

**TWO-DIMENSIONAL TRANSIENT SIMULATION
OF SEMICONDUCTOR DEVICES COUPLED
TO AN EXTERNAL CIRCUIT**

E. PALM and F. VAN de WIELE
Université Catholique de Louvain
Laboratoire de Microélectronique
Place du Levant 3, B-1348 Louvain-la-Neuve, Belgium

SUMMARY

In many practical situations of numerical semiconductor device simulation, especially in transient switching processes, a realistic analysis should take into account the interaction between the analysed device and the environment it is embedded in. The contact voltages can then no longer be considered, as is usually done, as "a priori" known boundary conditions for the basic partial differential equations (p.d.e.) describing the transport phenomena. They must rather be adjusted to satisfy the N (for an $(N+1)$ -terminal device) nonlinear equations of the external circuit. As the resolution algorithm influences directly the number of expensive numerical solutions of the p.d.e., it must be chosen very carefully.

The iterative algorithm proposed in this paper is the sequential $(N+1)$ -point secant method which essentially replaces the real current versus bias dependency by a linear relationship using the $N+1$ most recent approximations of the applied voltages and the corresponding currents. Each iteration requires only one resolution of the p.d.e., whereas, for example, a discrete Newton iteration costs $N+1$ times more. The method is able to handle efficiently resistive, reactive and nonlinear -analytically or even numerically modelled- components connected to the device under study. It is suitable for steady state and transient simulations in two or three space dimensions. It can be implemented without too much difficulty into existing device simulators, regardless of the particular numerical method used to solve the p.d.e. The asymptotic convergence rate of the secant method is superlinear and remains reasonably high for moderate values of N but, unfortunately, no theoretical guarantee of convergence can be given. Practical experience showed, however, that the method gives good results and the required precision was usually reached in 3 to 6 iterations.

The simulations of the turn-off of a gate controlled switch illustrates the use of the method. The impedances in gate and anode circuits definitely influence the observed current and voltage waveforms and may not be neglected.

1. INTRODUCTION

Since the first paper on numerical simulation of semiconductor devices, in 1964 [1], considerable effort has been and is being devoted to the improvement of existing methods or to the development of new, more powerful numerical techniques. Numerous publications witness of the progress realized in this area. From simplified 1D steady state problems, the application field has been widened to include 2D and even 3D, static and transient calculations, taking into account most of the relevant physical mechanisms through sophisticated physical models [e.g. 2-3].

Much less attention has, however, so far been paid to the problem of the interaction between the simulated device and the external circuit it is connected to. The basic partial differential equations (p.d.e.) describing the transport physics are commonly considered as a boundary value problem, where, at the metallic contacts, the carrier densities are fixed and the electric potential takes "a priori" known values, according to the prescribed bias conditions. The numerical resolution of the p.d.e. yields the distributions of internal variables such as free carrier concentrations, potential and current densities. After integration of the carrier flux, the terminal currents are obtained for given voltages. This corresponds to devices operating under ideal voltage sources and is often sufficient to yield useful simulation results. There are, however, many situations, where the interaction with a nonideal external circuit becomes essential for the device operation and can no longer be neglected. This is true for many switching processes, such as the forced turn-off of gate controlled power rectifiers. In the beginning of the switching process, the load impedance acts as a current source and keeps the anode current constant during the storage period. Afterwards, when the device impedance rises, the anode voltage increases sharply during the fall period of the anode current. Hence, at any moment, anode current and bias result from a balance between the gate controlled switch and the anode load circuit. The bias conditions are no longer "a priori" known, but must be adjusted to satisfy the nonlinear equations of the external circuit. The resolution algorithm for this apparently simple problem must be chosen very carefully, because it will require more or less, expensive numerical solutions of the p.d.e.

Already in 1977, Turgeon and Navon [4] presented simulation results of bipolar transistor switching with a reactive load, but little information was given about

the underlying algorithm. The approach developed by Mock [5-6] is attractive because it does not require any iterations for the circuit equations, but it is strongly related to the specific way the p.d.e. are solved. Recently, Grossmann and Hargrove [7] described a method where the circuit equations are inserted directly into the discretized p.d.c. and solved together with them. Once again, no additional iterative loop is necessary for the circuit equations, but a possible drawback is that the contact currents are evaluated by flux integration through a surface very close to the contact and hence laying in a heavily doped neutral region. In such regions, it is difficult to calculate accurately the current densities and, consequently, the terminal currents [8], in spite of highly accurate carrier densities and electric potential. Tomizawa et al. [9] used the discrete Newton's method to solve the circuit equations iteratively. This approach will be considered more in detail in section 3, because it has many similarities with our method which is based upon the sequential $(N+1)$ - point algorithm. It allows to handle linear -static or dynamic- or nonlinear loads, described either analytically or even numerically. In the following presentation, only the former situation is considered, but the same reasoning applies to several interconnected, numerically modelled semiconductor devices. In that case, each device is treated separately within every iteration cycle.

In the following section 2, the problem to be solved is stated in a fairly general way. The sequential $(N+1)$ -point secant method is described in section 3. It is compared to the related discrete Newton's method and basic theoretical convergence results are highlighted. The practical implementation of the method into an existing device simulation program is considered in section 4. The simulation of the turn-off of a gate controlled switch, presented in section 5, illustrates the use of the method for resistive circuit conditions and gives insight into the device internal phenomena during the switching. Section 6 concludes the present study.

2. STATEMENT OF THE PROBLEM

Let us consider a semiconductor device with $N+1$ metallic contacts, where N usually ranges from 1 to 3. One contact is grounded without loss of generality and the others are numbered from 1 to N . Figure 1 is a sketch of the analysed system composed of a semiconductor device and the external circuit.

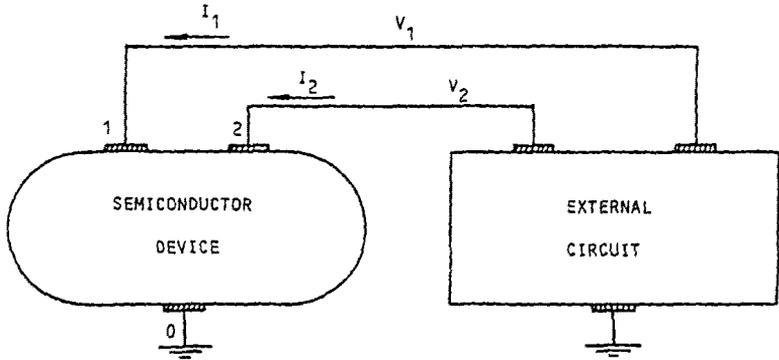


Fig.1 : Schematic representation of the analysed system ($N=2$).

For the semiconductor device, the current vs. voltage relationship can formally be written as :

$$\underline{I} = \underline{f}(\underline{V}) \quad (1)$$

where $\underline{I} = [I_1 \dots I_N]^t$, $\underline{V} = [V_1 \dots V_N]^t$ and $\underline{f} = [f_1 \dots f_N]^t$. Equation (1) must, however, not be taken as an explicitly given relationship. It rather is a symbolic representation of the complex sequence of operations necessary for the numerical simulation of an electronic component. The major part is the numerical resolution of the basic p.d.e., i.e. Poisson's equation and the carrier conservation laws, for given bias conditions \underline{V} . The specific method used to solve this boundary value problem (by finite elements or finite differences or any other approach) is not relevant at the present time. Once the internal distributions of the carrier densities and the potential are known, the current densities can be calculated. Integration of the current flow through surfaces separating the contacts, yields then the corresponding terminal currents.

Equation (1) still holds for transient calculations, but it has to be interpreted as a local relationship in time, assuming that partial time derivatives in the basic p.d.e. are expressed by discrete formulas and that the values of the dependent variables are known for all previous time instants.

The external circuit must satisfy a similar relationship :

$$\underline{I} = -\underline{g}(\underline{V}) \quad (2)$$

Normally, the external circuit is very simple, because the interest is focused upon the semiconductor device and the physical phenomena occurring within it. For instance, with resistive loads, we have :

$$g_i(\underline{V}) = g_i(V_i) = (V_i - E_i)/R_i \quad i = 1 \dots N \quad (3)$$

where R_i and E_i are respectively, the load resistance and the i th voltage source connected to the i th contact. The external circuit might, however, also be more complicated and contain dynamic or nonlinear elements described by analytical or even numerical models.

The question is now to find the bias conditions \underline{V} such that the currents leaving the external circuit equal the currents entering the semiconductor :

$$\underline{f}(\underline{V}) + \underline{g}(\underline{V}) = 0 \quad (4)$$

This apparently simple nonlinear problem requires a careful selection of the resolution algorithm in accordance with special features. The number of unknowns N is very small, but the function evaluation $\underline{f} + \underline{g}$ is very expensive. Furthermore, no explicit expressions of the derivatives can be given, but they can be estimated through discrete approximations. Consequently, the resolution method should converge very quickly, and the number of evaluations of the function (and the derivatives) per iteration should be as small as possible.

3. THE SEQUENTIAL (N+1)-POINT SECANT METHOD

To solve equation (4), one might think of Newton's iterative method :

$$\underline{d} = -(\underline{J}_f + \underline{J}_g)^{-1} (\underline{f}(\underline{V}) + \underline{g}(\underline{V})) \quad (5)$$

$$\underline{V}_{\text{new}} = \underline{V} + \underline{d}$$

where the Jacobian matrix $\underline{J} = \underline{J}_f + \underline{J}_g$ is the sum of the derivatives of \underline{f} and \underline{g} w.r.t. the components of \underline{V} . The latter are normally easily obtained, but no explicit formulas can be given for the former. Therefore, \underline{J}_f must be evaluated by other means. For instance, each component of \underline{V} can be increased individually by a small amount and the corresponding function values \underline{f} are used to approximate the derivatives by difference formulas. This discretized

Newton iteration has been successfully used by Tomizawa et al. [9], but it will have, in general, only a linear rate of convergence [10, p.186] and requires $N+1$ function evaluations per iteration.

Another approach is given by the general secant method [10, p.189 ff.] which replaces the Jacobian matrix by a secant matrix obtained in the following way. The exact current vs. voltage dependency $\underline{I} = \underline{f}(\underline{V})$ is replaced by a linear relationship using the function values at $N+1$ different points. Obviously, still $N+1$ function evaluations are, in general, necessary, but, if the points are properly selected, a second order convergence rate can be reached.

A very interesting special case is the sequential $(N+1)$ -point method [10, p.196], where the available function values at the $N+1$ most recent approximations of \underline{V} are used to set up the secant matrix. Only one function evaluation is sufficient to find the improved approximation of \underline{V} and hence to finish off the iteration. As the method has also a superlinear convergence rate, it appears to be particularly well suited for the present application. Although the convergence rate decreases with N , it remains reasonably high for the values of N of practical interest: it is 1.62, 1.47 and 1.38 for $N = 1, 2$ and 3 respectively [10, p.373]. Unfortunately, there is no theoretical guarantee of local convergence [10, p.379], but our experience showed that the method gives good results. A special problem is due to the fact that the secant matrix becomes singular, if the $N+1$ last approximations of \underline{V} are not in general position in R^N , i.e. if the N vectors issued from one of these points to the others are linearly dependent. This problem is however rarely encountered and can easily be dealt with by keeping the old secant matrix until the points are again in general position. The algorithm degenerates hence temporarily into a parallel chord method.

4. PRACTICAL IMPLEMENTATION

The procedure described in the previous section can be implemented into existing semiconductor device simulators without too much difficulty, independently of the particular numerical method used to solve the p.d.e. The method is suitable for steady state and transient calculations in two or three space dimensions, regardless of the internal resolution algorithm.

We have introduced the method into our 2D simulation program based upon a generalized finite difference discre-

tization scheme which handles combined triangular and rectangular grids [11]. The original code has been extended to include the time discretization, using a first order implicit predictor-corrector scheme [12]. The time step is adjusted automatically to achieve a constant prescribed local truncation error. The discrete equations are solved by a coupled resolution approach, based upon a full Newton linearization. The system matrix is assembled efficiently by a node by node technique which avoids multiple evaluation of the same quantities [11-12]. The resolution of the linear system is carried out using a successive line overrelaxation (SLOR) method which has been adapted to account for the particular matrix structure arising from irregular finite element like meshes. The line length is variable and the nonzero elemental 3x3 blocks are irregularly distributed outside of the three main diagonals. These features can be handled without too much added complexity, and the SLOR method is economic in both memory requirements and CPU time.

If the secant method is used for the circuit equations, an additional outer iteration loop for the bias conditions is introduced. At the start of the simulation, the secant matrix is not yet defined, because only one function value is available. The matrix is initialized using the discrete Newton's method. The first iteration needs hence N+1 function evaluations. This matrix is then kept unchanged for the following N-1 iterations, until N+1 function values are available. From then on, the secant matrix is updated at every iteration provided the resulting matrix is nonsingular, as explained in section 3.

The iteration loop is continued until convergence is reached, typically within 6 to 10 iterations for steady state calculations, depending on the accuracy of the initial guess. In the transient case, the required number of iterations ranges from 3 to 6, the former value corresponds to a smaller time step. It is not necessary to re-initialize the secant matrix when a new time step is started. The last matrix of the previous time step is simply taken over to the next one.

5. APPLICATION : THE TURN-OFF OF A GATE CONTROLLED SWITCH (GCS)

The forced switching-off of a gate controlled silicon rectifier is a typical application, where the interaction between the device and the load circuit is essential for the device operation. The considered structure of the gate controlled switch (GCS) is sketched in fig.2a. The doping profile along the x-axis is given in fig.2b, for

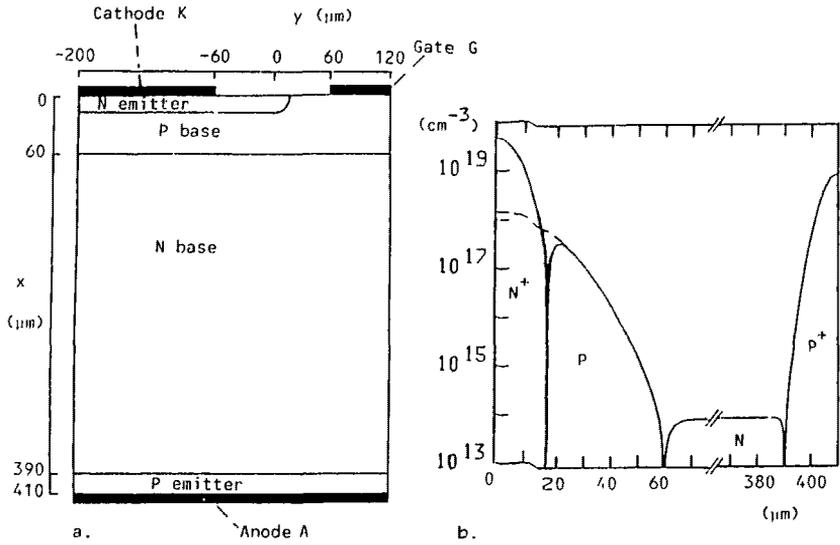


Fig.2 : Schematic representation of the GCS and the impurity distribution along the x-axis.

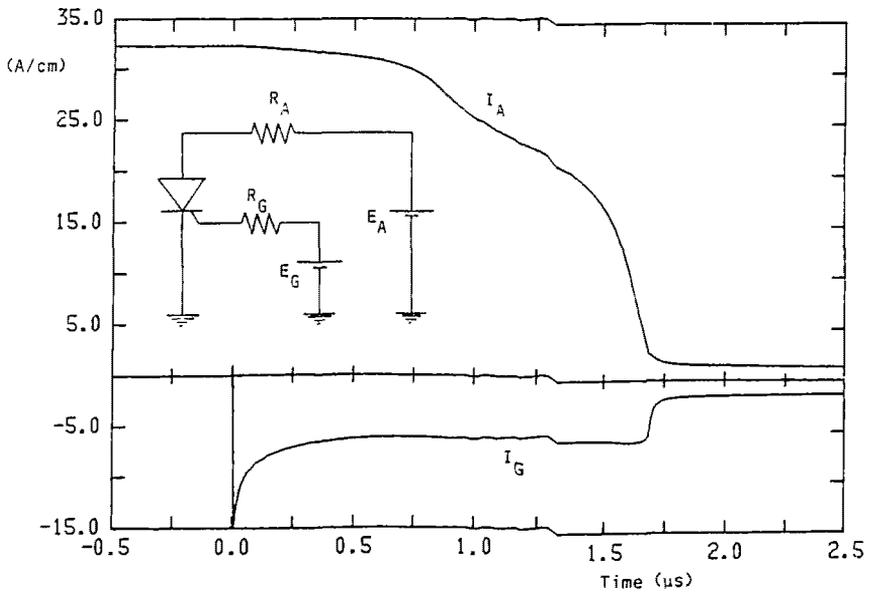


Fig.3 : Current wave forms for the simulated turn-off.
 $(R_A=3\Omega\text{cm}, R_G=0.5\Omega\text{cm}, E_A=100\text{V}, E_G=1.08\text{V if } t \leq 0, E_G=-10\text{V if } t > 0)$

the region below the cathode ($y \leq 0$) and below the gate ($y \geq 30 \mu\text{m}$). The physical models that have been used to complete the basic current conservation laws and Poisson's equation are those described by Engl and Dirks [13]. The expressions of the carrier mobilities account for impurity scattering, carrier-carrier scattering and carrier velocity saturation at high electric fields. Auger and Shockley-Read-Hall recombination, the latter with impurity dependent lifetimes [13] are included.

Fig.3 represents the current wave forms during the gate turn-off with the indicated resistive circuit. The simulation starts from the on-state ($V_{AK} = 3\text{V}$) without gate current. At $t = 0$, the gate source voltage changes abruptly from the open-circuit value 1.08V to -10V . The storage period with an (almost) constant anode current lasts $0.8 \mu\text{s}$. It is followed by the fall time corresponding to a rapidly decreasing anode current, from 0.8 to $1.7 \mu\text{s}$. Afterwards, once the cathode current has vanished, begins the tail period with slowly decaying anode and gate currents.

Typical convergence behaviour of the sequential secant method is represented in fig.4 where the residuals of both anode and gate circuit equations are plotted against the number of iterations, for two different time steps. The first lies in the early storage period ($0.04 \mu\text{s}$)

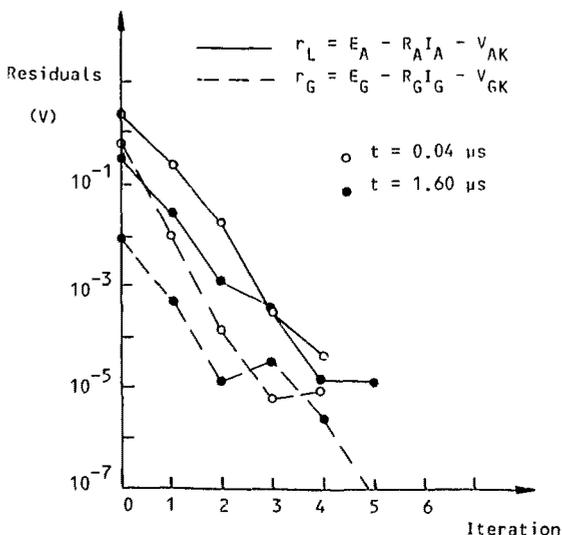


Fig.4 : Typical convergence behaviour of the secant method

and the second is in the fall period at $1.6 \mu\text{s}$. Residuals smaller than 10^{-4}V are usually reached in less than 5 iterations. An exception is the first time step, which in the present case required 11 iterations, mainly because the external gate voltage is changed abruptly.

The sudden change of gate bias is also responsible for the somewhat unusual negative peak of gate (and cathode) current, immediately after $t = 0$ (fig.3). It corresponds to the fast reverse recovery of the part of the cathode junction which is closest to the gate and which becomes suddenly reverse biased. This phenomenon is not observed if the gate source voltage changes as a slower ramp [14]. From then on, the gate current remains almost constant until the tail period. The storage time, in the present example, results from two contributions of comparable importance. The conducting area in the P-base is reduced by the so-called plasma-pinching mechanism (for about $0.5 \mu\text{s}$) which is followed by the desaturation of the remaining active part of the central junction ($0.3 \mu\text{s}$). These considerations are illustrated in fig.5 showing the electric potential distribution and current lines at different time instants. Hence, both mechanisms must be considered for an accurate evaluation of the storage time, though their relative importance may vary with the device parameters. After $0.5 \mu\text{s}$, only a narrow channel near the device center is still conducting carriers through the P-base, and 80% of the anode current flow through about $30 \mu\text{m}$. But even within this area, the current density is not uniform and increases towards the device center. The maximum value at $y = -200 \mu\text{m}$ of $12,000 \text{ A/cm}^2$ is reached at $t = 0.8 \mu\text{s}$, i.e. in the beginning of the fall period. This agrees with other results for the case that a high current is turned off [14]. In the present example, the anode current waveform during the fall time is, however, different and shows a sequence of fast (around $0.8 \mu\text{s}$), more slow (from 1 to $1.2 \mu\text{s}$) and, again, fast decrease. Similar behaviour has been observed on experimental devices, but a satisfactory explanation has not yet been found. At the end of the fall time, the rest of the cathode junction becomes reverse biased, and the cathode current is reversed for a short time, while the junction recovers, and drops then to zero. The gate current changes by the same amount and becomes equal to the anode current. The GCS behaves now like a P-N-P transistor in the active operation mode which has its base in high injection. The current decays at the rate the excess charge is removed from the N-base. The time constant of the tail period depends mainly on the high-level lifetime in the N-base. The current spreads out over the whole width of the N-base and, except for the highly doped part of the P-base, the device behaviour becomes essentially one-dimensional.

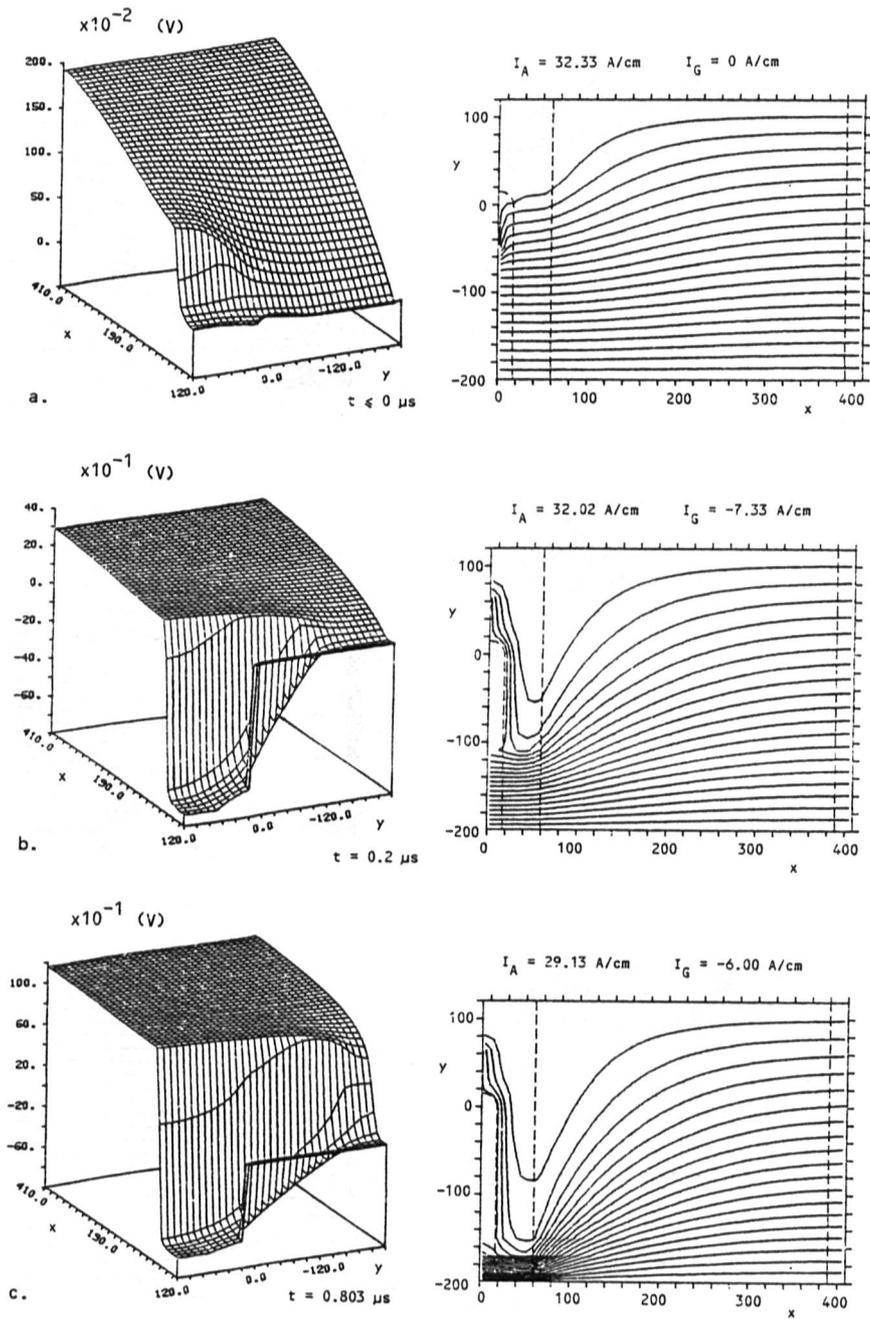


Fig.5 : The potential and the current lines at different moments of the turn-off.

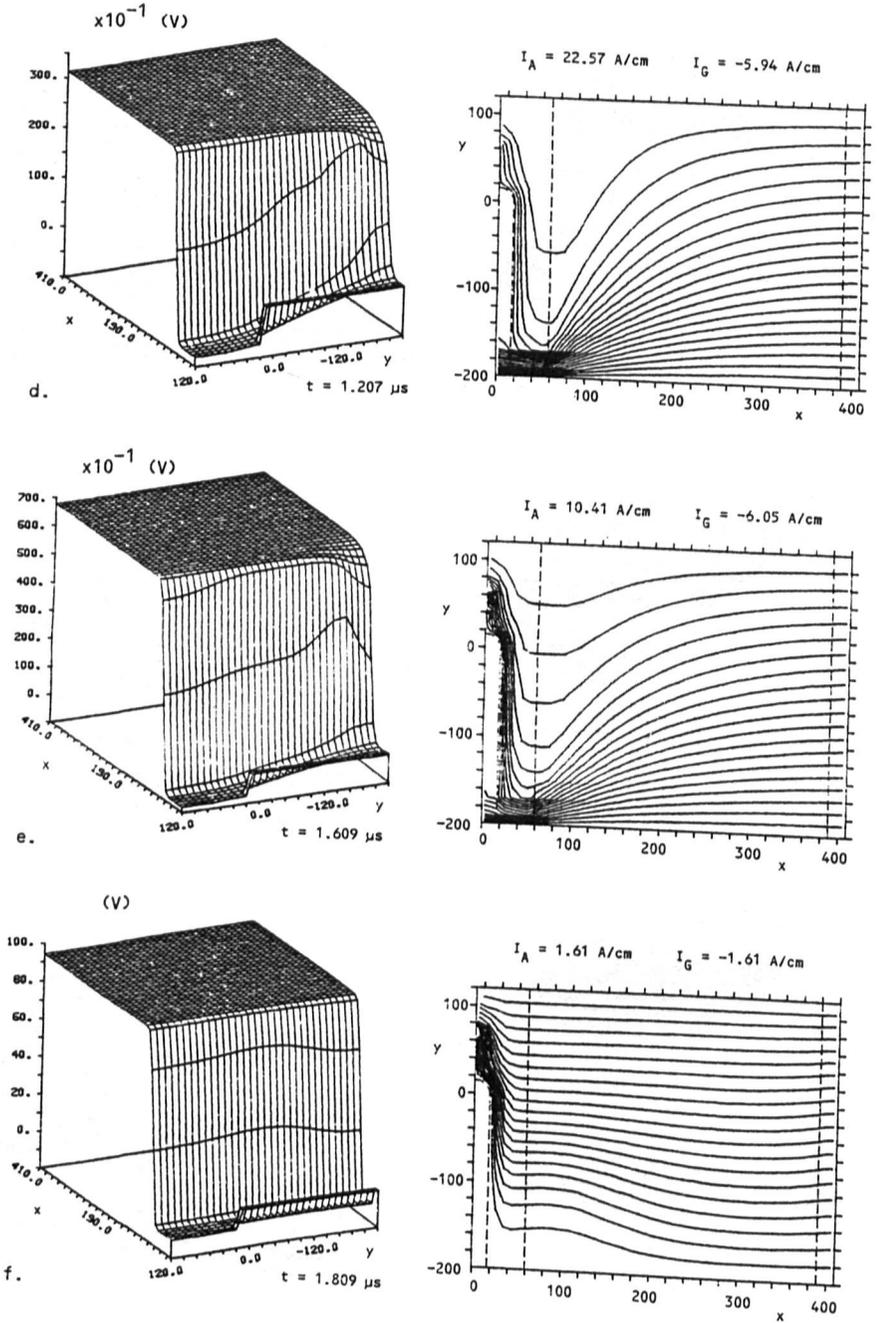


Fig. 5 (cont.)

6. CONCLUSION

The sequential (N+1)-point secant method has been presented as an attractive approach to handle the interaction between a numerically modelled semiconductor device and the external circuit it is connected to. In many situations, especially in transient switching processes, this interaction becomes essential for the device operation and may not be neglected. The turn-off of a gate controlled switch is a typical example of considerable practical interest. A two-dimensional simulation of the gate turn-off has illustrated the proposed algorithm which showed a good convergence behaviour. Though only resistive circuit conditions were considered in practice, the method is fundamentally more general. It is able to handle also linear dynamic or nonlinear elements in the external circuit. Even several connected numerically modelled semiconductor devices may be considered and they will be treated separately by the simulation program.

ACKNOWLEDGMENTS

The financial support for the research project on theoretical analysis of semiconductor power devices from I.R.S.I.A. (Institut pour l'Encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture) and from A.C.E.C. (Ateliers de Construction Electriques de Charle-roi) is gratefully acknowledged.

REFERENCES

- [1] GUMMEL, H.K.
"A Self-Consistent Iterative Scheme for One-Dimensional Steady-State Transistor Calculations"
IEEE Trans. Electron Dev. ED-11, p.455, 1964.
- [2] BROWNE, B.T. and MILLER, J.J.H. (ed).
"Numerical Analysis of Semiconductor Devices and Integrated Circuits", Proceedings of the NASECODE II Conference, held from 17th to 19th June 1981 in Dublin
Boole Press, Dublin, Ireland, 1981.
- [3] MILLER J.J.H. (ed.)
"NASECODE III", Proceedings of the NASECODE III Conference, held from 15th to 17th June 1983, in Galway
Boole Press, Dublin, Ireland, 1983.
- [4] TURGEON, L.J. and NAVON, D.H.
"Two-Dimensional Nonisothermal Carrier Flow in a Transistor Structure under Reactive Circuit Conditions"

- IEEE Trans. Electron Dev., ED-25, p.837, 1978.
- [5] MOCK, M.S.
"Time-Dependent Simulation of Coupled Devices"
In ref. [2], p.113, 1981.
- [6] MOCK, M.S.
"Analysis of Mathematical Models of Semiconductor Devices"
Boole Press, Dublin, Ireland, 1983.
- [7] GROSSMAN, B.M. and HARGROVE, M.J.
"Numerical Solution of the Semiconductor Transport Equations with Current Boundary Conditions"
IEEE Trans. Electron Dev. ED-30, p.1092, 1983.
- [8] POLSKY, B.S. and RIMSHANS, J.S.
"Numerical Simulation of Transient Processes in 2-D Bipolar Transistors"
Solid-State Electronics Vol.24, p.1081, 1981.
- [9] TOMIZAWA, M., YOKOYAMA, K., YOSHII, A., and SUDO, T.
"Two-Dimensional Device Simulator for Gate Level Characterization"
Solid-State Electronics Vol. 25, p.913, 1982.
- [10] ORTEGA, J.M. and RHEINBOLDT, W.C.
"Iterative Solution of Nonlinear Equations in Several Variables"
Academic Press, New York, 1970.
- [11] PALM, E. and VAN de WIELE, F.
"A General Finite Difference Formulation for Combined Triangular and Rectangular Grids with Application to a Gate Controlled Rectifier in the On-State",
In ref.[3], p.214, 1983.
- [12] PALM, E. and VAN de WIELE, F.
To be published.
- [13] ENGL, W.L. and DIRKS, H.
"Models of Physical Parameters", In : An Introduction to the Numerical Analysis of Semiconductor Devices and Integrated Circuits" (Ed. J.J.H. Miller), Lecture Notes of a Short Course held in association with the NASECODE II Conference,
Boole Press, Dublin, Ireland, p.42, 1981.
- [14] NAKAGAWA, A. and NAVON, D.H.
"A Time- and Temperature-Dependent Simulation of the GTO Turn-Off Process", In "International Electron Device Meeting IEDM 1982", Technical Digest, paper 18.5, IEEE, New York, p.496, 1982.
-