JOURNAL OF COMPUTATIONAL PHYSICS 90, 65–97 (1990)

Particle Simulations of the Semiconductor Boltzmann Equation for One-Dimensional Inhomogeneous Structures*

P. DEGOND AND F. GUYOT-DELAURENS

Centre de Mathématiques Appliquées, Unité Associée CNRS n° 756, Ecole Polytechnique, 91128 Palaiseau Cedex, France

Received January 24, 1989; revised July 6, 1989

This paper deals with the first application of the weighted particle method to the inhomogeneous semiconductor Boltzmann equation. In this method, the collision term is treated in a deterministic way. In a previous paper we reported the application of this method to a homogeneous case. This paper gives a detailed analysis of the inhomogeneous case. In particular, we investigate three different methods to perform the coupling of the particles trajectories with the Poisson equation. Two test problems are considered. The first one is "slightly inhomogeneous" and concerns damped plasma oscillations. The second one is "strongly" inhomogeneous and consists in the modelling of a one-dimensional N^+–N^-–N^+ structure, with a sharp doping discontinuity.

1. INTRODUCTION

We present numerical simulations of the semiconductor Boltzmann equation, using the weighted particle method [1, 2]. In this paper, one-dimensional inhomogeneous semiconductor structures are considered. The coupling of the Boltzmann and the Poisson equations is fully taken into account, but on the other hand, a simplified relaxation time model is considered for the collision operator. On the contrary, in a previous paper [2] (see also [3] for physical results), we provided a detailed analysis of the weighted particle method for physically realistic collision operators (e.g., polar optical and intervalley collision operators), but for a homogeneous structure in which the Poisson and the Boltzmann equations are decoupled. On the basis of these two investigations, a physically realistic numerical code for one-dimensional semiconductor structures will be developed soon.

Semiconductor device simulations are mostly based on the drift-diffusion equations (see [4] and references therein) or, for the most recent ones, on the hydrodynamic equations (see [5]). The drift-diffusion model is not suitable for the simulation of fast, submicron structures, because it does not take into account the gradients of the carrier temperature. The hydrodynamic model does, but, on the

* This work has been partially supported by the "Centre National d'Etudes des Télécommunications," under Grant 878B087 LAB/ICM/TOH and by the CNRS, under the "ATP-Mathématiques et Informatique."

0021-9991/90 $3.00

other hand, it uses some parameters (such as relaxation times for momentum or energy) which depend on the distribution function and which cannot be easily evaluated. Thus, it seems important to develop numerical codes based on the semiconductor Boltzmann equation, to overcome the difficulties encountered by the drift-diffusion and hydrodynamic models.

In particular situations, the semiconductor Boltzmann equation can be solved by analytical methods [6]. But, in most cases, a numerical solution is needed. The most widely used numerical method is certainly the Monte-Carlo method (cf. [7] and references therein), although other methods have been tried in particular geometries (cf. Rees' method [8]) or for particular collision operators (see the recent methods developped by Baranger [9] or Kuivalainen and Lindberg [10]). This paper is concerned with the analysis of new methods and algorithms which are somehow derived from the Monte-Carlo method, but which are expected to be an interesting alternative, as far as views of the distribution function and simulations of the transient regimes are concerned.

The weighted particle method was first introduced by G. H. Cottet, S. Mas-Gallic, and P. A. Raviart [11, 12], for viscous perturbations of the incompressible Euler equation. Then, the method was adapted to the treatment of collision terms in kinetic equations [1]. Its first application to the semiconductor Boltzmann equation has been done in [2], and an error analysis, relevant to this particular physical context, has been performed in [13]. Some physical applications can be found in [3]. In this method, the particles move along the characteristics of the convective (first-order differential) part of the equation, while the collision term varies the weights of the particles. The collision integral is evaluated by a discrete quadrature, where the particles play the role of the quadrature points.

In this paper, we try to optimize the coupling of the Poisson equation with the motion of the particles, in the presence of the sharp discontinuities of the ion background density which are typical of semiconductor device simulations. For that purpose, we have investigated three methods. The first one is the very classical "particle in cell" (PIC method [14, 15]. The second one is a modified PIC method where the grid density is computed by a quadrature of the distribution function, which itself is recovered from the particle distribution by interpolation. This method will be referred to as the "interpolation method." The last one uses an exact computation of the mutual Coulomb interaction between pairs of particles by means of the Green's function of Poisson's equation. These methods are detailed in Sections 3 and 4, after a brief description of the Boltzmann equation in Section 2.

Two test problems are considered in this paper. The first one deals with collision damped plasma oscillations and is intended to check the ability of the method to reproduce the high frequency behaviour of the semiconductor. The dispersion relation of plasma waves in presence of a relaxation time collision term is shortly derived and analyzed in Section 5. In Section 6, the corresponding numerical results are given. The comparisons show a very good agreement between the numerical and the analytical oscillation frequencies and damping rates, when the interpolation method is used. In the same situation, the PIC method develops mesh instabilities,

which are generated by the charge assignment procedure. We also investigate the influence of some numerical parameters.

The second test problem is a model for a real one-dimensional N^+–N^-–N^+ structure, with a very sharp discontinuity of the doping density ($N^- = 2.10^{15}\,\mathrm{cm}^{-3}$, while $N^+ = 10^{18}\,\mathrm{cm}^{-3}$). In this case, the Green's function method appears to be better than the two other methods. It allows the use of fewer particles and provides a good description of the lowly doped region, which is the physically interesting region. The numerical results which support these conclusions are detailed in Section 7.

2. The One-Dimensional Semiconductor Boltzmann Equation

We refer to [7] and references therein, for a detailed presentation of the physical aspects of the semiconductor Boltzmann equation. Our one-dimensional model for the semiconductor Boltzmann equation is written

$$\frac{\partial f}{\partial t} + v\,\frac{\partial f}{\partial x} - \frac{q}{m}\,E\,\frac{\partial f}{\partial v} = Q(f), \tag{1}$$

where $f(x, v, t)$ denotes the electron distribution function at position $x \in [0, L]$, velocity $v \in \mathbb{R}$, and time $t > 0$. q denotes the elementary charge, and m, the electron effective mass. The electric field $E(x, t)$, the electric potential $\phi(x, t)$ and the density $n(x, t)$ are related to f by the Poisson equation,

$$E(x, t) = -\frac{\partial \phi}{\partial x}\,(x, t) \tag{2}$$

$$-\frac{\partial^2 \phi}{\partial x^2} = \frac{q}{\varepsilon}\,(n_D(x) - n(x, t)) \tag{3}$$

$$n(x, t) = \int f(x, v, t)\,dv, \tag{4}$$

where the material permittivity ε and the doping density $n_D(x)$ are given. This set of equations has to be supplemented by an initial condition for (1),

$$f(x, v, 0) = f_0(x, v), \tag{5}$$

and boundary conditions for (1) and (3),

$$\begin{aligned} f(0, v, t) &= g_0(v, t) \qquad \text{for} \quad v \geqslant 0 \\ f(L, v, t) &= g_L(v, t) \qquad \text{for} \quad v \leqslant 0 \end{aligned} \tag{6}$$

$$\phi(0, t) = \phi_0(t); \qquad \phi(L, t) = \phi_L(t) \tag{7}$$

with f_0, g_0, g_L, ϕ_0, ϕ_L suitably given. Finally, the collision operator is written

$$Q(f)(x, v, t) = \int (s(x, v', v) f(x, v', t) - s(x, v, v') f(x, v, t)) \, dv'. \tag{8}$$

For our numerical experiments, we have taken a relaxation time model,

$$s(x, v, v') = \frac{1}{\tau} M(v'), \tag{9}$$

where τ is a constant relaxation time and $M(v)$ is the Maxwellian associated with the lattice thermal velocity v_{th}:

$$M(v) = \frac{1}{\sqrt{2\pi} v_{\text{th}}} \exp\left(-\frac{v^2}{2v_{\text{th}}^2}\right). \tag{10}$$

Therefore we neglect both the angular dependence of the scattering rates and the discrete character of the energy losses in the phonon interaction. This simplification is obviously unrealistic, but on the other hand, it allows analytic computations (see, in particular, Section 5), which are very useful to analyse the algorithm. The conclusion of the present study will be transposed to more realistic collision models in a forthcoming paper.

3. The Numerical Method Presentation

Our numerical method is based on the weighted particle method (cf. [1, 11–13] for other presentations and applications of the weighted particle method). The distribution function is approximated by a sum of Delta measures,

$$\begin{aligned} f(x, v, t) &\approx f_h(x, v, t) \\ &= \sum_{i=1}^{N} \omega_i f_i(t) \, \delta(x - x_i(t)) \otimes \delta(v - v_i(t)), \end{aligned} \tag{11}$$

where $x_i(t)$, $v_i(t)$, $f_i(t)$, and ω_i are respectively the position, velocity, weight, and control volume of the ith particle. They evolve in time according to

$$\frac{dx_i}{dt} = v_i; \qquad x_i(0) = x_i^0 \tag{12a}$$

$$\frac{dv_i}{dt} = -\frac{q}{m} E_i(t); \qquad v_i(0) = v_i^0 \tag{12b}$$

$$\frac{df_i}{dt} = Q_i(t); \qquad f_i(0) = f_i^0 \tag{13}$$

$$\omega_i(t) = \omega_i^0,$$

where $E_i(t)$ and $Q_i(t)$ are the approximations of the electric field and of the collision operator acting on the ith particle. The initial x_i^0, v_i^0, f_i^0, and ω_i^0 are chosen so that

$$f_0(x, v) \approx \sum_{i=1}^{N} \omega_i^0 f_i^0 \, \delta(x - x_i^0) \, \delta(v - v_i^0). \tag{14}$$

The approximations (11) and (14) are taken in the weak topology of measures. The time differential system (12), (13) can be solved by any classical scheme. In our computations, we used the order-2 Adams–Bashforth scheme.

To define $Q_i(t)$, we introduce a cutoff function $\zeta_\alpha(x)$ such that

$$\zeta_\alpha(x) = \zeta\left(\frac{x}{\alpha}\right); \qquad \zeta(-x) = \zeta(x); \qquad \int \zeta(x)\, dx = 1, \tag{15}$$

where ζ is a compactly supported function. We write (omitting the t-dependence of x_i and v_i):

$$\begin{aligned} Q(f)(x_i, v_i, t) &= \int [s(x_i, v', v_i) f(x_i, v', t) - s(x_i, v_i, v') f(x_i, v_i, t)] \, dv' \\ &= \iint [s(x', v', v_i) f(x', v', t) - s(x_i, v_i, v') f(x_i, v_i, t)] \, \delta(x' - x_i) \, dx' \, dv' \\ &\simeq \iint [s(x', v', v_i) f(x', v', t) - s(x_i, v_i, v') f(x_i, v_i, t)] \, \zeta_\alpha(x' - x_i) \, dx' \, dv' \\ &\simeq \sum_{j=1}^{N} [s(x_j, v_j, v_i) f_j(t) - s(x_i, v_i, v_j) f_i(t)] \, \zeta_\alpha(x_j - x_i) \, \omega_j. \end{aligned} \tag{16}$$

Therefore we let

$$Q_i(t) = \sum_{j=1}^{N} [s(x_j, v_j, v_i) f_j(t) - s(x_i, v_i, v_j) f_i(t)] \, \zeta_\alpha(x_j - x_i) \, \omega_j \tag{17}$$

In our computations, we use the "hat function" W_2 (see definition at (19)) as cutoff function ζ.

The method (17) for the computation of the collision operator has first been proposed by S. Mas-Gallic [1]. She proved an error estimate for the field free ($E = 0$) Boltzmann equation, showing that

$$|f(x_i, v_i, t) - f_i(t)| \leqslant C\left(\alpha^m + \left(\frac{h}{\alpha}\right)^k\right). \tag{18}$$

The first error term comes from the replacement of the delta measure by ζ_α in (16),

and m is related to some moment properties of ζ. The second term comes from the discrete quadrature (h is the mean interparticle distance and k is related to the smoothness of the cutoff and of the data). From (18), there is an optimal α, which minimizes the total error.

4. The Poisson Solver

For the approximation of $E_i(t)$, we considered three methods: first, the classical "particle in cell" (PIC) method [14, 15]; second, a new method, which relies on the interpolation of the distribution function; and third, the direct computation of the mutual Coulomb interaction between pairs of particles, using the Green's function of the Poisson's equation.

In the PIC method, one considers fixed equally spaced mesh points $X_m = m\,\Delta x$, $m = 0, ..., M_x$, and an interpolation function $W(x)$. The approximation of the density $n(x, t)$ can be defined at the mesh points n_m, by using the charge assignment procedure:

$$n_m(t) = \sum_{j=1}^{n} \omega_j f_j(t) \frac{1}{\Delta x} W\left(\frac{x_j(t) - X_m}{\Delta x}\right).$$

Then, one solves Poisson's equation by finite differences on the grid X_m (where ϕ_m and E_m denote approximations of $\phi(X_m)$ and $E(X_m)$):

$$\phi_{m+1} - 2\phi_m + \phi_{m-1} = -\frac{q\,\Delta x^2}{\varepsilon}(n_D(X_m) - n_m(t))$$

$$E_m = -(\phi_{m+1} - \phi_{m-1})/2\,\Delta x.$$

Then, the electric field is interpolated at the location of the ith particle:

$$E_i(t) = \sum_{m=0}^{M_x} E_m W\left(\frac{x_i(t) - X_m}{\Delta x}\right).$$

Generally, B-splines are used for the interpolation function W,

$$W = W_p = \chi^{*p}, \tag{19}$$

where χ is the characteristic function of $[-\frac{1}{2}, \frac{1}{2}]$. In our computations, we used the "hat-function" W_2. The PIC method has been widely used and many useful and interesting considerations are developed in [14–16]. The reader will also find an extended bibliography in these references. The error analysis has been performed in [17]; the error is similar to (18) with α replaced by Δx, and ζ by W.

To define the interpolation method, we introduce a 2-dimensional mesh ($X_m = m\,\Delta x$, $V_p = p\,\Delta v$) with $m = 0, ..., M_x$ and $p = -M_v, ..., 0, ..., M_v$, and a

2-dimensional interpolation function $W(x, v)$. A grid approximation $f_{m,p}$ of the distribution function is defined by:

$$f_{m,p} = \left(\sum_{j=1}^{N} W(x_j - X_m, v_j - V_p) f_j \omega_j \right) \Big/ \left(\sum_{j=1}^{N} W(x_j - X_m, v_j - V_p) \omega_j \right). \quad (20)$$

Since the volumes ω_j are independent of j, they cancel in formula (20). However, the denominator of (20) can be interpreted as the "local volume" of the jth particle. In this sense, this method is a close relative of Brackbill and Ruppel's method [18], which uses a similar "local volume" evaluation. Then, the grid density is recovered by the numerical integration of $f_{m,p}$ on the v mesh:

$$n_m = \sum_{p=-M_v}^{M_v} f_{m,p} \Delta v.$$

So far, the method is not conservative. We force the conservativity by letting

$$n'_m = \alpha n_m; \qquad \alpha = \left(\sum_{j=1}^{N} \omega_j f_j \right) \left(\sum_{m=0}^{M_x} n_m \Delta x \right)^{-1}$$

and taking n'_m as the grid density. The method proceeds exactly as in the PIC method. In our computations, we have taken $W(x, v) = W_2(x)\, W_2(v)$. As far as we know, the interpolation method (20) is new, and no error analysis has yet been performed. However, an error estimate similar to (18) should be expected.

The Green's function method relies on an exact representation of the mutual Coulomb interaction between particles. In this one-dimensional problem, this is particularly simple: first, the electric field generated by a one-dimensional particle (physically a charged plane) is a bounded function; second, a very simple "fast" algorithm is obtained once the particles are ordered by increasing positions. From a direct integration of (11) with respect to v, we get a particle representation of the charge density according to:

$$\rho(x, t) \approx \rho_h(x, t) = q \left[n_D(x) - \sum_{j=1}^{N} \omega_j f_j \delta(x - x_j(t)) \right]. \quad (21)$$

We denote by Q and $U(t)$ the following quantities:

$$Q = \int \rho_h(x, t)\, dx = q \left[\int n_D(x)\, dx - \sum_{j=1}^{N} \omega_j f_j(t) \right]$$

$$U(t) = \phi_L(t) - \phi_0(t).$$

From (21), we have the exact representation of the electric field by

$$E_h(x, t) = \frac{q}{\varepsilon} \left[\int K(x, y)\, n_D(y)\, dy - \sum_{j=1}^{N} \omega_j f_j(t)\, K(x, x_j(t)) \right] + \frac{1}{L} \left[\frac{Qx}{\varepsilon} - U(t) \right], \quad (22)$$

where $K(x, y)$ is given by

$$K(x, y) = -\frac{\partial G}{\partial x}(x, y)$$

and $G(x, y)$ is the Green's function of the Poisson equation with periodic boundary conditions:

$$-\frac{\partial^2}{\partial x^2} G(x, y) = \delta(x-y) - \frac{1}{L}; \qquad G(0, y) = G(L, y) = 0.$$

K has a jump at $x = y$ of magnitude 1 and, from symmetry considerations, we can let $K(y, y) = 0$. Thus the electric field $E_i(t)$ acting on particle i is given by

$$E_i(t) = \frac{q}{\varepsilon}\left[\int K(x_i(t), y)\, n_D(y)\, dy - \sum_{j=1}^{N} \omega_j f_j(t)\, K(x_i(t), x_j(t))\right] + \frac{1}{L}\left[\frac{Qx_i(t)}{\varepsilon} - U(t)\right]. \tag{23}$$

The second bracket and the convolution term are easy to compute. The most costly part of this method is the computation of the summation term.

In the collisionless case, the convergence of this method has been proven by Trocheris [19] in the topology of bounded measures provided with a "discrepancy" metric. Better convergence results in stronger topologies are available (Cottet and Raviart [20]) by use of a regularization procedure; we introduce a cutoff function $\zeta_\alpha(x)$ which satisfies the properties (15) (but the smoothing parameter α may be different for the collision term and for the electric field). We let

$$K^\alpha(x, y) = (K(\cdot, y) *_x \zeta_\alpha)(x), \tag{24}$$

where in (24), the periodic extension of K for x outside the interval $[0, L]$ is considered, so that K^α is again periodic with period L. Now the smoothed electric field reads

$$E_i^\alpha(t) = \frac{q}{\varepsilon}\left[\int K(x_i(t), y)\, n_D(y)\, dy - \sum_{j=1}^{N} \omega_j f_j(t)\, K^\alpha(x_i(t), x_j(t))\right] + \frac{1}{L}\left[\frac{Qx_i(t)}{\varepsilon} - U(t)\right]. \tag{25}$$

The Green's function method has not yet been widely used in practice, at least in semiconductor or plasma physics, because its computational cost grows quadratically with the number of particles. However, recently, Greengard and Rokhlin [21] have derived a fast algorithm with a linear complexity. Their algorithm solves the 2-dimensional periodic problem, by use of a multipole expansion. In one dimen-

sion, a more costly, but easier to implement algorithm can be obtained by use of a fast sorting algorithm. The problem is the evaluation of the term

$$F_i(t) = \sum_{j=1}^{N} \omega_j f_j(t)\, K(x_i(t), x_j(t)).$$

We suppose that the particles have been indexed by increasing positions and that there are no pairs of particles with the same position (this latter situation is exceptional and arises with probability zero). Thus, we suppose

$$i < j \Rightarrow x_i < x_j.$$

From the exact expression of K, we obtain

$$F_i(t) = S_i(t) + T_i(t)$$

$$S_i(t) = \sum_{j<i} \omega_j f_j(t) \left(\frac{x_j - x_i}{L} + \frac{1}{2} \right); \qquad T_i(t) = \sum_{j>i} \omega_j f_j(t) \left(\frac{x_j - x_i}{L} - \frac{1}{2} \right).$$

We have the recursion formulae,

$$S_i = S_{i-1} + A_{i-1} \left(\frac{x_{i-1} - x_i}{L} \right) + \frac{1}{2} \omega_{i-1} f_{i-1};$$

$$T_i = T_{i+1} + B_{i+1} \left(\frac{x_{i+1} - x_i}{L} \right) - \frac{1}{2} \omega_{i+1} f_{i+1},$$

with

$$A_i = \sum_{k=1}^{i} \omega_k f_k; \qquad B_i = \sum_{k=i}^{N} \omega_k f_k.$$

The computation of A_i, B_i, S_i, and T_i can be performed in order N operations. When the regularization procedure is used, we consider the electric field as as perturbation of the unsmoothed one. Since the cutoff function is compactly supported (again, we choose a B-spline W_1 or W_2), the perturbing electric field acting on a pair of particles of positions x_i and x_j is vanishing as soon as $|x_j - x_i| \geqslant A\alpha$, where A is such that Supp $\zeta \subset [-A, A]$. A preselection of such pairs of interacting particles is performed by localizing the particles with respect to a fixed mesh of mesh spacing A. This allows the computation of the perturbing electric field in order NM operations, where M is the mean number of particles per cell. The same algorithm is used to compute the collision operator (17). This is a rather classical procedure [14].

5. Plasma Oscillations; Analysis

The simulation of plasma oscillations is one of the most fundamental test problems for particle methods and has received a great deal of attention from plasma physicists, in the case of collisionless models [14–16]. In this paragraph, we analyse the dispersion relation of plasma waves, in presence of a relaxation time collision term (8), (9).

The dispersion relation of plasma waves is found by linearizing the Poisson–Boltzmann equation about a state of equilibrium with vanishing electric field [22]

$$f_{\text{eq}}(x, v) = n_D\, M(v) \qquad \text{with} \quad n_D \text{ independent of } x.$$
$$\phi_{\text{eq}}(x) = 0; \qquad E_{\text{eq}}(x) = 0.$$

The linearized equation is then solved by looking for plane wave solutions. Their frequency ω and their wave-vector k are related by the plasma dispersion relation,

$$D(k, \omega; \tau) = 0,$$

with

$$D(k, \omega; \tau) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{M'(v)\, dv}{v - z} + \frac{i}{\tau k} \int_{-\infty}^{\infty} \frac{M(v)\, dv}{v - z}. \tag{26}$$

Here, τ denotes the relaxation time (9) used in the collision operator (8); $\omega_p^2 = q^2 n_D / \varepsilon m$ is the plasma frequency and z is a complex phase velocity:

$$z = \frac{\omega}{k} + \frac{i}{\tau k}.$$

In the limit $\tau \to \infty$, we recover the classical plasma dispersion relation [22]. By expanding the Cauchy integrals in powers of v_{th}/z, and neglecting the terms of order $(v_{\text{th}}/z)^3$, we obtain

$$D(k, \omega; \tau) \simeq 1 - \frac{i}{\tau k v_{\text{th}}} \left(\frac{v_{\text{th}}}{z}\right) - \frac{1}{k^2 \lambda_D^2} \left(\frac{v_{\text{th}}}{z}\right)^2 = 0, \tag{27}$$

where $\lambda_D = v_{\text{th}}/\omega_p$ is the Debye length. The dispersion relation (27) is that of the classical damped oscillator. It thus exhibits two regimes with a threshold for $2\omega_p \tau = 1$:

First regime. Damped oscillations for $2\omega_p \tau \geqslant 1$, with (complex) frequency:

$$\omega = \omega_p \left[\pm \left(1 - \frac{1}{(2\omega_p \tau)^2}\right)^{1/2} - \frac{i}{2\omega_p \tau} \right] \tag{28}$$

Second regime. Exponential decay for $2\omega_p\tau \leqslant 1$ at rates:

$$\omega_{\pm} = -\frac{i}{2\tau}\,[1 \pm (1-(2\omega_p\tau)^2)^{1/2}]. \tag{29}$$

On Fig. 1, we plot the variations of $\omega_r = |\mathrm{Re}\,\omega|$ and $\omega_i = \mathrm{Im}\,\omega$ (for the first regime) and ω_+ and ω_- (for the second regime) as a function of ω_p. We have supposed that τ is constant, and τ^{-1} is chosen as a frequency unit. Indeed, in semiconductor physics, τ is almost constant (and related to the low field mobility μ by $\tau = q^{-1}m\mu$), whereas ω_p depends monotonously on the doping density n_D. Thus, varying ω_p is like varying the doping density, but in a normalized way so that the conclusions of the analysis are independent of the material parameters. For comparison, we also plot the dielectric frequency $\omega_d = \omega_p^2\tau$, found by linearization of the drift-diffusion model.

In the first regime (damped oscillations), the actual oscillation frequency ω_r is asymptotically (and very rapidly) equal to the plasma frequency ω_p as $\omega_p\tau$ increases. They differ in the range $\frac{1}{2} \leqslant \omega_p\tau \leqslant 1$, and ω_p quickly goes to 0 as $\omega_p\tau$ tends to the threshold value $\frac{1}{2}$. The damping rate remains constant, and, outside the range $\frac{1}{2} \leqslant \omega_p\tau \leqslant 1$, is more than one order of magnitude smaller than the oscillation frequency. Furthermore, the dielectric frequency ω_d strongly overestimates both time scales. The drift-diffusion model predicts a very sharp exponential decay where the actual behaviour is that of weakly damped high frequency oscillations.

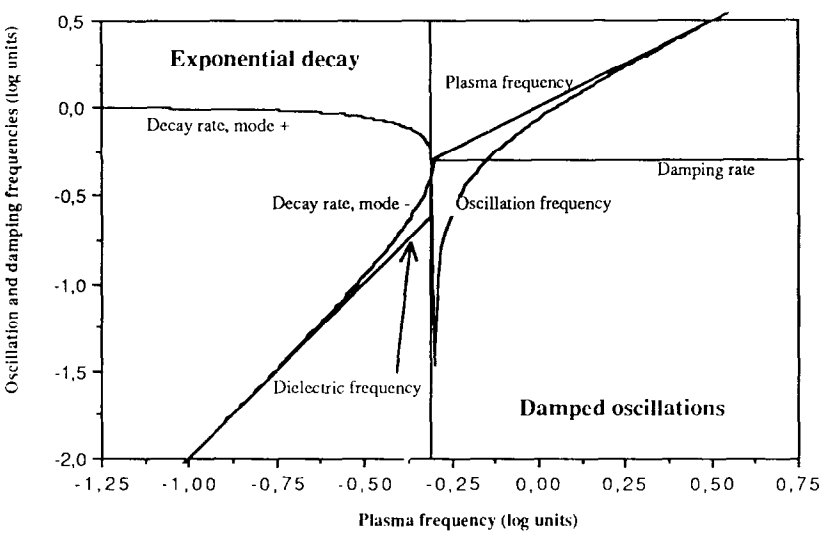

FIG. 1. Plots of $\omega_r = |\mathrm{Re}\,\omega|$ (oscillation frequency), $\omega_i = \mathrm{Im}\,\omega$ (damping rate) in the case 1 (damped oscillations); of ω_+ (decay rate, mode +) and ω_- (decay rate, mode −) in the case 2 (exponential decay); and of ω_p (plasma frequency) and ω_d (dielectric frequency) in log units, as a function of ω_p; the frequency unit is τ^{-1}.

TABLE I

Values of the Densities (in cm^{-3}) Corresponding to the Left and Right Bounds of the Transition Region and to the Threshold ($2\omega_p\tau = 1$), for Three Materials

	Left bound of *TR*	Threshold	Right bound of *TR*
Si	1.0×10^{15}	1.4×10^{15}	6.0×10^{15}
Ga As	4.7×10^{14}	7.0×10^{14}	2.7×10^{15}
In As	1.0×10^{14}	1.4×10^{14}	5.8×10^{14}

Note. The physical constants are taken from [23].

In the second regime (exponential decay), the dominant mode is ω_-, since its decay rate is much smaller than ω_+. It is very quickly asymptotically equal to the dielectric frequency ω_d as $\omega_p\tau$ tends to 0, which proves the validity of the drift-diffusion model in this regime. Outside the region $0.3 \leqslant \omega_p\tau \leqslant \frac{1}{2}$, the two damping rates ω_+ and ω_- have significantly different orders of magnitude. Inside this region, ω_- differs from ω_d, and both ω_+ and ω_- tend to $(2\tau)^{-1}$ as $\omega_p\tau$ tends to $\frac{1}{2}$.

From this analysis, we may define the "transition region" between the two regimes by $0.3 \leqslant \omega_p\tau \leqslant 1$. Outside this transition region, two very different time scales (ω_+ and ω_- on the one hand, ω_r and ω_i on the other hand) are present and lead to a stiff numerical problem. Values of the densities corresponding to the transition region for three semiconductor materials (Si, Ga As, In As) are given in Table I. These densities are quite low, and the range of the transition region is very narrow. Thus, in most situations, the electron gas is in a damped oscillation regime, and the validity of the drift-diffusion model (even for silicon) is highly questionable, at least for transient simulations. Table I also shows that the plasma frequency of the most highly doped region is the limiting time scale for many semiconductor simulations. Let us remark that a fluid model (in which the time derivatives of the momentum are not neglected) would have provided the same conclusions as the kinetic model, so far

Finally, this analysis does not account for the acoustic propagation of plasma waves, which can be found by expanding (27) to higher order terms. Also, the expansion (27) is not uniformly valid as $\tau \to \infty$. Indeed, when $\tau = \infty$, the residuals of the Cauchy integrals have to be added, which leads to Landau damping [22].

6. Plasma Oscillations: Simulations

For simulating plasma oscillations, periodic boundary conditions for f, ϕ, and E have been prescribed. The length L has been adjusted to be equal to one wavelength, so that $k = 2\pi$, and by the choice of suitable units, we let

$$\omega_p^2 = 1; \qquad n_D = 1; \qquad L = 1.$$

The only "free parameters" left in the problem are $\omega_p\tau$, $k\lambda_D = 2\pi v_{\text{th}}$, and the magnitude χ of the initial perturbation, such that the initial data are given by

$$f_0(x, v) = M(v)(1 + \chi \sin(2\pi x)); \qquad 0 \leqslant \chi \leqslant 1. \tag{30}$$

The object of the discussion is twofold. First, illustrate the considerations of the previous section and second, identify the effect of the numerical parameters on the accuracy of the computation.

On Figs. 2a and b, we plot the potential energy as a function of the time, for

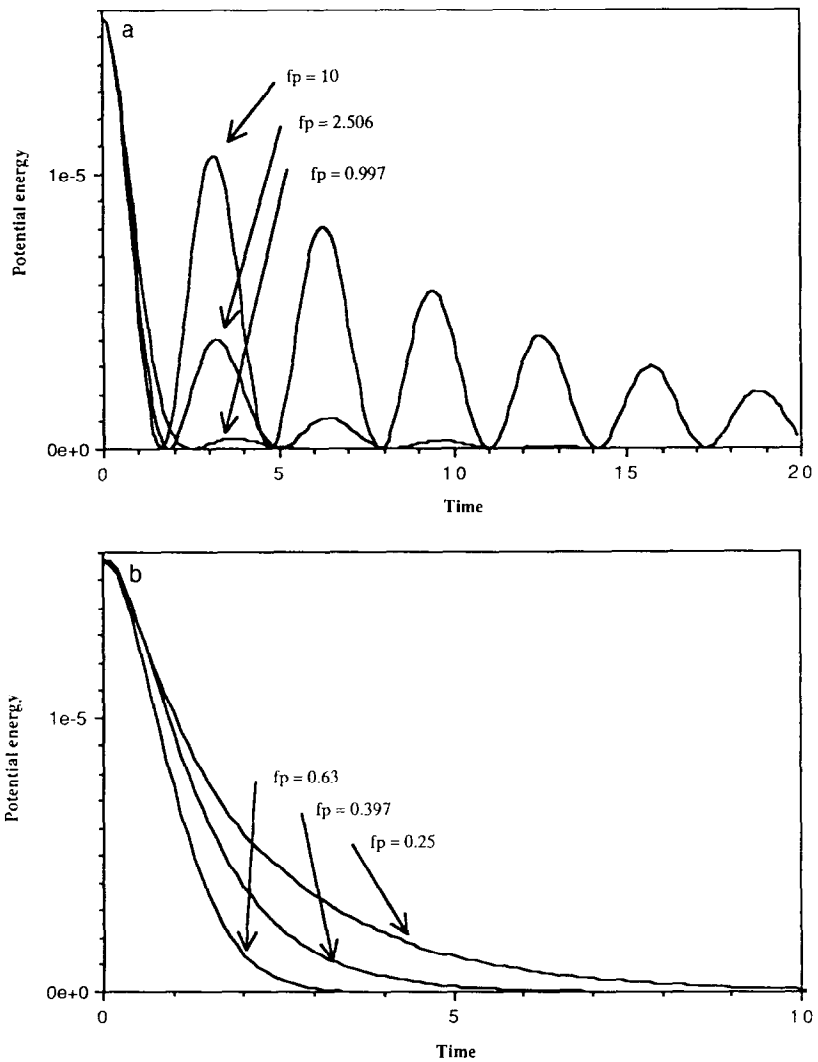

FIG. 2. Potential energy versus time for six values of $\omega_p\tau$ (denoted on the figure by fp), for the same values of $k\lambda_D = 0.1$ and $\chi = 0.05$.

TABLE II

Comparison between the Predicted and the Observed Half Period of the Oscillations for the Runs Displayed on Fig. 2a

$\omega_p\tau$	$T/2$ (observed)	$T/2$ (predicted)	Relative error (%)
10	3.13	3.146	0.5
2.506	3.2	3.206	0.2
0.997	3.65	3.631	0.5

several values of $\omega_p\tau$. In these simulations, the interpolation method has been used for the computation of the electric field. As $\omega_p\tau$ decreases from 10 ($\log_{10}(\omega_p\tau)=1.$) to 2.506 ($\log_{10}(\omega_p\tau)=0.4$), then to 0.997 ($\log_{10}(\omega_p\tau)=0.$), the oscillations are more strongly damped and their frequency slightly decreases, according to the prediction (28). In Table II we give the comparisons between the theoretical half period $T/2=\pi/\omega_r$ (from formula (28)) and the observed one (from Fig. 2). We do a similar comparison in Table III for the decrement per period, $\delta=\exp -\omega_i T$. The agreement is totally satisfactory.

For $\omega_p\tau=0.63$ ($\log_{10}(\omega_p\tau)=-0.2$), we are still in the oscillatory regime but just above the threshold with the exponential decay regime, and thus the second bounce of the oscillations is no more perceivable. Then, for $\omega_p\tau=0.397$ ($\log_{10}(\omega_p\tau)=-0.4$) and $\omega_p\tau=0.25$ ($\log_{10}(\omega_p\tau)=-0.6$), we step in the exponential decay zone. In that case, the perturbation of the density is a combination of the two eigenmodes $n_+ e^{-\omega_+ t}$ and $n_- e^{-\omega_- t}$. The initial amplitudes n_+ and n_- can be deduced from the continuity equation, and from the fact that f_0 has mean velocity zero (see (30)). They are connected by $n_+/n_-=\omega_-/\omega_+$. Then, the potential energy is given in suitable units by

$$U(t)=e^{-2\omega_- t}+\frac{2n_+}{n_-}e^{-(\omega_-+\omega_+)t}+\left(\frac{n_+}{n_-}\right)^2 e^{-2\omega_+ t}.$$

Since n_+/n_- is small the first term rapidly becomes dominant and the decay rate ω_- can be measured from the numerical simulation. In Table IV we display

TABLE III

Comparison between the Predicted and the Observed Decrement per Period of the Oscillations for the Runs Displayed on Fig. 2a

$\omega_p\tau$	δ (observed)	δ (predicted)	Relative error (%)
10	0.71	0.730	2.7
2.506	0.26	0.278	6.5
0.997	0.023	0.0262	12.2

TABLE IV

Comparisons between the Predicted and the Observed Decay Rates for the Runs Displayed on Fig. 2b

$\omega_p\tau$	ω_+	ω_-	n_+/n_-	Measured ω_-	Error on ω_- (%)
0.397	2.03	0.494	0.24	0.49	0.8
0.25	3.73	0.268	0.071	0.24	10.4

predicted values of ω_+, ω_-, and n_+/n_- according to formula (29) and the measured values of ω_- from Fig. 2b. The agreement is again very satisfactory.

Figure 3 is intended to illustrate the competition between Landau damping and collisional damping. The higher value of $k\lambda_D$ leads to a Landau damping constant $\omega_p\tau_L \simeq 20$ (see the formula for Landau damping in [22, p. 385]), whereas the collisional damping constant is $\omega_p\tau = 50$. The non-monotonous decay of the amplitude of the oscillation, as appearing on Fig. 3, is typical of nonlinear Landau damping [22, Section 10.7]. It originates in the trapped particles, moving at the same velocity as the wave phase velocity, which oscillate in the sinusoidal potential well, at the so-called bounce frequency ω_B. ω_B can be evaluated by approximating the sinusoidal well by a harmonic well, which gives (in our units):

$$\omega_B \simeq \sqrt{\chi}\,\omega_P. \tag{31}$$

As χ goes from 0.05 (Fig. 3.b) to 0.2 (Fig. 3a) the corresponding bounce frequency as measured on the figure, is multiplied by a factor 2.5, which satisfactorily agrees with (31). On the other hand, the measured absolute magnitude of the bounce period significantly differs from (31), but this is due to the harmonic approximation. After a few tens of plasma periods, the collisional damping takes over and the amplitude of the oscillations then decays monotonously.

We now comment on some comparisons between the PIC method and the interpolation method. In Fig. 4, we show the results of both methods on the same simulation cases, the data of which are summarized on Table V; (we recall that we denote by N the number of particles, and by Δx, Δv, the size of the cells in x and in v).

TABLE V

Simulation Parameters for Fig. 4

$\omega_p\tau$	∞ (Fig. 4a); 10 (Fig. 4b)	N	2000
$k\lambda_D$	0.314	$\Delta x/\lambda_D$	0.8
χ	0.05	$\Delta v/v_{th}$ (interpolation)	0.4

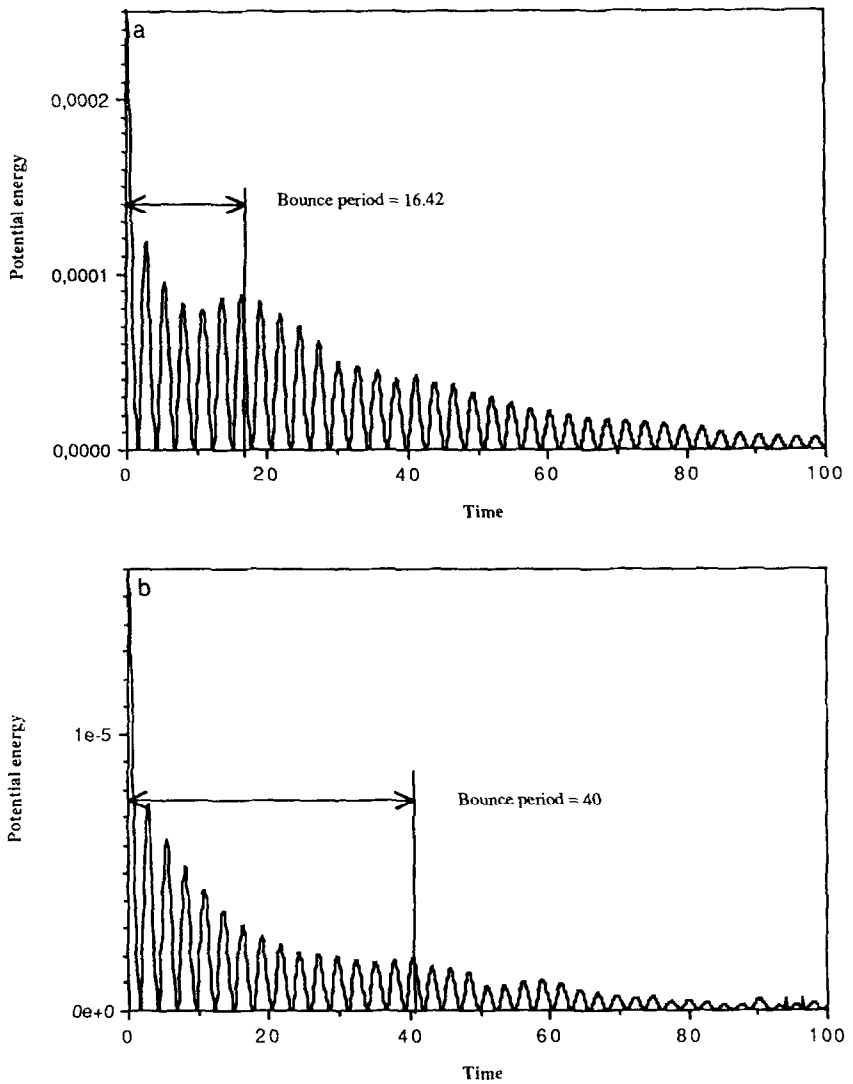

FIG. 3. Potential energy vs time for two values of χ. $\chi = 0.2$ (Fig. 3a); $\chi = 0.05$ (Fig. 3b) for $k\lambda_D = 0.3$ (i.e., $\omega_p \tau_L = 20$) and $\omega_p \tau = 50$.

The PIC method develops an instability after six plasma periods, whereas the interpolation method remains stable. We noticed that the development of the instability is delayed and its magnitude is decreased when the ratio N/M_x is increased. On the other hand, the magnitude of the instability is much higher in the collisional case (Fig. 4b) than in the collisionless case (Fig. 4a), which is surprising since the collision term should (at least partially) dissipate the energy of the instability. This instability is associated with mesh oscillations, which appear when

bunches of particles oscillate from one cell to the next. They are generated by the random fluctuations of the mesh-defined density, due to the charge assignment and field interpolation procedure. In the collisional case, once the physical oscillation is damped, the random fluctuations remain the only perturbation of the equilibrium, and mesh oscillations can develop freely. In the collisionless case, however, the physical plasma oscillations have constant energy and can interfere with the mesh oscillations, which may decrease the amplitude of the instability. Similarly, when the ratio N/M_x is increased, the fluctuations are decreased, and mesh oscillations cannot develop so easily. The interpolation method remains stable because the charge assignment procedure is less noisy than in the PIC method and does not allow the development of mesh oscillations.

Figure 4c shows comparisons of the mesh-defined density at time 45, for the PIC method and for the interpolation method. For reference, the amplitude of the initial perturbation is also shown. The mesh oscillations generated by the PIC method are clearly visible. The interpolation method is still noisy, but the amplitude of the fluctuations is smaller by a factor of 10.

We now examine the influence of the numerical parameters. Since this analysis is well known for the PIC method [14–16], we restrict our discussion to the interpolation method and to the evaluation of the collision operator. The effect of a variation of the smoothing length α for the discrete collision operator (17) is

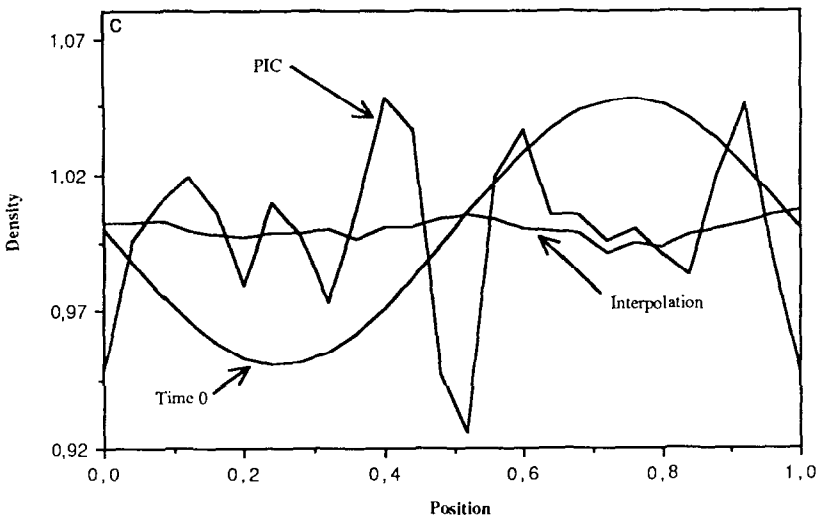

Figure 4—*Continued*

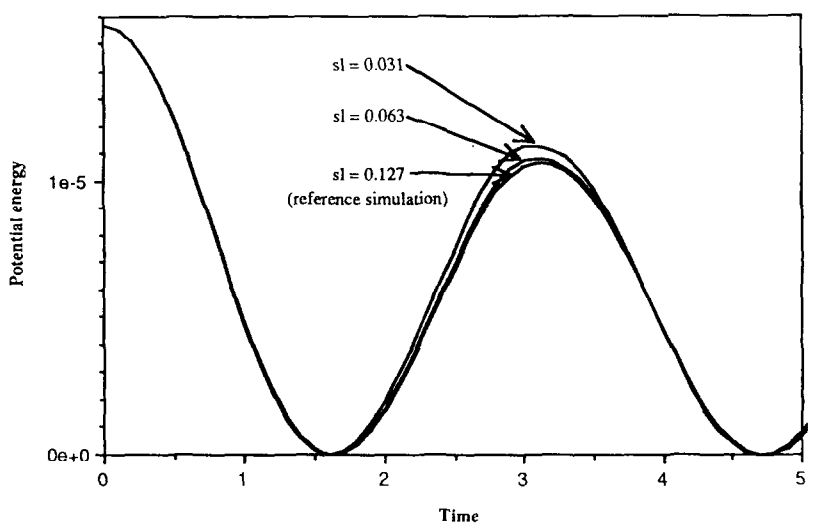

FIG. 5. Potential energy versus time for three values of the ratio α/λ_D (denoted by sl on the figures). The average number of particles per smoothing length is respectively 8 (for sl = 0.127), 4 (for sl = 0.063), and 2 (for sl = 0.031).

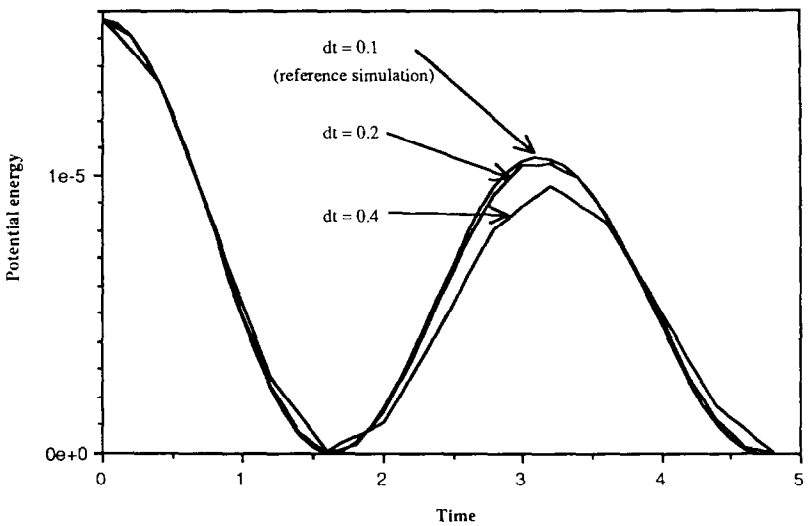

FIG. 6. Potential energy versus time for three values of the product $\omega_p \Delta t$ (denoted by dt on the figure).

PIC or interpolation solver. Therefore, any effort to improve the overall method should be directed towards the Poisson solver rather than towards the collision solver. This will even be more apparent in the context of the next example and will motivate the investigation of the Green's function Poisson solver.

7. AN N^+–N^-–N^+ STRUCTURE

This section deals with a test problem presented in [9] for which the doping profile is given by

$$n_D(x) = \begin{cases} n_D^+ & \text{for } 0 \leqslant x \leqslant x_1 \\ n_D^- & \text{for } x_1 \leqslant x \leqslant x_2 \\ n_D^+ & \text{for } x_2 \leqslant x \leqslant L \end{cases}$$

with n_D^+/n_D^- large. The typical value of this ratio in most numerical simulations is 500 [9, 24] or around (1000 in [25], 100 in [26], ...). As classical in semiconductor physics the behaviour of the device is dominated by the dynamics of the carriers in the N^- region [23] and the numerical methods should give a precise description of this region. This is a difficult problem because the numerical errors on the electron density in the N^+ regions reach the same orders of magnitude as the density itself in the N^- region, which strongly perturbs the overall potential. Another numerical problem is related to the very peaked profile of the electric field. If the trajectories solver is not accurate enough, a fast particle is very likely to jump over the peak of the electric field.

In Monte-Carlo methods, these problems are all the more difficult to solve that the particles have equal weights. Thus, the number of particles per unit length is in the same ratio as the doping densities. This "undersampling" of the N^- region needs to be balanced by averages over very long time intervals which do not allow a correct description of the transient regimes. In the weighted particle method however, the differences in the doping densities are absorbed in the differences of the weights, and the number of particles per unit length can be the same in both regions. By use of the Green's function solver, we can obtain a quite nice picture of the instantaneous electric field in the N^- region, for a very moderate number of particles.

The physical parameters are chosen according to Table VII. The lengths of the N^+ regions have been chosen smaller than in [9]. The thermal velocity corresponds to a 300 K temperature and 0.069 relative effective mass. The Debye length and the plasma frequency in the N^+ region are: $\lambda_D = 4.24 \times 10^{-3}\ \mu\text{m}$ and $\omega_p = 60.8\ 10^{12}\ \text{s}^{-1}$.

The distribution function is initialized by the equilibrium distribution function:

$$f_0(x, v) = M(v) \exp\left(\frac{q\phi(x)}{mv_{\text{th}}^2}\right), \tag{32}$$

where ϕ satisfies the semilinear elliptic equation:

$$-\frac{\partial^2 \phi}{\partial x^2} = n_D(x) - \exp\left(\frac{q\phi(x)}{mv_{\text{th}}^2}\right); \qquad \phi(0) = \phi(L) = 0. \tag{33}$$

Equation (33) is solved by finite differences and Newton's method. For all the simulations displayed in this section, we used periodic boundary conditions. Other boundary conditions were tested, but were found unstable.

The first series of tests are performed with a zero applied voltage and concern comparisons between the PIC method and the Green's function method, and between the regularized and unregularized Green's function method. For a zero applied voltage, the solution is stationary, equal to the equilibrium solution (32), and thus provides an easy bench mark. In this very inhomogeneous case, the interpolation method provided extremely bad results which we do not present here. Figure 7 shows comparisons between the unregularized Green's function method,

TABLE VII

Physical Parameters for Simulations Displayed in Section 7

$n_D^+ = 10^{18}\ \text{cm}^{-3}$	$n_D^- = 2.10^{15}\ \text{cm}^{-3}$
$x_1 = 0.4\ \mu\text{m}$	$x_2 = 0.8\ \mu\text{m}$
$L = 1.2\ \mu\text{m}$	$\tau = 0.29$ ps
$v_{\text{th}} = 258{,}000$ m/s	

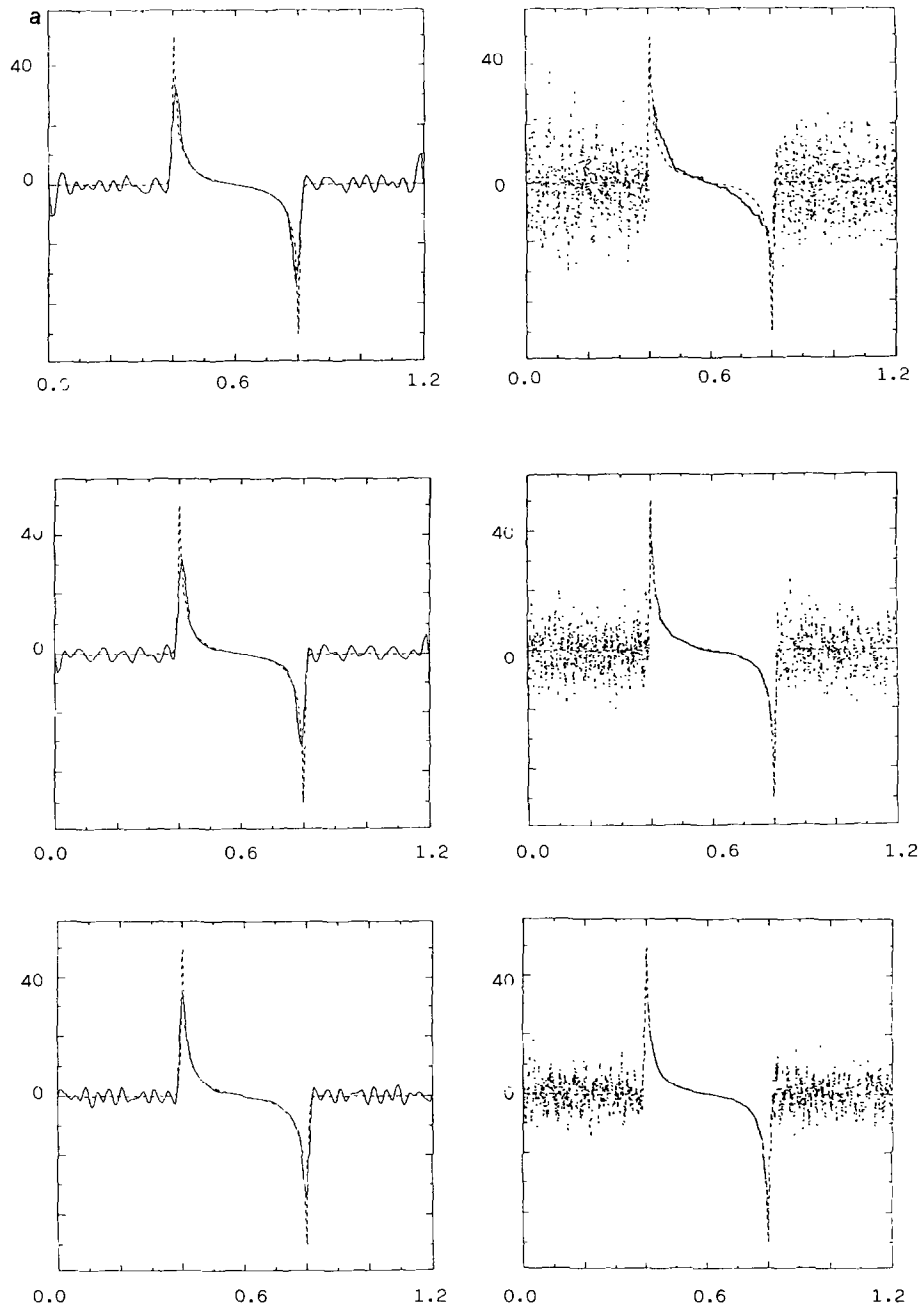

FIG. 7a. Electric field (in kV/cm) as a function of distance (in μm) at time $t = 1$ ps (corresponding to 100 time steps). Left figures: PIC method; right figures: Green's function method. Top figures: 3000 particles; middle: 9000 particles; bottom: 21,000 particles. The exact solution is shown in dashed lines.

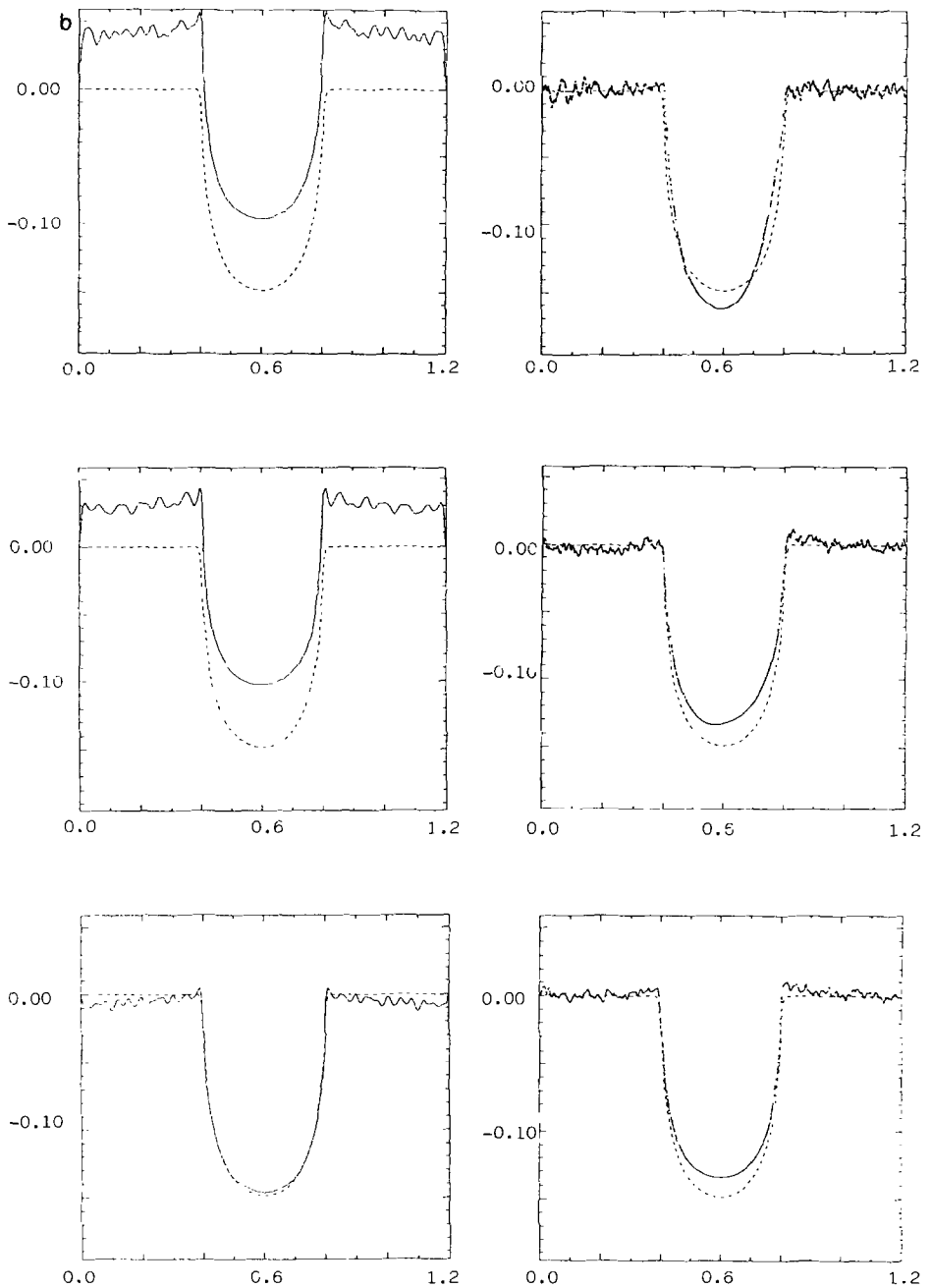

FIG. 7b. Potential (in V) as a function of distance (in μm) at time $t = 1$ ps. Left figures: PIC method; right figures: Green's function method. Top figures: 3000 particles; middle: 9000 particles; bottom: 21,000 particles. The exact solution is shown in dashed lines.

and the PIC method, for three values of the number of particles. The numerical values for these simulations are given in Table VIII.

Figure 7a represents the electric field as a function of the distance in the device after a 1 ps simulation. For the Green's function method, we represent the "particle electric field" (23), whereas for the PIC method, we represent the grid electric field E_m (see Section 4). The Green's function method appears to be quite precise inside the [illegible] description of the sharply peaked electric field. By contrast, the fluctuations are large in the highly doped region. These fluctuations are generated by the randomness of the positions of the particles in the phase space, which appears clearly on Fig. 8. Indeed, although the particles are initialized on a regular mesh, the flow is turbulent enough to randomize them after 1 ps. Therefore, the summation term in (23), which is the numerical quadrature of the integral of the electron density against the kernel K, is very near to a Monte-Carlo quadrature. The Monte-Carlo fluctuations are only apparant in the N^+ region because their amplitude is scaled by the doping densities. From Fig. 7a, we notice that this amplitude is decreased by a factor 2.5 when the number of particles is multiplied by a factor 7, which is consistent with the $N^{-1/2}$ rate of convergence of Monte-Carlo quadratures. The PIC method gives an oscillatory electric field. The oscillations are coherent, with a mesh wavelength: as the mesh size is decreased, the wavelength of the oscillations are decreased and their amplitude is increased. Furthermore, the peak of the electric field is smeared: the peak value is almost half of the exact one.

In our opinion, the fluctuations generated by the Green's function method do not affect much the dynamics of the particles. Indeed, these fluctuations are completely random, with a very short wavelength. Their influence along the trajectories of the particles has mean value zero. As pointed in [14, 15], it may happen that these fluctuations behave like an additional Fokker–Planck collision term in the equation. We believe that this effect is of little importance here and contrary to [14, 15], that it is not greatly reduced by smoothing (see the next comparisons). In the PIC method, the coherence of the oscillations react on the dynamics of the particles by creating bunches of particles which oscillate from one cell to the next ones and generates mesh oscillations.

Figure 7b displays the potential at time 1 ps. On the left, we represent the grid

TABLE VIII

Numerical Parameters for Fig. 7

N (particles)	3000	9000	21,000
Δx (PIC method; in μm)	8×10^{-3}	8×10^{-3}	6×10^{-3}
Number of part./cell (PIC)	20	60	105
α (collisions, in μm)	8×10^{-4}	2.66×10^{-4}	1.2×10^{-4}
Numb. of part./cutoff (coll.)	4	4	4.2
Time step (in fs)	10	10	10

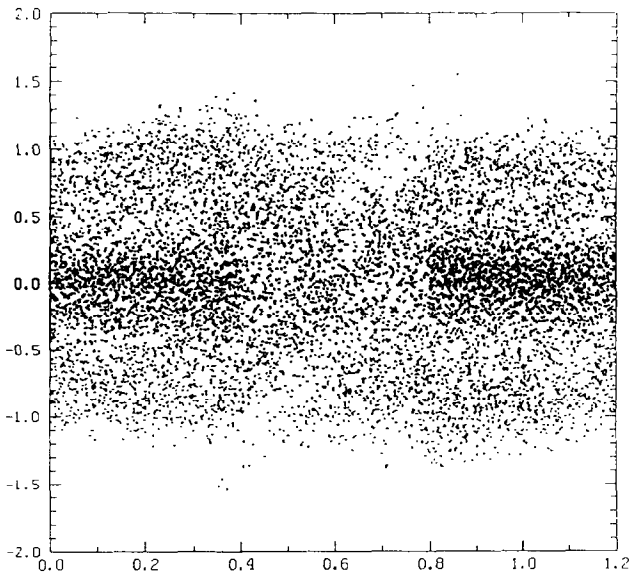

FIG. 8. Plots of the position of the particles in the phase space at time 1 ps; horizontal axis: distance in μm; vertical axis: velocity ($\times 10^6$ m/s). Zero applied voltage, Green's function method for the field; 9000 particles.

potential ϕ_m computed by the PIC method, and on the right, the potential given from the Green's function method by a numerical integration of the particle electric field E_i. With the Green's function method, the relative error on the barrier height is about 10 %, independently of the number of particles. This feature can be explained by two factors. First, the fluctuations of the potential in the N^+ region give rise to some "uncertainty" on the barrier height. Second, the numerical electric field is not as sharply peaked as the exact one. Since, the barrier height is exactly the area of the peak, it should be decreased proportionally. However, since in this method, the potential is independent of the trajectories solver, it is not clear if the discrepancy between the numerical and the exact potential really affects the motion of the particles.

The potentials given by the PIC method present similar oscillations as the electric field. Furthermore, with too few particles (3000 or 9000), a spurious potential barrier lies at the boundaries, which gives rise to large errors in the barrier height. However, with 21,000 particles, the PIC potential is approximately as good as the Green's function potential with 9000 particles, except that the barrier height is more accurate with the PIC method. When the mesh size of the PIC method is decreased, the overall shape of the potential improves, but at the expense of an increase of the fluctuation amplitude. This seems to indicate that better results would be obtained with two meshes: a coarser one for the charge assignment procedure and a finer one for the resolution of the Poisson equation. However, this procedure would lead to an increase of the numerical cost.

TABLE IX

CPU Time per Particle and per Time Step on Apollo DN-3000 (in ms)

Method \ N (particles)	3000	9000	21,000
PIC	1.89	1.86	1.87
Green's function	3.51	3.44	3.54

Comparisons of the CPU times per time step and per particle are given in Table IX. The Green's function method is twice as costly. For a given quality to be reached, the Green's function method needs approximately half the number of particles, but the overall computer time is the same. However, in real cases, the complexity of the collision operator is much greater, and the collision solver becomes the most time consuming step of the code. Since its complexity grows at least linearly with the number of particles, it may be interesting to choose a Poisson solver which allows the use of fewer particles for the same quality of results. Thus, we recommend the use of the Green's function Poisson solver.

We now turn to the comparisons of the unregularized and the regularized Green's function methods. Two examples of cutoff functions ζ are tested: the B-spline W_1 (see definition (19)) which generates a C^0 kernel K^α, and the B-spline W_2 which generates a C^1 kernel K^α. The numerical values of the smoothing parameter α are given in Table X. The other ones are those indicated in Table VIII. α has been chosen of the same order of magnitude as the mesh size Δx in the PIC method

The resulting electric fields are presented in Fig. 9. We notice a clear reduction of the noise amplitude in the N^+ region, of approximately a factor 2. The peak of the electric field is not smeared, which is surprising and interesting. Furthermore, the accuracy in the N^- zone is clearly improved, especially when fewer particles are used. Thus, the results are clearly improved by the smoothing procedure. On the other hand, the smoothness of the cutoff function is unimportant: the results are as good with W_1 as with W_2. The real drawback of the smoothing procedure is its cost. Table XI gives the CPU time per particle and per time step. The increase of

TABLE X

Smoothing Parameters α for the Simulations Displayed on Fig. 9 Using the Regularized Green's Function Method

N (particles)	3000	9000
α (regularization parameter in μm)	3.10^{-3}	3.10^{-3}
Number part./cutoff	15	45

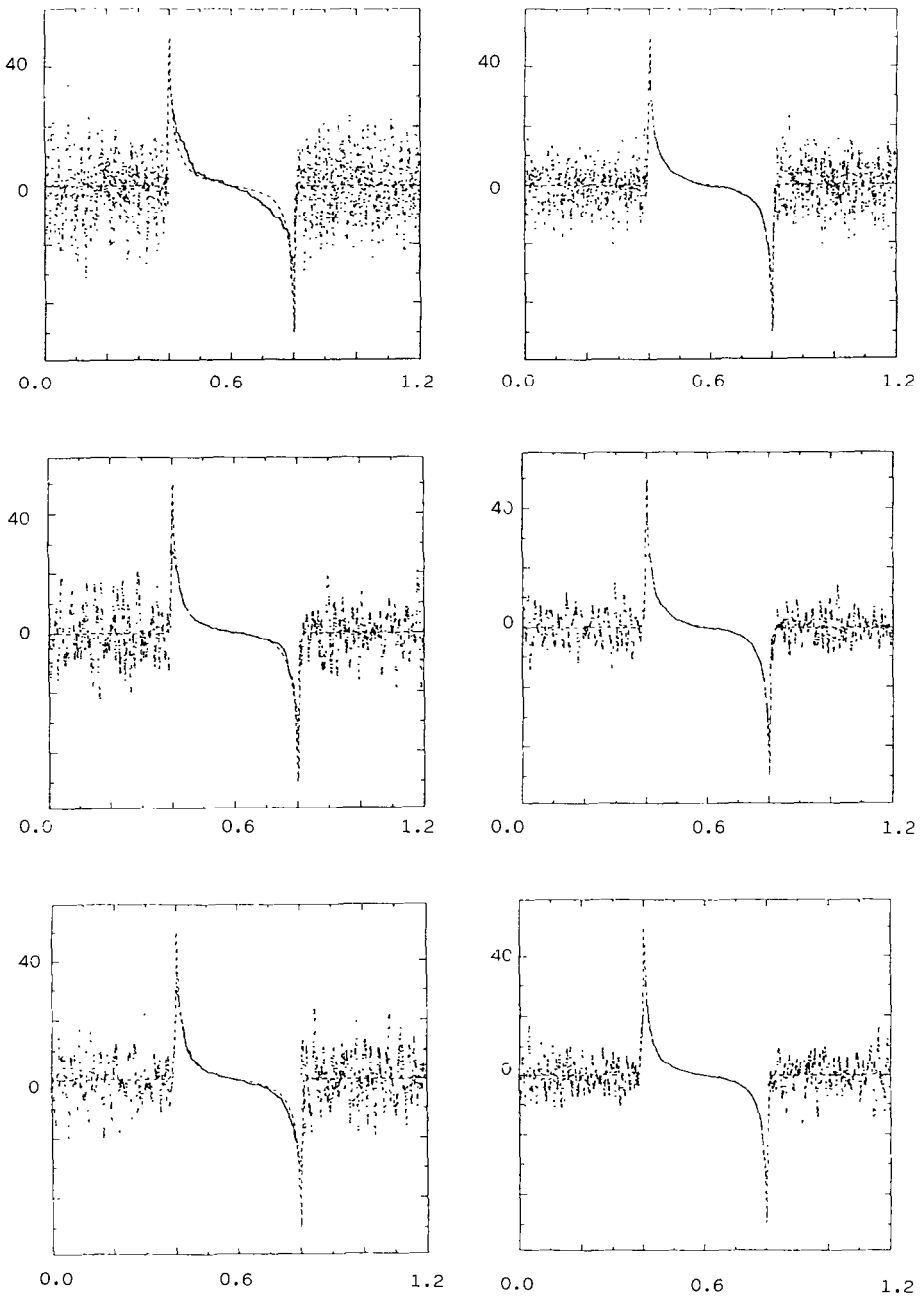

FIG. 9. Electric field (in kV/cm) as a function of distance (in μm) at time $t = 1$ ps. Left figures: 3000 particles; right figures: 9000 particles. Top figures: unregularized Green's function; middle: regularized Green's function, $\zeta = W_1$; bottom: regularized Green's function, $\zeta = W_2$. The exact solution is shown in dashed lines.

TABLE XI

CPU Time per Particle and per Time Step on Apollo DN-3000 (in ms) for the Regularized Green's Function Method

N (particles)	3000	9000
Unregularized	3.51	3.44
Regularized; cutoff $= W_1$	7.28	13.54
Regularized; cutoff $= W_2$	7.38	13.84

the CPU time due to the smoothing procedure depends on the number of particles per smoothing length, as a consequence of the localization procedure that is used to perform the regularization (see Section 4). The derivation of a "fast algorithm" for the regularized electric field (at least for simple cutoff functions like W_1) would greatly improve the method. This has not yet been performed.

A second series of tests are concerned with an applied bias $U(t) = \phi_L(t) - \phi_0(t) =$ 0.47 V, which corresponds to one of Baranger's simulations [9]. The initial datum is the equilibrium distribution function under zero applied voltage, given by (32)–(33). Thus, the simulation mimics the transient evolution of the structure under a sudden change of the applied bias at $t = 0$. The transient regime would probably be more realistic if this change was not so abrupt. The boundary conditions on the distribution function are still chosen periodic. For a long enough N^+ region, the specific form of the boundary condition is not very important, since the particles are thermalized before entering the active N^- region. Nevertheless, physically more realistic boundary conditions would allow the use of a shorter N^+ region, which would decrease the computer time. We have not found a convenient boundary condition for both the physical and numerical viewpoints yet.

An important diagnostic in semiconductor device simulations is the total current density $J(x, t)$ defined by

$$J(x, t) = J_{\text{part}}(x, t) + \varepsilon \frac{\partial E}{\partial t}(x, t),$$

where J_{part} is the "particle current,"

$$J_{\text{part}}(x, t) = -q \int f(x, v, t)\, v\, dv,$$

and $\varepsilon\, \partial E/\partial t$ is the displacement current. In this one-dimensional situation, J does not depend on x because it is divergence free. Thus it is equal to its mean value over the device:

$$J(t) = \frac{1}{L} \int_0^L J(x, t)\, dx = -\frac{q}{L} \iint f(x, v, t)\, v\, dx\, dv - \varepsilon \frac{\partial U}{\partial t}.$$