

## NUMERICAL STUDY OF HYPERBOLIC EQUATIONS WITH INTEGRAL CONSTRAINTS ARISING IN SEMICONDUCTOR THEORY\*

A. CARPIO<sup>†</sup>, P. HERNANDO<sup>‡</sup>, AND M. KINDELAN<sup>‡</sup>

**Abstract.** An efficient numerical scheme is described for the solution of certain types of nonlinear hyperbolic equations with an integral constraint which are used to model the Gunn effect in semiconductors with impurity capture. We analyze the stability and convergence properties of the scheme and present the results of numerical simulations. Depending on the value of the parameters defining the problem, a great variety of solutions are obtained, including periodic recycling of solitary waves and chaotic regimes.

**Key words.** Gunn effect, oscillations, chaos, numerical simulation, stability, convergence

**AMS subject classifications.** 35L70, 39A11, 65M06, 65M12, 65C20

**PII.** S0036142999360287

**1. Introduction.** Pattern formation and oscillatory phenomena involving recycling and motion of charge dipole waves have often been observed in semiconductors displaying nonlinear electrical conduction. These phenomena were first observed by Gunn [9] in bulk n-type GaAs for which the electron velocity is an N-shaped function of the electric field. When planar contacts are attached to an n-GaAs sample and an appropriate dc voltage bias is kept between them, there appear self-sustained oscillations of the current with frequencies in the microwave range. These oscillations are accompanied by periodic recycling and motion of charge density dipole waves (solitary waves of the electric field).

After Gunn's discovery, many materials were shown to have similar current self-oscillations, although the physical mechanisms causing the oscillations were often very different. However, all these materials had N-shaped current-voltage characteristics, which is an essential feature of the Gunn effect (see [2]). N-shaped current-voltage characteristics can appear due to impurity capture processes; such is the case in p-Ge [19] and many other semiconductors [16]. Precise measurements of the Gunn effect in p-Ge are reported in [14] and [18]. Experiments show that intermittency and spatiotemporal chaos are observed in addition to the usual time periodic oscillations [14].

The Gunn effect and other instabilities of the current have been studied by analyzing models of charge transport in semiconductors (see [16], [17]). For bulk semiconductor devices, charge transport may be described by the semiclassical Boltzmann equation or hydrodynamic or drift-diffusion models. Each class of models describes phenomena occurring at a different length and time scales [17]. For sufficiently large devices, hydrodynamic or drift-diffusion descriptions are appropriate and involve less computational cost.

---

\*Received by the editors August 16, 1999; accepted for publication (in revised form) October 9, 2000; published electronically February 28, 2001. This research was supported by DGES grants PB96-0663 and PB97-0088.

<http://www.siam.org/journals/sinum/39-1/36028.html>

<sup>†</sup>Departamento de Matemática Aplicada, Universidad Complutense de Madrid, 28040 Madrid, Spain (carpio@sunma4.mat.ucm.es).

<sup>‡</sup>Departamento de Matemáticas, Escuela Politécnica Superior, Universidad Carlos III de Madrid, Avenida de la Universidad 30, 28911 Leganés, Spain (pedroj@math.uc3m.es, kinde@ing.uc3m.es).

Drift-diffusion models hold for time and length scales in which the moment and energy distributions have relaxed. For the materials used in the experiments in [14] and [18], relevant time scales are those corresponding to impurity trapping. Such scales are so slow (milliseconds) that the displacement current is negligible, which means that the charge density distribution is quasi-stationary. Faster distributions (momentum and energy) have relaxed to a quasi-stationary state much earlier. Thus while one can argue in favor of the hydrodynamic model for short Gunn diodes (less than 5 microns long for a  $10^{15} \text{ cm}^{-3}$  doping, say), it is not realistic to do the same for a p-Ge device with a typical length of 1 to 14 mm and kHz frequency. In the case of self-oscillations in semiconductors with impurity capture, a drift-diffusion description is appropriate.

Some numerical studies are available for certain types of drift-diffusion models; see, for instance, [7], [12], [6]. Numerical studies of hydrodynamic [8], [13] and energy models [11] have also been carried out (mainly for GaAs). While most numerical analysis for semiconductors deals with Boltzmann, hydrodynamic, or drift-diffusion models, specific analysis for trap-dominated drift-diffusion models is lacking.

In p-Ge, carriers are holes and there are ionizable shallow acceptor impurities which act as capture centers and do not move. Equations describing charge transport consist of

- (i) Poisson equation (in one dimension) relating densities of holes and ionized impurities to the divergence of the electric field;
- (ii) a rate equation for the impurity density, including the processes of impact ionization by holes of neutral impurities and impurity-hole (capture) recombination.
- (iii) the charge continuity equation in which the flux of holes contains drift and diffusion terms. This equation is equivalent to Ampère's current balance law. In it, the total current density is equal to the sum of the drift-diffusion current due to holes and Maxwell's displacement current;
- (iv) appropriate bias, initial and boundary conditions of the contact regions of the semiconductor.

A model consisting of these equations and conditions is still much too complicated to analyze. It is possible to derive a simpler reduced model by singular perturbation methods. These methods exploit the fact that characteristic time scales of the ionization and capture processes are much longer than the drift time of free holes. As the ratio of these times goes to infinity, the following (nondimensional) model is derived [3], [19]:

$$(1.1) \quad \frac{\partial^2 E}{\partial x \partial t} + \mathcal{A}(E, J) \frac{\partial E}{\partial t} + \mathcal{B}(E, J) \frac{\partial E}{\partial x} + \mathcal{C}(E, J) \frac{dJ}{dt} + \mathcal{D}(E, J) = 0,$$

$$x \in (0, L), t > 0,$$

$$(1.2) \quad E(x, 0) = E_0(x), \quad x \in (0, L),$$

$$(1.3) \quad E(0, t) = J(t) \rho, \quad t \geq 0,$$

$$(1.4) \quad \int_0^L E(x, t) dx = \Phi \in R^+, \quad t \geq 0,$$

where  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ , and  $\mathcal{D}$  are known functions, depending on the semiconductor [15]. Typical plots for  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ , and  $\mathcal{D}$  are shown in Figure 1.1. The integral constraint (1.4) is the bias condition relating the applied voltage  $\Phi$  to the electric field  $E(x, t)$

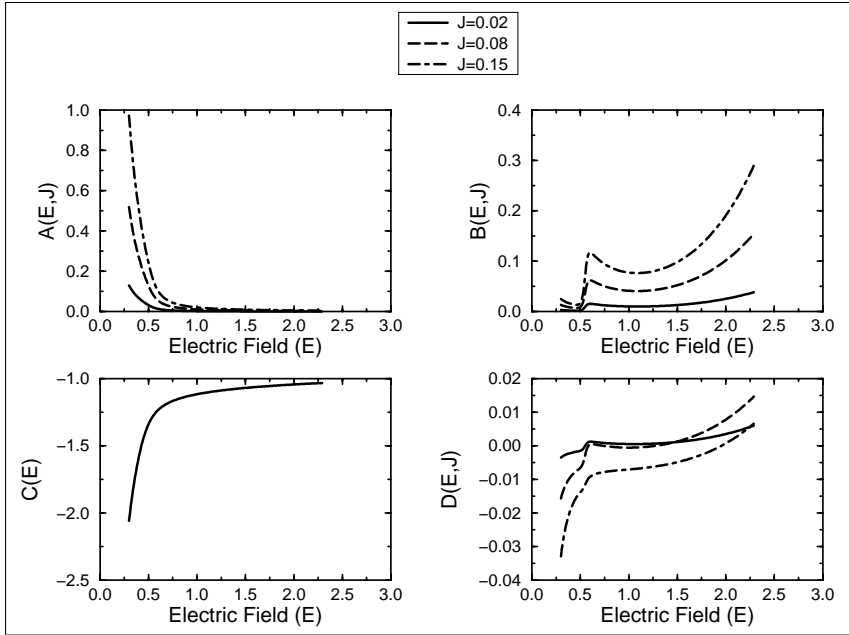


FIG. 1.1. Nondimensional functions  $\mathcal{A}(E, J)$ ,  $\mathcal{B}(E, J)$ ,  $\mathcal{C}(E)$ , and  $\mathcal{D}(E, J)$  used in the modeling of Gunn effect in  $p$ -Ge. The figure shows the values of  $\mathcal{A}(E, J)$ ,  $\mathcal{B}(E, J)$ , and  $\mathcal{D}(E, J)$  for three values of the current density  $J$ , 0.02 (low), 0.08 (medium), and 0.15 (high), which cover the range of values observed in the simulations.

throughout the sample. The scalars  $\rho$  (resistivity),  $\Phi$  (voltage), and  $L$  (length of the sample) are input parameters. Typically,  $\rho \in [7, 11]$ ,  $\Phi \in [500, 2000]$ , and  $L \in [1000, 5000]$ .

Solving these equations with appropriate initial data provides the time evolution of the current density,  $J(t)$ , and of the electric field distribution,  $E(x, t)$ . This reduced model has shown to be very useful to understand the generation and dynamics of oscillations in Germanium devices [1], [3], [5]. These types of systems exhibit a rich variety of solutions, including periodic and chaotic ones. An analytical and asymptotic study of nonlinear dynamics in this model has been carried out in [4].

In this paper we analyze an efficient numerical scheme to solve this type of problem with integral constraint. Some numerical simulations of these reduced equations have been presented in [1], [3] using progressive time differentiation. In our scheme, we use regressive time differentiation. We carry out a numerical analysis of stability and convergence and give sufficient conditions on the nonlinear coefficients for the scheme to work efficiently.

The mathematical structure of system (1.1)–(1.4) leads naturally to implicit schemes. The reasons for this are as follows:

- The higher order hyperbolic operator in (1.1) is written in characteristic form, so that space and time are characteristic variables. Thus initial and boundary data are given on characteristic curves  $t = 0$  and  $x = 0$ . Notice that the higher order term is mixed; it involves time and space derivatives.
- The boundary data at  $x = 0$  ( $E(0, t)$ ) is not known a priori. It is determined using the integral constraint (1.4), which couples all the values of  $E(x, t)$ ,  $x \in [0, L]$  at time  $t$ .

- The term  $\partial J/\partial t = (1/\rho) \partial E(0, t)/\partial t$  appearing in the differential equation (1.1) makes it *nonlocal*. The domain dependence properties and finite time propagation of hyperbolic operators are lost.

Under these conditions, when one discretizes (1.1)–(1.4) by using either progressive or regressive finite difference approximations for the time derivatives, a scheme is obtained where the new time level,  $E_i^{j+1}$ , is computed from the old level  $E_i^j$  by solving a system of equations. Although the resulting schemes are implicit, the peculiar structure of the matrices allows us to implement a fast resolution procedure.

We use our scheme to investigate oscillations and chaotic regimes in Germanium. Although the problem is one dimensional, it is very computationally intensive, and therefore efficient numerical methods are required. On the one hand, the very large nondimensional length of the samples of interest ( $L \sim 3500$ ) plus the accuracy desired leads to a large number of discretization points and, on the other hand, the interest in analyzing and characterizing chaotic solutions requires integration of the equations for a very large number of time steps. We also show some of the patterns observed for the current.

The paper is organized as follows. Section 2 describes the numerical scheme, the resulting equations, and the implementation of the method. Section 3 describes the analysis of the stability and convergence properties of the scheme, first for the linearized problem and then for the fully nonlinear problem. Section 4 presents the results of some specific numerical experiments and, finally, in several appendices we explain in more detail some technical results needed in the analysis of section 3.

**2. Description of the numerical scheme.** We consider the nonlinear hyperbolic problem (1.1)–(1.4). From Figure 1.1 we see that

$$(2.1) \quad \mathcal{A}, \mathcal{B} \geq 0, \quad \mathcal{C} \leq 0.$$

In what follows, we assume that (2.1) holds.

Notice that the characteristic curves of the second order operator are  $x = \text{constant}$ ,  $t = \text{constant}$ , so that for a point  $(x_0, t_0)$  the domain of dependence is  $0 \leq x_0, 0 \leq t \leq t_0$ . The characteristic curves of the first order operator solve  $x'(t) = \frac{\mathcal{A}}{\mathcal{B}}x(t)$ , so that they are contained in the region of dependence of the second order operator. In view of this, the problem is well posed given an initial condition and a boundary condition at  $x = 0$ . The integral condition on  $E$  may be satisfied adjusting the boundary condition through the extra unknown  $J(t)$ .

To solve (1.1) with boundary condition (1.3) and integral constraint (1.4) we use a uniform spatial grid with spacing  $h = L/N$  and time step  $\Delta t = k$  and use condition (1.3) to eliminate  $J$  in favor of  $E_0$ . Thus the unknowns of the problem are  $E_i^j \approx E(ih, jk)$ ,  $i = 0, 1, \dots, N$ , with the subscript  $i$  referring to space and the superscript  $j$  to time.

To approximate the time derivatives at time  $j + 1$  we use

$$E_t \quad \longrightarrow \quad \frac{\partial E_i^{j+1}}{\partial t} = \frac{2}{k}(E_i^{j+1} - E_i^j) - \frac{\partial E_i^j}{\partial t} + O(k^2),$$

which is equivalent to a second order, implicit, Runge–Kutta method (trapezoidal rule). The same approximation is used for  $J_t$ :

$$J_t \quad \longrightarrow \quad J^{j+1} = \frac{E_0^{j+1}}{\rho} \quad ; \quad \frac{dJ^{j+1}}{dt} = \frac{1}{\rho} \frac{\partial E_0^{j+1}}{\partial t} = \frac{2}{\rho k}(E_0^{j+1} - E_0^j) - \frac{1}{\rho} \frac{\partial E_0^j}{\partial t} + O(k^2).$$

We choose a second order backward difference approximation for the space derivatives,  $E_x$ , in the interior points ( $i = 2, 3, \dots, N$ ) because the use of central differences leads to instabilities. Thus

$$E_x \longrightarrow \frac{\partial E_i^{j+1}}{\partial x} = \frac{3E_i^{j+1} - 4E_{i-1}^{j+1} + E_{i-2}^{j+1}}{2h} + O(h^2).$$

The second derivative term,  $E_{xt}$ , combines the space and time derivative approximation just described. Thus

$$\begin{aligned} E_{xt} \longrightarrow \frac{\partial^2 E_i^{j+1}}{\partial x \partial t} &= \frac{1}{hk} \left( 3E_i^{j+1} - 4E_{i-1}^{j+1} + E_{i-2}^{j+1} - 3E_i^j + 4E_{i-1}^j - E_{i-2}^j \right) \\ &+ \frac{1}{2h} \left( -3 \frac{\partial E_i^j}{\partial t} + 4 \frac{\partial E_{i-1}^j}{\partial t} - \frac{\partial E_{i-2}^j}{\partial t} \right) + O(h^2, k^2). \end{aligned}$$

At  $i = 1$  we use a first order backward difference formula,

$$\frac{\partial E_1^{j+1}}{\partial x} = \frac{E_1^{j+1} - E_0^{j+1}}{h} + O(h),$$

$$\frac{\partial^2 E_1^{j+1}}{\partial x \partial t} = \frac{2}{hk} \left( E_1^{j+1} - E_0^{j+1} - E_1^j + E_0^j \right) - \frac{1}{h} \left( \frac{\partial E_1^j}{\partial t} - \frac{\partial E_0^j}{\partial t} \right) + O(h, k^2).$$

The use of this formula leads to a triangular matrix (see section 3) relating the values of the unknowns at time  $j + 1$  to their values at time  $j$ . The resulting system of equations can be solved very efficiently. A more accurate second order formula can also be used with slight changes in the solution procedure.

Introducing these formulas in (1.1), denoting  $\dot{\mathbf{E}}_i = \frac{\partial E_i^j}{\partial t}$  and multiplying the resulting finite difference equation by  $2\rho h k$  produces, for  $i = 2, \dots, N$ ,

$$\begin{aligned} (4Ch)E_0^{j+1} + \rho(6 + 4Ah + 3Bk)E_i^{j+1} - 4\rho(2 + Bk)E_{i-1}^{j+1} + \rho(2 + Bk)E_{i-2}^{j+1} \\ = (4Ch)E_0^j + 2\rho(3 + 2Ah)E_i^j - (8\rho)E_{i-1}^j + (2\rho)E_{i-2}^j + (2Chk)\dot{\mathbf{E}}_0 \\ (2.2) \quad + \rho(2Ahk + 3k)\dot{\mathbf{E}}_i - (4\rho k)\dot{\mathbf{E}}_{i-1} + (\rho k)\dot{\mathbf{E}}_{i-2} - 2\rho Dhk. \end{aligned}$$

Similarly, for  $i = 1$ ,

$$\begin{aligned} 2\rho(2 + 2Ah + Bk)E_1^{j+1} + (-4\rho - 2\rho Bk + 4Ch)E_0^{j+1} \\ = (-4\rho + 4Ch)E_0^j + 4\rho(1 + Ah)E_1^j + (-2\rho k + 2Chk)\dot{\mathbf{E}}_0 \\ (2.3) \quad = 2\rho(k + Ahk)\dot{\mathbf{E}}_1 - 2\rho Dhk. \end{aligned}$$

The pattern of grid points involved in the numerical scheme is shown in Figure 2.1. Notice that the discrete dependence domain contains the continuous dependence domain.

Since there are  $N + 1$  unknowns, an additional equation is needed which is obtained from (1.4):

$$(2.4) \quad E_0 + 2E_1 + \dots + 2E_{N-1} + E_N = \frac{2\Phi}{h}.$$

Equations (2.2)–(2.4) constitute a system of  $N + 1$  nonlinear algebraic equations which have to be solved at each time step to obtain the values of  $E_0, \dots, E_N$ .

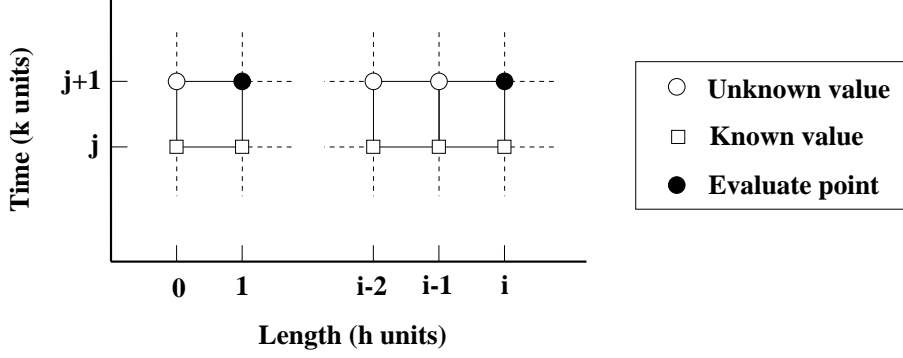


FIG. 2.1. Stencil of the finite difference operator used in the numerical scheme.

As an initial condition to start the scheme, we use  $E_n^0 = (\Phi/L)$ ,  $\dot{E}_n^0 = 0$ .

Notice that, since the characteristics are horizontal and vertical lines, we will not have the usual Courant–Friedrichs–Lewy stability and convergence condition  $(k/h) \leq (1/c)$ ,  $c$  being the slope of the characteristics. Instead, we will require  $h, k$  to be small enough. We will further assume that  $kN$  is bounded. Let us recall that  $hN = L$  is fixed.

To analyze the properties of the method it is convenient to rewrite it as

$$(2.5) \quad M^{j+1}E^{j+1} = \mathcal{M}^{j+1} \left( E^j + \frac{k}{2}\dot{E}^j \right) + F^{j+1},$$

where

$$E^j = \begin{pmatrix} E_1^j \\ \vdots \\ E_N^j \\ E_0^j \end{pmatrix}, \quad \dot{E}^j = \begin{pmatrix} \dot{E}_1^j \\ \vdots \\ \dot{E}_N^j \\ \dot{E}_0^j \end{pmatrix}, \quad F^j = \begin{pmatrix} -2\rho hk \mathcal{D}_1^j \\ \vdots \\ -2\rho hk \mathcal{D}_N^j \\ \frac{2\Phi}{h} \end{pmatrix},$$

$$M^{j+1} = \begin{pmatrix} T^{j+1} & v^{j+1} \\ u & 1 \end{pmatrix}, \quad \mathcal{M}^{j+1} = \begin{pmatrix} T^{j+1} & \nu^{j+1} \\ 0 & 0 \end{pmatrix},$$

$$T = \begin{pmatrix} 4\rho & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ -8\rho & 6\rho & 0 & 0 & \dots & 0 & 0 & 0 \\ 2\rho & -8\rho & 6\rho & 0 & \dots & 0 & 0 & 0 \\ 0 & 2\rho & -8\rho & 6\rho & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 2\rho & -8\rho & 6\rho \end{pmatrix},$$

$$T^{j+1} = T + 4\rho h \operatorname{diag}(\mathcal{A}_1^{j+1}, \dots, \mathcal{A}_N^{j+1}) + \frac{k}{2} \operatorname{diag}(\mathcal{B}_1^{j+1}, \dots, \mathcal{B}_N^{j+1})T,$$

$$\mathcal{T}^{j+1} = T^{j+1} - \frac{k}{2} \operatorname{diag}(\mathcal{B}_1^{j+1}, \dots, \mathcal{B}_N^{j+1})T,$$

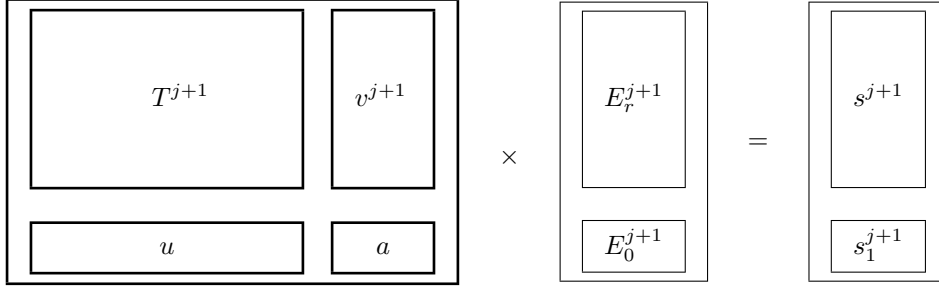


FIG. 2.2. Structure of the system.

$$v^{j+1} = \begin{pmatrix} 4\mathcal{C}_1^{j+1}h - 2\rho\mathcal{B}_1^{j+1}k - 4\rho \\ 4\mathcal{C}_2^{j+1}h + \rho\mathcal{B}_2^{j+1}k + 2\rho \\ 4\mathcal{C}_3^{j+1}h \\ \dots \\ 4\mathcal{C}_N^{j+1}h \end{pmatrix}, \quad \nu^{j+1} = v^{j+1} - \begin{pmatrix} -2\rho\mathcal{B}_1^{j+1}k \\ \rho\mathcal{B}_2^{j+1}k \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad u = \begin{pmatrix} 2 \\ 2 \\ \dots \\ 2 \\ 1 \end{pmatrix}^T.$$

For short, we denote  $D_A = \text{diag}(\mathcal{A}_1^{j+1}, \dots, \mathcal{A}_N^{j+1})$  and  $D_B = \text{diag}(\mathcal{B}_1^{j+1}, \dots, \mathcal{B}_N^{j+1})$ . We also set  $b^{j+1} = (-2\rho\mathcal{B}_1^{j+1}k, \rho\mathcal{B}_2^{j+1}k, 0, \dots, 0)$ . To solve (2.5) it is convenient to take advantage of the triangular structure of matrix  $T^{j+1}$ . Following the method proposed in [1], if the system is partitioned as in Figure 2.2, its solution can be obtained by solving two linear systems defined by the same matrix,  $T^{j+1}$ , with different right-hand sides. In fact,

$$(2.6) \quad T^{j+1}y^{j+1} = s^{j+1} \quad \rightarrow \quad \boxed{y^{j+1}},$$

$$(2.7) \quad T^{j+1}z^{j+1} = v^{j+1} \quad \rightarrow \quad \boxed{z^{j+1}},$$

and since

$$(2.8) \quad aE_0^{j+1} + uE_r^{j+1} = s_1^{j+1},$$

$$(2.9) \quad T^{j+1}E_r^{j+1} + v^{j+1}E_0^{j+1} = s^{j+1},$$

these two equations can be used to solve for  $E_r^{j+1}$  and  $E_0^{j+1}$  in terms of  $y^{j+1}$ ,  $z^{j+1}$ . From (2.9) and (2.7)

$$T^{j+1}E_r^{j+1} + T^{j+1}z^{j+1}E_0^{j+1} = T^{j+1}y^{j+1} \quad \Rightarrow \quad E_r^{j+1} = y^{j+1} - z^{j+1}E_0^{j+1},$$

and introducing this value in (2.8),

$$(2.10) \quad E_0^{j+1} = \frac{s_1^{j+1} - uy^{j+1}}{a - uz^{j+1}},$$

$$(2.11) \quad E_r^{j+1} = y^{j+1} - z^{j+1}E_0^{j+1}.$$

In the present case,  $a = 1$ ,  $s_1^{j+1} = 2\Phi/h$ , and

$$s^{j+1} = T^{j+1} \left( E_r^j + \frac{k}{2}\dot{E}_r^j \right) + \nu^{j+1} \left( E_0^j + \frac{k}{2}\dot{E}_0^j \right) + F_r^{j+1}.$$

The subindex  $r$  means that we are removing the  $N + 1$ th coordinate from the vectors, i.e.,  $E_r^j = (E_1^j, \dots, E_N^j)$ .

Let us rewrite these expressions in order to make explicit which is the matrix relating the  $j + 1$ th time step to the  $j$ th step.

We introduce the following matrices:

$$Z^{j+1} = \frac{z^{j+1}u^t}{1 - u \cdot z^{j+1}}, N^{j+1} = I + Z^{j+1}, Q^{j+1} = \frac{k}{2}(T^{j+1})^{-1}D_B T.$$

Then we may express  $E^{j+1}$  in terms of  $E^j$  as follows:

$$\begin{aligned} E_0^{j+1} &= \frac{1}{1 - u \cdot z^{j+1}}(u \cdot (T^{j+1})^{-1}b^{j+1}) \left( E_0^j + \frac{k}{2}\dot{E}_0^j \right) \\ &\quad + \frac{1}{1 - u \cdot z^{j+1}}(u \cdot Q^{j+1}) \left( E_r^j + \frac{k}{2}\dot{E}_r^j \right) \\ (2.12) \quad &\quad + \frac{u \cdot (T^{j+1})^{-1}F_r^{j+1}}{1 - u \cdot z^{j+1}} + E_0^j + \frac{k}{2}\dot{E}_0^j, \end{aligned}$$

$$\begin{aligned} E_r^{j+1} &= (I - N^{j+1}Q^{j+1}) \left( E_r^j + \frac{k}{2}\dot{E}_r^j \right) + N^{j+1}(T^{j+1})^{-1}F_r^{j+1} \\ (2.13) \quad &\quad - N^{j+1}(T^{j+1})^{-1}b^{j+1} \left( E_0^j + \frac{k}{2}\dot{E}_0^j \right). \end{aligned}$$

Notice that  $u \cdot z^{j+1}$  stands for the scalar product of two vectors, whereas  $z^{j+1}u^t$  denotes the matrix obtained as a product of a vector  $N \times 1$  and a vector  $1 \times N$ .

### 3. Analysis of the scheme.

**3.1. The iteration matrix.** In order to analyze the stability and convergence properties of the scheme, it is convenient to express the level  $j + 1$  in terms of the level 0.

Since the time derivatives are related as follows,

$$\dot{E}_i^j = \frac{2}{k}(E_i^j - E_i^{j-1}) - \dot{E}_i^{j-1} = \frac{2}{k}[E_i^j - 2E_i^{j-1} + 2E_i^{j-2} - \dots + (-1)^j E_i^0] + (-1)^j \dot{E}_i^0$$

and  $\dot{E}_i^0 = 0$ , we may rewrite (2.12) and (2.13) as

$$\begin{aligned} E_r^{j+1} &= (I - N^{j+1}Q^{j+1})[2E_r^j - 2E_r^{j-1} + 2E_r^{j-2} - \dots + (-1)^j E_r^0] \\ &\quad - N^{j+1}(T^{j+1})^{-1}b^{j+1}[2E_0^j - 2E_0^{j-1} + 2E_0^{j-2} - \dots + (-1)^j E_0^0] \\ (3.1) \quad &\quad + N^{j+1}(T^{j+1})^{-1}F_r^{j+1}, \end{aligned}$$

$$\begin{aligned} E_0^{j+1} &= \frac{1}{1 - u \cdot z^{j+1}}[u \cdot (T^{j+1})^{-1}b^{j+1}][2E_0^j - 2E_0^{j-1} + 2E_0^{j-2} - \dots + (-1)^j E_0^0] \\ &\quad + \frac{1}{1 - u \cdot z^{j+1}}(u \cdot Q^{j+1})[2E_r^j - 2E_r^{j-1} + 2E_r^{j-2} - \dots + (-1)^j E_r^0] \\ (3.2) \quad &\quad + \frac{u \cdot (T^{j+1})^{-1}F_r^{j+1}}{1 - u \cdot z^{j+1}} + [2E_0^j - 2E_0^{j-1} + 2E_0^{j-2} - \dots + (-1)^j E_0^0] \end{aligned}$$



The vector  $E^{j+1}$  is given by

$$E^{j+1} = R^{j+1} [2E^j - 2E^{j-1} + \dots + (-1)^j E^0] + K^{j+1}$$

with

$$R^{j+1} = \begin{pmatrix} I - N^{j+1}Q^{j+1} & -N^{j+1}(T^{j+1})^{-1}b^{j+1} \\ \frac{1}{1-u \cdot z^{j+1}} (u \cdot Q^{j+1}) & 1 + \frac{1}{1-u \cdot z^{j+1}} [u \cdot (T^{j+1})^{-1}b^{j+1}] \end{pmatrix},$$

$$K^{j+1} = \begin{pmatrix} N^{j+1}(T^{j+1})^{-1}F_r^{j+1} \\ \frac{u \cdot (T^{j+1})^{-1}F_r^{j+1}}{1-u \cdot z^{j+1}} \end{pmatrix}.$$

Therefore,

$$(3.3) \quad \begin{aligned} E^{j+1} &= R^{j+1}(2R^j - I) \dots (2R^1 - I)E^0 + 2R^{j+1}K^j + K^{j+1} \\ &\quad + 2 \sum_{l=0}^{j-1} R^{j+1}(2R^j - I) \dots (2R^{j-l} - I)K^l. \end{aligned}$$

To analyze the convergence and stability properties of this scheme, a detailed study of all the matrices involved in it is needed. This is done in Appendix A.

**3.2. The linear problem.** We consider the linear problem in which the functions  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  are only functions of  $x$  and  $t$ . We also assume that those functions are bounded with bounded derivatives and such that  $\mathcal{A}(x, t) \geq 0$ ,  $\mathcal{B}(x, t) \geq 0$ ,  $\mathcal{C}(x, t) \leq 0$ . Let  $\Lambda > 0$  be such that

$$\max_{[0, L] \times [0, \infty)} \{|\mathcal{A}|, |\mathcal{B}|, |\mathcal{C}|, |\mathcal{D}|, |\mathcal{B}'|\} \leq \Lambda.$$

The discretization of this problem leads to system (2.5), where  $\mathcal{A}_i^j = \mathcal{A}(ih, kj)$ ,  $\mathcal{B}_i^j = \mathcal{B}(ih, kj)$ ,  $\mathcal{C}_i^j = \mathcal{C}(ih, kj)$ ,  $\mathcal{D}_i^j = \mathcal{D}(ih, kj)$  are known sequences. Therefore,  $E^{j+1}$  is computed explicitly from  $E^j$  by means of (2.13), (2.12) or from  $E^0$  by means of (3.3).

Let us study the stability properties of scheme (3.3) when the coefficients do not depend on  $E_i^j$ . From (3.3) we get

$$\begin{aligned} \|E^{j+1}\|_2 &\leq \|R^{j+1}\|_2 \|2R^j - I\|_2 \dots \|2R^1 - I\|_2 \|E_0\|_2 + \|K^{j+1}\|_2 \\ &\quad + 2 \|R^{j+1}\|_2 \|K^j\|_2 + 2 \sum_{l=0}^{j-1} \|R^{j+1}\|_2 \|2R^j - I\|_2 \dots \|2R^{j-l} - I\|_2 \|K^l\|_2. \end{aligned}$$

In Appendix A we show that for  $h, k$  small enough, with  $kN$  bounded, we have  $\|R^{j+1}\|_2 < \gamma$ ,  $\|2R^{j-l} - I\|_2 < \gamma$ ,  $0 < l < j$  with  $0 < \gamma < 1$  and  $\|K^j\|_2 \leq G \|\mathcal{D}\|_\infty$  with  $G > 0$  for all  $j$ . Then (cf. [10])

$$(3.4) \quad \|E^{j+1}\|_2 \leq \gamma \|E_0\|_2 + \frac{1 + \gamma}{1 - \gamma} G \|\mathcal{D}\|_\infty.$$

On the other hand, the discretization of (1.1) given by (2.5) is consistent. Let  $E(x, t)$  be the solution of (1.1), which is bounded, with bounded derivatives. We define  $\mathcal{E}_i^j = E(ih, jk)$ . Then  $\mathcal{E}_i^j$  satisfies (2.5) with an error

$$(3.5) \quad e^{j+1} = M^{j+1} \mathcal{E}^{j+1} - \mathcal{M}^{j+1} \left( \mathcal{E}^j + \frac{k}{2} \dot{\mathcal{E}}^j \right) - F^{j+1} = hkO(h^p + h^q), \quad p, q > 1.$$

More precisely, since we made approximations of order 2 for the derivatives, except for the spatial derivatives at  $i = 1$  where we chose order 1, the vector  $e^{j+1}$  has all its components of order  $hk(h^2 + k^2)$  except one of order  $hk(h + k^2)$ .

Using the first order approximation at  $i = 1$  results in a triangular structure which simplifies the analysis in Appendix A, which leads to the stability bound (3.4). It is easy to show that using the second order approximation at  $i = 1$  results in a small perturbation to the previous result, so that (3.4) holds again and in (3.5)  $p = q = 2$ .

Now, if  $E_i^j$  is the solution to (2.5), following the same proof as for (3.4) and assuming  $k, h$  small and  $kN$  bounded we get for some  $C > 0$

$$\|E_i^j - \mathcal{E}_i^j\|_2 \leq C(h^p + k^q)$$

with  $p, q$  as above. That is, when all approximations are of order 2, we get convergence of order 2.

**3.3. The nonlinear problem with smooth bounded nonlinearities.** We assume now that the coefficients  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  are functions of  $E$ . We furthermore assume that the functions  $\mathcal{A}(E), \mathcal{B}(E), \mathcal{C}(E), \mathcal{D}(E)$  are globally bounded and have bounded derivatives:

$$\max(\|\mathcal{A}\|_\infty, \|\mathcal{A}'\|_\infty, \|\mathcal{B}\|_\infty, \|\mathcal{B}'\|_\infty, \|\mathcal{C}\|_\infty, \|\mathcal{C}'\|_\infty, \|\mathcal{D}\|_\infty, \|\mathcal{D}'\|_\infty) \leq \Lambda.$$

We keep the assumptions on the signs:  $\mathcal{A}, \mathcal{B} \geq 0$  and  $\mathcal{C} \leq 0$ .

The discretization of this problem leads to system (2.5), where  $\mathcal{A}_i^j = \mathcal{A}(E_i^j)$ ,  $\mathcal{B}_i^j = \mathcal{B}(E_i^j)$ ,  $\mathcal{C}_i^j = \mathcal{C}(E_i^j)$ ,  $\mathcal{D}_i^j = \mathcal{D}(E_i^j)$  are known sequences. Therefore,  $E^{j+1}$  is computed solving the implicit nonlinear system (2.5).

**3.3.1. Resolution of the nonlinear system.** To compute  $E_i^{j+1}$  given  $E^j$  we solve

$$(3.6) \quad M(E^{j+1})E^{j+1} = \mathcal{M}(E^{j+1}) \left( E^j + \frac{k}{2} \dot{E}^j \right) + F(E^{j+1}),$$

where  $M(E^{j+1}), \mathcal{M}(E^{j+1}), F(E^{j+1})$  are defined as in section 2.3 but with coefficients involving  $\mathcal{A}(E^{j+1}), \mathcal{B}(E^{j+1}), \mathcal{C}(E^{j+1}), \mathcal{D}(E^{j+1})$ . We follow an iterative procedure.

First, we solve (3.6) with coefficients  $\mathcal{A}(E^j), \mathcal{B}(E^j), \mathcal{C}(E^j), \mathcal{D}(E^j)$  and get a solution  $E^{j+1,1}$ .

For  $n > 1$ , we iterate and solve (3.6) with coefficients  $\mathcal{A}(E^{j+1,n-1}), \mathcal{B}(E^{j+1,n-1}), \mathcal{C}(E^{j+1,n-1}), \mathcal{D}(E^{j+1,n-1})$ . In this way, we get a sequence of solutions  $E^{j+1,n}$ .

Then we prove that  $E^{j+1,n}$  converges to a solution  $E^{j+1}$  of (3.6). We leave the details of the convergence for Appendix B, where we also prove that

$$\begin{aligned} |E_i^{j+1}| &\leq C(\Lambda, E^0), |E_i^{j+1} - E_{i-1}^{j+1}| \leq C(\Lambda)h, \\ \left| E_i^{j+1} - \left( E_i^j + \frac{k}{2} \dot{E}_i^j \right) \right| &\leq C(\Lambda)k, |\dot{E}_i^{j+1}| \leq C(\Lambda). \end{aligned}$$

These bounds are important in Appendix C to conclude that the norm of the Jacobian matrices are small.

**3.3.2. Uniqueness and speed of convergence.** Let us prove uniqueness of solution for (3.6) and compute the speed of convergence of the iterative schemes. Keeping the notation of section 3.2, we see that

$$E^{j+1} = R(E^{j+1}) \left( E^j + \frac{k}{2} \dot{E}^j \right) + K(E^{j+1}).$$

Let us assume that we have another solution  $\mathcal{E}^{j+1}$ :

$$\mathcal{E}^{j+1} = R(\mathcal{E}^{j+1}) \left( E^j + \frac{k}{2} \dot{E}^j \right) + K(\mathcal{E}^{j+1}).$$

Then

$$E^{j+1} - \mathcal{E}^{j+1} = (R(E^{j+1}) - R(\mathcal{E}^{j+1})) \left( E^j + \frac{k}{2} \dot{E}^j \right) + K(E^{j+1}) - K(\mathcal{E}^{j+1})$$

and

$$\begin{aligned} \|E^{j+1} - \mathcal{E}^{j+1}\|_2 &\leq \text{Max}_\xi \|DK(\xi)\|_2 \|E^{j+1} - \mathcal{E}^{j+1}\|_2 \\ &+ \sum_l \text{Max}_\xi |DR^l(\xi) \left( E^j + \frac{k}{2} \dot{E}^j \right)|_2 \|E^{j+1} - \mathcal{E}^{j+1}\|_2. \end{aligned}$$

From the estimates on the derivatives of  $R$  and  $K$  (C.1), (C.2), we see that for  $h, k$  small

$$(3.7) \quad \|E^{j+1} - \mathcal{E}^{j+1}\|_2 \leq C(h, k) \|E^{j+1} - \mathcal{E}^{j+1}\|_2$$

with  $0 < C(h, k) < 1$ . Thus  $E^{j+1} = \mathcal{E}^{j+1}$  if  $E^j = \mathcal{E}^j$  and  $\dot{E}^j = \dot{\mathcal{E}}^j$ . Since the initial data are the same for both, we conclude by iterating that they coincide at all the subsequent levels.

A similar computation yields

$$\begin{aligned} \|E^{j+1, n+1} - E^{j+1, n}\|_2 &\leq \text{Max}_\xi \|DK(\xi)\|_2 \|E^{j+1, n} - E^{j+1, n-1}\|_2 \\ &+ \text{Max}_\xi \left\| DR(\xi) \left( E^j + \frac{k}{2} \dot{E}^j \right) \right\|_2 \|E^{j+1, n} - E^{j+1, n-1}\|_2, \end{aligned}$$

so that using (C.1), (C.2) from Appendix C,

$$\|E^{j+1, n+1} - E^{j+1, n}\|_2 \leq C(h, k)^n \|E^{j+1, 1} - E^{j+1, 0}\|_2$$

with  $C(h, k) < 1$ . We have fast geometrical convergence.

**3.3.3. Stability and convergence.** Since in the previous schemes we always have  $|E_i^j| \leq C(\Lambda, L)$ , the  $L^\infty$  stability is guaranteed. We assume in the following that all the approximations for the derivatives are of order 2 (see comments after (3.5)).

Let  $\mathcal{E}$  be a solution of the continuous nonlinear problem. We denote  $\mathcal{E}_i^j = \mathcal{E}(ih, jk)$ . Then

$$\mathcal{E}^{j+1} = R(\mathcal{E}^{j+1}) \left( \mathcal{E}^j + \frac{k}{2} \dot{\mathcal{E}}^j \right) + K(\mathcal{E}^{j+1}) + hO(h^2 + k^2).$$

Proceeding as in section 3.1, we can express the level  $j+1$  in terms of the level 0 for both the solution  $E_i^j$  provided by our scheme and the “discretized” continuous solution  $\mathcal{E}_i^j$ :

$$(3.8) \quad \begin{aligned} E^{j+1} &= R(E^{j+1})(2R(E^j) - I) \cdots (2R(E^1) - I)E^0 + 2R(E^{j+1})K(E^j) \\ &+ K(E^{j+1}) + 2 \sum_{l=0}^{j-1} R(E^{j+1})(2R(E^j) - I) \cdots (2R(E^{j-l}) - I)K(E^l), \end{aligned}$$

$$\begin{aligned}
 \mathcal{E}^{j+1} &= R(\mathcal{E}^{j+1})(2R(\mathcal{E}^j) - I) \cdots (2R(\mathcal{E}^1) - I)\mathcal{E}^0 + K(\mathcal{E}^{j+1}) + hO(h^2 + k^2) \\
 &\quad + 2 \sum_{l=0}^{j-1} R(\mathcal{E}^{j+1})(2R(\mathcal{E}^j) - I) \cdots (2R(\mathcal{E}^{j-l}) - I)(K(\mathcal{E}^l) + hO(h^2 + k^2)) \\
 (3.9) \quad &\quad \quad \quad + 2R(\mathcal{E}^{j+1})(K(\mathcal{E}^j) + hO(h^2 + k^2))
 \end{aligned}$$

taking  $\dot{E}^0 = 0 = \dot{\mathcal{E}}^0$ . Therefore, since  $\mathcal{E}^0 = E^0$ ,

$$\begin{aligned}
 &E^{j+1} - \mathcal{E}^{j+1} \\
 (3.10) \quad &= \left( R(E^{j+1})(2R(E^j) - I) \cdots (2R(E^1) - I) - R(\mathcal{E}^{j+1})(2R(\mathcal{E}^j) - I) \cdots \right. \\
 &\quad \left. (2R(\mathcal{E}^1) - I) \right) \mathcal{E}^0
 \end{aligned}$$

$$\begin{aligned}
 (3.11) \quad &+ 2 \sum_{l=0}^{j-1} \left( R(E^{j+1})(2R(E^j) - I) \cdots (2R(E^{j-l}) - I) - R(\mathcal{E}^{j+1})(2R(\mathcal{E}^j) - I) \cdots \right. \\
 &\quad \left. (2R(\mathcal{E}^{j-l}) - I) \right) K(E^l) + 2 \left( R(E^{j+1}) - R(\mathcal{E}^{j+1}) \right) K(E^j)
 \end{aligned}$$

$$\begin{aligned}
 (3.12) \quad &+ 2 \sum_{l=0}^{j-1} R(\mathcal{E}^{j+1})(2R(\mathcal{E}^j) - I) \cdots (2R(\mathcal{E}^{j-l}) - I) \left( K(E^l) - K(\mathcal{E}^l) \right) \\
 &\quad + \left( K(E^{j+1}) - K(\mathcal{E}^{j+1}) \right) + 2R(\mathcal{E}^{j+1})(K(E^j) - K(\mathcal{E}^j))
 \end{aligned}$$

$$\begin{aligned}
 (3.13) \quad &- hO(h^2 + k^2) - 2 \sum_{l=0}^{j-1} R(\mathcal{E}^{j+1})(2R(\mathcal{E}^j) - I) \cdots (2R(\mathcal{E}^{j-l}) - I) hO(h^2 + k^2) \\
 &\quad - 2R(\mathcal{E}^{j+1}) hO(h^2 + k^2).
 \end{aligned}$$

We recall that for  $k, h$  small, there exists  $0 < \gamma < 1$  such that

$$(3.14) \quad \|R(E^j)\|_2 < \gamma, \|R(\mathcal{E}^j)\|_2 < \gamma, \|2R(E^j) - I\|_2 < \gamma, \|2R(\mathcal{E}^j) - I\|_2 < \gamma.$$

The terms in (3.13) are bounded by  $C(\Lambda, L) \frac{1+\gamma}{1-\gamma} (h^2 + k^2)$ . Using the results in Appendix C, the terms in (3.12) are bounded by  $C(\Lambda, L) k \frac{1+\gamma}{1-\gamma} \max_i \|E^i - \mathcal{E}^i\|_2$ . The difference of matrix products in (3.10) can be bounded using the identity

$$(3.15) \quad a_j a_{j-1} \cdots a_2 a_1 - b_j b_{j-1} \cdots b_2 b_1 = \sum_{l=0}^{j-1} a_j a_{j-1} \cdots (a_{j-l} - b_{j-l}) b_{j-l-1} \cdots b_1$$

with  $a_j = R(E^j), b_j = R(\mathcal{E}^j)$  and  $a_i = 2R(E^i) - I, b_i = 2R(\mathcal{E}^i) - I$  for  $1 \leq i < j$ . Using (3.14), (3.15) and repeating the procedure in Appendix C, we bound the terms in (3.10) by  $C(\lambda, L) k \gamma^{j-1} \max_i \|E^i - \mathcal{E}^i\|_2$ . Analogously, we bound the sum of terms in (3.11) by  $C(\lambda, L) k \frac{1+\gamma}{1-\gamma} \max_i \|E^i - \mathcal{E}^i\|_2$ . Thus, for some positive constants  $C_1, C_2$ ,

$$(3.16) \quad (1 - kC_1) \max_j \|E^j - \mathcal{E}^j\|_2 \leq C_2 (h^2 + k^2).$$

We conclude that the convergence is of order 2.

**4. Numerical results.** As an example of the results obtained, we have used the method described in section 2 to carry out numerical simulations for the case of

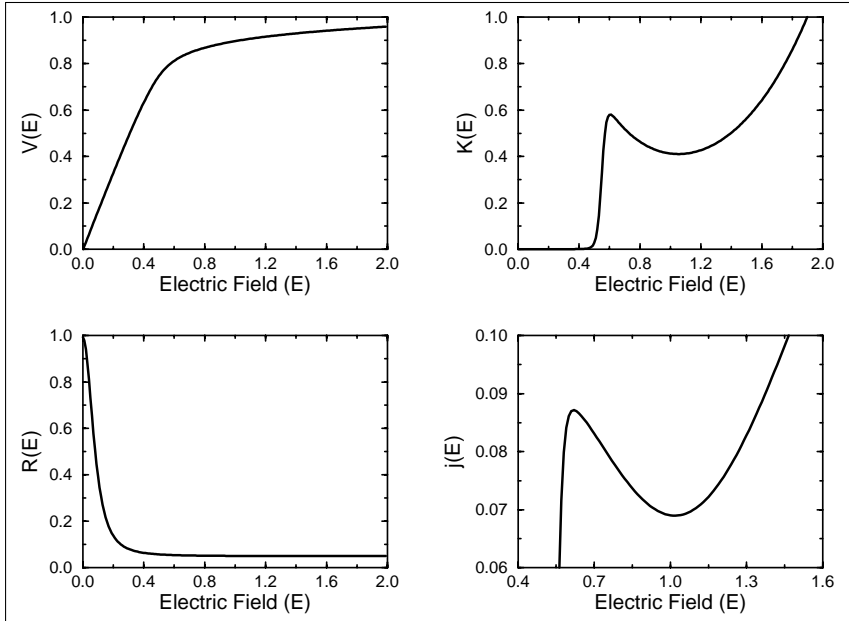


FIG. 4.1. Nondimensional functions  $V(E)$ ,  $K(E)$ ,  $R(E)$ , and  $j(E)$  used in the modeling of the Gunn effect in  $p$ -Ge.

ultrapure  $p$ -type Ge [3], [4]. The functions defining this problem are

$$(4.1) \quad \begin{aligned} \mathcal{A}(E) &= \frac{JV'(E)}{V(E)^2}, \mathcal{B}(E) = \frac{K(E) + R(E)}{V(E)}, \mathcal{C}(E) = \frac{-1}{V(E)}, \\ \mathcal{D}(E) &= \frac{K(E) + R(E)}{V(E)^2} J(j(E) - J), j(E) = \left( \frac{\alpha K(E)}{K(E) + R(E)} - 1 \right) V(E) \end{aligned}$$

with  $J = E(0, t) / \rho_0$ . The behavior of functions  $K, R, V, j$  is shown in Figure 4.1, and introducing these values in (4.1), the behavior of  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  as a function of  $E$  and  $J$  is deduced. This behavior is shown in Figure 1.1 for typical values of  $J$ .

The system behavior depends fundamentally on the parameters  $\rho$  and  $\Phi/L$ . The nondimensional contact resistivity  $\rho$  regulates the rate of generation of the waves in the anode (left side of the spatial domain). The  $\Phi/L$  parameter regulates the size and velocity of the propagating waves. Thus each point on the  $\rho - \Phi/L$  plane corresponds to a particular solution. For small values of  $\Phi/L$  an ohmic solution is observed (stationary field distribution). For values of  $\Phi/L$  above a critical value, different types of dynamic solutions result. For most points in this region, complex periodic solutions are generated and there are small areas in the plane which correspond to chaotic solutions with complex transitions of different kinds.

Notice that the coefficient functions defined in (4.1) are not globally bounded. However, this difficulty can be avoided by a typical truncation argument. We want to use our schemes to track numerically the pulses described in [4]. These pulses are such that  $0 < E_{min} \leq E \leq E_{max}$  with  $E_{min}, E_{max}$  constants given in [4].

Therefore, we truncate our coefficient functions to appropriate constant values outside the interval  $(E_{min} - \epsilon, E_{max} + \epsilon)$ ,  $\epsilon > 0$  small in such a way that the truncated coefficients  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  are globally bounded, with bounded derivatives, and coincide with the original coefficients in the interval where we expect to find the pulse. The

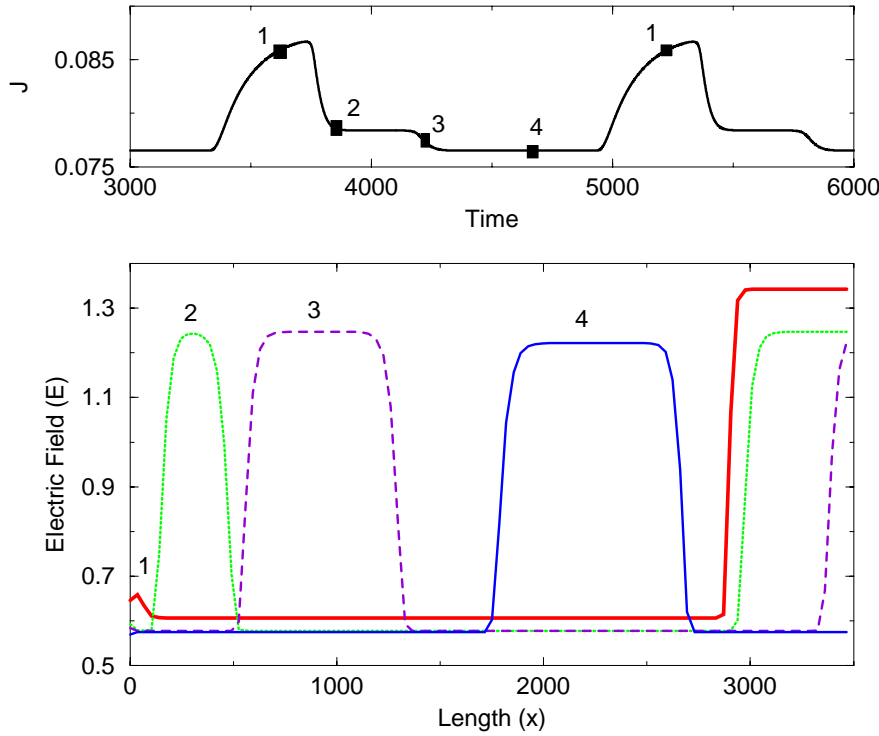


FIG. 4.2. Periodic cycle in the current intensity and electric field distribution for the parameters of the Figure 4.3.

solution of the problem with truncated coefficients coincides with the solution of the original problem. Thus write down the numerical algorithms for the truncated problem and the stability and convergence results established in section 3 are guaranteed.

Figures 4.2 and 4.3 show the results obtained with the numerical method described above in the case in which the solution exhibits a periodic behavior. Figure 4.2 shows the field distribution at four instants of a complete cycle. At time (1) a wave is disappearing at the right boundary and a new wave appears in the left boundary. At time (2) the new wave grows very fast, leading to a fast decrease in the current density, while the old wave continues disappearing. At time (3) the new wave is increasing its area at a slow rate to compensate for the area loss of the disappearing wave. At time (4) the wholly developed wave appears. Figure 4.3 shows the graphic representation of a complete simulation including several cycles. Notice that the small space step necessary to get the desired accuracy ( $h = 0.5$ ) and the large length of the sample ( $L = 3500$ ) leads to a space discretization requiring  $N = 7000$  grid points.

Figure 4.4 shows the results obtained for values of the parameters that produce a more complex behavior apparently chaotic. In this case, the numerical results are very sensitive to the spatial step  $h$  and temporal step  $k$ , so that it is necessary a strict control of the precision of the method. It should be pointed out that in order to characterize these chaotic solutions, the transition between chaotic and periodic solutions and the location in the parameter space of chaotic solutions, it is necessary to compute very long simulations ( $t \sim 100000$  in nondimensional time units). In fact, there are regions in the parameter space where periodic solutions are observed but

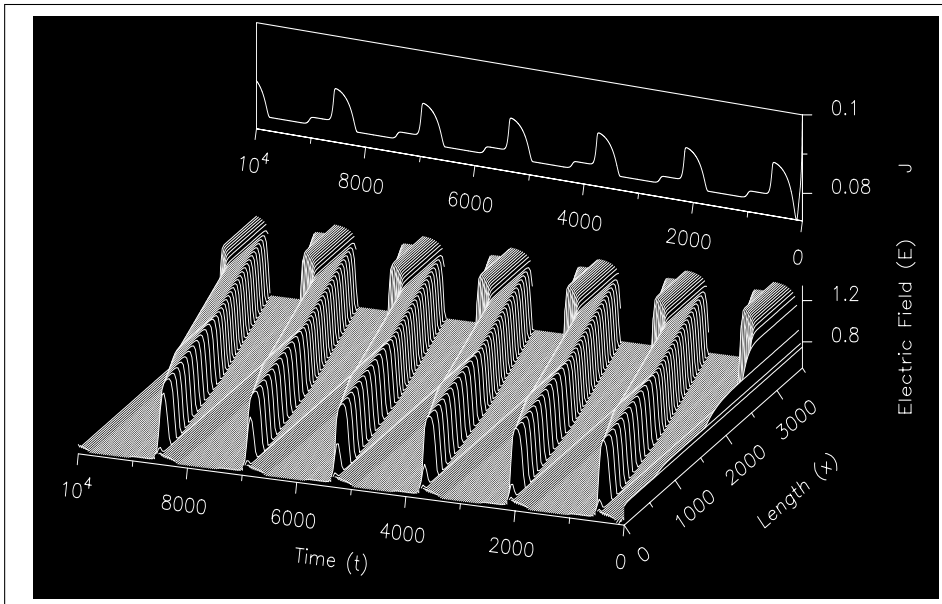


FIG. 4.3. Graphic representation of the numerical results obtained in the simulation for  $L = 3500$ ,  $E_0 = 0.75$ ,  $\rho = 7.5$ ,  $\alpha = 1.21$  and space and time steps  $h = 0.5$ ,  $k = 0.1$ . The 3D graphic located in the lower part of the figure shows the space-time evolution of the electric field,  $E$ . The 2D graphic located in the upper part of the figure shows the temporal variation of the current density,  $J$ .

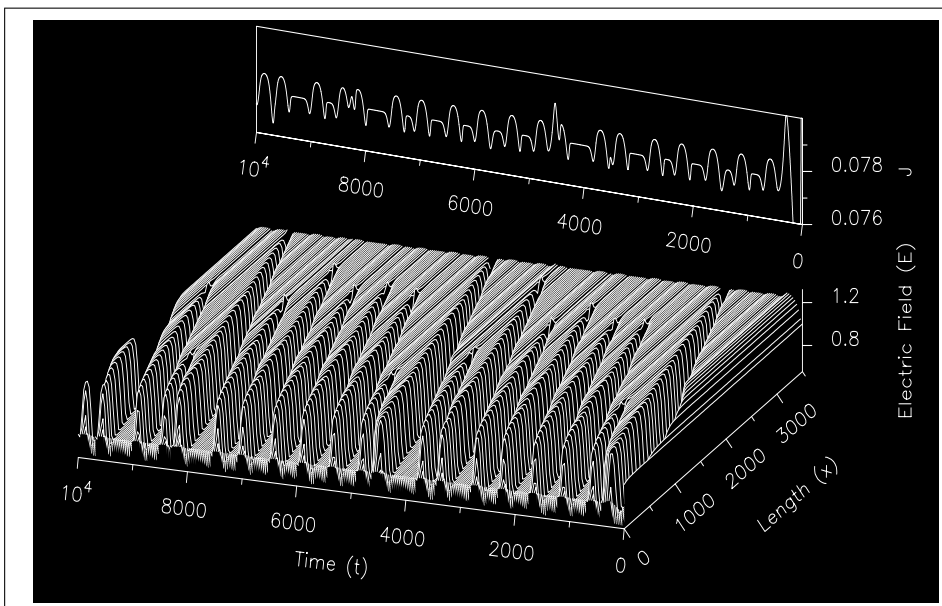


FIG. 4.4. Graphic representation of the numerical results obtained in the simulation for  $L = 3500$ ,  $E_0 = 1.0$ ,  $\rho = 10.0$ ,  $\alpha = 1.21$  and space and time steps  $h = 0.5$ ,  $k = 0.1$ . The graphics layout is the same as that of Figure 4.3.

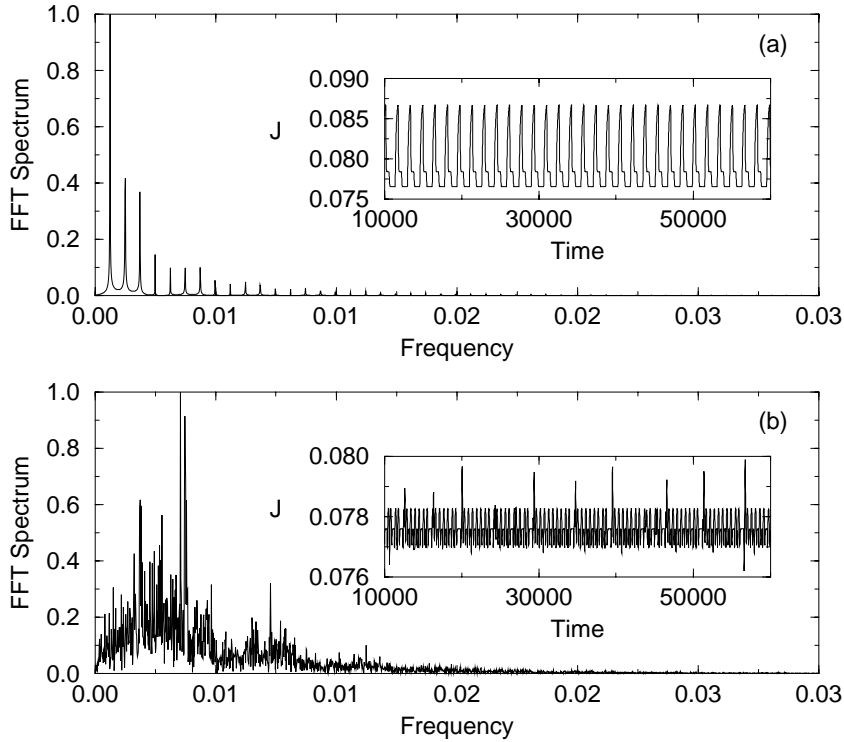


FIG. 4.5. FFT spectrum of the series  $J-t$ : (a) corresponds to the simulations of Figure 4.3 and (b) corresponds to the simulation of Figure 4.4. One can observe the periodicity of (a) (Lyapunov exponent of  $-0.000321$ ) and the chaotic appearance of (b) (Lyapunov exponent of  $0.000512$ ). In both figures, the FFT spectrum is in normalized units and the frequency is in arbitrary units.

with very long periods ( $\sim 10000$ ), and to distinguish these solutions from the chaotic ones it is necessary to include at least 10 time periods. For the time step and total time used in Figure 4.4 the number of time steps needed is  $N_t = 100000$ .

Figure 4.5 shows the FFT spectrum of the runs corresponding to Figures 4.3 and 4.4 for a simulation time of 60000, where the initial transients have been removed. The spectrum of case (b) appears chaotic and in fact this can be confirmed computing the maximum Lyapunov exponent,  $-0.000321$ , in the case (a) and  $0.000512$  in the case (b).

The computational complexity for each iteration and each time step of the algorithm can be estimated as follows:

- Matrix element computation: this is the most computationally intensive part, especially due to the evaluation of functions  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  at each grid point, which includes computation of exponentials, logarithms, and arc tangents. This requires approximately  $115 * N$  floating point operations (*flops*). Once these functions are evaluated, computation of the matrix elements requires approximately  $60 * N$  *flops*. Thus a total of  $175 * N$  *flops* are carried out.
- Back substitution to solve two triangular systems: this is very fast since only  $4 * N$  *flops* are needed
- Residual calculation: this computation is needed to check convergence and is computationally analogous to the matrix element computation.



Thus the computational cost per iteration and time step is  $354 * N \text{ flops}$  (2.5 megaflops for  $N = 7000$ ). The number of iterations required at each time step for convergence is very small. Generally 1–3 iterations suffice, except at certain times when a new wave is generated or when two waves coalesce, in which case up to 10 iterations are needed. Considering an average of three iterations per time step, the total computational cost for a particular simulation (Cost  $\approx 354 N N_{iterations} N_{steps}$ ) is of the order of  $710^{11} \text{ flops}$ . Assuming that the program is run on a system with a performance of 10 megaflops per second, a total of 20 hours are needed per run. These estimates show that although the problem is one dimensional, it is rather computationally intensive and therefore it is important to optimize the numerical algorithm.

### Appendix A. Study of the matrices.

**A.1. Comparison principles for  $T^{j+1}$ .** Let us recall that  $T^{j+1} = (I + \frac{k}{2}D_B)T + 4\rho hD_A$ , with  $D_A, D_B, T$  as in section 2. We assume the coefficients satisfy (2.1) in section 3.  $T^{j+1}$  is a lower triangular matrix. It is obviously nonsingular, with spectral radius smaller than 1 for the values of  $\rho$  we work with.

We need to be able to give sharp bounds on the solutions of systems of the type  $T^{j+1}x = b$ . Typically,  $b = b^{j+1}, v^{j+1}$ . The bounds will be obtained as a consequence of the following series of comparison principles.

**LEMMA A.1.** *Let us consider the system  $(I + \frac{k}{2}D_B)Tx = b$ . If  $b_1 \geq 0, b_2 + 2b_1 \geq 0$  and  $b_n \geq 0, n > 2$ , then  $x_n \geq 0, x_n > x_{n-1}, x_n \leq \frac{3}{2} \sum_{k=0}^N \frac{\bar{b}_k}{6\rho}$ , with  $\bar{b}_n = \frac{b_n}{6\rho + 3\rho\mathcal{B}_nk}, n > 1$  and  $\bar{b}_1 = \frac{b_1}{4\rho + 2\rho\mathcal{B}_1k}$ .*

*More generally, we have the maximum principle  $(I + \frac{k}{2}D_B)Tx = b, (I + \frac{k}{2}D_B)Ty \leq b \implies y \leq x$ .*

*Proof.* The  $x_n$  are computed as solutions of the difference equation  $x_n = \frac{\bar{b}_n}{6\rho} + \frac{4}{3}x_{n-1} - \frac{1}{3}x_{n-2}, n > 2$  starting with  $x_1 = \frac{\bar{b}_1}{4\rho}, x_2 = \frac{\bar{b}_2}{6\rho} + \frac{4}{3}x_1$ .  $\square$

Now for  $n > 1$  we set

$$(A.1) \quad \alpha_n = 6\rho + 3\rho\mathcal{B}_nk, \beta_n = 8\rho + 4\rho\mathcal{B}_nk, \gamma_n = 2\rho + \rho\mathcal{B}_nk$$

and  $\alpha_1 = 4\rho + 2\rho\mathcal{B}_1k$ . Notice that  $\frac{\beta_n}{\alpha_n} = \frac{4}{3}, \frac{\gamma_n}{\alpha_n} = \frac{1}{3}$  and  $\beta_n = \alpha_n + \gamma_n$ . We also set  $\delta_n = \frac{4\rho\mathcal{A}_nh}{\alpha_n}$ . The solution  $x_n$  to  $T^{j+1}x = b$  is given by

$$(A.2) \quad \begin{aligned} (1 + \delta_1)x_1 &= \frac{b_1}{\alpha_1}, (1 + \delta_2)x_2 = \frac{b_2}{\alpha_2} + \frac{4}{3}x_1, \\ (1 + \delta_n)x_n &= \frac{b_n}{\alpha_n} + \frac{4}{3}x_{n-1} - \frac{1}{3}x_{n-2}, n > 2. \end{aligned}$$

**LEMMA A.2.** *Let us consider the sequence  $x_n$  defined by*

$$(A.3) \quad \begin{aligned} (1 + \delta)x_1 &= \frac{b_1}{\alpha_1}, (1 + \delta)x_2 = \frac{b_2}{\alpha_2} + \frac{4}{3}x_1, \\ (1 + \delta)x_n &= \frac{b_n}{\alpha_n} + \frac{4}{3}x_{n-1} - \frac{1}{3}x_{n-2}, n > 2 \end{aligned}$$

*with  $\delta > 0$  small. If  $b_n \geq 0$  (resp.,  $b_n \leq 0$ ), then  $x_n \geq 0$  (resp.,  $x_n \leq 0$ ). The condition  $b_2 \geq 0$  (resp.,  $b_2 \leq 0$ ) may be relaxed to  $3(1 + \delta)b_2 + 4b_1 \geq 0$  (resp.,  $\leq 0$ ).*

*Proof.* The solutions of the homogeneous difference equation  $(1 + \delta)y_n - \frac{4}{3}y_{n-1} + \frac{1}{3}y_{n-2} = 0$  are given by

$$(A.4) \quad y_n = A r_+(\delta)^{n-3} + B r_-(\delta)^{n-3}, n > 2,$$

where  $r_-(\delta) \sim \frac{1}{3} + \frac{1}{6}\delta$  and  $r_+(\delta) = 1 - \frac{3}{2}\delta$ .

For a right-hand side  $b_1, b_2, 0, \dots, 0$ , we compute  $y_1, y_2$  and then use (A.4) to compute  $A$  and  $B$  so that we know  $y_n, n > 2$ . If  $b_1, b_2 > 0$ , then  $A > 0, B < 0$  and  $y_2 > y_1 > 0$ . For large  $n$ ,  $y_n - y_{n-1} \sim A(r_+ - 1)r_+^{n-4} < 0$ . However, we cannot have  $y_n < y_{n-1}$  and  $y_n < 0$  at the same time so that  $y_n$  remains always positive. More precisely,  $0 < y_n < A(b_1, b_2)r_+^{n-3}$ .

For a right-hand side  $0, \dots, 0, b_{n_0}, 0, \dots, 0$ , we have  $y_n = 0, n < n_0, y_{n_0} = \frac{b_{n_0}}{\alpha_{n_0}(1+\delta)}$  and we compute  $y_n, n > n_0$  from (A.4) starting from  $y_{n_0-1}, y_{n_0}$ . Again, if  $b_{n_0} > 0$ ,  $A > 0, B < 0$ , and  $0 \leq y_n \leq A(b_{n_0})r_+^{n-(n_0+1)}$  for  $n > n_0$ .

The sequence  $x_n$  is obtained adding the solutions  $y_n$  for the right-hand sides  $b_1, b_2, 0, 0, \dots$  and  $0, \dots, 0, b_{n_0}, 0, \dots, 0, n_0 > 2$ . Thus lemma A.2 follows.  $\square$

LEMMA A.3. *We assume that  $0 \leq \underline{\delta} \leq \delta_k \leq \bar{\delta}$ . Let  $x_n$  be defined by (A.2). Let  $\underline{x}_n$  be the solution to (A.3) with  $\delta = \underline{\delta}$  and  $\bar{x}_n$  the solution to (A.3) with  $\delta = \bar{\delta}$ , both with right-hand side  $\frac{b_n}{\alpha_n}$ . If  $b_n \geq 0$  and  $\bar{\delta}$  is small enough, then  $x_n > 0$ . Moreover,*

$$(A.5) \quad \bar{x}_n \leq x_n \leq \underline{x}_n.$$

When  $b_n < 0$ , the signs are reversed in the above inequalities. These comparison principles also hold when  $b_2 > 0$  (resp.,  $b_2 < 0$ ) is replaced by  $3\alpha_1(1+\delta_1)b_2 + 4\alpha_2b_1 > 0$  (resp.,  $< 0$ ). This implies that  $(T^{j+1})^{-1}$  is a positive matrix.

*Remark.* We may take  $\delta = \frac{2}{3} \max(\mathcal{A})\bar{h}$  for  $h \leq \bar{h}$ . For the lower bound, we may take either the uniform bound  $\underline{\delta} = 0$  or  $\underline{\delta} = \frac{2}{3+c} \min(\mathcal{A})h$ .

*Proof.* Assume  $b_n \geq 0$ . Then we have  $0 < x_1 < x_2$ . Assume that  $x_n > 0$  for  $n \leq n_0$ . Then

$$(A.6) \quad (1 + \bar{\delta})x_n = \frac{4}{3}x_{n-1} - \frac{1}{3}x_{n-2} + \frac{b_n}{\alpha_n} + (\bar{\delta} - \delta_n)x_n.$$

By the comparison principle  $x_n \geq \bar{x}_n$  for  $n \leq n_0$ . In a similar way,  $x_n \leq \underline{x}_n$  for  $n \leq n_0$ . Let us see that  $x_{n_0+1} > 0$  and by induction we get (A.5) for all  $n$ . Now

$$(A.7) \quad (1 + \delta_{n_0+1})x_{n_0+1} \geq x_{n_0} + \frac{1}{3}(x_{n_0} - x_{n_0-1}).$$

To ensure  $x_{n_0+1} > 0$  we need  $x_{n_0-1} - x_{n_0} < 3x_{n_0}$ . Since  $x_{n_0-1} - x_{n_0} \leq \underline{x}_{n_0} - \bar{x}_{n_0}$ , it is enough to prove  $\underline{x}_{n_0-1} < 4\bar{x}_{n_0}$ . This follows from the explicit expressions as solutions of difference equations for  $\underline{x}_n$  and  $\bar{x}_n$  when  $\bar{\delta}$  is small enough.  $\square$

**A.2. Order of magnitude of  $N^{j+1}$ .** We recall that

$$N^{j+1} = I + \frac{1}{1 - u \cdot z^{j+1}} \begin{pmatrix} 2z_1^{j+1} & \dots & 2z_1^{j+1} & z_1^{j+1} \\ \dots & \dots & \dots & \dots \\ 2z_N^{j+1} & \dots & 2z_N^{j+1} & z_N^{j+1} \end{pmatrix},$$

where  $z^{j+1}$  is the solution to  $T^{j+1}z^{j+1} = v^{j+1}$ . From the previous maximum principles, it is clear that

$$(A.8) \quad z_n^{j+1} < 0 \quad \forall n, \quad \frac{1}{1 - u \cdot z^{j+1}} > 0, \quad |u \cdot z^{j+1}| \geq |u \cdot \bar{z}|,$$

where  $\bar{z}$  is the solution for  $\bar{\delta} = h \max(\mathcal{A})$  with right-hand side  $(-\alpha_1, \gamma_2, 0, \dots, 0)$ . Since  $\sum \bar{z}_n \sim cte \sum r_+^n \sim \frac{cte}{h \max \mathcal{A}}$ , we see that

$$(A.9) \quad 0 < \frac{1}{1 - u \cdot z^{j+1}} < cte h \|\mathcal{A}\|_\infty.$$

**A.3. Order of magnitude of  $Q^{j+1}$ .** This matrix is lower triangular with spectral radius less than 1 for  $k$  small. In view of its definition, we expect  $Q^{j+1}$  to be almost diagonal.

LEMMA A.4. *The diagonal elements of  $Q^{j+1}$  are of order  $k$ . The nondiagonal elements are of order  $kh$ .*

*Proof.* The columns  $q_i$  of  $Q^{j+1}$  are solutions to the systems  $T^{j+1}q_i = c_i$ , where the vectors  $c_i$  are columns of  $\frac{k}{2}D_B T$ .

Let  $c_i = (0, \dots, 0, 3\rho k \mathcal{B}_i, -4\rho k \mathcal{B}_{i+1}, \rho k \mathcal{B}_{i+2}, 0, \dots, 0)^t$  be one of such columns. Then  $q_i$  is given by  $q_{n,i} = 0$  for  $n < i$  and

$$\begin{aligned} (\alpha_i + \delta_i)q_{i,i} &= 3\rho k \mathcal{B}_i, \\ (\alpha_{i+1} + \delta_{i+1})q_{i+1,i} &= \beta_{i+1}q_{i,i} - 4\rho k \mathcal{B}_{i+1}, \\ (\alpha_{i+2} + \delta_{i+2})q_{i+2,i} &= \beta_{i+2}q_{i+1,i} - \gamma_{i+2}q_{i,i} + 2\rho k \mathcal{B}_{i+2}, \\ \left(1 + \frac{\delta_n}{\alpha_n}\right)q_{n,i} &= \frac{4}{3}q_{n-1,i} - \frac{1}{3}q_{n-2,i} \quad n > i + 1. \end{aligned}$$

Developing the coefficients at the node  $i$  in powers of  $h, k$ , we see that  $q_{i,i} \sim \frac{k}{2}\mathcal{B}_i(1 - \frac{k}{2}\mathcal{B}_i - \frac{2}{3}h\mathcal{A}_i)$ ,  $q_{i,i+1} \sim -\frac{4}{9}hk\mathcal{A}_i\mathcal{B}_i - \frac{2}{3}hk\mathcal{B}'_i$  and  $q_{i,i+2} \sim -\frac{61}{27}hk\mathcal{A}_i\mathcal{B}_i - \frac{5}{9}hk\mathcal{B}'_i$ . From this, we conclude that  $|q_{n,i}| \leq \underline{q}_{n,i} \leq C(\mathcal{A}_i, \mathcal{B}_i)hk$ ,  $n > i + 2$ , where  $\underline{q}_{n,i}$  is the solution to (A.3) with  $\delta = \underline{\delta} = 0$ , zero right-hand side and starting from  $|q_{i+1,i}|, |q_{i+2,i}|$ .  $\square$

**A.4. Order of magnitude of  $Z^{j+1}Q^{j+1}$ .** We first remark that

$$Z^{j+1}Q^{j+1} = \frac{1}{1 - u \cdot z^{j+1}} \begin{pmatrix} z_1^{j+1}S_1 \dots z_1^{j+1}S_1 \\ \dots \\ z_N^{j+1}S_1 \dots z_N^{j+1}S_N \end{pmatrix},$$

where  $S_i = 2(q_{i,i} + \dots + q_{N-1,i}) + q_{N,i} \sim C(\mathcal{A}, \mathcal{B})kL$ . On the other hand, using Lemmas A.1–A.3

$$(A.10) \quad \frac{z_k S_i}{1 - u \cdot z} \sim hkC(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)$$

since  $|z_k| \leq \underline{z}_k \sim \max(|C|)L + C(\alpha_1, \beta_1)$  for  $\underline{\delta} = 0$ .

**A.5. Norm of  $I - N^{j+1}Q^{j+1}$ .** Products of matrices of this type will appear when we express the level  $j + 1$  in terms of the level 0. In order to have  $E^{j+1}$  bounded as  $j$  tends to infinity, we need to see that the norm of these matrices is smaller than 1 for a certain norm.

LEMMA A.5. *For  $h, k$  small enough with  $kN$  bounded*

$$(A.11) \quad \|I - N^{j+1}Q^{j+1}\|_2^2 = \rho((I - N^{j+1}Q^{j+1})(I - (Q^{j+1})^t(N^{j+1})^t) < 1.$$

*Proof.* We decompose

$$\begin{aligned} &(I - N^{j+1}Q^{j+1})(I - (Q^{j+1})^t(N^{j+1})^t) \\ &= (I - Q^{j+1})(I - Q^{j+1})^t + (I - Q^{j+1})(Q^{j+1})^t(Z^{j+1})^t \\ &\quad + Z^{j+1}Q^{j+1}(I - (Q^{j+1})^t) + Z^{j+1}Q^{j+1}(Q^{j+1})^t(Z^{j+1})^t. \end{aligned}$$

We must prove that the spectral radius of this matrix is smaller than 1. By Gershgorin's theorem [10, page 135], the eigenvalues lie in balls centered in the diagonal elements and with radius either the sum of the modulus of the elements in the same column or the sum of the modulus of the elements in the same row.

Let us consider the different matrix products  $P = RS$  involved.

If  $R = (I - Q^{j+1})$ ,  $S = R^t$ ,

$$(A.12) \quad |p_{i,j}| \leq C(\mathcal{A}, \mathcal{B}) \sum_{l < \min(i,j)} k^2 h^2 \leq C(\mathcal{A}, \mathcal{B}) k^2 h L,$$

$$(A.13) \quad |p_{i,i}| \leq \left(1 - \frac{k}{2} \mathcal{B}_i\right)^2 + C(\mathcal{A}, \mathcal{B}) k^2 h L.$$

If  $R = (I - Q^{j+1})$ ,  $S = (Q^{j+1})^t (Z^{j+1})^t$  or  $R = Z^{j+1} Q^{j+1}$ ,  $S = (I - (Q^{j+1})^t)$ ,

$$(A.14) \quad |p_{i,j}| \leq C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L) k^2 h, \quad |p_{i,i}| \leq C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L) (k^2 h L + kh).$$

If  $R = Z^{j+1} Q^{j+1}$ ,  $S = R^t$ ,

$$(A.15) \quad |p_{i,j}| \leq C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L) k^2 h, \quad |p_{i,i}| \leq C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L) k^2 h.$$

Adding up, we see that we have a matrix with diagonal elements of order  $(1 - \frac{k}{2} \mathcal{B}_i)^2 + C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L) (k^2 h + kh)$  and nondiagonal elements of order  $C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L) k^2 h$ . Therefore, the eigenvalues lie in balls with centers  $(1 - k \mathcal{B}_i)$  and radius of order  $\sum k^2 h L C(\mathcal{A}, \mathcal{B})$ , that is,  $C(\mathcal{A}, \mathcal{B}) k^2 L^2$ . For  $k$  small enough, the spectral radius is of order  $\max(1 - k \mathcal{B}_i) < 1$ . Thus the norm-2 of these matrices is less than 1.  $\square$

**A.6. Norm of  $R^{j+1}$ ,  $2R^{j+1} - I$ .** We have the following lemma.

LEMMA A.6. *For  $h, k$  small enough with  $kN$  bounded  $\|R^j\|_2 = \alpha < 1$ ,  $\|2R^j - I\|_2 < \alpha < 1$ , and  $\|K^j\|_2 \leq G\|\mathcal{D}\|_\infty$ .*

*Proof.* By Lemma A.5, the result holds for  $I - N^{j+1} Q^{j+1}$ . In  $R^{j+1}$ , we have one more column and row. Let us estimate their size.

The comparison principle for  $T^{j+1}$  allows us to bound the solution of  $T^{j+1}x = b^{j+1}$  by the solution of  $(T^{j+1} - 4\rho h D_A)y = b^{j+1}$ , which is bounded by  $C(\Lambda, L)k$ . Since the elements of  $Z^{j+1}$  are bounded by  $C(\Lambda, L)h$ , we conclude that

$$(A.16) \quad \|N^{j+1} (T^{j+1})^{-1} b^{j+1}\|_\infty \leq C(\Lambda, L)k.$$

From (A.8), (A.9) we know that  $\frac{u \cdot (T^{j+1})^{-1} b^{j+1}}{1 - u \cdot z^{j+1}} < 0$  and  $0 < \frac{1}{1 - u \cdot z^{j+1}} < C(\Lambda, L)h$ . Thus,  $\|\frac{u \cdot (T^{j+1})^{-1} b^{j+1}}{1 - u \cdot z^{j+1}}\|_\infty \leq C(\Lambda, L)k$  and  $0 < 1 - C(\Lambda, L)k < 1 + \frac{u \cdot (T^{j+1})^{-1} b^{j+1}}{1 - u \cdot z^{j+1}} < 1$ . Analogously, we obtain  $\|K^j\|_2 \leq G\|\mathcal{D}\|_\infty$ .

Finally, since the nondiagonal elements of  $Q^{j+1}$  are of order  $kh$  and the diagonal elements are of order  $h$ , we get  $\|\frac{1}{1 - u \cdot z^{j+1}} (u \cdot Q^{j+1})\|_\infty \leq C(\mathcal{A}, \mathcal{B}, L)kh$ . In view of these estimates on the terms of the extra row and column, we obtain  $\|R^{j+1}\|_2 < \alpha < 1$ . Now  $2R^j - I$  has a similar structure to  $R^j$  and  $\|2R^j - I\|_2 < \alpha$  again.  $\square$

## Appendix B. Bounds on discrete derivatives.

LEMMA B.1. *Given a vector  $E^0$ , we can compute a sequence  $E^{j+1}$  such that*

$$M(E^{j+1})E^{j+1} = \mathcal{M}(E^{j+1}) \left( E^j + \frac{k}{2} \dot{E}^j \right) + F(E^{j+1}).$$

Moreover,

$$(B.1) \quad |E_i^{j+1}| \leq C(\Lambda, E^0), |E_i^{j+1} - E_{i-1}^{j+1}| \leq C(\Lambda)h,$$

$$(B.2) \quad \left| E_i^{j+1} - \left( E_i^j + \frac{k}{2} \dot{E}_i^j \right) \right| \leq C(\Lambda)k, |\dot{E}_i^{j+1}| \leq C(\Lambda).$$

*Proof.* Let us assume that we have computed  $E^j$  and

$$\begin{aligned} |E_i^j| &\leq C(\Lambda, E^0), |E_i^j - E_{i-1}^j| \leq C(\Lambda)h, \\ |E_i^j - (E_i^{j-1} + \frac{k}{2} \dot{E}_i^{j-1})| &\leq C(\Lambda)k, |\dot{E}_i^j| \leq C(\Lambda). \end{aligned}$$

We solve  $M(E^j)E^{j+1,1} = \mathcal{M}(E^j)(E^j + \frac{k}{2}\dot{E}^j) + F(E^j)$ . Then

$$\begin{aligned} T(E^j) &\left( \left( E_r^{j+1,1} - E_r^j - \frac{k}{2} \dot{E}_r^j \right) + z(E^j) \left( E_0^{j+1,1} - E_0^j - \frac{k}{2} \dot{E}_0^j \right) \right) \\ &= -Q(E^j)(E_r^j + \frac{k}{2} \dot{E}_r^j) - b(E^j)(E_0^j + \frac{k}{2} \dot{E}_0^j) - 2\rho hk \mathcal{D}(E^j). \end{aligned}$$

Using  $|\dot{E}_i^j| \leq C(\Lambda)$  and  $|E_i^j - E_{i-1}^j| \leq C(\Lambda)h$  we see that the right-hand side is of order  $k^2 + kh$ . Therefore,

$$\left| \left( E_r^{j+1,1} - E_r^j - \frac{k}{2} \dot{E}_r^j \right) + z(E^j) \left( E_0^{j+1,1} - E_0^j - \frac{k}{2} \dot{E}_0^j \right) \right| \leq C(\Lambda)k.$$

Thus

$$(B.3) \quad \left| E_0^{j+1,1} - E_0^j - \frac{k}{2} \dot{E}_0^j \right| \leq C(\Lambda)k$$

and for  $i > 0$   $|E_i^{j+1,1} - E_i^0 - \frac{k}{2} \dot{E}_i^j| \leq C(\Lambda)k$ . On the other hand, by (3.4), we see that  $|E_i^{j+1,1}| \leq C(\Lambda, E^0)$ . Now from

$$\begin{aligned} &(4\rho + 2\rho\mathcal{B}(E_1^j)k)(E_1^{j+1,1} - E_0^{j+1,1}) \\ &= 4\rho\mathcal{A}(E_1^j)h(E_1^j - E_1^{j+1,1}) - 2\rho\mathcal{D}(E_1^j)hk, \\ (6\rho + 3\rho\mathcal{B}(E_i^j)k)(E_i^{j+1,1} - E_{i-1}^{j+1,1}) &- (2\rho + \rho\mathcal{B}(E_i^j)k)(E_{i-1}^{j+1,1} - E_{i-2}^{j+1,1}) \\ &= 4\rho\mathcal{A}(E_i^j)h(E_i^j - E_i^{j+1,1}) - 2\rho\mathcal{D}(E_i^j)hk \end{aligned}$$

and from (B.3), it follows that  $|E_i^{j+1,1} - E_{i-1}^{j+1,1}| \leq C(\Lambda)h$ .

Assuming we know  $E^{j+1,n-1}$ , we compute the solution  $E^{j+1,n}$  to

$$M(E^{j+1,n-1})E^{j+1,n} = \mathcal{M}(E^{j+1,n-1}) \left( E^j + \frac{k}{2} \dot{E}^j \right) + F(E^{j+1,n-1}).$$

As before, we obtain

$$\begin{aligned} |E_i^{j+1,n}| &\leq C(\Lambda, E^0), |E_i^{j+1,n} - E_{i-1}^{j+1,n}| \leq C(\Lambda)h, \\ \left| E_i^{j+1,n} - \left( E_i^j + \frac{k}{2} \dot{E}_i^j \right) \right| &\leq C(\Lambda)k, |\dot{E}_i^{j+1,n}| \leq C(\Lambda). \end{aligned}$$

Taking to the limit as  $n \rightarrow \infty$  we conclude the proof.  $\square$

**Appendix C. Bounds on Jacobian matrices.**

LEMMA C.1. For  $K$  and  $R$  defined in section 3.1, we have

$$(C.1) \quad \text{Sup}_\xi \|DK(\xi)\|_2 \leq C(\Lambda, L)k,$$

$$(C.2) \quad \text{Sup}_\xi \|DR^l(\xi)\|_2 \leq C(\Lambda, L)k^2h, \quad l = 1, \dots, N.$$

*Proof.* We estimate the different terms appearing in these matrices.

(1) Term  $N(E)(T(E))^{-1}F_r(E)$ .

First, remark that  $c = (T(E))^{-1}F_r(E)$  is a solution to  $T(E)c = F_r(E)$  and  $|c| \leq C(\mathcal{A}, \mathcal{B}, \mathcal{D}, L)k$ . Therefore,  $c_{E_i}(E)$  is a solution to  $T(E)c_{E_i} = (F_r)_{E_i} - T_{E_i}c$ . This right-hand side is  $(0, \dots, 0, -2\rho hk\mathcal{D}'(E_i) + 4\rho k\mathcal{B}'(E_i)c_{i-2} - \rho k\mathcal{B}'(E_i)c_{i-1} - 3\rho k\mathcal{B}'(E_i)c_i - 4\rho h\mathcal{A}'(E_i)c_i, 0, \dots, 0)$ . From the comparison results in Appendix A we conclude that  $\|c_{E_i}\|_\infty \leq C(\mathcal{A}, \mathcal{B}, \mathcal{D}, L)(hk + k^2)$ , assuming  $hL, kL \leq \text{constant}$ .

We now differentiate  $N(E)$  with respect to  $E_i$  and get

$$(C.3) \quad \frac{u \cdot z(E)_{E_i}}{(1 - u \cdot z(E))^2} z(E)u^t + \frac{1}{1 - u \cdot z(E)} z(E)_{E_i}u^t.$$

As before,  $z(E)_{E_i}$  is a solution to  $T(E)z_{E_i} = v(E)_{E_i} - T_{E_i}z$ . This right-hand side is  $(0, \dots, 0, 4kC'(E_i) + \rho k\mathcal{B}(E_i)z_{i-2} - \rho k\mathcal{B}(E_i)z_{i-1} - 3\rho k\mathcal{B}(E_i)z_i - 4\rho h\mathcal{A}(E_i)z_i, 0, \dots, 0)$  plus a term of order  $\rho k\mathcal{B}'(E_i)$  if  $i = 1, 2$ . Thus  $\|z_{E_i}\|_\infty \leq C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)k$ . The elements of  $N_{E_i}$  are therefore of order  $C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)kh$ . In conclusion, the components of the vector  $(N(E)(T(E))^{-1}F_r(E))_{E_i}$  are of order  $C(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, L)(k^2 + kh)$ .

(2) Term  $\frac{u \cdot (T(E))^{-1}F(E)_r}{1 - u \cdot z(E)}$ .

Reasoning as in (1),  $|\frac{u \cdot (T(E))^{-1}F(E)_r}{1 - u \cdot z(E)}|_{E_i} \leq C(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})(k^2 + kh)$ .

(3) Term  $N(E)(T(E))^{-1}b(E)$ .

Since  $c = (T(E))^{-1}b(E)$  solves  $T(E)c = b(E)$ , we see that  $T(E)c_{E_i} = b_{E_i} - T_{E_i}c$ . Therefore,  $\|c_{E_i}\|_\infty \leq C(\mathcal{A}, \mathcal{B}, L)k^2$  for  $i \neq 1, 2$  and  $\|c_{E_i}\|_\infty \leq C(\mathcal{A}, \mathcal{B}, L)k$  for  $i = 1, 2$ .

We conclude that  $(N(E)(T(E))^{-1}b(E))_{E_i} \leq C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)k^2$  except for  $i = 1, 2$ , where we have order  $k$ .

(4) Term  $\frac{1}{1 - u \cdot z(E)}(u \cdot (T(E))^{-1}b(E))$ .

In this case  $|\frac{1}{1 - u \cdot z(E)}(u \cdot (T(E))^{-1}b(E))|_{E_i} \leq C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)k^2$  except for  $i = 1, 2$ , where we have order  $k$ .

(5) Term  $(I - N(E)Q(E))(E^{j-1} + \frac{k}{2}\dot{E}^{j-1})$ .

The columns  $c$  of  $Q(E)$  are solutions of the systems  $T(E)c = l$ , where  $l$  are columns of  $\frac{k}{2}\mathcal{D}_B(E)T$ . Therefore, the derivatives of the columns,  $c_{E_i}$ , are solutions to  $T(E)c_{E_i} = l_{E_i} - T_{E_i}c$ . The vector  $l_{E_i}$  is zero except for the columns  $c$  in the positions  $i, i-1, 1-2$ . We conclude that the elements of  $Q(E)_{E_i}$  are of order  $kh$ , except the elements in the columns  $i, i-1, i-2$ , which are of order  $k$ . Let us set  $e = E^{j-1} + \frac{k}{2}\dot{E}^{j-1}$ . From Appendix B, we know that  $e_i - e_{i-1} \sim C(\Lambda)h$ . Therefore,  $[Q(E)e]_{E_i}$  is of order  $kh^2$ , except for the columns  $i, i-1, i-2$  which are of order  $kh$ .

Now  $((I - N(E)Q(E))e)_{E_i} = -N(E)_{E_i}Q(E)e - N(E)(Q(E)e)_{E_i}$ . The elements of  $N(E)_{E_i}Q(E)e$  are of order  $k^2h$ . The elements of  $N(E)(Q(E)e)_{E_i}$

are of order  $kh^2$ , except for three columns. In conclusion, the elements of  $((I - N(E)Q(E))e)_{E_i}$  are of order  $C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)kh^2$ , except three columns of order  $kh$ .

(6) Term  $\frac{1}{1-u \cdot z(E)}(u \cdot Q(E))$ .

The elements of  $(\frac{1}{1-u \cdot z(E)}(u \cdot Q(E)))_{E_i}$  are of order  $C(\mathcal{A}, \mathcal{B}, \mathcal{C}, L)kh$ .

From (1), (2) we obtain (C.1). From (3), (4), (5), (6) we get (C.2).  $\square$

LEMMA C.2. *It holds that*

$$(C.4) \quad \|K(E^j) - K(\mathcal{E}^j)\|_2 \leq C(\Lambda, L)k\|E^j - \mathcal{E}^j\|_2,$$

$$(C.5) \quad \left\| (R(E^j) - R(\mathcal{E}^j)) \left( E^{j-1} + \frac{k}{2} \dot{E}^{j-1} \right) \right\|_2 \leq C(\Lambda, L)k^2\|E^j - \mathcal{E}^j\|_2.$$

*Proof.* It follows from Lemma C.1 and

$$(C.6) \quad K_i(E^j) - K_i(\mathcal{E}^j) = \nabla K_i(\xi)(E^j - \mathcal{E}^j),$$

$$\left( (R(E^j) - R(\mathcal{E}^j)) \left( E^{j-1} + \frac{k}{2} \dot{E}^{j-1} \right) \right)_i = \sum_l (R_{il}(E^j) - R_{il}(\mathcal{E}^j)) \left( E_l^{j-1} + \frac{k}{2} \dot{E}_l^{j-1} \right)$$

$$(C.7) \quad = \sum_l \nabla \left[ R_{il}(\xi) \left( E_l^{j-1} + \frac{k}{2} \dot{E}_l^{j-1} \right) \right] (E^j - \mathcal{E}^j). \quad \square$$

**Acknowledgment.** We thank Professor L. L. Bonilla for suggesting the problem and for fruitful discussions.

#### REFERENCES

- [1] M. J. BERGMANN, S. W. TEITSWORTH, L. L. BONILLA, AND I. R. CANTALAPIEDRA, *Solitary-wave conduction in p-type Ge under time-dependent voltage bias*, Phys. Rev. B, 53 (1996), pp. 1327–1335.
- [2] L. L. BONILLA AND I. R. CANTALAPIEDRA, *Universality of the Gunn effect: Self-sustained oscillations mediated by solitary waves*, Phys. Rev. E(3), 56 (1997), pp. 3628–3632.
- [3] L. L. BONILLA, I. R. CANTALAPIEDRA, M. J. BERGMANN, AND S. W. TEITSWORTH, *Onset of current oscillations in extrinsic semiconductors under dc voltage bias*, Semicond. Sci. Technol., 9 (1994), pp. 599–602.
- [4] L. L. BONILLA, P. J. HERNANDO, M. KINDELAN, M. A. HERRERO, AND J. J. L. VELÁZQUEZ, *Asymptotics of the trap-dominated Gunn effect in p-type Ge*, Phys. D, 108 (1997), pp. 168–190.
- [5] L. L. BONILLA, P. J. HERNANDO, M. KINDELAN, AND F. PIAZZA, *Determination of EL2 capture and emission coefficients in semi-insulating n-GaAs*, Appl. Phys. Lett., 74 (1998), pp. 988–990.
- [6] F. BREZZI, L. D. MARINI, AND P. PIETRA, *Numerical simulation of semiconductor devices*, Comput. Methods Appl. Mech. Engrg., 756 (1989), pp. 493–514.
- [7] Z. X. CHEN AND B. COCKBURN, *Analysis of a finite-element method for the drift-diffusion semiconductor-device equations—The multidimensional case*, Numer. Math., 71 (1995), pp. 1–28.
- [8] E. FATEMI, J. JEROME, AND S. OSHER, *Solution of the hydrodynamic device model using high-order nonoscillatory shock capturing algorithms*, IEEE Trans. Comput. Aided Des. Integr. Circuits Syst., 10 (1991), pp. 232–244.
- [9] J. B. GUNN, *Microwave oscillations in III-IV semiconductors*, Solid State Comm., 1 (1963), pp. 88–91.
- [10] E. ISAACSON AND H. B. KELLER, *Analysis of Numerical Methods*, John Wiley, New York, 1966.
- [11] J. W. JEROME AND C.-W. SHU, *Energy transport systems for semiconductors: Analysis and simulation*, in Nonlinear Analysts, V. Lakshmikantham, ed., de Gruyter, Berlin, 1996, pp. 3835–3846.
- [12] J. W. JEROME, *The approximation problem for drift-diffusion systems*, SIAM Rev., 37 (1995), pp. 552–572.

- [13] A. JÜNGEL AND P. PIETRA, *A discretization scheme for a quasi-hydrodynamic semiconductor model*, Math. Models Methods Appl. Sci., 7 (1997), pp. 935–955.
- [14] A. M. KAHN, D. J. MAR, AND R. M. WESTERVELT, *Dynamics of space-charge domains in ultrapure Ge*, Phys. Rev. Lett., 68 (1992), pp. 369–372.
- [15] L. REGGIANI, ED., *Hot-Electron Transport in Semiconductors*, Springer-Verlag, Berlin, 1985.
- [16] E. SCHÖLL, *Nonequilibrium Phase Transitions in Semiconductors*, Springer-Verlag, Berlin, 1987.
- [17] M. P. SHAW, V. V. MITIN, E. SCHÖLL, AND H. L. GRUBIN, *The Physics of Instabilities in Solid State Electron Devices*, Plenum Press, New York, 1992.
- [18] S. W. TEITSWORTH, R. M. WESTERVELT, AND E. E. HALLER, *Nonlinear oscillations and chaos in electrical breakdown in Ge*, Phys. Rev. Lett., 51 (1983), pp. 825–828.
- [19] S. W. TEITSWORTH, M. J. BERGMAN, AND L. L. BONILLA, *Space charge instabilities and nonlinear waves in extrinsic semiconductors*, in Nonlinear Dynamics and Pattern Formation in Semiconductors and Devices, F. J. Niedernostheide, ed., Springer, Berlin, 1995, pp. 46–69.