Comp. & Maths. with Appls. Vol. 9, No. 2, pp. 293–305, 1983
Printed in Great Britain.

0097–4943/83/020293–13$03.00/0
Pergamon Press Ltd.

NUMERICAL MODELING OF A MOVING BOUNDARY DIFFUSION/DRIFT PROBLEM

H. ROHDIN†

Department of Electrical Engineering, Washington University, St. Louis, MO 63130, U.S.A.

(*Received June* 1982)

Communicated by E. Y. Rodin

Abstract—We have modeled numerically the redistribution by diffusion and drift of impurities during epitaxial growth of semiconductors. We use a Crank–Nicholson scheme with dynamically adjusted time increment. The coupling between the redistribution of the charged impurities and the electric field is accounted for by solving the nonlinear Shockley–Poisson equation for an arbitrary doping profile by means of a quasi-linearization scheme. The combination of large gradients and a moving boundary necessitates a dynamically adjusted nonuniform mesh in the finite difference schemes. Except for some occasionally occurring spurious ripples, which are accounted for by stability arguments, the results are physically real and can explain some experimentally observed features [1].

1. INTRODUCTION

The present paper describes in detail the numerical methods used in a recently published simulation of impurity redistribution during epitaxial growth of semiconductors [1]. In Ref. [1] we describe the physical situation in detail, discuss the assumptions made in the model and give two extensive examples of typical cases of epitaxy.

At the relatively high temperatures typical of epitaxial growth the charged impurities present in the substrate and those incorporated in the grown layer will redistribute due to diffusion and drift in the built-in electric field in the semiconductor. Frequently impurities diffuse in a reactive fashion (e.g. by a substitutional–interstitial mechanism [2]). The redistribution can generate impurity and carrier density profiles that are unexpected and that may be unfavorable for certain applications. Reactive diffusion has not been modeled here, but it seems in principle straightforward to extend the treatment to include it. Even without this complication, however, the simulation involves several numerical problems. The equation for the electrostatic potential is nonlinear and has a forcing term which is the total charge density associated with the impurities. This means that there is a two-way coupling between the electric field and the redistribution. We also have to allow for a nonuniform mesh which has to be variable due to the moving boundary (the growth surface). These and other aspects of the numerical method will be discussed and examples of growth simulations will be given.

2. FORMULATION OF THE PROBLEM

The governing equation for the concentration $N_k(x, t)$ of the kth impurity I_k is the continuity equation which for simple diffusion and drift due to an electric field $E(x, t)$ (the position x being measured from the moving growth surface) reads:

$$\frac{\partial N_k}{\partial t} = D_k \frac{\partial^2 N_k}{\partial x^2} - z_k B_k D_k \frac{\partial}{\partial x}(EN_k) - v \frac{\partial N_k}{\partial x}; \quad k = 1, 2, \ldots, k_m. \tag{1}$$

D_k is the diffusion coefficient for I_k. z_k is the number of elementary charges (with sign) that I_k is carrying. B_k is a numerical factor in the Einstein relation for the ionic mobility for I_k [3]. v is the growth velocity.

The electric field is given by

$$E = -\frac{\partial \psi}{\partial x},$$

†Present address: Hewlett-Packard Laboratories, 1501 Page Mill Road, Palo Alto, CA 94304, U.S.A.

where $\psi(x, t)$ is the electrostatic potential governed by the Shockley–Poisson equation[4]

$$\frac{\partial^2 \psi}{\partial x^2} = 2 \sinh(\psi_0 + \psi) - Q. \tag{2}$$

Q is the net doping:

$$Q(x, t) = \sum_k z_k N_k(x, t).$$

We treat the substrate as semi-infinite with uniform impurity concentrations in the bulk. ψ_0 is then given by

$$\psi_0 = \sinh^{-1}\left(\frac{Q(\infty, t)}{2}\right)$$

where $Q(\infty, t) = Q(\infty, 0)$ independent of t. Numerically, we will study a finite region

$$0 \le x \le x_m$$

such that there is essentially no redistribution or field for $x > x_m$ at any time studied.

We have an initial value problem for the concentrations:

$$N_k(x, 0) = N_{0k}(x); \quad 0 \le x \le x_m; \quad k = 1, 2, \ldots, k_m.$$

The boundary conditions are

$$-D_k \frac{\partial N_k}{\partial x}(0, t) + \{z_k B_k D_k E(0, t) + v(t)\} N_k(0, t) = v(t) N_{ak}(t) \qquad k = 1, 2, \ldots, k_m \tag{3a}$$

$$N_k(x_m, t) = N_k(x_m, 0) = N_{0k}(x_m) \tag{3b}$$

$$\frac{\partial \psi}{\partial x}(0, t) = -\frac{\sigma_s}{\epsilon} \tag{4a}$$

$$\psi(x_m, t) = 0. \tag{4b}$$

N_{ak} is the intended doping with impurity No. k of the epitaxial layer. σ_s is a surface charge density associated with surface states and ϵ is the dielectric constant.

3. NUMERICAL SOLUTION

The problem is solved numerically by stepwise (in time) calculations of the impurity concentrations using the field and concentrations from the last step, and average values of the growth velocity v and doping N_{ak} over the present time increment. As illustrated in the flowchart (Fig. 1) and discussed in detail below, the time increment is automatically adjusted.

Each step is associated with two two-point boundary condition problems: calculation of the field due to the present impurity concentration and the redistribution from the latter during the step, due to this field, diffusion and growth. The potential (and from it, the field) is calculated with a quasilinearization scheme[5] and the subsequent redistribution is calculated by a Crank–Nicholson scheme ([6], p. 17). Both involve tridiagonal matrix equations arising from finite difference approximations of the derivatives at discrete mesh points. The equations are solved by Gauss elimination.

A method for dynamical mesh adjustment is developed and a discussion of errors and stability follows.

3.1 *Automatic adjustment of time increment*

As indicated in Fig. 1, the criterion for sufficiently small time increment Δt is that the redistribution obtained directly from t to $t + \Delta t$ is the same (to within a preset tolerance) as that

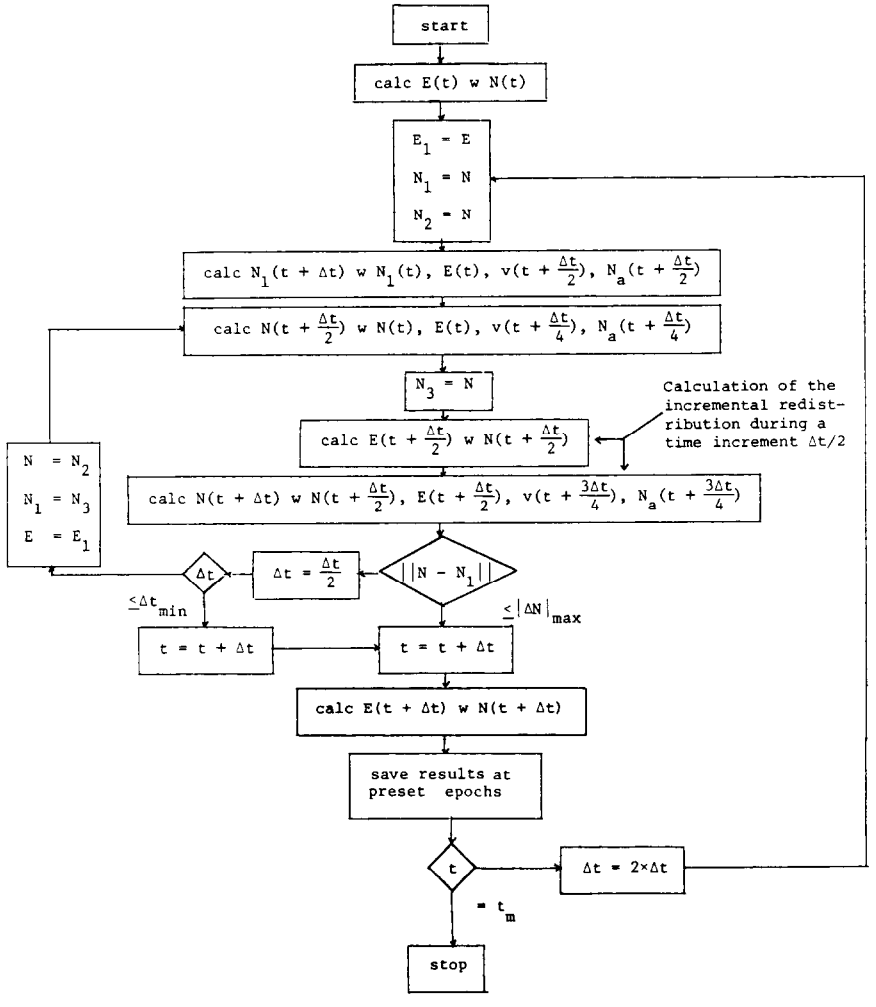

Fig. 1. Flow chart showing schematically the calculation of the time evolution of the impurity concentrations and electric field. The x dependence has for this purpose not been explicitly indicated while the time is indicated after the variables. "calc" stands for "calculate". "w" stands for "with", meaning "using". t = variable containing the time up to which the calculation has been done; E = vector variable containing the electric field at its mesh points; N = array containing the concentrations for the different impurity types at their respective mesh points; v = variable containing the growth velocity; N_a = vector variable containing the intentional doping; Δt = variable containing the attempted time increment; t_m = variable containing the maximum t; N_i = storage array for N, $i = 1, 2, 3$; E_1 = storage vector for E; $|\Delta N|_{max}$ = variable containing the maximum error in N; Δt_{min} = variable containing minimum Δt.

obtained in two substeps, t to $t+\Delta t/2$ and $t+\Delta t/2$ to $t+\Delta t$. If it is not, the time increment is halved unless it gets smaller than a minimum allowed. If it is, the subsequent time increment is initially doubled. In this way an automatic adjustment of the time increment is implemented.

3.1.1 *Choice of minimum time increment.* In the stepwise calculation of the redistribution the result should, in principle, become more accurate the smaller the time increment Δt is picked. For small diffusion constants and/or large growth velocities, very small Δt are necessary to get convergence. However, we cannot allow too small a Δt, primarily because of excessive computer costs. But also: the diffusion theory is macroscopic, single jump events for the atoms cannot be expected to be accounted for accurately. During a minimum time increment Δt_{min} at least the dominant redistribution process should displace the atoms some distance, considered to be the minimum macroscopic distance. Here we write this na, where a is the atomic jump distance and n is some number, typically in the range 10–100.

$$\max_m \{v_m\} \Delta t_{min} = na$$

v_m is the velocity of the atoms. m stands for the different mechanisms diffusion, drift and growth:

$$v_{\text{diff}} = \frac{D}{a}$$

$$v_{\text{drift}} = \frac{qEa}{kT} v_{\text{diff}}$$

Typically $[(qEa)/kT] \ll 1$ so that the minimum time increment becomes

$$\Delta t_{\min} = \frac{na}{\max\left\{v_{\text{growth}}, \frac{D}{a}\right\}}$$

3.1.2 *Comparison of two discretized functions.* To check whether two functions $f^{(1)}$ and $f^{(2)}$ known at discrete points should be considered identical, the straightforward way is to look at the differences of the two functions at the mesh points associated with one of the functions. However, in regions of locally steep gradients, the difference between the functions can be large even though the two function graphs are close together. This is often encountered in the problem at hand. Instead of allowing just a deviation $|\Delta f|_{\max}$ between the functions, we allow a deviation $|\Delta f|_{\max}$ in the function values *or* a deviation $|\Delta x|_{\max}$ between the function graphs in the x direction. A simple way of estimating the deviation $|\Delta x|$ between two graphs is indicated in Fig. 2. The criterion for small enough Δt becomes

$$|\Delta f_i| \leq |\Delta f|_{\max} \quad \textit{or} \quad |\Delta f_i| \leq |f_i'|\,|\Delta x|_{\max}, \quad i = 1, 2, \ldots, i_m - 1$$

where f_i' is the derivative of the function known at x_i at this point. The advantage of this criterion is that it allows us to take larger time increments in the calculation of the redistribution.

It seems reasonable to choose $|\Delta x|_{\max}$ as the minimum macroscopic distance na mentioned in Section 3.1.1. $|\Delta f|_{\max}$ can be chosen as a small number relative to the maximum value of the function. Due to error coupling it may be necessary to pick $|\Delta f|_{\max}$ with more care (see Section 3.5).

3.2 *Quasilinearization of the Shockley–Poisson equation*

The Shockley–Poisson equation (2) is very nonlinear. A quasilinearization solution[5] was used. This means that we made successive approximations $\psi^{(n)}$ of the potential ψ, and obtained

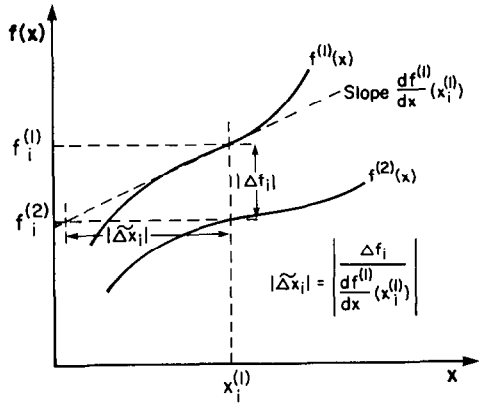

Fig. 2. A simple way of estimating the deviation in the x-direction between two discretized functions. Here we are looking at the ith mesh point of the first function. The estimate is $|\tilde{\Delta x}_i|$.

a better approximation $\psi^{(n+1)}$ by solving the linearized equation

$$\frac{\partial^2 \psi^{(n+1)}}{\partial x^2} = 2 \sinh(\psi_0 + \psi^{(n)}) + 2 \cosh(\psi_0 + \psi^{(n)})(\psi^{(n+1)} - \psi^{(n)}) - Q(x) \tag{5}$$

The quasilinearization method is well suited for this problem for three reasons:

(1) $\sinh(\psi_0 + \psi)$ is monotone in ψ, so even if the linear correction term $\cosh(\psi_0 + \psi^{(n)})(\psi^{(n+1)} - \psi^{(n)})$ for small values of n may be a poor approximation of the entire nonlinear Taylor correction

$$\sum_{m=1}^{\infty} \frac{1}{m!} \left[\frac{\mathrm{d}^m}{\mathrm{d}y^m} \sinh y\right]_{y=\psi_0+\psi^{(n)}} (\psi^{(n+1)} - \psi^{(n)})^m,$$

it always has the correct sign. Sufficiently many iterations should thus give an arbitrarily small error.

(2) Since we make the calculations for discrete times spaced by a small time increment, the change in ψ is moderate, so that few iterations are necessary, except at $t = 0$.

(3) Initially, at $t = 0$, it is obvious that a reasonable first guess is the charge-neutrality potential ψ_{CHN}

$$\psi^{(0)} = \psi_{\mathrm{CHN}} = \sinh^{-1}\left(\frac{Q(x)}{2}\right) - \psi_0. \tag{6}$$

This is a rather good guess since the mobile carriers in a semiconductor tend to compensate changes in the doping. Significant deviation of the correct potential from ψ_{CHN} occur close to charged boundary surfaces and in regions of steep doping gradients.

3.3 *Crank–Nicholson scheme for the continuity equation*

To solve the continuity equation (1) we employ a scheme similar to the Crank–Nicholson scheme for the simple diffusion equation described in [6], p. 17. This means that we approximate the l.h.s. of (1) by a finite difference

$$\frac{\partial N_k}{\partial t}\left(x, t + \frac{\Delta t}{2}\right) \approx \frac{N_k(x, t + \Delta t) - N_k(x, t)}{\Delta t}$$

where Δt is the time increment discussed above. The r.h.s. of (1) is also evaluated at $t + \Delta t/2$. In this linear approximation of $N_k(x, t')$ in $t \le t' \le t + \Delta t$ we get

$$N_k\left(x, t + \frac{\Delta t}{2}\right) = \frac{1}{2}\{N_k(x, t + \Delta t) + N_k(x, t)\}$$

and similar expressions for its derivatives. This makes the scheme implicit with improved stability ([6], p. 64). However, we do have to settle for the approximation

$$E\left(x, t + \frac{\Delta t}{2}\right) \approx E(x, t) \tag{7}$$

since we do not know $E(x, t + \Delta t)$ before $N_k(x, t + \Delta t)$, $k = 1, 2, \ldots, k_m$, are known.

With these approximations the continuity equation becomes (for each time increment):

$$\begin{aligned} & D_k \frac{\partial^2}{\partial x^2} N_k(x, t + \Delta t) - z_k B_k D_k \frac{\partial}{\partial x}\{E(x, t) N_k(x, t + \Delta t)\} \\ & \quad - v\left(t + \frac{\Delta t}{2}\right) \frac{\partial}{\partial x} N_k(x, t + \Delta t) - \frac{2}{\Delta t} N_k(x, t + \Delta t) \\ & \quad = - D_k \frac{\partial^2}{\partial x^2} N_k(x, t) + z_k B_k D_k \frac{\partial}{\partial x}\{E(x, t) N_k(x, t)\} \\ & \quad + v\left(t + \frac{\Delta t}{2}\right) \frac{\partial}{\partial x} N_k(x, t) - \frac{2}{\Delta t} N_k(x, t). \end{aligned} \tag{8}$$

The r.h.s. is of course known from the last step. The boundary condition (3a) has to be evaluated at $t + \Delta t/2$ in a similar fashion.

3.4 *Finite difference approximation with nonuniform mesh*

For each step the two equations (5) and (8) with boundary conditions (3) and (4) respectively are solved by the method of finite differences. As indicated in Section 2, we choose an upper boundary x_m effectively at infinity such that there is negligible redistribution and field for $x \geq x_m$. The specific choice depends on the diffusion constant, the Debye length, growth velocity and growth time. Since the function values are prescribed at x_m (see (3b) and (4b)), the choice can be checked after the completion of the calculation by looking at the gradients of the functions at $x = x_m$ and deciding whether they are sufficiently small.

The functions we are dealing with are almost constant in most of the studied region, but typically, have steep gradients in a few narrow regions. This necessitates a nonuniform mesh. The fact that the material is growing means that the steep-gradient regions in general will move, and we have to have a dynamically modified mesh.

For a uniform mesh $\{x_i\}_{i=1}^{i_m}$ with spacing h, the first and second derivatives of a function $f(x)$ are usually approximated by the finite differences ([6], p. 6).

$$f_i' = \frac{df}{dx}(x_i) = \frac{f_{i+1} - f_{i-1}}{2h}$$

$$f_i'' = \frac{d^2f}{dx^2}(x_i) = \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2}$$

where $f_i = f(x_i)$. This is really the result one gets if one fits a parabola locally through the points (x_{i-1}, f_{i-1}), (x_i, f_i) and (x_{i+1}, f_{i+1}).

If we have a nonuniform mesh with the varying spacing

$$h_i = x_{i+1} - x_i,$$

and make a local parabola approximation $\tilde{f}_i(x)$ of $f(x)$ in the region $x_{i-1} \leq x \leq x_{i+1}$ (Fig. 3)

$$f(x) \approx \tilde{f}_i(x) = f_i + f_i'(x - x_i) + \frac{1}{2} f_i''(x - x_i)^2,$$

the approximations for the derivatives become:

$$f_i' = -\frac{h_i}{h_{i-1}(h_i + h_{i-1})} f_{i-1} + \frac{h_i - h_{i-1}}{h_i h_{i-1}} f_i + \frac{h_{i-1}}{h_i(h_i + h_{i-1})} f_{i+1}$$

$$f_i'' = \frac{2}{h_{i-1}(h_i + h_{i-1})} f_{i-1} - \frac{2}{h_i h_{i-1}} f_i + \frac{2}{h_i(h_i + h_{i-1})} f_{i+1}.$$

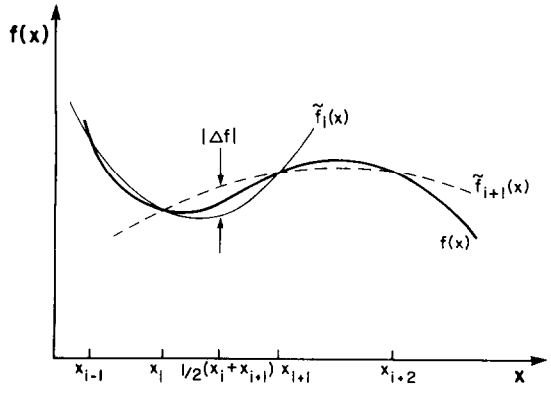

Fig. 3. Two adjacent local parabola approximations $\tilde{f}_i(x)$ and $\tilde{f}_{i+1}(x)$ of the discretized function $f(x)$. The maximum absolute difference between the two in their common region of validity $[x_i, x_{i+1}]$ occurs at the midpoint.

3.4.1 *Interpolation.* Since we will vary the mesh dynamically and we use the old function values and derivatives in the calculation, we have to interpolate these to the new mesh points. To do this we can use the local parabola approximation above. However, at any point x in the studied region we have *two* approximations of the function (Fig. 3). In the first and last interval, $[x_1 = 0, x_2]$ and $[x_{i_m-1}, x_{i_m} = x_m]$ respectively, one of the approximations, $\tilde{f}_1(x)$ and $\tilde{f}_{i_m}(x)$ respectively, is a straight line. To resolve this ambiguity, we take a weighted average of $\tilde{f}_i(x)$ and $\tilde{f}_{i+1}(x)$ for a point x in the interval $[x_i, x_{i+1}]$:

$$f(x) \approx \tilde{\tilde{f}}_{i,i+1}(x) = \frac{1}{h_i}\{(x_{i+1} - x)\tilde{f}_i(x) + (x - x_i)\tilde{f}_{i+1}(x)\}$$

This is a cubic interpolation scheme readily obtained from our (quadradic) local parabola approximations. It gives a unique value for $f(x)$ and a well behaved continuous derivative $(\mathrm{d}f/\mathrm{d}x)$:

$$\frac{\mathrm{d}\tilde{\tilde{f}}_{i,i+1}}{\mathrm{d}x}(x_i) = \frac{\mathrm{d}\tilde{\tilde{f}}_{i-1,i}}{\mathrm{d}x}(x_i).$$

3.4.2 *Dynamical mesh adjustment.* The fact that we have two approximations for the function $f(x)$ at any point x in $[x_1, x_{i_m}]$ is actually very useful, since it gives us a means to decide when to modify the mesh. The maximum absolute difference $|\Delta f|$ between the two approximations occurs at the midpoint. We impose a *criterion for adding a mesh point* between two existing ones, namely

$$|\Delta f| > |\Delta f|_{\max}.$$

If this is so, an additional mesh point is created at $x = \frac{1}{2}(x_i + x_{i+1})$. We go through all intervals in this fashion, do the Gauss elimination (with the old function values etc. interpolated to the new mesh points) and repeat the procedure until no further mesh points need to be added.

Now, there is an additional complication in this problem since the physical boundary is moving in the laboratory frame (due to growth). Without consideration of this, there would be regions left with unnecessarily fine meshes. We solve this by imposing a criterion *for removing a mesh point*, namely:

$$|\Delta f| < \epsilon |\Delta f|_{\max}.$$

ϵ is chosen considerably less than 1 but not so small that removal of excess mesh points does not occur. We typically chose $\epsilon = 0.01$.

In summary: we have an interval or a "window"

$$[\epsilon |\Delta f|_{\max}, |\Delta f|_{\max}]$$

and if $|\Delta f|$ falls within the window, no mesh modification is made.

3.4.3 *Calculation of incremental redistribution.* Calculation of the redistribution over a time increment consists of two steps: updating the electric field (solve (5)) and calculating the incremental redistribution with this field (solve (8)). In Fig. 1, one such step is indicated.

In the nonuniform mesh the finite difference approximations of the two equations (5) and (8) with the boundary conditions (3) and (4), respectively, become tridiagonal matrix equations of the form:

$$\begin{bmatrix} b_2 & -c_2 & 0 & 0 \ldots\ldots & \\ -a_3 & b_3 & -c_3 & 0 \ldots\ldots & \\ 0 & -a_i & b_i & -c_i & 0 \\ 0 \ldots\ldots\ldots 0 & & -a_{i_m-1} & b_{i_m-1} & \end{bmatrix} \begin{bmatrix} f_2 \\ f_3 \\ f_i \\ f_{i_m-1} \end{bmatrix} = \begin{bmatrix} d_2 \\ d_3 \\ d_i \\ d_{i_m-1} \end{bmatrix} \tag{9}$$

where f_i is the unknown function value at the ith mesh point. This is easily solved by Gauss elimination ([6], p. 20).

The step starts by calculating the electric field due to the total net doping at time t resulting from the last incremental redistribution. f_i is the value of $\psi^{(n+1)}$, i.e. the $(n+1)$th iteration of the potential, at the ith mesh point. The other matrix and vector elements in (9) are known and contain the mesh distances, the values of the doping, and the last iteration of the potential, intentional doping at time $t+\Delta t/2$. As is clear from equation (8), the electric field just result converges. The potential from the last step is used as the zeroth iteration, except at $t=0$ when the charge neutrality potential (6) is used.

In the subsequent calculation of the redistribution of I_k $(k=1,2,\ldots,k_m)$, f_i is the value of N_k at the ith mesh point at time $t+\Delta t$. In this case the other matrix and vector elements in (9) contain, in addition to the mesh distances, the time increment and the growth velocity and intentional doping at time $t+[(\Delta t)/2]$. As is clear from equation (8), the electric field just calculated and the concentration from the previous step, both interpolated to the present mesh points, appear as well.

In each such step the meshes for the field and concentration are dynamically adjusted as described.

3.5 *Error coupling*

Since the whole treatment is numerical we have to allow certain errors in the potential, field and concentrations. For the purpose of studying the coupling of errors, we approximate the potential by the charge neutrality potential

$$\psi \approx \psi_{\mathrm{CHN}} = \sinh^{-1}\left(\frac{Q(x)}{2}\right) - \psi_0$$

For the electric field we get:

$$E \approx -\frac{1}{\sqrt{(Q^2+4)}}\frac{\mathrm{d}Q}{\mathrm{d}x}.$$

A variation δQ in the net doping and a variation $\delta(\mathrm{d}Q/\mathrm{d}x)$ in the slope of the net doping give the variations

$$\delta\psi \approx \frac{\delta Q}{Q+\sqrt{(Q^2+4)}}$$

and

$$\delta E \approx \frac{1}{\sqrt{(Q^2+4)}}\left\{\frac{Q(\mathrm{d}Q/\mathrm{d}x)}{Q^2+4}\,\delta Q - \delta\left(\frac{\mathrm{d}Q}{\mathrm{d}x}\right)\right\}$$

in the potential and field respectively. The largest errors in the field and potential are thus expected to occur when the doping is low. However, for net doping less than the intrinsic electron concentration, an error in the doping does not affect the potential and field.

If we have a majority dopant, an error in its concentration will cause an error in the field, which will cause an error in the rest of the dopant concentrations. In the cases we have studied so far, we have allowed a relative error of 10^{-3} for each dopant (relative to the maximum concentration of the dopant). This will allow large absolute error for a majority dopant, sometimes as large as the minority dopant concentration. To be able to calculate accurately changes in conductivity etc., we would have to decrease the allowed error for the majority dopant.

3.6 *Stability analysis*

For some impurities ripples appear in the concentration close to the substrate-layer interface. Sometimes ripples appear in the field/potential as well. The examples below will illustrate the phenomenon. The ripples do not seem to appear first in the field, and this was verified by a field-free trial run. The ripples appear first in the concentrations, and under

conditions of intermediate doping concentrations, are transferred by error coupling to the field. This indicates that the ripples are due to some tendency towards instability in the Crank–Nicholson scheme of solving the equation

$$\frac{\partial N}{\partial t} = D\frac{\partial^2 N}{\partial x^2} - v\frac{\partial N}{\partial x}. \tag{10}$$

To get an idea why the phenomenon occurs, we make a simplified stability analysis assuming: (1) uniform mesh distances h ($M+2$ mesh points) (2) concentration $N=0$ at the two boundaries. Locally the meshes are approximately uniform, and the phenomenon is highly localized close to the interface, well away from the boundaries, once the growth has started.

With these assumptions we can use the treatment in [6], p. 60, and extend it to our problem. The tridiagonal $M \times M$-matrix equation in the Crank–Nicholson scheme with a time increment Δt for the diffusion-growth problem (10) reads, with the two assumptions above

$$(2\mathbf{I} - \mathbf{A})\bar{N}_{j+1} = (2\mathbf{I} + \mathbf{A})\bar{N}_j$$

where $\mathbf{I}$ is the $M \times M$ unit matrix,

$$\mathbf{A} = \begin{bmatrix} -2r & r-s & 0 & 0 \ldots\ldots\ldots 0 \\ r+s & -2r & r-s & 0 \ldots\ldots\ldots 0 \\ 0 \ldots & r+s & -2r & r-s \quad 0 \quad \ldots 0 \\ 0 \ldots\ldots\ldots\ldots 0 & & & r+s \quad -2r \end{bmatrix},$$

$$r = \frac{D\Delta t}{h^2},$$

$$s = \frac{v\Delta t}{2h}$$

and $\bar{N}_j$ is the column vector containing the mesh point concentrations at epoch no. j.

We can write the matrix equation

$$\bar{N}_{j+1} = \mathbf{B}\bar{N}_j, \quad \text{where } \mathbf{B} = (2\mathbf{I} - \mathbf{A})^{-1}(2\mathbf{I} + \mathbf{A})$$

The errors $\bar{e}_j$ will satisfy the same equation. For the scheme to be stable, i.e. for the errors not to grow, all eigenvalues of $\mathbf{B}$ must be less than unity in magnitude. An eigenvalue λ_B can be expressed in terms of the corresponding eigenvalue λ_A for $\mathbf{A}$ as

$$\lambda_B = \frac{2+\lambda_A}{2-\lambda_A}.$$

So for stability we must have:

$$Re\,(\lambda_A) < 0$$

The eigenvalues of $\mathbf{A}$ are according to [6], p. 65

$$\lambda_A = -2r + 2\sqrt{(r^2 - s^2)}\cos\frac{m\pi}{M+1}; \quad m = 1, 2, \ldots, M$$

For $r \geq s$, these are real and negative so that the scheme is stable; errors die out exponentially. For $s > r$, the eigenvalues are complex

$$\lambda_A = -2r + i2\sqrt{(s^2 - r^2)}\cos\frac{m\pi}{M+1}; \quad m = 1, 2, \ldots, M$$

We get *oscillatory errors* which, under the assumptions made, should die out. This is essentially what is observed: At a time t an error occurs at an x-value corresponding approximately to the thickness of the epitaxial layer. Initially it seems to grow, perhaps due to the fact that we actually have nonuniform dynamically modified mesh distances. After a short time, however, the oscillations do die out. Studying time evolution, it looks like a propagating wave packet whose amplitude grows slightly, possibly due to the reasons mentioned and to some coupling to the field. To avoid oscillations we should have

$$s \leq r$$

or

$$h \leq \frac{2D}{v}.$$

For a typical growth situation

$$D = 10^{-14}\ \mathrm{cm^2/s} = 10^{-18}\ \mathrm{m^2/s}$$

$$v = 10\ \mu\mathrm{m/h} = 2.8 \times 10^{-9}\ \mathrm{m/s}$$

which leads to

$$h \leq 7.2\ \text{Å}.$$

We cannot possibly have a uniform mesh distance of 7.2 Å over several microns. This illustrates why we must have a nonuniform mesh which must be fine close to the interface where the described phenomenon occurs. Since the interface moves, this region of fine mesh must also move. This is accomplished with the dynamical mesh adjustment which will be demonstrated in the examples below.

4. EXAMPLES AND COMMENTS

As an example of the calculation of the redistribution we show some results from the two 1000 K GaAs growth simulations in [1]. The first (Case I) represents growth of a lightly doped n-type layer on a heavily doped n-type substrate. The whole picture involves four impurities, two in the substrate and two incorporated in the epitaxial layer, in each case a donor and an acceptor. Here we will only show the redistribution of donors and acceptors incorporated in the layer (Figs. 5 and 6) and the electric field (Fig. 4). All quantities are given in a normalized

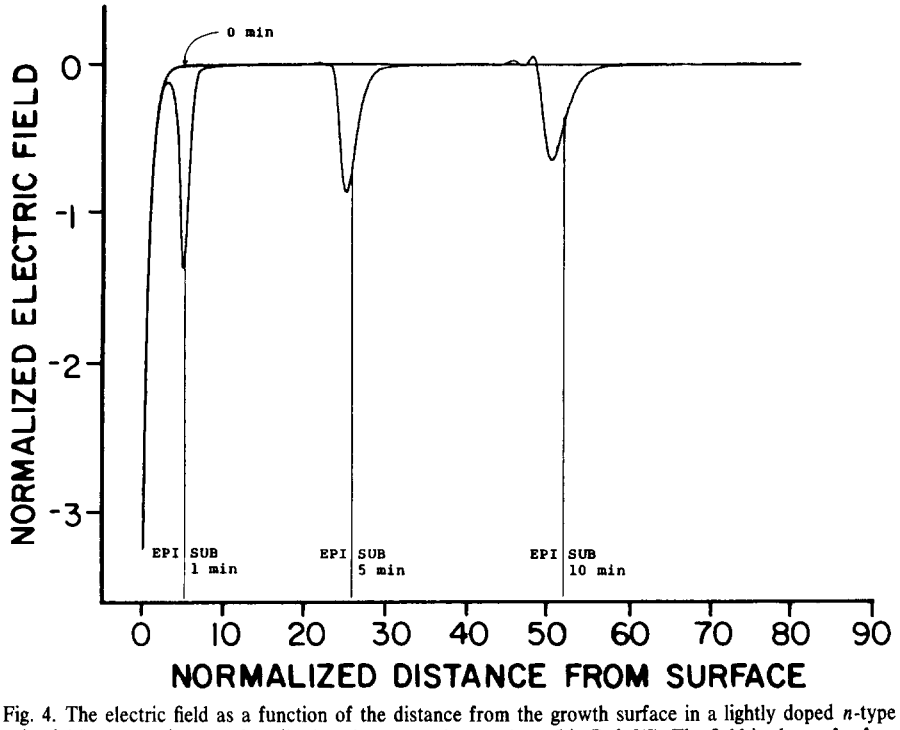

Fig. 4. The electric field as a function of the distance from the growth surface in a lightly doped n-type epitaxial layer growing on a heavily doped n-type substrate (case I in Ref. [1]). The field is shown for four instances: $t = 0, 1, 5$ and 10 minutes.

form [1]. For the donors we have: $z = +1$, $D = 0.057$ and $B = 1.27$; for the acceptors: $z = -1$, $D = 0.57$ and $B = 1.27$. The growth velocity is 5.15 in normalized units.

The electric field at the surface is fixed by negative surface states. At the interface, a negative electric field peak develops, mainly because of the substrate donor which is the majority dopant. The peak gets smoothed by the redistribution. The ripples that occur are due to ripples originating in the substrate donor concentration as discussed in the last section. These ripples are similar to those of the incorporated donors in Fig. 5. In this figure we also see how the electric field causes a pileup of these donors to the left of the interface; the acceptors (Fig. 6) exhibit a dip there and pile up to the right of the interface. Note that the ripples are smaller for the acceptors as one would expect since the diffusion constant is an order of magnitude larger.

In Fig. 7 the number of time increments between each time epoch is plotted for a similar growth simulation (Case II in [1]). The automatic adjustment of time increment was discussed in Section 3.1 and we see that at first, when the initially sharp interface separates from the growth surface, the time increment has to be very small. Once the growth continues and the interface gets smoother, the time increment can be increased. A factor that limits this increase is the approximation (7) made for the electric field during the time increment. The change in the field and potential over the time increment is therefore small and very few iterations, typically 1–2, are necessary in the field calculation. Point 2 in Section 3.2 is borne out.

In Fig. 7 we have also plotted the total number of mesh points for the four impurities and the field/potential for the same simulation, as a function of time. This illustrates how the dynamical mesh adjustment works. We see that in all cases the number stabilizes after a moderate or small increase. This indicates that even though the interfacial region of high mesh point density moves to the right, the dynamical mesh modification sees to it that no region to the left of the interface is left with unnecessarily small mesh distances.

It is important to choose the initial mesh with care. Initially the redistribution and the field will be largest at the surface (unless there is some unusual initial substrate doping). It is thus necessary to have a finer mesh at the surface ($x = 0$). The h_i were picked so that

$$h_{i+1} = 2h_i; \quad i = 1, 2, \ldots, i_m - 1.$$

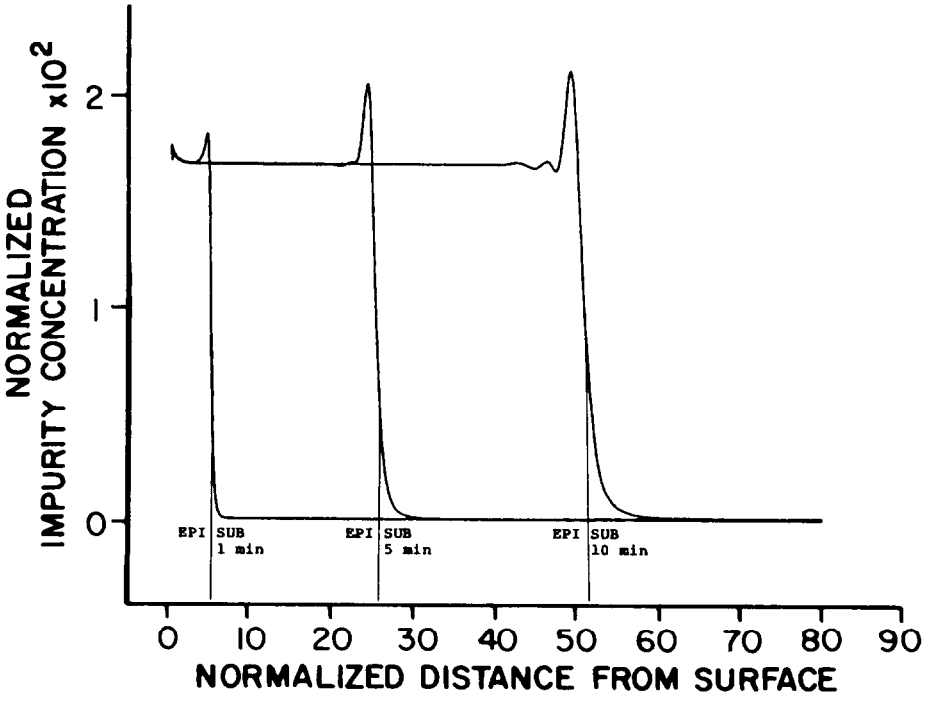

Fig. 5. The concentration of donors incorporated in the epitaxial layer as a function of the distance from the growth surface. Same growth case as described in the text and in the caption to Fig. 4. The concentration is shown for three instances: $t = 1, 5$ and 10 minutes.

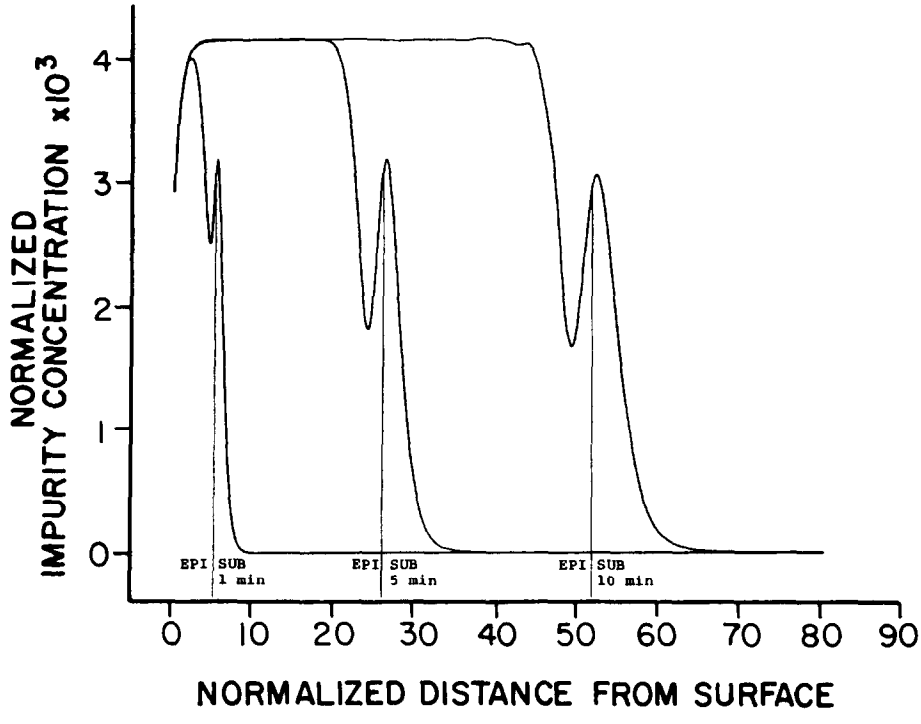

Fig. 6. The concentration of acceptors incorporated in the epitaxial layer as a function of the distance from the growth surface. Same growth case as in Figs. 4 and 5. The concentration is shown for three instances: $t = 1, 5$ and 10 minutes.

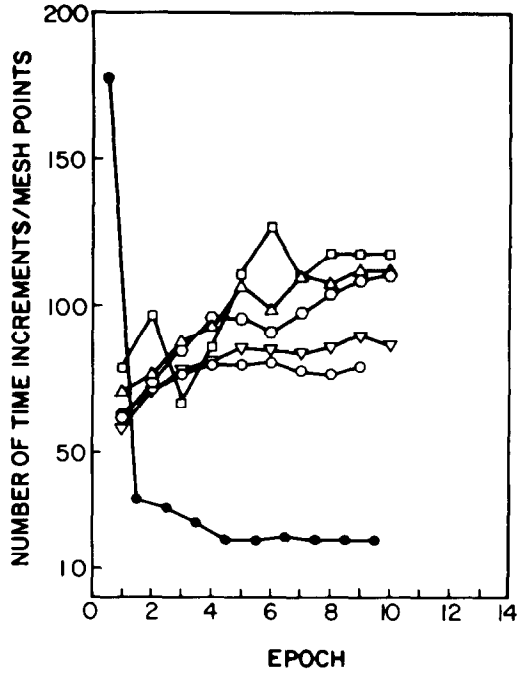

Fig. 7. Number of time increments (●) between each epoch and the time evolution of the number of mesh points (open symbols) for the growth case II in [1]. ○, slow substrate donor (No. 1 in [1]; □, fast substrate donor (No. 2 in [1]); △, slow incorporated donor (No. 3 in [1]); ▽, incorporated acceptor (No. 4 in [1]); ○, electric field/potential.

For the field, h_1 was picked on the order of or smaller than a Debye length. For the redistribution, h_1 was picked small enough so that the initial redistribution at the surface during the initial time increment Δt_i is not negligible. It is suggested that a user make a trial run for a very short time to see whether the redistribution seems to start properly. If it does not, a modification of h_1 and Δt_i will probably be helpful; at least this has been the case so far in the runs that we have performed.

Acknowledgements—The author is grateful to Dr. M. W. Muller for valuable advice throughout the course of this work and to Dr. D. L. Elliott for enlightening discussions of numerical techniques. The work was supported by ONR Contract N00014-79-C-0840.

REFERENCES

1. H. Rohdin, M. W. Muller and C. M. Wolfe, Impurity redistribution during epitaxial growth. *J. Electron. Mater.* **11**, 517 (1982).
2. H. C. Casey, Jr., Diffusion in the III–V compound semiconductors. *Atomic Diffusion in Semiconductors* (Edited by D. Shaw). Plenum Press, New York (1973).
3. L. A. Girifalco, *Atomic Migration in Crystals*, p. 99. Blaisdell, New York (1964).
4. W. Shockley, *Electrons and Holes in Semiconductors*, p. 306. Van Nostrand, New York (1950).
5. L. Fox and D. F. Mayers, *Computing Methods for Scientists and Engineers*, p. 207. Clarendon Press, Oxford (1968).
6. G. D. Smith, *Numerical Solution of Partial Differential Equations*. Oxford University Press, New York (1965).