Two dimensional simulation of MOS transistors

THE-RC 65339

Eindhoven University of Technology

Computing Centre Note 28

Two dimensional simulation of MOS transistors

Ir. M. van der Woude

December 1985

Two dimensional simulation of MOS transistors

M. van der Woude

Eindhoven University of Technology

ABSTRACT

In this report an overview is given of the results of the research on two dimensional simulation of MOS transistors that has been performed by the author from 1978 to 1981. In 1981 the program CADDET became available for the simulation of the steady state behaviour of various two dimensional field effect devices. Some results of the program CADDET will be discussed that have been made visible by means of a graphical postprocessor.

1. Introduction

In [1] the physical and mathematical backgrounds are given for the numerical analysis of two dimensional semiconductor structures. The following set of equations has to be solved:

the Poisson equation:

$$\Delta \psi = -\rho = -(p - n - N_{dope}) \tag{1.1}$$

and the continuity equations:

$$\frac{\partial n}{\partial t} = -U + \nabla \cdot (-n\,\mu_n\,\nabla \psi + D_n\,\nabla n) \tag{1.2a}$$

$$\frac{\partial p}{\partial t} = -U + \nabla \cdot (+p\,\mu_p\,\nabla \psi + D_p\,\nabla p) \tag{1.2b}$$

In these equations ψ, n and p are functions of co-ordinates x, y and time t; ψ stands for the electrostatic potential, ρ is the charge density, n and p are the electron and hole concentrations, N_{dope} is the net impurity concentration, i.e. acceptor minus donor concentration. U is the generation/recombination rate, μ_n and μ_p are the effective electron and hole mobilities, which may be functions of $\nabla \psi$, D_n and D_p are the diffusion constants for electrons and holes.

Scaling

The function ψ is scaled with

$$\psi_{char} = V_T = \frac{kT}{q} \tag{1.3}$$

The concentrations n , p and N_{dope} are scaled with N_{char}, the absolute impurity concentration in the channel. The mobilities μ_n and μ_p are scaled with

$$\mu_{char} = v_{char} \times \frac{l_{char}}{\psi_{char}} \tag{1.4}$$

in which v_{char} is the electron saturation velocity and

$$l_{char} = \left| \frac{\epsilon_0 \epsilon_{Si}}{q} \frac{\psi_{char}}{N_{char}} \right|^{1/2} \tag{1.5}$$

the external Debye length in the channel. The diffusion constants D_n and D_p are scaled with

$$D_{char} = \mu_{char} \times \psi_{char} \tag{1.6}$$

Finally x and y are in units l_{char} and the time t is in units

$$t_{char} = \frac{l_{char}}{v_{char}} \tag{1.7}$$

The boundary conditions of the Poisson equation (see also fig. 1) are the following:

- at the contacts of source , gate and drain and backgate ψ is prescribed by the applied voltage:
- across the Si/SiO_2 interface the normal derivative of $\epsilon\psi$ is continuous:
- at the other boundaries the normal derivative of ψ is zero.

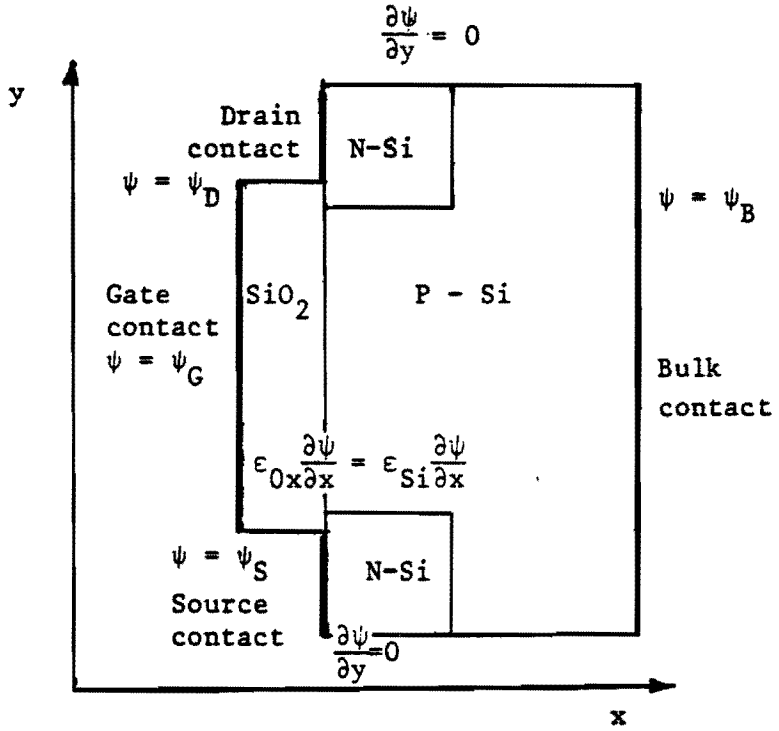

fig. 1. Geometry of MOS transistor with boundary conditions.

The boundary values of the continuity equations are:

- at metal contacts $n = n_c$, $p = p_c$ with n_c and p_c the equilibrium values of electron and hole concentrations;
- at isolating boundaries the normal components of the electron and hole current are zero, so here the boundary conditions are of mixed type. The electron and hole current density is given by:

$$J_n = -q \left(n \mu_n \nabla \psi - D_n \nabla n \right) \tag{1.8a}$$

$$J_p = -q \left(p \mu_p \nabla \psi + D_p \nabla p \right) \tag{1.8b}$$

Equations (1.1) and (1.2) define the transient problem. When the partial derivatives with respect to time in (1.2a) and (1.2b) are set to zero these equations describe the stationary behaviour of semiconductor devices. So we get the steady state problem defined by (1.1) and:

$$\nabla .(-n\,\mu_n\,\nabla\,\psi + D_n\,\nabla\,n) - U = 0 \qquad (1.9a)$$

$$\nabla .(+p\,\mu_p\,\nabla\,\psi + D_p\,\nabla\,p) - U = 0 \qquad (1.9b)$$

2. Numerical Solution methods

One dimensional problem

At the start of this research a program has been made, based on the one dimensional forms of (1.1) and (1.9). With this program a stationary one dimensional semiconductor structure, i.e. a structure which is large in two of the three dimensions, e.g. an MOS capacity, has been analyzed. Some results with this program have been described in [2]. One of the conclusions from this one dimensional program was that the convergence rate became very low when the carrier concentration in the channel became high due to inversion. Therefore it was decided to make a program called MOS2 which could solve the equations (1.1), (1.2a) and (1.2b) for the transient case in two dimensions. It was hoped that this program could also find steady state solutions as a limit for large t of the transient problem.

Two dimensional problem

To solve the set of equations (1.1) - (1.2) numerically in two dimensions one has to discretize them in time as well as in space. For the space discretization there exist in general two methods, the finite difference method (FDM) and the finite element method (FEM).

Since the finite element method is more flexible in handling non rectangular geometries it was decided to choose for the Poisson equation the finite element method. For the continuity equation this method is only adequate when so called flux conserving elements are used, which is rather intricate. It was decided to use for these equations the Scharfetter-Gummel discretization method. This method is relatively simple and with it the current can be calculated accurately and also non rectangular geometries can be solved. We will in the next sub sections discuss both methods of discretization.

2.1 The finite element method

To derive the finite element discretization for the Poisson equation (1.1) we divide the region of interest into (not necessarily rectangular) polygons called finite elements. We shall assume for simplicity that the elements are rectangular and that they are numbered

$$1, 2, \ldots, LL$$

that the nodes are numbered

$$1, 2, \ldots, NN$$

and that node k has co-ordinates x_k, y_k (See fig. 2).

We try to find an approximate solution $\hat{\psi}$ of the form:

$$\hat{\psi}(x, y) = \sum_{1}^{NN} \Psi_k\,\phi_k(x, y) \qquad (2.1)$$

with

$$\Psi = (\Psi_1, \Psi_2, \ldots, \Psi_{NN})^T \tag{2.2}$$

the vector of $\hat{\psi}$-values on the gridpoints. So we have:

$$\Psi_k = \hat{\psi}(x_k, y_k) \tag{2.3}$$

The functions $\phi_k (x, y)$ are called global base functions. We will assume that the function $\phi_k (x, y)$ is continuous and has value 0 on all elements except for the elements surrounded by node k, on these elements it is linear in x and in y. On node k it has value one and on all other nodes it is zero. Such a function has the form of a roof. Since the basefunctions ϕ_k are linear in x and y we call the elements also linear. Equation (2.1) states that $\hat{\psi}(x, y)$ is obtained by interpolation between the approximate values Ψ_k of ψ at the nodes using the functions ϕ_k as interpolation functions.

By requiring the innerproduct of equation (1.1) with each ϕ_k corresponding to a free node, i.e. a node whose ψ-value is not prescribed, to be zero, one can derive (see e.g. Strang and Fix [13]) that the vector R of global residuals has to be zero:

$$R = \iint_G ((\nabla \phi^T)^T \nabla \hat{\psi} - \phi \rho) dG = 0 \tag{2.4}$$

in which

$$\phi = (\phi_1(x, y), \phi_2(x, y), \ldots, \phi_{NN}(x, y))^T$$

and G is the domain of ϕ. By solving (2.4) we obtain the values of the unknown components of Ψ, given ρ and the boundary conditions. To solve (2.4) we decompose first R into vectors of local residuals and then assemble these local residuals. After the assembly process a set of linear equations has to be solved.

The global base functions ϕ_k that are not zero on element l are denoted as the vector $\phi^l(x, y)$. This vector has four components for linear rectangular elements. Similarly a vector of local Ψ-values ψ^l may be defined as those components of Ψ that correspond to the nodes of element l.

The local vector of residuals for element l may be defined as:

$$r^l = \iint_{E^l} ((\nabla \phi^{l^T})^T \nabla \hat{\psi} - \phi^l \rho) dG \tag{2.5}$$

where E^l is the area of element l. By observing that on element l

$$\hat{\psi}(x, y) = \phi^l(x, y)^T \psi^l \tag{2.6}$$

and by defining the matrix H^l as the Jacobian of ϕ^l:

$$H^l = (\nabla \phi^{l^T})^T \tag{2.7}$$

it follows that the local vector of residuals may be written as:

$$r^l = \iint_{E^l} (H^l H^{l^T} \psi^l - \phi^l \rho) dG \tag{2.8}$$

With the element matrix K^l defined by:

$$K^l = \iint_{E^l} H^l H^{l^T} dG \tag{2.9}$$

and the element vector b^l defined by:

$$b^l = \iint_{E^l} \phi^l \rho \, dG \tag{2.10}$$

the local vector of residuals (2.8) may be rewritten as:

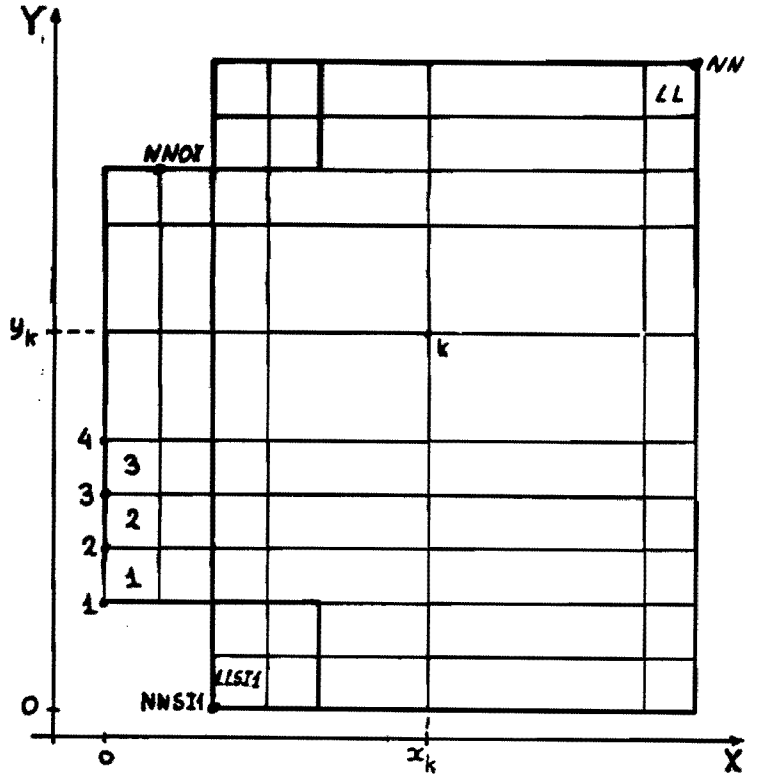

fig. 2. Finite element grid with rectangular elements.

$$r^l = K^l \psi^l - b^l \tag{2.11}$$

The local vectors of residuals r^l of each element E^l may be assembled to the global residual vector R. To do this we note that one component of R, say R_k can be written as the sum of those components r^l_m of local residual vectors r^l that correspond to the global base function ϕ_k. After assembly we have:

$$K_s \Psi - b_s = 0 \tag{2.12}$$

with K_s the so called structure matrix assembled from the element matrices K^l and b_s the structure vector assembled from the element vectors b^l.

It can be shown that since each K^l is symmetric and positive definte, K_s is symmetric and positive definite, so no singularity problems will occur when calculating Ψ with (2.12). K_s can be stored efficiently, e.g. for a rectangular region it has only nine nonzero diagonals of which five differ from each other.

2.2 The Scharfetter-Gummel discretization

To derive the Scharfetter-Gummel discretization we omit indices n and p and write the continuity equation (1.2a) as:

$$\frac{\partial n}{\partial t} = -\nabla \cdot F \tag{2.13}$$

assuming for simplicity no generation/recombination. The electron flux F may be written as:

$$F = n\,v - D\,\nabla\,n \tag{2.14}$$

with v the electron velocity, a known function of $\nabla\,\psi$. Usually v is written as

$$v = \mu_e \nabla\,\psi$$

For low electric field strengths the effective mobility μ_e may be considered independent of the electric field, for high field strengths v saturates to the saturation velocity so μ_e decreases nonlinearly with the absolute value of $\nabla\,\psi$. Around each node p of the finite element grid in the Si-region we define a nodal area A_p (see fig. 3). The area A_p is bounded by the perpendicular bisectors of the grid edges incident with node p.

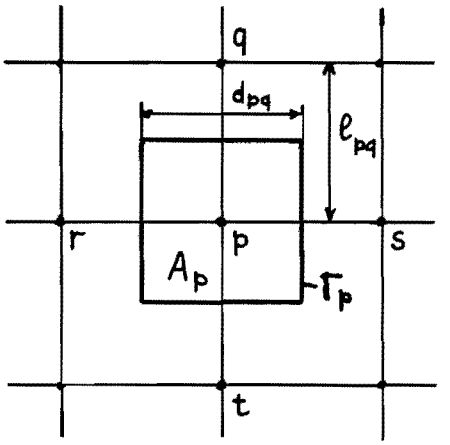

fig. 3. Nodal area A_p with boundary Γ_p used with Scharfetter-Gummel discretization.

By integration of (2.13) over A_p and applying the divergence theorem one can derive (with Γ_p the boundary of A_p):

$$\iint_{A_p} \frac{\partial n}{\partial t}\,dG = -\int_{\Gamma_p} F\,.\,\vec{e}_n\,d\,\Gamma \tag{2.15}$$

The edges of A_p are called transversal edges.

With each edge e_{pq} of the FEM grid a transversal edge t_{pq} is associated. The length of e_{pq} is denoted by l_{pq}, that of t_{pq} by d_{pq}. To approximate the line integral in (2.15) we first approximate the value of the flux F on the transversal edge t_{pq}. Assume v and F have constant values v_{pq} and F_{pq} on t_{pq} and D is constant. Then (2.14) may be written as:

$$F_{pq} = n\,v_{pq} - D\,\frac{d\,n}{ds} \tag{2.16}$$

By writing ds explicitly and integrating along t_{pq} we find:

$$F_{pq} = \frac{n_p\,v_{pq}\,e^{\frac{v_{pq}\,l_{pq}}{D}} - n_q\,v_{pq}}{e^{\frac{v_{pq}\,l_{pq}}{D}} - 1} \tag{2.17}$$

with n_p and n_q the values of n at the nodes p and q. We approximate now the left side of (2.15) by the midpoint rule and find (note that $v_{qp} = -v_{pq}$):

$$\frac{d\,n_p}{d\,t} = \frac{1}{A_p} \sum_{i\,=\,q,r,s,t} a_{pi}\,n_i - a_{ip}\,n_p \tag{2.18}$$

with

$$a_{pq} = \frac{d_{pq}\, v_{pq}}{e^{\frac{v_{pq}\, l_{pq}}{D}} - 1} \tag{2.19}$$

By combining equation (2.18) for all nodes with unknown electron concentration we get the vector equation:

$$\frac{d\,\vec{n}}{d\,t} = M_n\,\vec{n} \tag{2.20a}$$

with $\vec{n}$ the vector of electron concentrations on the nodes of the finite element grid. Similarly we can derive a vector equation for the hole concentrations:

$$\frac{d\,\vec{p}}{d\,t} = M_p\,\vec{p} \tag{2.20b}$$

The components of the matrices M_n and M_p are dependent on the local electron resp. hole velocity, so on $\nabla\,\psi$, on the lengths of the edges and the transversal edges and on the nodal areas.

2.3 Time discretization

In sections (2.1) and (2.2) it has been shown that by space discretization the partial differential problem (1.1) - (1.2) can be approximated by the equations (2.12) and (2.20) which may be written as:

$$K_s\,\Psi = b_s(\vec{n}\,,\vec{p}\,) \tag{2.21}$$

$$\frac{d\,\vec{n}}{d\,t} = M_n(\Psi)\,\vec{n} \tag{2.22a}$$

$$\frac{d\,\vec{p}}{d\,t} = M_p(\Psi)\,\vec{p} \tag{2.22b}$$

In these equations

- Ψ, $\vec{n}$ and $\vec{p}$ are the vector values of the electrostatic potential, the electron and hole concentration at the NN gridpoints. Note that Ψ, $\vec{n}$ and $\vec{p}$ are functions of time.
- The structure matrix K_s is a symmetric positive definite $NN * NN$ band matrix with constant components, the value of which depend only of the mask geometry.
- The space charge or structure vector b_s is a $1 * NN$ vector whose components depend on the mask geometry and the space charge, so they depend also on time.
- M_n , M_p are structurally symmetric band matrices with component values depending on Ψ and therefore also on time. From (2.18) one can derive that M_n nd M_p have five nonzero diagonals consisting of at most NN - NN_{ox} components where

$$1, 2, \ldots, NN_{ox}$$

are the nodes inside the oxide region.

To solve (2.21) - (2.22) one can use backward time differences, e.g. backward Euler, to transform the differential equations into a set of nonlinear equations:

$$K_s\,\Psi(t + \Delta t) - b_s(\vec{n}\,(t + \Delta t)\,,\vec{p}\,(t + \Delta t)) = 0 \tag{2.23}$$

$$\vec{n}(t + \Delta t) - \vec{n}(t) - \Delta t\ M_n(\Psi(t + \Delta t))\vec{n}(t + \Delta t) = 0 \qquad (2.24a)$$

$$\vec{p}(t + \Delta t) - \vec{p}(t) - \Delta t\ M_p(\Psi(t + \Delta t))\vec{p}(t + \Delta t) = 0 \qquad (2.24b)$$

These sets of nonlinear equations may be solved iteratively be linearizing them to one system of linear equations. However much computer time and memory is required. This method is called the full implicit method [3].

One can also consider the equations (2.21) - (2.22) as three separate equations and find for each time step a solution by iteratively solving each of the three equations separately. As has been shown by Reiser [3] one has to solve (2.22a) and (2.22b) with a half implicit method since explicit methods lead to extremely small time steps.

Full implicit methods are unconditionally stable, half implicit methods are conditionally stable. For the backward Euler scheme the half implicit method becomes:

$$K_s\ \Psi(t) - b_s(\vec{n}(t), \vec{p}(t)) = 0 \qquad (2.25)$$

$$\vec{n}(t + \Delta t) - \vec{n}(t) - \Delta t\ M_n(\Psi(t))\vec{n}(t + \Delta t) = 0 \qquad (2.26a)$$

$$\vec{p}(t + \Delta t) - \vec{p}(t) - \Delta t\ M_p(\Psi(t))\vec{p}(t + \Delta t) = 0 \qquad (2.26b)$$

The systems (2.26) can be transformed to two sets of linear equations with matrices M_n and M_p whose components depend on Ψ, consequently the matrix components have to be recalculated each time step. During one time step the following loop has to be iterated:

loop:

step 1: calculate $\Psi(t)$

step 2: calculate $\vec{n}(t + \Delta t)$

step 3: calculate $\vec{p}(t + \Delta t)$

step 4: $t := t + \Delta t$

end of loop.

During each of step 1, 2 or 3 a linear matrix equation has to be solved, besides in step 1 the right hand vector has to be calculated and in steps 2 and 3 a matrix and a right hand vector has to be calculated.

To gain experience, it was decided to first implement in the program MOS2 a time discretization method based on the half implicit scheme (2.25) - (2.26). In section 2.4 an outline of the program will be given.

2.4 The program MOS2

In this section we give in pseudo code an overview of the program MOS2, which is based on the theory as discussed in the previous sections.

```
BEGIN
    input;
      { of device type, device geometry, grid geometry,
        terminal voltages, dope profile }
    scale;
      { according to (1.3) - (1.7) }
    mesh generation;
      { The mesh data structure is created according to fig. 2.
        Nodes in Ox and Si:
                    1 . . NNOX,NNSI1 . . NN
        Elements:
                    1 . . LLSI1 . . LL
        Edges:
                    1 . . EE
        Coordinates of nodes:
                    X, Y [1 . . NN]
        Valency of nodes:
                    VAL [1 . . NN]
        Numbers of adjacent nodes of each node:
                    INCMAT [1 . . NN, 1 . . 4]
        Nodes belonging to each element:
                    ELMN [1 . . LL, 1 . . 4]
      }
    fill boundary arrays;
      { The boundary nodes are nodes with prescribed Ψ values
        (nodes of contacts):
        Node numbers belonging to set of boundary nodes:
                    BNODE [1 . . NBN]
        Prescribed Ψ-values of these nodes:
                    BNVAL [1 . . NBN]
        Boolean array indicating when a node has (FALSE) or has not (TRUE)
        a prescribed value:
                    FREE [1 . . NN]
      }
    assemble structure matrix K_s ;
    choleski_decompose K_s ;
    calculate geometric data;
      { according to fig. 3:
        lengths of edges:
                    e_pq
        lengths of transversal edges:
                    d_pq
        nodal areas:
                    A_p
        for each node:
                    outgoing and incoming edges
      }
```

initialize $\vec{n}$, $\vec{p}$, $\vec{N}_{dope}$, Ψ ;
determine integration time-step Δt ;
determine print interval Δt_{print} ;
determine max number of print steps NNPR;
{the values of Δt , Δt_{print} and NNPR may be read in)
NPR := 0;
{NPR is the number of the present print interval
NITER8 := 0;
{NITER8 is the number of the present iteration time step}

CONVERGED := FALSE;

{start of print interval loop:}

WHILE NPR < NNPR and not CONVERGED DO
BEGIN
NPR := NPR + 1;
TPRNT := Δt_{print} * NPR;

{start of time integration loop:}

REPEAT
NITER8 := NITER8 + 1;
assemble structure vector b_s ;
Ψ_{old} := Ψ ;
solve $\Psi = K_s^{-1} * b_s$;
CONVERGED := $| \Psi - \Psi_{old} |_\infty \leq eps$;
TITER8 := TITER8 + Δt ;
calculate M_n , M_p ;
calculate right-hand vectors $\Delta\vec{n}$, $\Delta\vec{p}$;
{ solve with CG method: }
$\Delta\vec{n}$:= $M_n^{-1} * \Delta\vec{n}$;
$\Delta\vec{p}$:= $M_p^{-1} * \Delta\vec{p}$;
$\vec{n}$:= $\vec{n}$ + $\Delta\vec{n}$;
$\vec{p}$:= $\vec{p}$ + $\Delta\vec{p}$;
UNTIL CONVERGED or TITER8 $\geq$ TPRNT;
print TITER8, Ψ, $\vec{n}$, $\vec{p}$;
END;
END.

Remarks

- The structure matrix K_s is assembled by calculating for each element its element matrix K^l and adding it to the initially zero matrix K_s. The calculation of K^l depends on the form of the base functions. For base functions of rectangular elements that are linear in x and y, the integral (2.9) can be solved and the components of K^l are simple expressions of the length and width of the elements. K_s is a symmetric positive definite bandmatrix with bandwidth BW equal to the number LY of gridlines in y-direction. It can be decomposed employing Choleski decomposition.
- The structure vector b_s is assembled by calculating for each element the element vector b^l and adding it to the initially zero vector b_s. The components of b^l are calculated according to (2.10) by employing Gausspoint integration. Since they are dependent on ρ they have to be recalculated each time step.
- For the calculation of the components of the matrices M_n and M_p one needs (see fig. 3) according to (2.18) and (2.19) for each edge its length, transversal length and for each node its nodal area. To calculate the average value of the electric field on each transversal edge one needs for each node a list of outgoing and of incoming edges.
- The matrix equations for the calculation of electron and hole calculation can be solved efficiently by using the conjugate gradient method. The matrices M_n and M_p are structurally symmetric with nonzero diagonals at distances 1 and LY of both sides of the main diagonal, they have to be recalculated each time-step (see appendix).

Conclusions.

The results with this program were not very encouraging. As soon as the carrier concentration in the channel became higher, the time step had to become smaller to prevent instabilities (oscillations) in the solution. These results are in agreement with Reiser's stability analysis. He shows that the integration time step has to satisfy

$$\Delta t \leqslant \frac{\epsilon_{Si}\,\epsilon_0}{q}\,\frac{1}{\mu_n N_{max}} \tag{2.27}$$

where N_{max} is the maximum carrier concentration in the device. Since the concentration becomes very high in the channel the allowable time-step becomes correspondingly small, resulting in large amounts of CPU time.

Nevertheless some useful conclusions could be drawn from the experiences with the program MOS2, which was a relatively simple program, due to its dependence on a rectangular device geometry. These conclusions may be summarized as follows:

- When employing FEM analysis it is possible to write a program that can analyze two dimensional MOS structures of any form. However the programming task becomes much simpler when the devices are restricted to rectangular cross-sections. For arbitrary geometries it may be favourable to employ existing mesh generators which are in use in mechanical analysis, e.g. TRIQUAMESH [4].
- By restricting the mesh geometry one can arrange that the matrix equations that result after assembly and after linearization of nonlinear equations become bandmatrices with small bands. These equations can be solved efficiently by the conjugate gradient method (CG) [5], especially when the matrices are positive definite. In our case the matrices resulting from the continuity equations were not positive definite, yet a variant of the CG-method worked well (see appendix), except when the solution strongly oscillated due to instabilities.
- The solution method as outlined above is not well suited for MOS structures, since it suffers from instabilities which can only be circumvented by employing extremely small time steps. It has been shown by Mock [6] that the instabilities are due to the decoupling of the three sets of equations.

- It looks that the full implicit method will result in huge requirements of computer resources, both of time and memory.
- A combination of the FEM discretization method together with Mock's stream function approach (see section 2.5) may be a good choice for simulation of MOS-structures with non-rectangular shapes.

A modified version of the program MOS2 has been used successfully to calculate the breakdown characteristics of a diffused mesa diode [7].

2.5 Other methods

A lot of work has been done in the field of two and also three dimensional analysis of semiconductor structures. The reader is referred to the proceedings of the NASECODE conferences which have been held in Ireland each two years since 1979 (see [1]). We discuss here into some more detail the work of Mock, since this seems a promising approach.

The Steady state problem

The steady state problem has been formulated in section 1. as the system of equations (1.1), (1.9):

$$\Delta \psi = -\rho = -(p - n - N_{dope}) \tag{1.1}$$

$$\nabla .(-n\,\mu_n\,\nabla \psi + D_n\,\nabla n) - U = 0 \tag{1.9a}$$

$$\nabla .(+p\,\mu_p\,\nabla \psi + D_p\,\nabla p) - U = 0 \tag{1.9b}$$

It has been reported by Buturla and Cottrell [8] that also with these equations convergence problems may arise when they are discretized to 3 uncoupled sets of equations which are solved iteratively. In the same paper a solution method is considered in which the system (1.1), (1.9) is discretized to one set of linear equations, which is about 3 times as large as each of the previous uncoupled equations, this method is called the coupled method. The coupled method has better convergence properties, but requires lots of computer resources.

A better approach seems to be that of Mock [9]. Mock's method is based on 'stream functions'. The method has been implemented in the program CADDET [10], which is available for testing (see section 3). Therefore we will discuss some of its principles.

Essential for this method is that the carrier generation/recombination term in (1.9a) and (1.9b) may be neglected and that the Einstein relation holds, i. e. (in unscaled quantities):

$$D = \frac{kT}{q}\mu \tag{2.28}$$

Further it is assumed (for computational economy only) that the hole current may be neglected, which is usually physically reasonable for n-channel devices. In that case the hole concentration may be represented by

$$p = p_0 e^{-\frac{q}{kT}(\psi - \psi_{BG})} \tag{2.29}$$

where p_0 is the equilibrium hole concentration at the backgate (i.e. bulk) contact and ψ_{BG} is the prescribed backgate potential.

Equation (1.9a) may in that case be written as

$$\nabla .\vec{J} = 0 \tag{2.30}$$

with $\vec{J}$ the total (electron) current density vector, which may be written as:

$$\vec{J} = (J_x , J_y)^T = \mu e^{\psi} \nabla (e^{-\psi} n) \tag{2.31}$$

From vector calculus we know that a vector field $\vec{J}$ with zero divergence has a vector potential $\vec{\theta}$. In this case it is easy to show that we may consider $\vec{\theta}$ to be a vector in z-direction. We will call this z-component of $\vec{\theta}$ the stream function θ and we may write:

$$J_x = J\frac{\partial\theta}{\partial y} \tag{2.32a}$$

$$J_y = J\frac{\partial\theta}{\partial x} \tag{2.32b}$$

with θ scalar and $0 \leqslant \theta \leqslant 1$ and J the source drain current. Mock shows in [9] that θ satisfies the elliptic partial differential equation:

$$\frac{\partial}{\partial x}\frac{e^{-\psi}}{\mu}\frac{\partial\theta}{\partial x} + \frac{\partial}{\partial y}\frac{e^{-\psi}}{\mu}\frac{\partial\theta}{\partial y} = 0 \tag{2.33}$$

with boundary conditions $\theta = 1$ on the oxide interface and $\theta = 0$ on the other isolating boundaries and with normal derivative of θ is zero on contact boundaries. By integrating equation (2.32a) from source (S) to drain (D) along the oxide boundary we have for the source drain current J:

$$J = \frac{(n_D e^{-\psi_D} - n_S e^{-\psi_S})}{\int_S^D \frac{e^{-\psi(x,0)}}{\mu}\frac{\partial\theta}{\partial y}\,dx} \tag{2.34}$$

with ψ_S and ψ_D the prescribed source and drain potential and n_S and n_D the equilibrium source and drain electron concentration.

Knowing once J and ψ the electron concentration n can be calculated by the line integral:

$$n(x,y) = e^{\psi(x,y)}\left| n_S e^{-\psi_S} + J\int_S^{x,y}(\nabla \times \theta)\,d\vec{l}\right| \tag{2.35}$$

From (1.1), (2.33), (2.34) and (2.35) an iteration scheme can be constructed with which the steady state problem can be solved effectively. In [9] Mock reports that the convergence speed becomes slow when the dopes in source and drain regions are high. He indicates how in that case the calculations can be speeded up. Our experience with CADDET indicate also good computational properties of the method.

The time dependent problem

The stream function approach is essentially based on the zero divergence property of the current density. By observing that in the time dependent case the total current density, i. e. the sum of electron current density, hole current density and dielectric current density, has also the zero divergence property, Mock has designed, a solution method based on stream functions [11]. He has constructed the program "SIFCOD" according to this method, which has been used for over a year and which is commercially available [12].

3. Results with the program CADDET.

The program CADDET has been used at the Eindhoven University of Technology since 1981. It is capable of solving the steady state problem for various types of MOS transistors. We have performed some calculations on two rectangular n-channel MOS transistors with channel lengths of 1 and 2 μ. All other parameters were the same for both transistors, viz.:

- channelwidth 1050 μ (for historical reasons)
- source and drain diffusion depth 0.2 μ
- source and drain donor impurity concentration (uniform) $10^{+19}\ cm^{-3}$
- bulk acceptor impurity concentration (uniform) 7.0 $10^{+15}\ cm^{-3}$
- backgate voltage 0 Volt

In fig. 4 a and b we show the I_{DS} / V_{DS} curves for these devices as calculated by CADDET. All curves of fig. 4 seem reasonably in agreement with practice, except for the curve for $V_{GS} = 5\ V$ in fig. 4b. From calculations with other V_{DS} increments, it follows that the value of the calculated current has a relative error of about 10%, especially in the case of high currents. For low currents the error becomes smaller.

To show the short channel effects we have calculated in fig. 5 I_{DS} as a function of V_{GS} for values of V_{DS} of 1 and 2 Volts for both transistors in the sub threshold range. It is clear that the slope of these curves is much smaller for the transistor with the shorter channel. The results of fig. 5 are in agreement with the theory on short channel effects.

Finally we show in fig. 6 a three dimensional plot of the potential distribution, the electron distribution and a contourplot of the electron distribution.

These pictures have been made with the graphical postprocessor for the CADDET program [14].

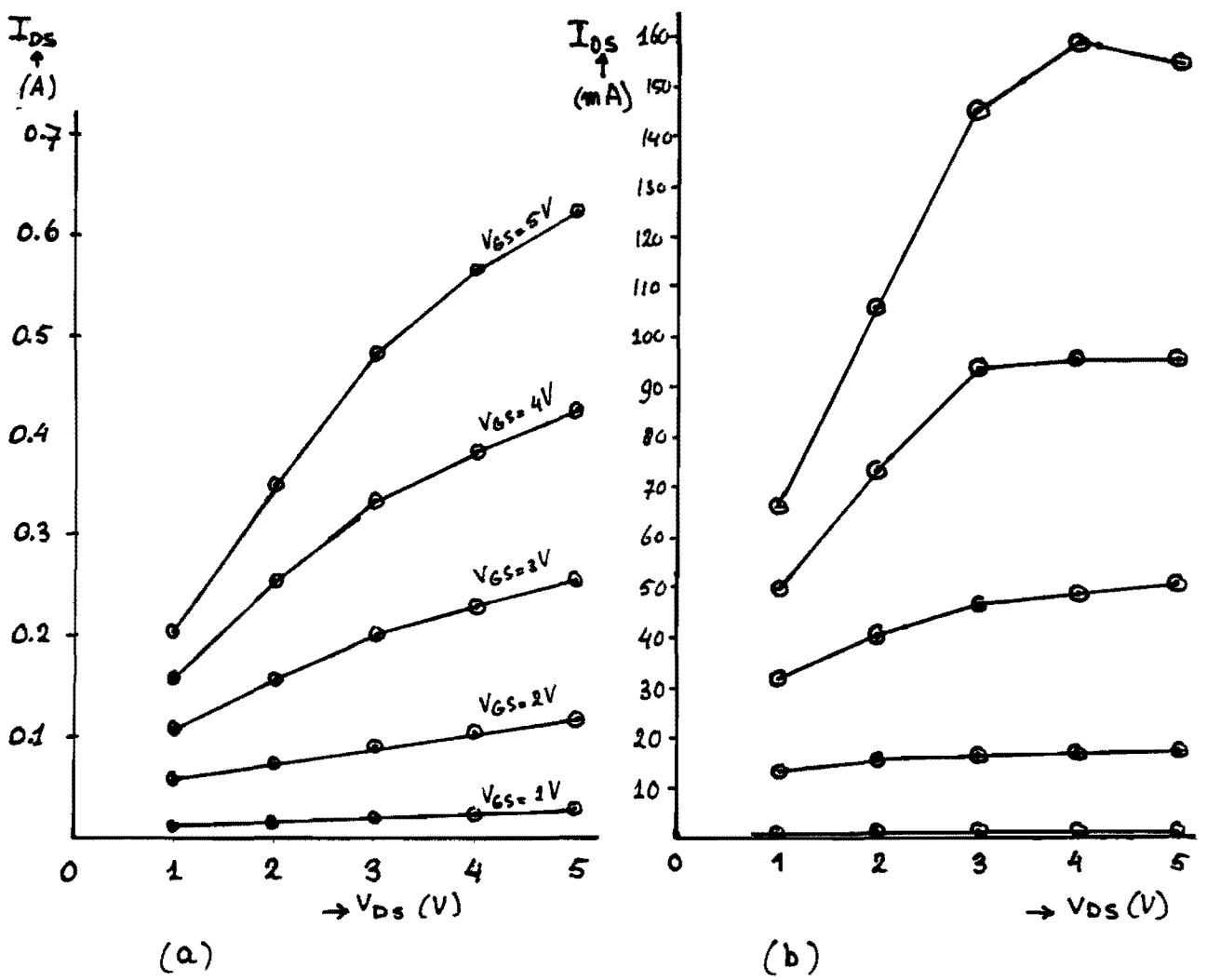

fig. 4. I_{DS} / V_{DS} **curves for MOS transistors with channel lengths of 1** μ **(a) and 3** μ **(b).**

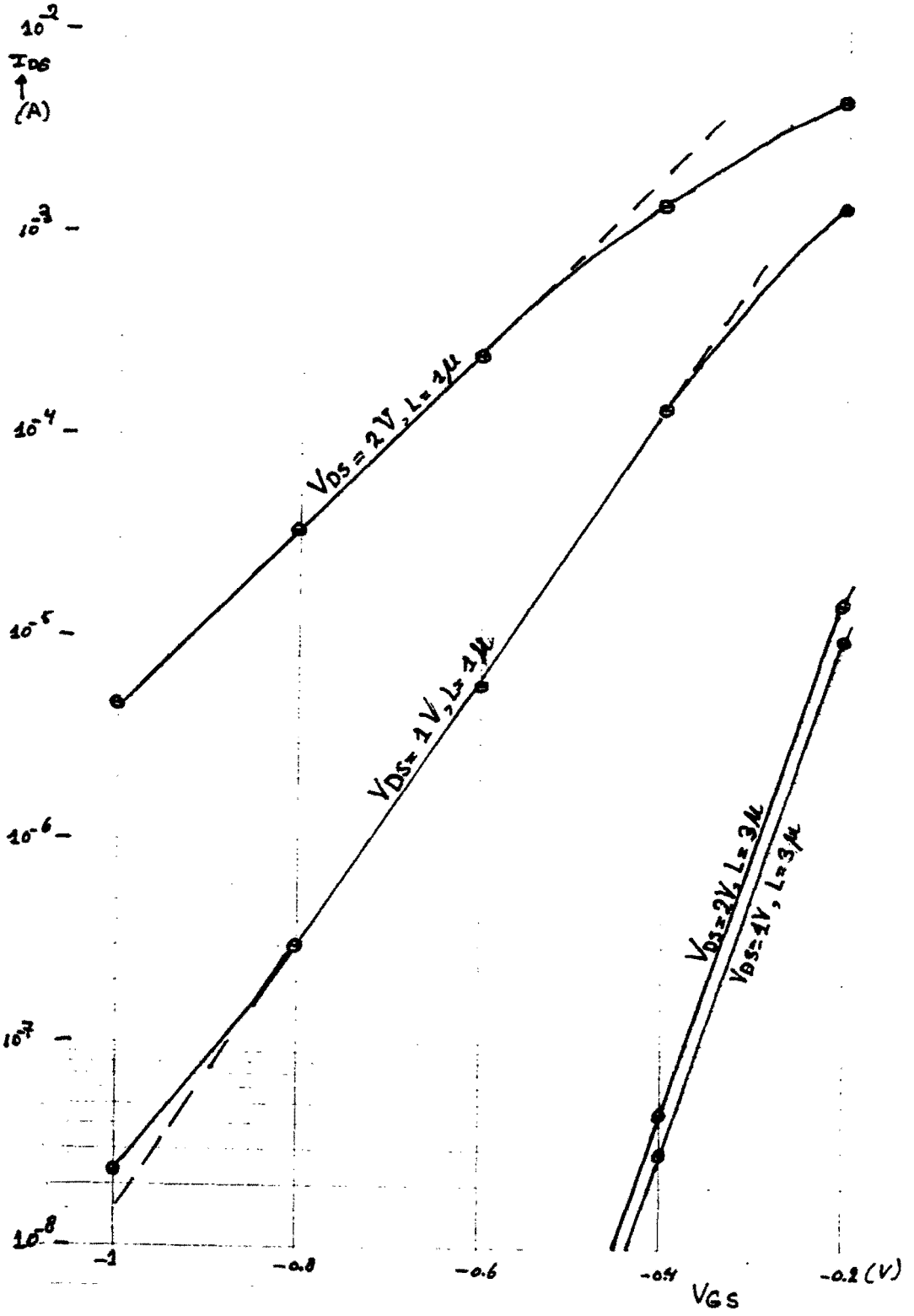

fig. 5. I_{DS} / V_{GS} **curves in the sub threshold range for** V_{DS} **= 1 and 2 V and** $l_{channel}$ **= 1** μ **and 3** μ.

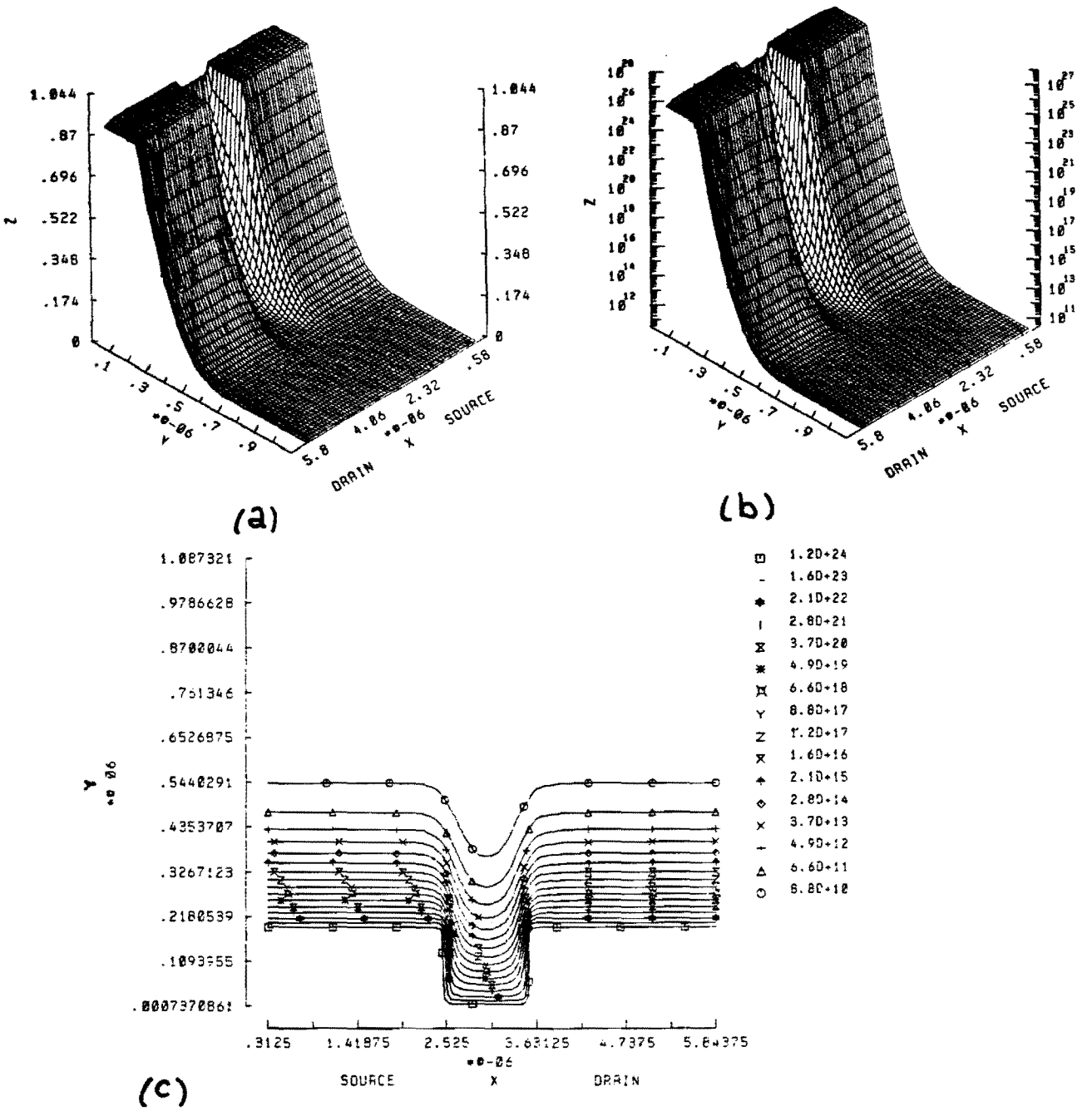

fig. 6. 3D-potential distribution (a), 3D-electron distribution (b) and contourplot of electron distribution(c) for a 1 μ MOS transistor with V_{DS} = 0 V and V_{GS} = 4 V.

4. Conclusions

We have discussed a method for two dimensional simulation of MOS devices, which is based on FEM space discretization for the Poisson equation and on Scharfetter-Gummel space discretization method for the continuity equations. The method has been programmed for simple rectangular MOS structures. From these experiences it follows that though the FEM method is basically more flexible then the FDM (Finite Difference Method) space discretization method, the programming effort becomes considerably more complex when arbitrary non-rectangular geometries have to be calculated (e.g. V-shaped MOS transistors). A problem with the examined method was that its convergence properties are poor in case of high carrier concentrations. These convergence problems do not occur with the program CADDET which is based on the stream function approach of Mock. From the results with the program CADDET it follows that especially high currents are not calculated accuretely. Small channel current behaviour can be simulated with CADDET, the results are in agreement with theory.

Appendix: A generalization of the conjugate gradient method.

The matrices M_n and M_p that occur in equations (2.20) are sparse, non-singular and non-symmetric with nonzeros only on the main diagonal and on the diagonals on distances ± 1 and $\pm LY$ from the main diagonal. This follows from (2.18). Since the conjugate gradient method (CG) is only applicable to symmetric matrices, we used a generalized form of CG as proposed by Kershaw [5], which worked well for the cases we tried.

Suppose we have to solve

$$A\,x = y \tag{A.1}$$

The CG method consists of an initialisation step followed by a series of iterations. The iteration process runs fast for matrices which are near to the unity matrix. Therefore the matrix equation (A.1) is first preconditioned as follows. We first find matrices L and U which satisfy

$$L\,U \approx A$$

by performing a "normal" LU-decomposition an the nonzero elements of A and forcing those entries in L and U to zero, whose corresponding positions in A are zero. Then we try to solve the preconditioned system:

$$L^{-1}A\,U^{-1}(U\,x) = L^{-1}y \tag{A.2}$$

which has of course the same solution as (A.1)

Given an initial guess x_0 of the solution of (A.2), Kershaw's method is as follows:

Start:

$$r_0 := y - A\,x_0; \tag{A.3}$$

$$p_0 := (U^T U)^{-1}A^T (L\,L^T)^{-1}r_0; \tag{A.4}$$

$$i := 0;$$

REPEAT

$$a_i := \frac{(r_i\,,(L\,L^T)^{-1}r_i)}{(p_i\,,U^T U\,p_i)}; \tag{A.5}$$

$$x_{i+1} := x_i + a_i\,p_i\,; \tag{A.6}$$

$$r_{i+1} := r_i - a_i\,A\,p_i\,; \tag{A.7}$$

$$b_i := \frac{(r_{i+1}\,,(L\,L^T)^{-1}r_{i+1})}{(r_i\,,(L\,L^T)^{-1}r_i)}; \tag{A.8}$$

$$p_{i+1} := (U^T U)^{-1}A^T (L\,L^T)^{-1}r_{i+1} + b_i\,p_i\,; \tag{A.9}$$

UNTIL $|\,a_i\,p_i\,| < eps$

During the iteration process a number of matrix-vector multiplications, backsubstitutions and innerproduct calculations have to be performed, resulting in a relatively simple algorithm. The convergence properties of this algorithm were very well, except when the matrix became ill-conditioned, due to instabilities in the time integration.

It should be noted that the programming becomes more intricate when the sparsity is less regular. In that case also the required computer time and memory will grow.

References

1. M. van der Woude, "Two dimensional analysis of MOS transistors", Computing Centre Note 2, Eindhoven University of Technology, 1980.
2. M. van der Woude, " The Calculation of Small Signal Parameters and Sensitivities from the Spatial Analysis of Semiconductor Devices", in: Numerical Analysis of Semiconductor Devices", Proceedings of the NASECODE I Conference, Dublin, June 1979, pp. 296-298.
3. M. Reiser, "A two dimensional Numerical FET model for DC, AC and Large Signal Analysis", IEEE Transactions on Electron Devices, Vol. ED-20,No. 1, Jan. 1973, pp. 35-45.
4. A. J. G. Schoofs e.a., "TRIQUAMESH Gebruikershandleiding" (in Dutch), THE-rapport WE 78-01, Eindhoven University of Technology, Jan. 1980.
5. D. S. Kershaw, "The Incomplete Choleski - Conjugate Gradient Method for the Iterative Solution of Systems of Linear Equations", Journal of Computational Physics, Vol. 21, 1977, pp. 41-65.
6. M. S. Mock, "The charge neutral approximation and time dependent simulation", in: "Numerical Analysis of Semiconductor Devices", Proc. of NASECODE I Conference, Dublin, June 1979, pp. 120-135.
7. P. W. Lam, "Breakdown characteristics of a diffused Mesa Diode", Training report Philips International Institute, May 1981. Eindhoven University of Technology.
8. E. M. Buturla and P. E. Cotrell, "Simulation of semiconductor transport using coupled and decoupled solution techniques", Solid State Electronics, Vol. 23, 1980, pp. 331-334.
9. M. S. Mock, "A two dimensional mathematical model of the insulated gate field effect transistor", Solid State Electronics, Vol. 16, 1973, pp. 601-609.
10. "Computer Aided Device Design in Two Dimensions", Internal Documentation on the CADDET program, Hitachi Central Research Laboratory.
11. M. S. Mock, "Analysis of Mathematical Models of Semiconductor Devices", Boole Press, Dublin 1983.
12. M. S. Mock, private communication, 1985.
13. G. Strang and G. J. Fix, "An analysis of the Finite Element Method", Prentice Hall Inc. Englewood Cliffs, N.J.
14. J. H. M. van der Vorst, "Construction of a graphical postprocessor to the CADDET program", Eindhoven University of Technology, Department of Electrical Engineering Report EEA/320/9/84, October 1984.