TK 155.264

KFKI-1981-13

M.A. SZUHÁR

TWO-DIMENSIONAL MOS TRANSISTOR SIMULATION

Hungarian Academy of Sciences

CENTRAL RESEARCH INSTITUTE FOR PHYSICS

BUDAPEST

KFKI-1981-13

TWO-DIMENSIONAL MOS TRANSISTOR SIMULATION

M.A. Szuhár

Central Research Institute for Physics
H-1525 Budapest 114, P.O.B. 49, Hungary

HU ISSN 0368 5330
ISBN 963 371 790 6

ABSTRACT

A two-dimensional mathematical model for MOS transistors is reviewed and the general, physical and computational aspects of the problem described. Boltzmann statistics are used to formulate the basic differential equation system, whose discretization is made by finite-difference scheme. The coupled equation system is solved by the Gummel algorithm using the strongly implicit procedure in the inner loop. Finally, an error analysis is given which is very helpful in determining probable accuracy and in meeting the problems encountered when choosing a reasonable mesh system.

АННОТАЦИЯ

Статья занимается математическим моделированием МОП транзисторов в двумерном пространстве. Рассмотрены общие физические и математические стороны проблемы. Больцмановская статистика используется при описании системы фундаментальных дифференциальных уравнений, дискретизация которых осуществляется с помощью конечных дифференций. Связанная система уравнений решается с помощью алгорифма Гумеля, а для решения линейной системы уравнений используется SIP метод. В конце статьи дается анализ ошибок, который служит для определения ожидаемой точности и оказывает помощь в выборе подходящей сетевой системы.

KIVONAT

A dolgozat MOS tranzisztorok kétdimenziós matematikai modellezésével foglalkozik. Tárgyalja a probléma általános, fizikai és számitástechnikai aspektusait. A differenciális alapegyenletek felirásánál a Boltzmann statisztikát használja, a diszkrétizációt pedig a véges differenciák segitségével valósitja meg. A csatolt egyenletrendszer megoldására a Gummel algoritmus kerül felhasználásra a belső ciklusban pedig a Stone féle SIP módszert alkalmazzuk. A dolgozat végén egy hibaanalizist találunk, amely alkalmas a várható hiba becslésére és utmutatást ad a megfelelő rácsrendszer kiválasztásához.

I. INTRODUCTION

Sophisticated LSI integrated circuit design falls mainly into three broad categories, viz. system design, circuit design, device design, and it is hardly believable that somebody setting out from the basic physical aspects can finish, within a reasonable time, the project even of a medium scale integrated circuit. It is reasonable therefore, to divide the whole design into several steps in which the problem is solved up to a certain phase. A model is then created which can serve as an input to the next step. Such models could be the basic circuit models /gate, operational amplifier model, etc/ for system design; lumped device models /Ebers-Moll, charge-control model, etc./ for circuit design; and the basic mathematical description of device operation /one-dimensional, two-dimensional description with different simplifications and boundary conditions/ for device design. This work deals with the last of these i.e. with the mathematical simulation of semiconductor device operation by solving the basic differential equation system. To simplify the description we will show this type of modelling through MOS transistor simulation which is perhaps the most up - to-date simulation problem.

The earliest publications about the two-dimensional mathematical modelling of MOS transistors were issued in the first half of the 1970's [1,2], and interest in the subject is still growing. The mathematical modelling as was noticed earlier, sets off from the basic differential equation system of semiconductor devices and, taking into account the given boundary conditions, solves the equation system for the variables describing unambiguously the physical state of the device under the given conditions.

The appearance and rapid adaption of this kind of simulation can be explained by the fact that traditional device design based on quite a few simplifications could not describe many important physical processes taking place in sophisticated semiconductor devices. But these mainly two-dimensional processes could be observed and accurately calculated by mathemetical simulation. This kind of simulation allowed, for example, the understanding of some of the subthreshold and punch - through problems appearing in MOS structures. Futhermore, since mathematical modelling links the technoligical variables with the electrical device parameters it can serve as the theoretical substitution for device technology.

II. PHYSICAL ASPECTS OF SIMULATION

Physical processes taking place in semiconductor devices can be described by the Poisson /1/ and the two continuity equations /2,3/ as stated by Schockley [3] and Roosbroeck [4] at the dawn of semiconductor device research. Normalized, dimensionless forms of these equations are

$$\nabla^2\psi = n - p - DOP \qquad /1/$$

$$\frac{\partial n}{\partial t} = \nabla J_n - R_n \qquad /2/$$

$$\frac{\partial p}{\partial t} = \nabla J_p - R_p \qquad /3/$$

where:
- ψ - electrostatic potential
- n - electron concentration
- p - hole concentration
- DOP - doping level
- J_n - electron current density
- J_p - hole current density
- R_n, R_p - recombination-generation rate for electrons and holes respectively

Here, the Poisson equation connects the divergence of the electric field with the space charge, and the transport equation express the simple fact that the change of charge concentration in time can occur only by the change of current density in space or by generation and recombination processes. However, since the transient problems are beyond the scope of this work we can simplify the continuity equations /2,3/ to the so called transport equations /4,5/:

$$\nabla J_n = R_n \qquad /4/$$

$$\nabla J_p = -R_p \qquad /5/$$

The current densities of the transport equation can be described in detail, as follows /for simplicity, written for electrons only/:

$$J_n = -q\mu_n \cdot n \cdot \nabla \left| \frac{1}{q} E_F(n) \right| \qquad /6/$$

where: μ_n - electron mobility

q - electron charge

$E_F/n/$ - Fermi level as a function of electron concentration.

Further,

$$E_F(n) = q \cdot V - E_c - \xi(\frac{n}{N_c}) \qquad /7/$$

where: V - externally applied potential

E_c - the energy at the bottom of the conduction band

$\xi(\frac{n}{N_c})$ - the Fermi-Dirac distribution function

N_c - effective density of states in the conduction ban.

Now, if we presume that the semiconductor described is nondegenerate, then the Fermi-Dirac statistics can be substituted by the Boltzmann statistics, and

$$E_F(n) = q \cdot V - E_c - q\frac{kT}{q} \ln(\frac{n}{N_c}) \qquad /8/$$

where: k - Boltzmann's constant

T - absolute temperature

kT/q - thermal potential

In possession of these if we define the quasi-Fermi potential for electrons /9/ and for holes /10/, using the normalized form of equations further on:

$$\phi_n = \psi - \ln(n) \qquad /9/$$

$$\phi_p = \psi + \ln(p) \qquad /10/$$

we can express the current densities by

$$J_n = -\mu_n \cdot n \cdot \nabla \phi_n \qquad /11/$$

$$J_p = -\mu_p \cdot p \cdot \nabla \phi_p \qquad /12/$$

If we express $\nabla\phi$ in more detail we are able to obtain the traditional description of current densities, i.e. the current density as the sum of a drift and a diffusion component. Now if we substitute /11/ and /12/ into the basic equation system /1,4,5/ we can achieve a formulation,

$$\nabla^2\psi = n - p - DOP$$

$$\nabla(-\mu_n \cdot n \cdot \nabla\phi_n) = R_n \qquad /13/$$

$$\nabla(-\mu_p \cdot p \cdot \nabla\phi_p) = R_p \qquad /14/$$

which, expressing ϕ_n and ϕ_p through /9/ and /10/, can serve as the fundamental equation system for mathematical simulation. In this case we have only three unknown variables and three independent equations so that the equation system gives a unique solution.

The /1,13,14/ equation system can be rearranged in another way, expressing n and p through ϕ_n and ϕ_p:

$$\nabla^2\psi = \exp(\psi)\cdot\Phi_n - \exp(-\psi)\cdot\Phi_p - DOP \qquad /15/$$

$$\nabla(\mu_n \cdot \exp(\psi)\cdot\nabla\Phi_n) = R_n \qquad /16/$$

$$\nabla(-\mu_p \cdot \exp(-\psi)\cdot\nabla\Phi_p) = R_p \qquad /17/$$

where from the practical point of view, it was reasonable to use the exponential of quasi-Fermi potentials as unknowns, that is:

$$\Phi_n = \exp(-\phi_n) \qquad /18/$$

$$\Phi_p = \exp(\phi_p) \qquad /19/$$

In this form the three unknowns are ψ, Φ_n and Φ_p. From the physical point of view the two formulations are identical of course, but from the computational side they cause entirely different problems. For example, the first formulation does not imply diagonally dominant matrixes for numerical solution techniques which would be very important for the iterative solution procedures, but on the other hand it eliminates the overflow and undrflow problems during computation. The second formulation ensures the diagonal dominance but has no protection against the underflow and overflow problems.

We considered the second formulation to be better for our purposes and thus solved the /15,16,17/ equation system for the variables, ψ, Φ_n and Φ_p. For our purposes we could assume that the channel width was much greater than the channel length eliminating the side effects in a middle section so that the MOS transistor could have been accepted as a two-dimensional device and the /15,16,17/ equation system had to be solved in two dimensions, naturally under appropriate boundary conditions. Then, after solving the equation system the obtained three two-dimensional surfaces /ψ, Φ_n,Φ_p/ can serve as the basis for determining electron or hole concentration profiles, electron or hole current densities or any other device parameters.

The mobilities appearing in the transport equations were calculated using the Caughey-Thomas model [5] , thus the dependence of the mobilities on impurity concentration and field were taken into account.

$$\mu_p(DOP) = 47.7 + \frac{447.3}{1+(N_T/6.3\cdot 10^{16})^{0.76}} \qquad /20/$$

$$\mu_n(DOP) = 65 + \frac{1265}{1+ (N_T/8.5\cdot 10^{16})^{0.72}} \qquad /21/$$

$$\mu(DOP,E) = \frac{\mu(DOP)}{\left(1+\left|\frac{E\cdot\mu(DOP)}{V_m}\right|^S\right)^{1/S}} \qquad /22/$$

where: N_T - total impurity concentration

E - electric field

V_m, S - constants, values of which can be found in the literature [5].

The bulk recombination rate R is assumed to follow the Hall-Shockley--Read recombination law [6] as corresponding to a set of uniformly distributed recombination centres with a single energy level in the centre of the bandgap:

$$R = \frac{\Phi_n\Phi_p - 1}{T_p[\Phi_n\exp(\psi)+1] + T_n[\Phi_p\exp(-\psi)+1]} \qquad /23/$$

where: T_n, T_p - lifetime of mobile charges.

A similar expression could have been utilized for the surface recombination rate along the oxide layer, but was only taken into account through the flat-band voltage $/V_{FB}/$.

To make the task correct the differential equation system should be completed by the boundary conditions. As explained earlier, this task is essentialy two-dimensional thus the equation system should be solved within the ABCDA area */Fig. 1/* with the following boundary conditions:

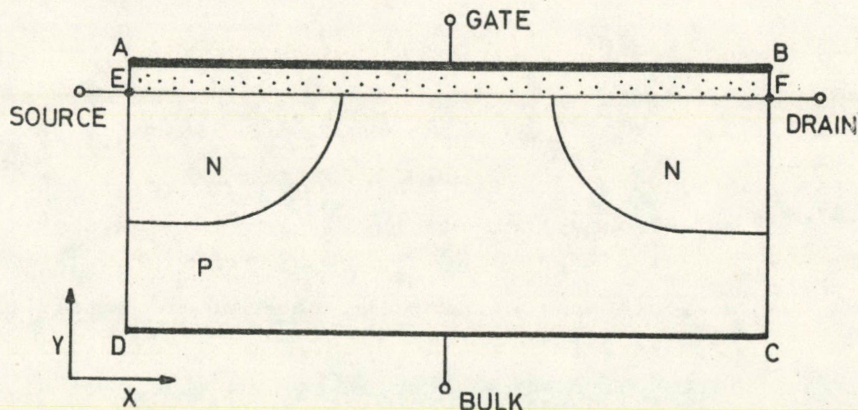

Fig. 1 Two-dimensional section of an MOS transistor

- along the contacts /AB and CD lines, E and F points/, infinite surface recombination velocities are assumed or, in other words, the carriers are assumed to be in thermodynamic equilibrium and the charge neutrality condition is valid, thus

$$\psi = \psi_{therm.eq.} + V_{applied}$$

$$\Phi_n = \Phi_p^{-1} = \exp /V_{applied}/ \qquad /24/$$

where the externally applied voltage, is given in normalized units and the thermal equilibrium potential is:

$$\psi_{therm.eq.} = \sinh^{-1} \frac{N_s}{2n_i}$$

where N_s - the doping level near the contact point. To speak about Φ_n and Φ_p out of the area EFDCE obviously has no sense.

- along the lines ED and FC in *Fig. 1*, the normal components of hole and electron current densities and of the electric field are zero, yielding

$$\frac{\partial \psi}{\partial x} = \frac{\partial \Phi_n}{\partial x} = \frac{\partial \Phi_p}{\partial x} = 0 \qquad /25/$$

- along the line EF the normal components of hole and electron current density are zero, thus

$$\frac{\partial \Phi_n}{\partial y} = \frac{\partial \Phi_p}{\partial y} = 0 \qquad /26/$$

- along the line BF assuming homogenous oxide layer without any charge in it the electric field is constant

$$\frac{\partial^2 \psi}{\partial y^2} = 0 \qquad /27/$$

III. MATHEMATICAL AND COMPUTATIONAL ASPECTS OF SIMULATION

As was stated in the previous section a second order nonlinear elliptical partial differential equation system /15,16,17/ in self- adjoint form should be solved for given boundary conditions /24,25,26,27/ for variables ψ, Φ_n and Φ_p. Because of the nonlinear character of the equations this task is only seldom solved exactly, mainly in trivial cases. Thus a numerical solution technique is to be applied, i.e. the equation system should be discretized on the given ABCDA area and solved only at discret points with the help of an

iteration algorithm convergence which, in the same way as the actual existence of the solution can only be proved in practice.

The equations were discretized by the five-point finite difference technique [7] using an automatically generated or externally given nonlinear mesh system. The choice of optimum mesh system is a rather complicated task, but one can say that an acceptable mesh system is one that has not too big and unhomogenous jumps of the electrostatic potential on the obtained potential surface. Thus, with the help of discretization the differential equations were transformed to linear ones, matrixes of which are sparse matrixes with nonzero terms only in five diagonals. But we would have an extra problem with the descretized Poisson equation if we solved it for ψ, namely we would achieve a transcendental equation with ψ linear at one side and exponential at the other. In view of this, it is reasonable to solve the Poisson equation not for ψ itself, but for the error in ψ between two successive iterations /δ/

$$\delta = \psi_{NEW} - \psi_{OLD} \qquad /28/$$

Thus using the Taylor expansion the Poisson equation should first be transformed to a form linear in δ [8]

$$\nabla^2\delta-(n+p)_{OLD}\cdot\delta = -\nabla^2\psi_{OLD}+(n-p)_{OLD}-DOP-O(\delta^2) \qquad /29/$$

and discretized only after this.

The matrix elements of the linearized and discretized Poisson equation for a particular i,j point of the mesh system will be:

$$a_{i,j-1} = 2.0/k_{j-1}(k_{j-1}+k_j)$$

$$a_{i-1,j} = 2.0/h_{i-1}(h_{i-1}+h_i)$$

$$a_{i,j} = -\ 2.0/h_{i-1}\cdot h_i-2.0/k_{j-1}\cdot k_j-(n_{i,j}+p_{i,j})_{OLD} \qquad /30/$$

$$a_{i,j+1} = 2.0/k_j(k_{j-1}+k_j)$$

$$a_{i+1,j} = 2.0/h_i(h_{i-1}+h_i)$$

$$b_{i,j} = (-\nabla^2\psi_{OLD})_{i,j}+(n_{i,j}-p_{i,j})_{OLD}-DOP_{i,j}$$

where: a,b - the terms in the i,j row of the given

$A\delta = b$ linear equation

h,k - mesh spacings as shown in *Fig. 2.*

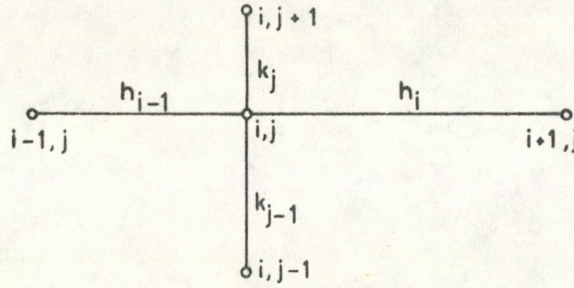

Fig. 2 Part of the nonlinear mesh system

The matrix terms of the discretized electron transport equation for a particular i,j point of the nonlinear mesh system will be:

$$C_{i,j-1} = 2.0 \cdot \mu_{i,j-1/2} \cdot \exp(\psi_{i,j-1/2})/k_{j-1}(k_{j-1}+k_j)$$

$$C_{i-1,j} = 2.0\ \mu_{i-1/2,j} \cdot \exp(\psi_{i-1/2,j})/h_{i-1}(h_{i-1}+h_i)$$

$$C_{i+1,j} = 2.0\ \mu_{i+1/2,j} \cdot \exp(\psi_{i+1/2,j})/h_i(h_{i-1}+h_i)$$

$$C_{i,j+1} = 2.0\ \mu_{i,j+1/2} \cdot \exp(\psi_{i,j+1/2})/k_j(k_{j-1}+k_j) \qquad /31/$$

$$C_{i,j} = -C_{i,j-1} - C_{i-1,j} - C_{i+1,j} - C_{i,j+1}$$

$$d_{i,j} = R_{i,j}$$

where: c,d - the terms in the i,j row of the given $C\Phi_n = D$ linear equation

$\mu_{i,j-1/2} \cdots \mu_{i,j+1/2}$ - parameters at the midpoint of that given in the suffix section /*Fig.3*/.

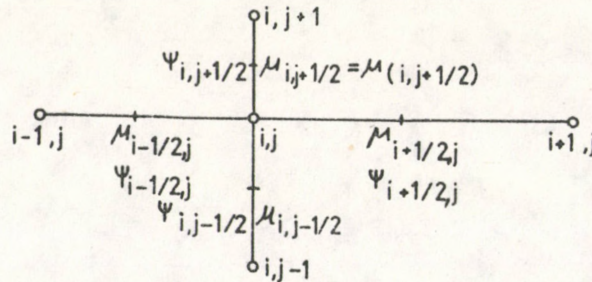

Fig. 3 Part of the nonlinear mesh system

The /30,31/ definitions are valid, of course, only at the interior points of the mesh system. For the boundary points they should be modified by the usual finite difference tricks [9].

As a result of discretization we have three sets of equations all of which consist of $n_i \cdot n_j$ linear equations /where n_i is the number of columns and n_j the number of rows in the mesh system/.
Since $n_i \cdot n_j \geq 500$ it is easy to understand that the choice of the solution technique is a very critical matter. The Gaussian elimination, which would be the simplest technique, cannot be used due to the accumulation of truncation errors not only decreasing the available accuracy but often leading even to divergency. A further reason for this elimination technique not being acceptable is that it would strongly increase the needed main store capacity due to filling the sparse matrix with undesirable terms.

However, since the coefficcient matrix for all three equations is diagonally dominant and positive definite, they can be solved by effective iterative processes such as SOR /successive over-relaxation method/, SLOR /successive line over-relaxation method/ [9], SIP /strongly implicit procedure/ [10], etc. We tried all three solution techniques and, have noticed that:

- in spite of simplicity the SOR method is applicable only for one--dimensional cases because of the slow convergence;
- the SLOR method had a good convergence rate at near optimal acceleration factor, but to find this factor one often needs more computational effort than to solve the original equation system;
- the SIP procedure, though it needed more store capacity and CPU time for one iteration than the SOR or SLOR methods,always ensured fast convergence even if the coefficient matrix was not definitely diagonally dominant and the acceleration factor was far enough from optimum.

Therefore the SIP procedure - which is essentially a factorization technique - was used for the solution of linear equation sets and suitable results were generally achieved in three iteration steps.

However, the separate solution of the linearized and discretized differential equations does not mean the end of the task since these are self-adjoint equations and solving one of them for one of the variables in an inner iteration loop is possible only if it is assumed that the other two equations are solved and the other two variables known. To meet this problem, an outer iteration loop is to be created and using, for example, the Gummel algorithm [11] the variables should be corrected step by step so that the final solution can be achived for all of them. In this case the linearized Poisson equation dominated the overall convergence behaviour

and nearly quadratic convergence results for low current levels. For higher current, the mutual coupling between the equations becomes stronger and the convergence slows down.

Because of the self-adjoint character of the equation system it is not reasonable to force a very accurate solution in the inner loop since this result is based anyway on a rough assumption viz. that the exact value of the other two variables is known. Accordingly, an exit criterion is placed on the calculations which allows at most 10% relative error in the inner cycle. In spite of this constraint the Poisson equation can be solved with less than 0.01 kT/q /<260μV/ absolute error and the transport equations with less than 0.5% relative error in at most 25 outer iteration cycles.

To prevent divergence several limitations were placed on the computation process and to ensure fast convergence an outer acceleration factor was used on the basis of Fadeev's work [12].

Figures 4 and *5* show the algorithm of the solution procedure.

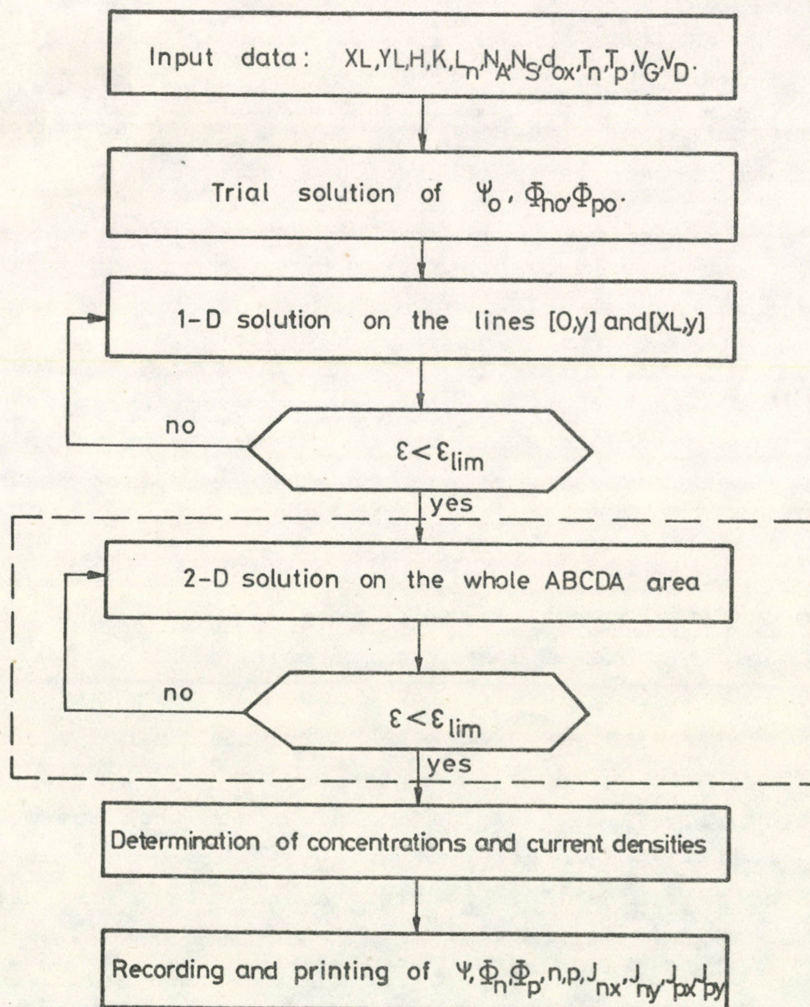

Fig. 4 Algorithm of the whole solution procedure

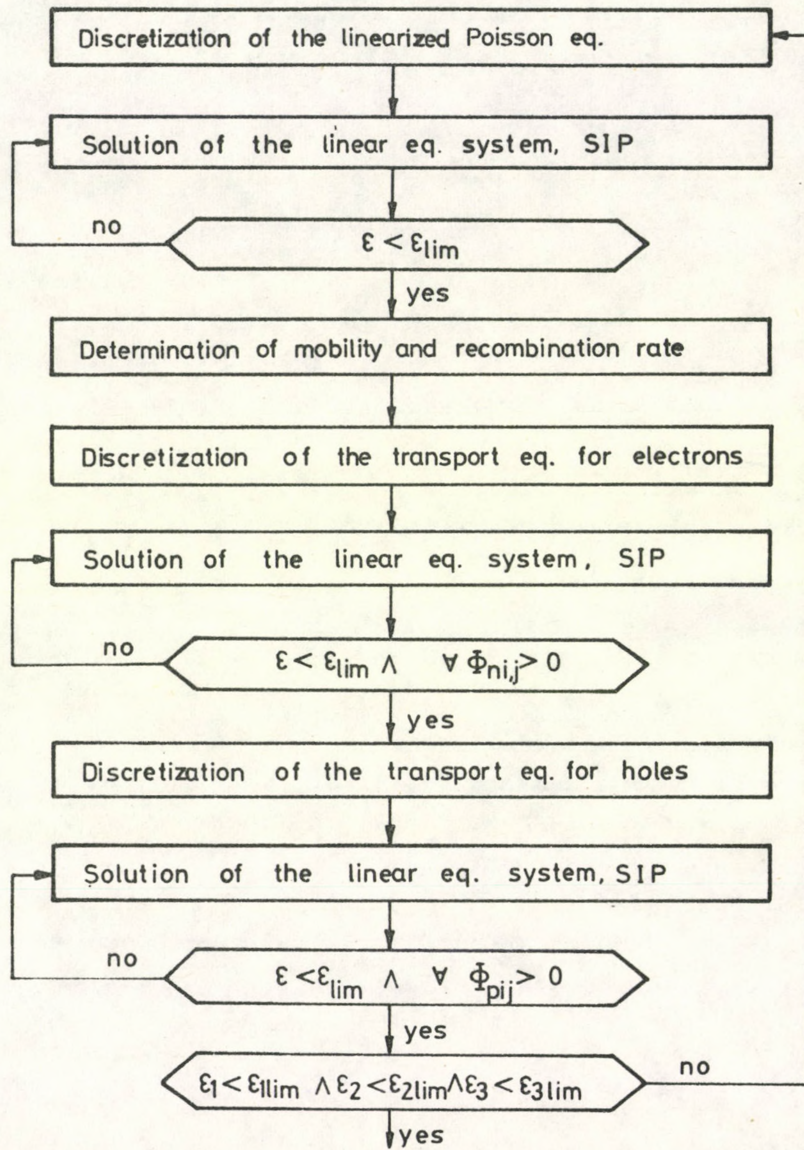

Fig. 5 Algorithm of the outer loop, framed by the dotted line in Fig. 4

Notation of the input and output parameters /Fig. 4/:

XL	- length of the examined section of the MOS transistor
YL	- depth of the examined section of the MOS transistor
H	- vector of mesh line spacings in the X direction
K	- vector of mesh line spacings in the Y direction
L_n	- diffusion length
N_A	- bulk concentration
N_S	- surface concentration of the diffusion layer
d_{ox}	- gate oxide thickness
T_n, T_p	- lifetime of electrons and holes respectively
V_G	- the externally applied gate voltage
V_D	- the externally appoied drain voltage
ψ	- electrostatic potential
Φ_n	- exponential quasi-Fermi potential for electrons
Φ_p	- exponential quasi-Fermi potential for holes
n	- electron concentration
p	- hole concentration
J_{nx}, J_{ny}	- electron current density in the X and Y directions respectively
J_{px}, J_{py}	- hole current density in the X and Y directions respectively

IV. ERROR ANALYSIS

A. TRUNCATION ERRORS

Consider first the errors introduced by the finite difference; for simplicity, in a single dimension only. These errors result from the truncation of the Taylor series expansion when approximating the differential operators.

The ∇^2 operator - when approximated by the matrix - gives an error τ_p. Using the three nonuniformly spaced points a,b and c shown in *Fig. 6*, $\nabla_x^2 f$ is approximated by Taylor series.

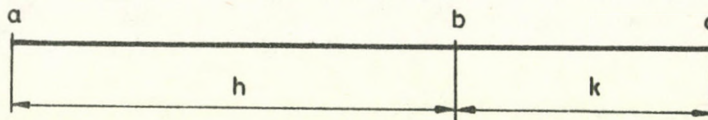

Fig. 6 Part of a 1-D mesh system

$$kf(a) = k\cdot\left(f(b) - hf'(b) + \frac{h^2}{2!}f''(b) - \frac{h^3}{3!}f'''(b) + \ldots\right) \qquad /32/$$

$$h\cdot f(c) = h\left(f(b) + kf'(b) + \frac{k^2}{2!}f''(b) + \frac{k^3}{3!}f'''(b) + \ldots\right) \qquad /33/$$

Summing /32/ and /33/ gives:

$$f''(b) = \frac{2f(a)}{h(h+k)} + \frac{2f(c)}{k(h+k)} - \frac{2f(b)}{h\cdot k} + \tau_E \qquad /34/$$

where:

$$\tau_E = \frac{h-k}{3!}f'''(b) + \frac{h^2+k^2-h\cdot k}{4!}f''''(b) + O(h^3) + O(k^3) \qquad /34/$$

Thus, if $h = k$ $\quad \tau_E = O(h^2)$

if $h >> k$ $\quad \tau_E = O(h)$ /36/

Generally it is impossible to fix an upper bound on f''' and f''''. However, in regions where function f is virtually constant these derivatives will vanish, and even with a comparatively coarse mesh the error will remain small.

For example, in the model the potential surface is constant or varies linearly in the neutral or charge-free regions. Thus, a coarse mesh may be used here without loss of accuracy in the potential solution.

The self-adjoint $\nabla e(x)\nabla$ when approximated by the matrix gives an error vector τ_c. Using the nonuniformly spaced points a,b and c, $\nabla_x(e\nabla_x f)$ is approximated at b.

The function e(x) is assumed to be defined at the implicit mid points p and q */Fig. 7/* for this analysis. The error in this approximation is considered later.

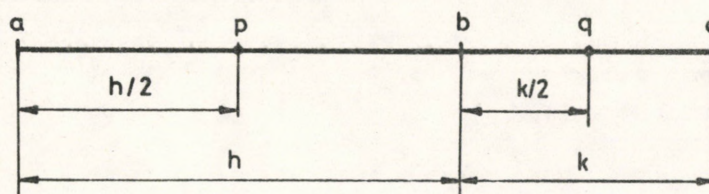

Fig. 7 Part of a 1-D mesh system

Using the Taylor series:

$$e(p)\cdot f'(p) = e(p)\left|\frac{f(b)-f(a)}{h} - \frac{h^2}{8\cdot 3!}f'''(p)+O(h^4)\right| \qquad /37/$$

$$e(q)f'(q) = e(q)\left|\frac{f(c)-f(b)}{k} - \frac{k^2}{8\cdot 3!}f'''(q)+O(k^4)\right| \qquad /38/$$

Now define:

$$g(x) = e(x)\nabla_x f(x) \qquad /39/$$

and again using the Taylor series:

$$g'(b) = \frac{2g(p)-2g(q)}{h+k} - \frac{h-k}{4\cdot 2!}g''(b) - \frac{h^2+k^2-h\cdot k}{8\cdot 3!}g'''(b)+O(h^3) \qquad /40/$$

Substituting /37/ and /38/ in /40/ gives:

$$g'(b) = \frac{2}{h+k}\left|e(q)\frac{f(c)-f(b)}{k} - e(p)\frac{f(b)-f(a)}{h}\right| + \tau_c$$

where:

$$\tau_c = e(p)\frac{h^2}{8\cdot 3!}f'''(p) - e(q)\frac{k^2}{8\cdot 3!}f'''(q) - \frac{h-k}{4\cdot 2!}g''(b) - \frac{h^2+k^2-h\cdot k}{8\cdot 3!}g'''(b) \qquad /41/$$

and again:

$$\text{if} \quad h = k \qquad \tau_c = O(h^2)$$

$$\text{if} \quad h \gg k \qquad \tau_c = O(h) \qquad /42/$$

In this analysis it is assumed that the function $e(x)$ is defined at points p and q. However, in the model the function is approximated for example, as the geometric average of the values at the adjacent mesh points, that is,

$$e(p) = \sqrt{e(a) \cdot e(b)} \qquad /43/$$

If constant mobility is assumed ($\mu_n = \mu_p = 1.0$), then

$$e(x) = \exp(\psi(x)) \qquad /44/$$

Using the Taylor series expansion it can be proved, that

$$\psi(p) = \frac{\psi(a)-\psi(b)}{2} - \left| \frac{h^2}{4 \cdot 2!} \psi''(p)+O(h^4) \right| \qquad /45/$$

Now define

$$E = \frac{h^2}{4 \cdot 2!}\psi''(p)+O(h^4) \qquad /46/$$

then through /43/; /44/ and /45/

$$\psi(p) = \ln(\sqrt{e(a) \cdot e(b)})-E \qquad /47/$$

Thus,

$$e(p) = \sqrt{e(a) \cdot e(b)} \cdot \exp(-E) \qquad /48/$$

Now define a relative error ε:

$$\varepsilon = \frac{e(p)-\sqrt{e(a) \cdot e(b)}}{\sqrt{e(a) \cdot e(b)}} \qquad /49/$$

then

$$\varepsilon = \exp(-E)-1 = -E+\frac{E^2}{2!} - \frac{E^3}{3!} + \dots . \qquad /50/$$

thus

$$\varepsilon = \frac{h^2}{4 \cdot 2!}\psi''(p)+O(h^4) \qquad /51/$$

Again the error is $O(h^2)$. However, an indication of the order of magnitude of the relative error can be obtained using Poisson's equation.

The Poisson equation in 1-D is:

$$\psi''(x) = -\rho(x) \qquad /52/$$

Thus

$$\varepsilon = - \frac{h^2}{4\cdot 2!}\rho(p)+O(h^4) \qquad /53/$$

This error will therefore be significant only in the space charge regions at the p-n junctions. Here the net space charge term is large due to the dipole effect. The mesh about these regions should therefore be fine. Conversely, in the quasi-space charge neutral regions the mesh may be quite coarse and the approximation /43/ will still be accurate.

b. ERRORS IN THE CURRENT DENSITY CALCULATIONS

Current densities are calculated from equation /54/. For simplicity the subscripts n and p have been dropped and only one coordinate direction is considered:

$$J = \mu\cdot\exp(\psi)\nabla_x\Phi \qquad /54/$$

Consider first the errors introduced by truncating the iterative procedure after a finite number of iterations. Thus, if $\bar{\phi}_E$ is the exact solution to the discretized equations, then if the iterative procedure is truncated it will introduce an error $\bar{\varepsilon}$ in $\bar{\phi}^{(m)}$, where $\bar{\phi}^{(m)}$ is the solution after m iterations

$$\bar{\varepsilon} = \bar{\phi}_E - \bar{\phi}^{(m)} \qquad /55/$$

Current densities are calculated from the local gradient of $\Phi^{(m)}$.

$$\nabla_x\Phi^{(m)}_{i+1/2} = \Phi^{(m)}_{i+1/2}\cdot\nabla_x\phi^{(m)}_{i+1/2} = \left[\frac{\phi_{i+1} - \phi^{(m)}_i}{\delta x} + O(\delta x^2)\right]\cdot\Phi^{(m)}_{i+1/2} \qquad /56/$$

If we ignore the second order terms in δx the error in /56/ is given by ξ_i which is related to the truncation errors ε_i by:

$$\xi_i = \frac{\varepsilon_{i+1} - \varepsilon_i}{\delta x}\cdot\Phi^{(m)}_{i+1/2} \qquad /57/$$

Let us now define a relative error ξ^R by:

$$\bar{\xi}^R = \bar{\xi}/\frac{\delta\bar{\phi}}{\delta x}\cdot\bar{\Phi} \qquad /58/$$

Then

$$\xi^R_i = \frac{\varepsilon_{i+1} - \varepsilon_i}{\phi^{(m)}_{i+1} - \phi^{(m)}_i} \qquad /59/$$

A bound on ε_i is now required. If the iterative procedure is continued until the maximum absolute change in ϕ between iterations is δ, then ε_i is defined by the sum of the absolute sequence produced by truncating the procedure.

Thus for all i :

$$/\varepsilon_i/ \leq c/\delta/ \qquad /60/$$

Where c is a constant determined by the rate of convergence. Provided that the rate of convergence is good c will be near unity and an upper bound may thus be given for the relative error

$$/\bar{\xi}^R/ \leq \frac{2/\delta/}{\min[\phi_{i+1}^{(m)} - \phi_i^{(m)}]} \qquad /61/$$

According to /61/ the largest errors will occur in the current density calculation in the neutral regions where the ϕ surface is flat. But from the other side the electrostatic potential and thus the quasi-Fermi potentials may be calculated most accurately just at these regions. Consider, for example, a calculation where the results show that the quasi-Fermi potential changes by only 10^{-4} kT/q between mesh points. Thus, to obtain an upper bound of 1% in the current density the iterative procedure must be continued until $\delta \leq 10^{-6}$kT/q at these regions which can easily be achieved.

There are also errors involved in approximating the gradient of an exponential by the central finite difference method. Unlike the above-mentioned truncation errors, these errors will occur in regions of high potential gradient.

Consider $\Phi_x = \exp(\phi(x))$; using Taylor series expansion an error term may be obtained for the central finite difference approximation to the gradient

$$\nabla\Phi(x) = \frac{\Phi(x+\delta x/2) - \Phi(x-\delta x/2)}{\delta x} + \varepsilon \qquad /62/$$

where

$$\varepsilon = -2\cdot\left|\frac{\delta x^2}{8\cdot 3!}\Phi'''(x) + \frac{\delta x^4}{32\cdot 5!}\Phi'''''(x) + \ldots\right| \qquad /63/$$

In the general case, bounds cannot be placed on the nth derivatives of Φ. However, if it is assumed that $\phi(x)$ is piecewise linear in x, all the second order or higher derivatives in $\phi(x)$ will vanish.

Thus the nth derivative of $\Phi(x)$ is given by

$$\Phi_{(x)}^{(n)} = \left(\frac{\delta\phi}{\delta x}\right)^n \Phi(x) \qquad /64/$$

and

$$\varepsilon = -2\cdot\Phi(x)\cdot\left|\frac{\delta x^2}{8\cdot 3!}\left(\frac{\delta\phi}{\delta x}\right)^3 + \frac{\delta x^4}{32\cdot 4!}\left(\frac{\delta\phi}{\delta x}\right)^5 + \dots\right| \qquad /65/$$

If a relative error ε^R is defined by

$$\varepsilon^R = \varepsilon / \frac{\delta\phi}{\delta x}\,\Phi(x) \qquad /66/$$

then

$$\varepsilon^R = -\left|\frac{\delta\phi^2}{24} + \frac{\delta\phi^4}{1920} + \frac{\delta\phi^6}{645\ 120} + \dots\right| \qquad /67/$$

According to this calculation a difference of just 5 kT/q in ϕ between mesh points gives a relative error of over 100%. However, because of the self consistent schme used, a high accuracy in one region of the device will imply a high accuracy elsewhere. Even so, this kind of relative error can serve as a mesh replacement criterion and shows that a very fine mesh should be used in regions of high potential gradient.

V. SUMMARY

This project has made available an MOS transistor model taking into account all the two-dimensional effects. This alone represents a great achievement in the field and forges one more link in the computer aided design of large integrated circuits. In addition to this, analysis of the results obtained will give a better understanding of device behaviour.

The only significant problem with this kind of modelling is the relatively high consumption of "computational power". Our computer program for MOS transistor simulation demands, for example, at about 200k main store capacity and 180-360 sec CPU time at an EC-1040 computer for determining the physical state of a transistor at a given work point, even though we strove to use sophisticated computational procedures and achieved an efficient and practically proved computer program.

However, taking into account all the advantages and all the problems one should accept that the right to existence of this kind of device simulation is already justified and since the computer and the computational technique are developing rapidly and the demands upon divice design are growing at least as rapidly, a fast increase in the significance of mathematical simulation can be expected in the future.

ACKNOWLEDGEMENTS

The author is indebted to the Semiconductor Department of the Industrial Research Institute for Electronics /HIKI/, Budapest, especially to L. Ugray, for allowing this work to be started at Southampton University, England, and to the Central Research Institute for Physics /KFKI/-especially to the Computer Department- for enabling it to be completed. He is grateful to T. Mohacsy for the critical review of the manuscript.

This work was sponsored as part of the Integrated Circuit Programme of the Solid State Physics Department at the Central Research Institute for Physics, Budapest.

REFERENCES

[1] M.S. Mock: A two-dimensional mathematical model of the insulated-gate field-effect transistor, Solid-State Electronics, 1973. Vol. 16 pp. 601-609

[2] D. Vandorpe et al: An accurate two-dimensional numerical analysis of the MOS transistor, Solid-State Electronics, 1972. Vol. 15 pp. 547-557

[3] W. Shockley: Theory of p-n junctions in semiconductors and p-n junction transistors, Bell.Syst.Tech.J., 1949. Vol. 29 pp. 435-489

[4] W.Van. Roosbroeck: Theory of the flow of electrons and holes in Ge and other semiconductors, Bell.Syst.Techn.J., 1950. Vol. 29 pp. 560-607

[5] D.M. Caughey, R.E. Thomas: Carrier mobilities in Si empirically related to doping and field, Proc. IEEE Letters, 1967. Vol. 55 pp. 2192-2193

[6] W. Shockley, W.T. Read: Statistics of the recombination of holes and electrons. Phys.Rev., 1952. Vol. 87 pp. 835-842

[7] G.D. Smith: Numerical solution of partial differential equations, OUP Oxford, 1965

[8] H.K. Gummel: Selfconsistent iterative scheme for one-dimensional steady state transistor calculations, IEEE Trans. Electron Devices, 1964. Vol. ED-11 pp. 455-465

[9] R.S. Varga: Matrix iterative analysis, Prentice-Hall, New York, 1962

[10] H.L. Stone: Iterative solution of implicit approximations of multidimensional partial differential equations, SIAM J.Numer.Anal., 1968. Vol. 5 pp. 530-558

[11] J.W. Slotboom: Computer-aided two-dimensional analysis of bipolar transistor, IEEE Trans.Electron Devices, 1973. Vol. ED-20 pp. 669-679

[12] A.K. Fagyejev: Vücsiszlityelnüje metodü linyejnoj algebrü, Fizmatgiz., Moszkva, 1960.

63.126

Kiadja a Központi Fizikai Kutató Intézet
Felelős kiadó: Krén Emil
Szakmai lektor: Mohácsy Tibor
Nyelvi lektor: Harvey Shenker
Gépelte: Beron Péterné
Példányszám: 450 Törzsszám: 81-110
Készült a KFKI sokszorositó üzemében
Felelős vezető: Nagy Károly
Budapest, 1981. február hó