hat{n}$ and $\hat{p}$ in (3.12) and therefore not on λ. For two dimensions /27/ shows that the choice $\gamma(s) = s \cdot \coth(s)$ is necessary for uniform convergence. Exponentially fitted schemes like the Scharfetter-Gummel scheme have been analyzed by Kellog /21/, /20/ and Doolan /12/ (for different classes of problems). The reason for the uniform

convergence of these schemes is that inside the pn-junction layers the interface conditions (3.16)-(3.19) are satisfied automatically if $|\mathrm{grad}\psi|$ is large and the gridspacing is not $O(\lambda)$.

The results for Finite Difference schemes suggest that a similar approach (like the exponentially fitted schemes) should be used in the case of Finite Elements. This fact has been intuitively observed by Engel /13/ for the one-dimensional case. A modeling group at IBM has tried to make use of the Scharfetter-Gummel scheme for Finite Elements in two and three space dimensions /5/, /4/, /8/. However, we have the impression that their approach needs still quite a bit of analysis, although it has been used effectively by other modelists too e.g. /31/. Macheck /23/ has tried to develop a more rigorous discretisation for Finite Elements using exponentially fitted shape functions. He uses classical bilinear shape functions for ψ and

$$\begin{aligned}
\alpha_1(x,y) &= [1 - \varphi_1(x,y)]\cdot[1 - \varphi_2(x,y)] \\
\alpha_2(x,y) &= \varphi_1(x,y) \cdot[1 - \varphi_2(x,y)] \\
\alpha_3(x,y) &= \varphi_1(x,y) \cdot \varphi_2(x,y) \\
\alpha_4(x,y) &= [1 - \varphi_1(x,y)]\cdot \varphi_2(x,y)
\end{aligned} \tag{4.10}$$

for u, and

$$\begin{aligned}
\beta_1(x,y) &= [1 - \sigma_1(x,y)]\cdot[1 - \sigma_2(x,y)] \\
\beta_2(x,y) &= \sigma_1(x,y) \cdot[1 - \sigma_2(x,y)] \\
\beta_3(x,y) &= \sigma_1(x,y) \cdot \sigma_2(x,y) \\
\beta_4(x,y) &= [1 - \sigma_1(x,y)]\cdot \sigma_2(x,y)
\end{aligned} \tag{4.11}$$

for v, where

$$\begin{aligned}
\varphi_1(x,y) &= f(x,\tfrac{\partial\psi}{\partial x}) \qquad & \varphi_2(x,y) &= f(y,\tfrac{\partial\psi}{\partial y}) \\
\sigma_1(x,y) &= f(x,-\tfrac{\partial\psi}{\partial x}) \qquad & \sigma_2(x,y) &= f(y,-\tfrac{\partial\psi}{\partial y})
\end{aligned} \tag{4.12}$$

with: $f(x,a) = (\exp(ax)-1)/(\exp(a)-1)$ (4.13)

The advantage of these shape functions is that they accomodate nicely the layer behaviour of the solution. They degenerate into the ordinary bilinear shape functions when the electric potential is constant. In order to be able to switch from coarse to fine grid spacing in different subdomains transition elements have to be used /23/. However, no theoretical investigations have been carried out so far to analyse the uniform convergence properties of this method.

4.2 Grid Construction

Since subregions of strong variation of ψ, n and p alternate with regions where these quantities behave smoothly (i.e. their gradients are small) different meshsizes are mandatory in these subregions. The discretisation scheme should be able to switch locally from a coarser to a finer grid. However, the question arises which criteria should be used to generate the mesh. If the user of a simulation program has to define his elements or nodes a priori as input parameters, this can

perhaps be done by experience /6/. If - as it is the case for modern user oriented programs - an adaptive mesh selection is desired mathematically formulated criteria, are a "sine qua non". Generally such criteria should satisfy two conditions. Firstly they should not cause the program to construct more gridpoints/elements than necessary to achieve a certain accuracy. Secondly they should guarantee that a prescribed relative accuracy δ is really achieved once they are satisfied. A usual way to design adaptive mesh refinement procedures is to equidistribute the local truncation error of the discretisation scheme. In the case of Finite Differences this error is proportional to the meshsize and the third and fourth derivatives of ψ, n and p. Markowich /27/ however showed that it is practically not possible to equidistribute this quantity. In the case of a simple MOS-transistor $O(\delta^{-2}\lambda^{-2})$ gridpoints would be required. On the other hand the singular perturbation analysis shows that the solution of the difference scheme approximates the solution of the reduced problem (3.13)-(3.15) even if this criterion is not satisfied inside the layer regions. Therefore the quantity to be equidistributed is the discretisation error of Poisson's equation (i.e. the partial derivatives of the space charge times the meshsizes). This equidistribution can be relaxed inside the junction layers by e.g. simply limiting the number of gridpoints there.

5. A GLIMPSE ON RESULTS

As an illustrative example a relatively simple structure, a two dimensional diode, is chosen. Fig.2 shows the doping profile as birds-eye-view plot. A substrate with $10^{14}cm^{-3}$ acceptor concentration and an exponentially graded n-region with $10^{19}cm^{-3}$ maximum doping is assumed. The initial mesh is automatically generated from the doping profile and the geometry definition. The simulation domain (device geometry) is a square of 100μm times 100μm size. At the n-region an ohmic contact with length 20μm is assumed. The substrate is fully contacted. The initial mesh for a Finite Boxes program is shown in Fig.3 and for a Finite Element program in Fig.4. The point allocation is identical for both representations. The grid consists of 121 points versus 178 when all gridlines are extended throughout the device. This clearly demonstrates the advantage of the Finite Boxes approach. In Finite Element representation one has to deal with 80 rectangular elements and 17 transition elements which consist of 51 triangles.

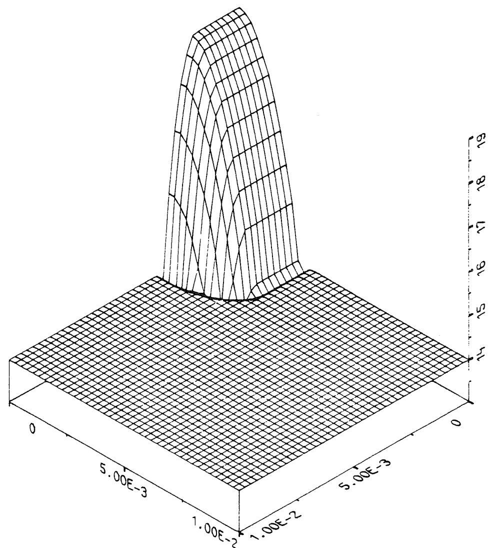

Fig.2 Doping distribution

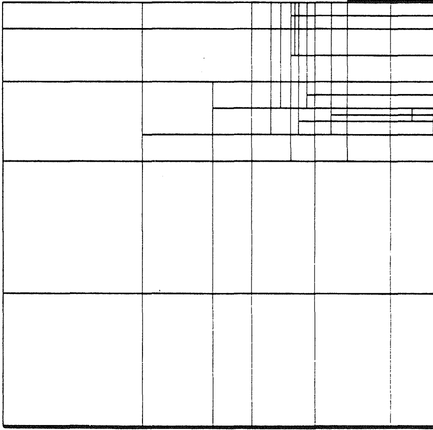

Fig.3 Finite Boxes initial mesh

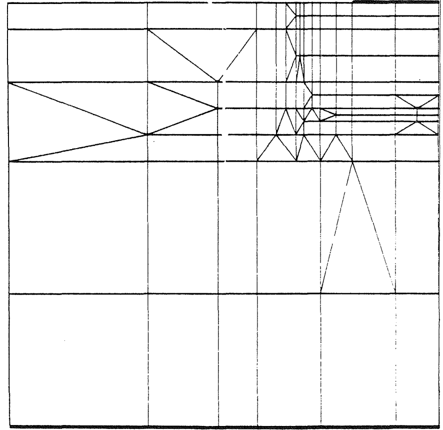

Fig.4 Finite Element initial mesh

Fig.5 shows the final grid for an operating condition of 0.7V forward bias in Finite Boxes representation. This mesh is obtained after several adaption processes using the criteria given in chapter 4. It consists of 270 points (versus 480 for the classical approach). In Fig.6 the electron density distribution is drawn. From this plot one nicely can deduce the effects of high injection. E.g. the substrate is flooded with carriers. Fig.7 shows the magnitude of the electron current density. The peak value is about 180 A/cm^2. The sharply pronounced peak which exists at the transition of the Dirichlet boundary condition to the Neumann boundary condition corresponds to a singularity of the carrier densities. Physically interpreted this effect is well known as contact-corner-current-crowding.

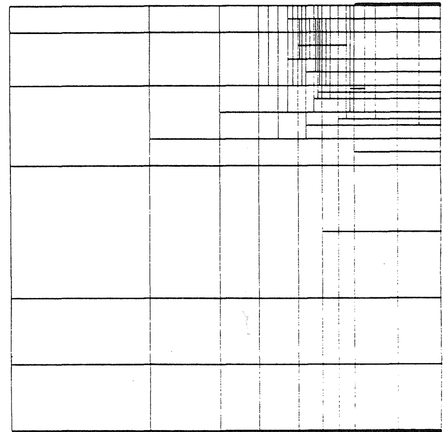

Fig.5 Final mesh U=0.7V

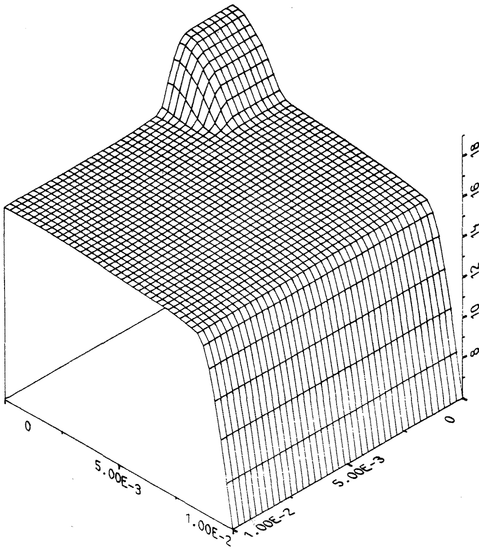

Fig.6 Electron density U=0.7V

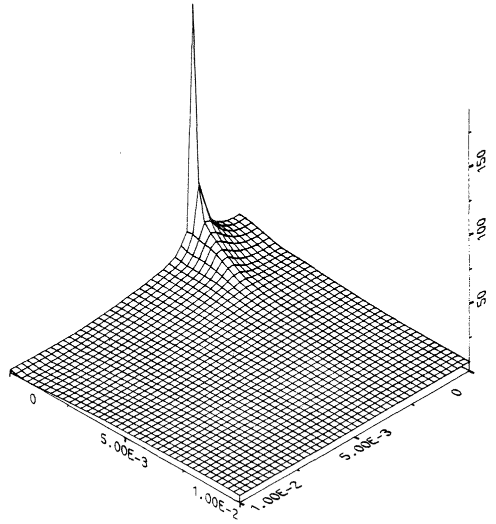

Fig.7 Electron current U=0.7V

Fig.8 shows the final grid for an operating condition of -20V (reverse) bias in Finite Element representation. This mesh consists of 363 points (625 for classical Finite Differences) which correspond to 277 rectangular elements and 41 transition elements (123 triangles). The electron density for this operating point is given in Fig.9. One nicely observes the depletion region and the typical shape of the drop of the electron density in that region owing to thermal generation. In Fig.10 the magnitude of the electron current density is drawn. The singularity at the contact corner is, although it still exists, not so pronounced. Note that there are about seven orders of magnitude difference in the peak value compared to Fig.10.

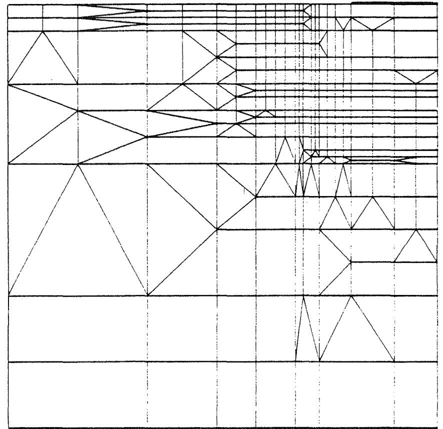

Fig.8 Final mesh U=-20V

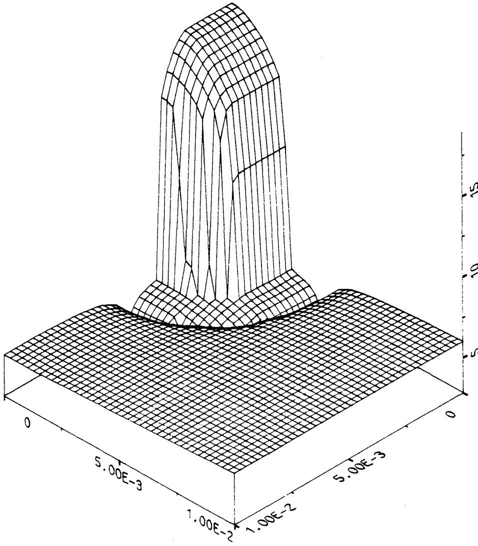

Fig.9 Electron density U=-20V

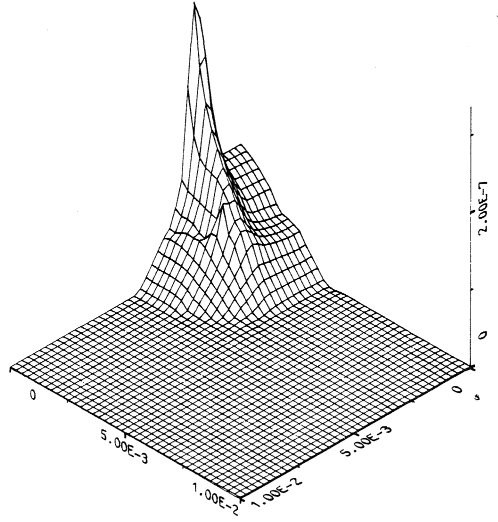

Fig.10 Electron current U=-20V

6. CONCLUSION

In this paper we have presented an analysis of the steady state semiconductor equations and the impact of this analysis on the design of device simulation programs. By appropriate scaling we have transformed the semiconductor equations into a singularly perturbed elliptic system with nonsmooth data. Information obtained from the singular perturbation analysis has been used to investigate stability and convergence of discretisation schemes with particular emphasis on the adaptive construction of efficient grids. An example has demonstrated the power and flexibility a device simulation program can achieve when using the information we have presented for program design.

ACKNOWLEDGMENT

This work has been supported by the "Fond zur Förderung der wissenschaftlichen Forschung" (Projekt Nr. S22/11), by the Research Lab. of SIEMENS AG Munich and by the United States Army (Contract No. DAAG-29-80-C-0041). Many helpfull discussions with Dr.P.Markowich and Prof.H.Pötzl who also read critically the manuscript are very appreciated. The authors wish to thank Dipl.Ing. D. Schornböck and the whole staff of the computer center for the excellent computer access. The computer plots have been provided by Dipl.Ing.A.Franz and Dipl.Ing.G.Franz.

REFERENCES

/1/ Adler M.S., **"A Method for Achieving and Choosing Variable Density Grids in Finite Difference Formulations and the Importance of Degeneracy and Band Gap Narrowing in Device Modeling"**, Proc.NASECODE I Conf., pp.3-30, 1979.

/2/ Adler M.S., **"A Method for Terminating Mesh Lines in Finite Difference Formulations of the Semiconductor Device Equations"**, Solid-State Electron., Vol.23, pp.845-853, 1980.

/3/ Bank R.E., Rose D.J., **"Parameter Selection for Newton-like Methods Applicable to Nonlinear Partial Differential Equations"**, SIAM J.Numer.Anal., Vol.17, pp.806-822, 1980.

/4/ Buturla E.M., Cotrell P.E., Grossman B.M., Salsburg K.A., Lawlor M.B.,McMullen C.T, **"Three-Dimensional Finite Element Simulation of Semiconductor Devices"**, Proc.Int.Solid-State Circuits Conf., pp.76-77, 1980.

/5/ Buturla E.M., Cotrell P.E., **"Simulation of Semiconductor Transport Using Coupled and Decoupled Solution Techniques"**, Solid-State Electron., Vol.23, pp.331-334, 1980.

/6/ Buturla E.M., Cottrell P.E., Grossman B.M., Salsburg K.A., **"Finite-Element Analysis of Semiconductor Devices: The FIELDAY Program"**, IBM J.Res.Dev., Vol.25, pp.218-231, 1981.

/7/ Chamberlain S.G., Husain A., **"Three-Dimensional Simulation of VLSI MOSFET's"**, Proc.Int.Electron Devices Meeting, pp.592-595, 1981.

/8/ Cotrell P.E., Buturla E.M., **"Two-Dimensional Static and Transient Simulation of Mobile Carrier Transport in a Semiconductor"**, Proc.NASECODE I Conf., pp.31-64, 1979.

/9/ DeMari A., **"An Accurate Numerical Steady-State One-Dimensional Solution of the P-N Junction"**, Solid-State Electron., Vol.11, pp.33-58, 1968.

/10/ DeMari A., **"An Accurate Numerical One-Dimensional Solution of the P-N Junction under Arbitrary Transient Conditions"**, Solid-State Electron., Vol.11, pp.1021-2053, 1968.

/11/ Deuflhard P., Heindl G., **"Affine Invariant Convergence Theorems for Newtons Method and Extensions to Related Methods"**, SIAM J.Numer.Math., Vol.16, pp.1-10, 1979.

/12/ Doolan E.P., Miller J.J.H., Schilders W.H.A., **"Uniform Numerical Methods for Problems with Initial and Boundary Layers"**, Boole Press, Dublin, 1980.

/13/ Engl W.L., Dirks H., **"Numerical Device Simulation Guided by Physical Approaches"**, Proc.NASECODE I Conf., pp.65-93, 1979.

/14/ Fife P.C., **"Semilinear Elliptic Boundary Value Problems with Small Parameters"**, Arch.Rational Mech.Anal., Vol.52, pp.205-232, 1973.

/15/ Franz A.F., Franz G.A., Selberherr S., Ringhofer Ch., Markowich P., **"FINITE BOXES - a Generalization of the Finite Difference Method Suitable for Semiconductor Device Simulation"**, presented at IEEE/SIAM conf.Numerical Simulation of VLSI Devices, 1982.

/16/ Frey J., **"Transport Physics for VLSI"**, in: Introduction to the Numerical Analysis of Semiconductor Devices and Integrated Circuits, pp.51-57, Boole Press, Dublin 1981.

/17/ Frey J., **"Physics Problems in VLSI Devices"**, in: Introduction to the Numerical Analysis of Semiconductor Devices and Integrated Circuits, pp.47-50, Boole Press, Dublin 1981.

/18/ Gummel H.K., **"A Self-Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculations"**, IEEE Trans.Electron Devices, Vol.ED-11, pp.455-465, 1964.

/19/ Hess K., **"Ballistic Electron Transport in Semiconductors"**, IEEE Trans.Electron Devices, Vol.ED-28, pp.937-940, 1981.

/20/ Kellog R.B., **"Analysis of a Diffference Approximation for a Singular Perturbation Problem in Two Dimensions"**, Proc.BAIL I Conf., pp.113-117, Boole Press, Dublin, 1980.

/21/ Kellog R.B., Han Houde, **"The Finite Element Method for a Singular Perturbation Problem Using Enriched Subspaces"**, Report BN-978, University of Maryland, 1981.

/22/ Kennedy D.P., O'Brien R.R., **"Two-Dimensional Mathematical Analysis of a Planar Type Junction Field-Effect Transistor"**, IBM J.Res.Dev., Vol.13, pp.662-674, 1969.

/23/ Machek J., Selberherr S., **"A Novel Finite-Element Approach to Device Modelling"**, presented at IEEE/SIAM Conf.Numerical Simulation of VLSI Devices, 1982.

/24/ Manck O., Engl W.L., **"Two-Dimensional Computer Simulation for Switching a Bipolar Transistor Out of Saturation"**, IEEE Trans.Electron Devices, Vol.ED-22, pp.339-347, 1975.

/25/ Marhsak A.H., Shrivastava R., **"Law of the Junction for Degenerate Material with Position-Dependent Band Gap and Electron Affinity"**, Solid-State Electron., Vol.22, pp.567-571, 1979.

/26/ Markowich P.A., Ringhofer Ch.A., Selberherr S., Langer E., **"A Singularly Perturbed Boundary Value Problem Modelling a Semiconductor Device"**, Report 2388, MRC, University of Wisconsin, 1982.

/27/ Markowich P.A., Ringhofer Ch.A., Selberherr S., **"A Singular Perturbation Approach for the Analysis of the Fundamental Semiconductor Equations"**, presented at IEEE/SIAM conf.Numerical Simulation of VLSI Devices, 1982.

/28/ Markowich P.A., Ringhofer Ch.A., **"An Asymptotic Analysis of Single PN-Junction Devices"**, Report xxxx, MRC, University of Wisconsin, 1982.

/29/ Mock M.S., **"On Equations Describing Steady-State Carrier Distributions in a Semiconductor Device"**, Comm.Pure and Appl.Math., Vol.25, pp.781-792, 1972.

/30/ Mock M.S., **"A Time-Dependent Numerical Model of the Insulated-Gate Field-Effect Transistor"**, Solid-State Electron., Vol.24, pp.959-966, 1981.

/31/ Price C.H., **"Two-Dimensional Numerical Simulation of Semiconductor Devices"**, Dissertation, Stanford University, 1980.

/32/ Scharfetter D.L., Gummel H.K., **"Large-Signal Analysis of a Silicon Read Diode Oscillator"**, IEEE Trans.Electron Devices, Vol.ED-16, pp.64-77, 1969.

/33/ Selberherr S., Schütz A., Pötzl H., **"Two Dimensional MOS-Transistor Modeling"**, NATO ASI on Process and Device Simulation for MOS-VLSI Circuits, 1982.

/34/ Thornber K.K., **"Current Equations for Velocity Overshoot"**, IEEE Electron Devive Lett., Vol.EDL-3, pp.69-70, 1982.

/35/ VanOverstraeten R.J., DeMan H.J., Mertens R.P., **"Transport Equations in Heavy Doped Silicon"**, IEEE Trans.Electron Devices, Vol.ED-20, pp.290-298, 1973.

/36/ VanRoosbroeck W.V., **"Theory of Flow of Electrons and Holes in Germanium and Other Semiconductors"**, Bell Syst.Techn.J., Vol.29, pp.560-607, 1950.

/37/ VanVliet K.M., **"The Shockley-Like Equations for the Carrier Densities and the Current Flows in Materials with a Nonuniform Composition"**, Solid-State Electron., Vol.23, pp.49-53, 1980.

/38/ Vasilev'a A.B., Butuzov V.F., **"Singularly Perturbed Equations in the Critical Case"**, translated Report 2039, MRC, University of Wisconsin, 1978.

/39/ Yoshii A., Kitazawa H., Tomizawa M., Horiguchi S., Sudo T., **"A Three-Dimensional Analysis of Semiconductor Devices"**, IEEE Trans.Electron Devices, Vol.ED-29, pp.184-189, 1982.