JOURNAL OF COMPUTATIONAL PHYSICS 61, 51–63 (1985)

A Method for Simulation of Vlasov Systems with Discontinuous Distribution Functions

WILLIAM S. LAWSON

Electronics Research Laboratory, University of California, Berkeley, Berkeley, California 95720

Received April 10, 1984; revised October 16, 1984

A method is presented for simulating non-periodic (or periodic) Vlasov systems with discontinuous distribution functions. The method is a hybrid of a standard Vlasov method and the Waterbag method; the discontinuity is integrated seperately from the distribution function. Sample runs and particle simulation comparison runs are included.

INTRODUCTION

Most simulations of plasma systems have, until recently, assumed periodic boundary conditions. While this is satisfactory for some problems, there is a growing interest in the simulation of problems having non-periodic boundary conditions. Particle simulation methods have been adapted to these problems [1], but particle methods are inherently noisy. Finite difference methods are not noisy, and have been applied to non-periodic problems [2, 3, 4], but none of these applications have addressed the central problem of discontinuities in the Vlasov distribution function. These discontinuities are inherent in almost all non-periodic boundary conditions. Presented here is a method of surmounting the difficulties inherent in using a finite difference grid method to simulate a non-periodic Vlasov system. The method is not as flexible as a particle method, but the accuracy is comparable, and the particle shot noise is eliminated.

This method bears a superficial resemblance to the Marker-and-Cell method [5], but is quite different both in philosophy and implementation.

PROBLEM DEFINITION AND APPROACH

The problem to be solved is the integration of the 1-dimensional Vlasov–Poisson system in time over a spatial domain of length L, and an infinite velocity domain. Specifically, to solve

$$\frac{\partial f_s}{\partial t} + v\frac{\partial f_s}{\partial x} + \frac{q_s E(x,t)}{m_s}\frac{\partial f_s}{\partial v} = 0 \tag{1}$$

0021-9991/85 $3.00

and

$$\frac{\partial^2 \phi}{\partial x^2} = -\frac{1}{\varepsilon_0} \sum_s \int_{-\infty}^{\infty} f_s(x, v, t)\, dv, \tag{2}$$

where

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

with appropriate boundary conditions on ϕ. (Note that the charge of each species is absorbed into the distribution function.) For simulation purposes, the velocity domain is cut off at $\pm v_{\max}$. The boundary conditions on the distribution function used are that the incoming distributions at $x = 0$, $v > 0$, and $x = L$, $v < 0$, are time-independent (or at worst are slightly perturbed around some time-independent distributions). The potential difference between $x = 0$ and $x = L$ is also taken to be time-independent (or slightly perturbed). The reasons for the rather strict boundary conditions will be explained shortly.

To simplify the explanation of the method, only one species is considered here, and the incoming distribution function at $x = L$ is taken to be zero (i.e., no particles coming from the $x = L$ side). The main simulation difficulty is numerical, and arises because the distributions entering from the two ends are independent of each other, and will therefore almost always give rise to a discontinuity in $f(x, v)$ where they meet. In Fig. 1, for example (a particle simulation of an electron diode), the distribution function jumps from a large value above a certain characteristic, to zero below it. If a finite difference method were used on this distribution function directly, the discontinuity would wreak havoc. Depending on the method used, the discontinuity would either be diffused out over many grid cells, or cause overshoots resulting in large negative values for the distribution function. This difficulty with the discontinuity is solved by extending the distribution function smoothly over the entire x–v domain, and following the position of the discontinuity separately. Since

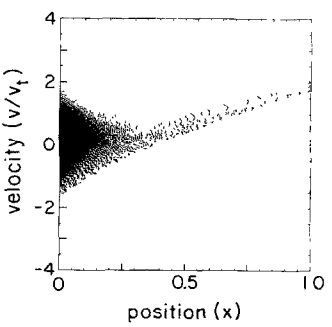

FIG. 1. Phase space plot of electrons in a single species diode with space-charge limited current. Distribution function jumps suddenly to zero.

the distribution function is now smooth, ordinary finite difference methods can be used. The extended part of the distribution function is then ignored in calculating the charge density, and so does not affect the physics.

To ensure that no large phase space gradients form in this extended distribution function—defeating the purpose of the extension—the boundary condition on the artificial part of the distribution function at $x = L$ must be chosen carefully. The simplest boundary conditions are a fixed voltage difference ϕ between $x = 0$ and $x = L$, and fixed incoming distribution functions, with the additional constraint that

$$f(L, -v) = f(0, \sqrt{v^2 - 2q\phi/m}) \qquad \text{for} \quad v^2 \geqslant 2q\phi/m. \tag{3}$$

The incoming distributions must, of course, be smooth. These conditions guarantee that the extended distribution will be continuous everywhere for the steady state. One must then hope that the time-dependent case is not too badly behaved. (It will be seen that some bad behavior can be tolerated.) For the specific boundary conditions chosen for the sample implementation (half-Maxwellian at both ends) condition (3) can be written

$$f(L, -v) = f(0, v) \exp(-q\phi(L)/kT). \tag{4}$$

The position of the discontinuity is followed by a string of test particles, as in the Waterbag model [6]. When two test particles drift too far apart, a new one is created between them. The entire scheme is diagrammed in Fig. 2.

Experience with the sample code has shown some pitfalls and limitations which seem obvious in retrospect. For instance, in many situations, despite choosing good boundary conditions, large gradients in phase space develop as time progresses. These cannot be dealt with in the same way as the test particles deal with the

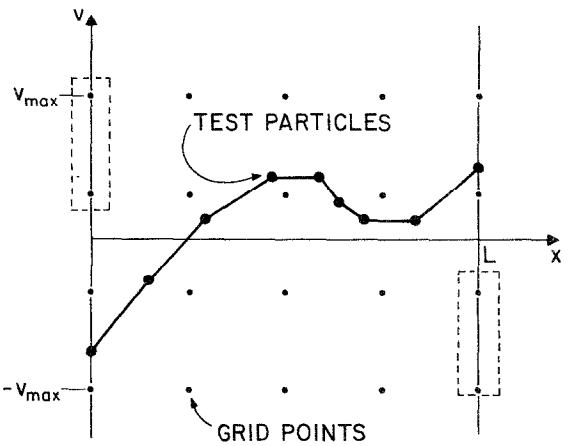

FIG. 2. Grid layout and test particle scheme. Incoming distribution funcion is specified at grid points in dashed boxes.

primary discontinuity, so they will give rise to numerical diffusion, and thus be a source of inaccuracy. There appears to be little that can be done about this. Another problem arises when one wants to vary the voltage across the domain in time. In order to prevent large gradients in phase space of the kind just mentioned, it is necessary to vary the incoming boundary conditions at $x = L$, but how to vary them so as to minimize the gradients in phase space is unclear. Some of these gradients (ripples, for instance) may be diffused away harmlessly, but others may cause unacceptable inaccuracy. No work has been done to resolve this problem.

Sample Implementation

The Vlasov equation integrator used in this implementation of the method is that used by Cheng and Knorr [7]. It is completely analogous to the leap-frog particle mover:

$$f^{p+1/2}(x, v) = f^{p}\left(x, v - \frac{qE^{p}(x)}{m}\Delta t\right), \tag{5}$$

$$f^{p+1}(x, v) = f^{p+1/2}(x - v\,\Delta t, v), \tag{6}$$

where $f(x, v)$ is the Vlasov distribution function, p is the time index, and x, v, and E are the position, velocity, and electric field respectively. This scheme yields second-order accuracy in time, and reduces the integration to two interpolations, which can be performed by any of several methods.

Many interpolation methods were tried, and it was discovered (empirically) that some numerical diffusion was necessary for stability. The best methods are those which have diffusion only when large phase-space gradients occur. When there is insufficient numerical diffusion, these large gradients apparently interact with the x–v grid to produce a numerical instability. Large phase space gradients also cause accuracy problems because of the inaccuracy of the test particle integrator. A minor shift in the position of the test particles near such a gradient can result in a major error when the charge density is computed. Introducing some numerical diffusion at steep gradients solves both these problems with a minimum error in other quantities, such as current and kinetic energy.

The interpolation schemes finally settled on are two, three, and four point *off-center* interpolations, and are shown below. They are first, second, and third order accurate, respectively:

$$\begin{aligned} f_i^{p+1} &= f_i^p - u(f_{i+1}^p - f_i^p) \qquad \text{if} \quad u < 0, \\ &= f_i^p - u(f_i^p - f_{i-1}^p) \qquad \text{if} \quad u \geqslant 0; \end{aligned} \tag{7}$$

$$\begin{aligned} f_i^{p+1} &= f_i^p - u\left(2f_{i+1}^p - \frac{3}{2}f_i^p - \frac{1}{2}f_{i+2}^p\right) + \frac{u^2}{2}(f_{i+2}^p - 2f_{i+1}^p + f_i^p) \qquad \text{if} \quad u < 0, \\ &= f_i^p - u\left(\frac{3}{2}f_i^p - 2f_{i-1}^p + \frac{1}{2}f_{i-2}^p\right) + \frac{u^2}{2}(f_i^p - 2f_{i-1}^p + f_{i-2}^p) \qquad \text{if} \quad u \geqslant 0; \end{aligned} \tag{8}$$

$$
\begin{aligned}
f_i^{p+1} &= f_i^p - u\left(f_{i+1}^p - \frac{1}{2}f_i^p - \frac{1}{3}f_{i-1}^p - \frac{1}{6}f_{i+2}^p\right) \\
&\quad + \frac{u^2}{2}(f_{i+1}^p - 2f_i^p + f_{i-1}^p) \\
&\quad - \frac{u^3}{6}(f_{i+2}^p - 3f_{i+1}^p + 3f_i^p - f_{i-1}^p) \qquad \text{if} \quad u < 0, \\
&= f_i^p - u\left(\frac{1}{3}f_{i+1}^p + \frac{1}{2}f_i^p - f_{i-1}^p + \frac{1}{6}f_{i-2}^p\right) \\
&\quad + \frac{u^2}{2}(f_{i+1}^p - 2f_i^p + f_{i-1}^p) \\
&\quad - \frac{u^3}{6}(f_{i+1}^p - 3f_i^p + 3f_{i-1}^p - f_{i-2}^p) \qquad \text{if} \quad u \geqslant 0;
\end{aligned}
\tag{9}
$$

where $u = v\,\Delta t/\Delta x$ or $u = (q/m)\,E\,\Delta t/\Delta v$. These formulae are derived simply by fitting straight lines, parabolas, and cubic curves through the given points. In practice, the three point scheme seems to be most economical.

The first scheme (Eq. (7)) models (to first order in Δx, Δv and Δt) the equation

$$
\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} + \frac{qE}{m}\frac{\partial f}{\partial v} = \frac{1}{2}\left[|v|\,\Delta x - v^2\,\Delta t\right]\frac{\partial^2 f}{\partial x^2} + \frac{1}{2}\left[\left|\frac{qE}{m}\right|\Delta v - \left(\frac{qE}{m}\right)^2\Delta t\right]\frac{\partial^2 f}{\partial v^2}. \tag{10}
$$

If $|v| > \Delta x/\Delta t$ or $|qE/m| > \Delta v/\Delta t$, the diffusion coefficients will become negative, and the method can be expected to become unstable. (These two stability criteria can also be easily derived from a Fourier analysis of the difference equations.)

Equations (8) and (9) are much more difficult to put into differential equation form because the diffusion terms ($\partial^2 f$) cancel out, leaving only terms which are second order or higher in Δx, Δv, and Δt, and involve third and higher derivatives of f. These equations have not been worked out. It is worth pointing out that the leap-frog splitting scheme is only of second order in Δt, so while the four-point interpolation may be more accurate than the three-point interpolation, it will not increase the *overall* order of the method over the 3-point scheme.

It is also worth mentioning the methods which were tried which did not work. Derivative methods (linear interpolations with the slope of the line determined by approximating the derivative) were tried for 2-, 3-, and 4-point approximations, and were found to be unstable. This is not surprising, since they are strictly unstable in the periodic or infinite grid case. Also tried were a centered 3-point interpolation, and a cubic spline interpolation. The reasons for the failure of these methods are less clear. These methods both failed because odd–even modes developed in regions where the distribution function was becoming very steep. The failure of the cubic spline is probably due to too little diffusion. The odd–even mode appeared at exactly the place where diffusion should have been occurring to remove

detail too small for the grid to resolve. Since the 4-point interpolation method should be more accurate than the 3-point centered method, the failure of the 3-point centered method is probably not due to insufficient diffusion, but to its symmetry in a fundamentally asymmetric problem. When the method is used, information can flow both ways with equal ease on the numerical grid, whereas the solution being modeled allows information to flow in only one direction. This is admittedly an insufficient explanation of the failure of the 3-point centered method, but it is a useful hypothesis for future investigation.

The test particle integrator is the usual leap-frog integrator:

$$v^{p+1/2} = v^{p-1/2} + \frac{qE^p(x)}{m}\Delta t \tag{11}$$

$$x^{p+1} = x^p + v^{p+1/2}\Delta t. \tag{12}$$

It is this particle integrator which dictates the use of the splitting scheme for the Vlasov integrator. The test particles must stay in step with the distribution function, or else significant systematic errors will result. A different Vlasov integrator could be used, but it might be difficult to decide what particle integrator would be appropriate. This combination of particle and Vlasov integrators is simple, fast, and second-order accurate in time.

As time passes, the test particles on the discontinuity will generally drift apart, and new ones must be created to maintain detail. The method used for positioning these new particles in the sample implementation is linear bisection between those particles which have drifted too far apart. A more sophisticated method could be used, such as fitting an ellipse through a set of four particles, but the time and complexity required make this unattractive; care must be taken to make sure that a change of scale will not produce different results, since the x and v axes have different units. It should also be possible to find some curvature criterion for introducing new particles which would introduce them where they would be most needed.

The charge integrator, which integrates the distribution function in v from the boundary of test particles to $v_{\max}$, uses the trapezoidal rule. The lower limit of integration is taken to be the crossing of the line of test particles with any given x-grid position. While the maximum spacing of the test particles could be any distance, the charge integration routine is simplified if the particles are never more than one x-grid cell apart; so, in the sample code, the maximum allowed distance between particles is precisely the grid spacing.

Because of the non-periodic boundary conditions and the limit in velocity, the boundary conditions deserve some comment. At the outgoing boundaries ($x=0$, $v<0$ and $x=L$, $v>0$) there are no problems, since all the interpolation methods are off-center in the upstream direction; however, at the incoming boundaries ($x=0$, $v>0$ and $x=L$, $v<0$), an interpolation method cannot be used where it would need points outside the domain. Therefore, a less accurate interpolation method must be used at these boundaries. The effect of this can easily be seen in the jump in the ballistic current diagnostic at the left-hand edge in the sample runs

(Fig. 3). This effectis certainly removable if it is a problem. An extra column of grid points, either used as a guard cell or placed between the first two columns of grid points, would probably smooth out the jump.

The velocity limit at $v = \pm u_{\max}$ presents an entirely different problem, and has no definitive solution. In the sample code, interpolations which require grid points outside the velocity domain assume an exponential fall-off. This method is apparently quite successful, and all but guarantees that the distribution function remains positive. It breaks down when any significant dynamics reaches the high velocity boundary, but in this case the problem must be rerun with a larger $v_{\max}$ anyway.

Sample Runs

The results of a sample run with a single particle species are shown in Fig. 3. The domain is initially vacuum, and the ends of the domain are shorted. The single species (electrons) is injected from the left end starting at $t = 0$, and it rapidly fills the simulation domain. Throughout the simulation $\omega_p \Delta t$, which is a function of position, is much less than 1. The diagnostics shown are contour plots of phase space (the heavy line is the discontinuity marked by the test particles—the distribution below this line is non-physical), charge density, and ballistic current density. Although it cannot be seen from the snapshots in Fig. 3, this simulated diode rapidly comes to a steady state with no overshoot, and no oscillation of the potential minimum, in contrast to particle simulations of similar problems [1, 8, 9]. This indicates that the oscillations observed in particle simulations are noise driven oscillations (which may nonetheless be quite physical), and not the result of an instability. It might be argued that numerical diffusion might wash out such an oscillation in the Vlasov simulation, but most of the important changes occur in the position of the test particle boundary (which by definition does not diffuse), and over a large scale in the distribution function, which should not be strongly affected by slight diffusion. The conclusion that the oscillations are noise driven is also supported by the roughly $1/\sqrt{N}$ dependence of the noise amplitude noted in [1] (where N is the number of particles).

The steady state solution for this problem can be accurately calculated numerically [10]. The steady-state current, according to theory, should be 9.116 throughout the simulation domain, so the current diagnostic represents a very sensitive test of the accuracy of the method. This is the closest approach to a test of conservation of momentum that has been tried. Conservation of energy has not been tested, nor has any attempt been made to develop a theory for conservation of energy and momentum. As can be seen from the final current density plot in Fig. 3, the simulation agrees with the theory to within 0.4%. The steady-state electrostatic potential (not shown) agrees even more closely than the current density.

The plots in Fig. 3 also point out the necessity for some numerical diffusion. The exact Vlasov solution of this problem contains a wall in the distribution function at the discontinuity (not to be confused with the discontinuity itself), which would go

to zero thickness without ever declining in height. Naturally a finite grid simulation method cannot cope with this kind of detail, and any interpolation method which does not have some diffusion when gradients in phase space become large must lead to a numerical instability. (Cubic spline interpolation is unstable, presumably for this reason.) Furthermore, this wall becomes physically uninteresting beyond a certain point, as the Vlasov equation is spatially averaged to begin with, so only a certain scale of detail is relevant. The diffusion allows the wall to fade away with a minimum of inaccuracy.

While the grid cannot support any detail smaller than the grid spacing, the string of test particles can. Thus, even though filamentation within the gridded dis-

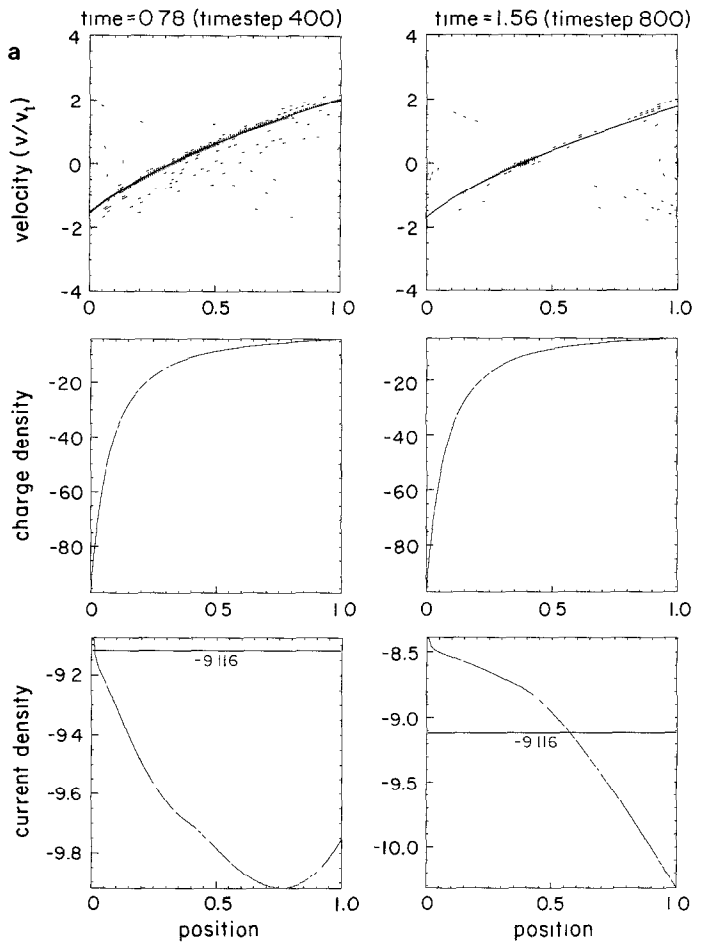

FIG. 3. (a) Results of a sample run: steady state current should be a constant 9.116 according to Langmuir's theory; (b) Results of a sample run (continued).

tribution function is smoothed out, the filamantation of the discontinuity is still observable. This property could conceivably prove to be useful.

The speed of this method is comparable to the speed of particle methods if one considers grid points to be roughly equivalent to particles. The test run shown was run on a 129×128 grid for 1600 time steps. The total time for the run was 30 s of CRAY time. The time for moving and maintaining the test particles on the discontinuity is about equal to the time required to advance the distibution function when there are 64 grid values of velocity.

Comparison runs were made using the newly developed particle code PDW1 (PDW is an acronym for Plasma Device Workshop, a workshop at the University

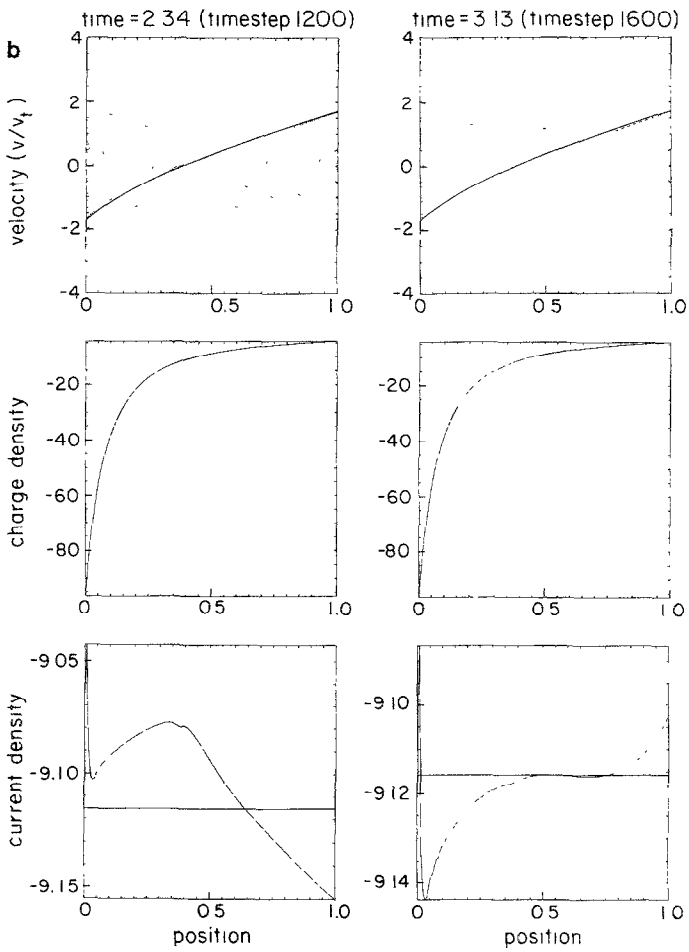

FIG. 3—*Continued.*

of California, Berkeley during the spring quarter of 1983 for which the code was written). Such particle codes have been used extensively, and are well understood [11]. The x grid has 128 grid cells, and the number of particles in the particle simulation is roughly three quarters of the number of grid points in the Vlasov simulation. PDW1 is fully vectorized (running at roughly 2 μs/particle/timestep), yet took over 100 s for this comparison run. The results of the particle run are shown in Fig. 4, and are in excellent agreement with the Vlasov simulation. The inaccuracy of the Vlasov simulation is, however, smaller than the noise level of the particle scheme, despite using a quiet loading scheme (bit-reversing) for the particle

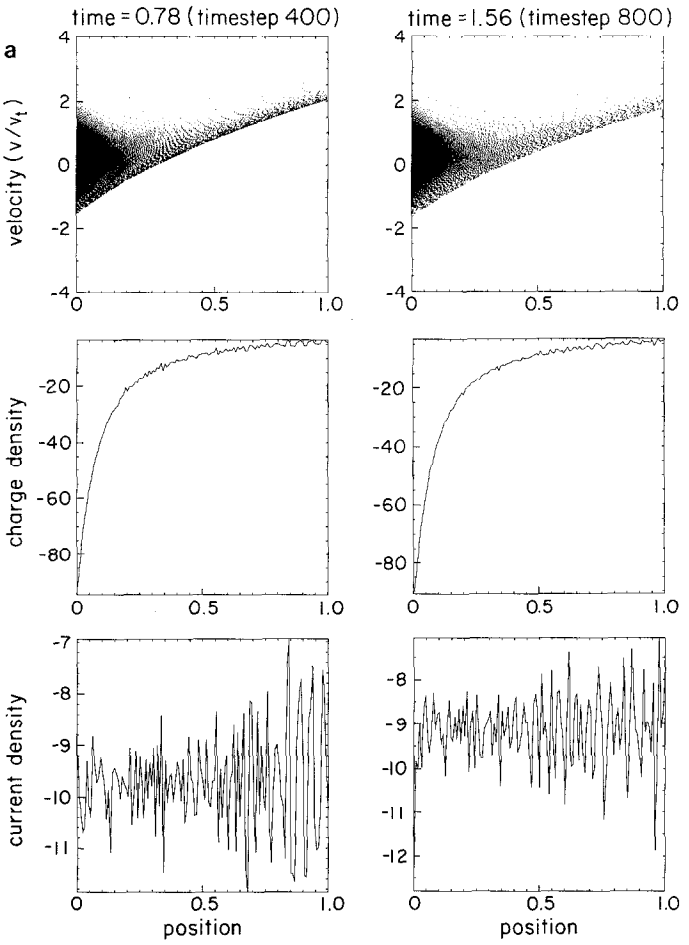

FIG. 4. (a) Particle code run of problem in Fig. 3, for comparison. (Note that the noise level of the final current density is much higher than the inaccuracy of the GASBAG result.) (b) Particle code run (continued).

run. This is a major advantage in tests of linearized theory in which one wants to look at small perturbations of a steady state.

Figure 5 is a comparison of the total kinetic energy histories of the particle and Vlasov runs. Here the particle noise is drastically reduced (since all of the particles are summed over), and a systematic deviation can be seen. The particle code is probably most accurate here, with the ineaccuracy of the Vlasov method stemming from the numerical diffusion near the potential minimum.

Some runs have been made with two species, both with the Vlasov code, and the particle code. Figure 6 shows a pair of such runs for a thermal plasma which

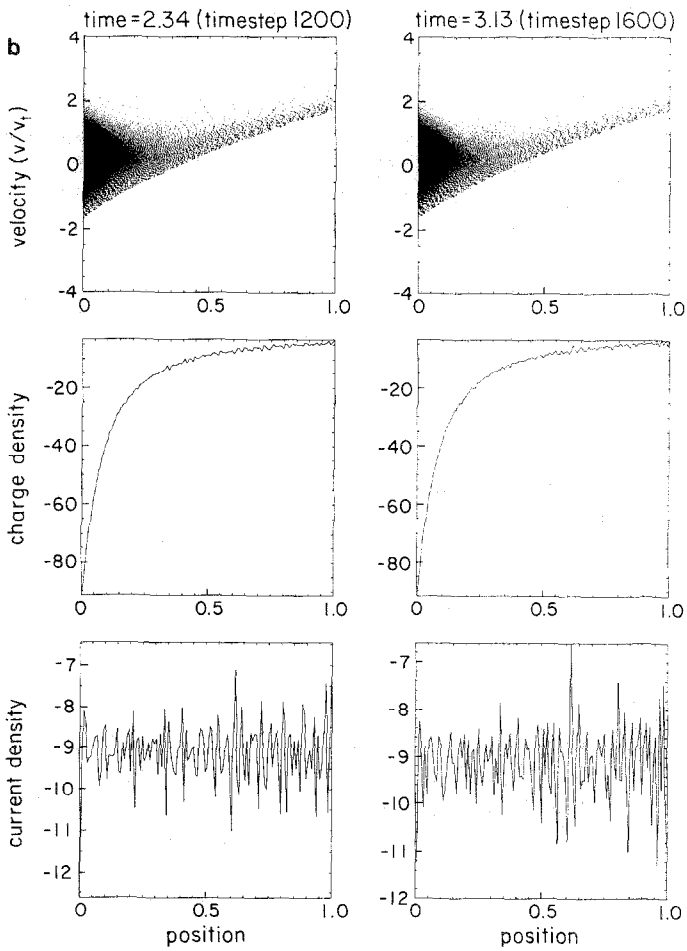

FIG. 4—*Continued.*

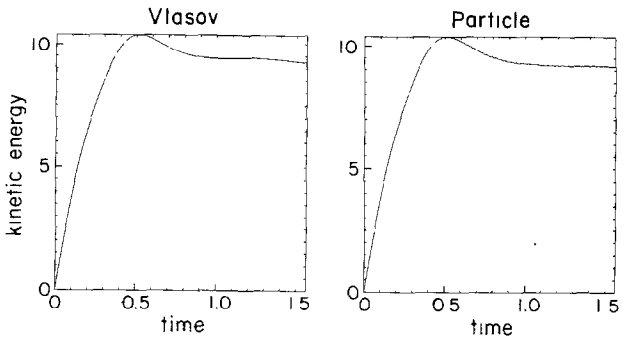

FIG. 5. Comparison of kinetic energy histories for both GASBAG and PDW1. Note the 3% difference between time 0.7 and 1.5.

eneters from the left-hand wall into a vacuum. The mass ratio is $m_i/m_e = 4$, and the external circuit is a short. Both simulations show an instability which starts with a wavelength roughly the length of the simulation, and then shrinks accordion-style to the final state shown with roughly three wavelengths in the simulation. At this point, the distribution breaks, forming long filaments around holes in phase space. Partial simulations run for longer times show that these filaments become very convolved; however, only a relatively small number of particles are involved, and they have little effect on the electric field. The agreement between the two codes is not

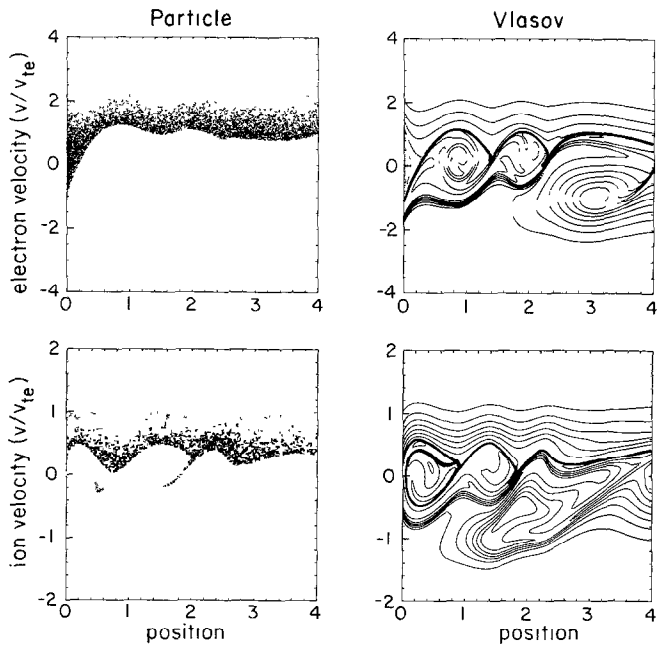

FIG. 6. Phase space plots of a plasma after expanding into a box from one side, for both Vlasov and particle methods.

perfect, and many more grid points were needed to get the Vlasov result (513×128), but the chief virtue of the Vlasov method is the low noise, and not high accuracy. If low noise is a must, then it is easier to use a fine mesh Vlasov code to get it than to use a huge number of particles.

SUMMARY

A method has been developed which allows the use of finite difference grid techniques to follow the evolution of non-periodic Vlasov systems having jump discontinuities. The method is numerically stable, fast, and is free of the noise problems associated with particle simulation methods. Sample runs were shown to be in excellent agreement with both theoretical results, and independent particle simulation results. The method is somewhat less flexible than particle methods with respect to boundary conditions (it requires boundary conditions which are at most slightly perturbed about some steady state), but future work may find ways of relaxing this requirement. Most importantly, it allows some problems (e.g., linearized and perturbation problems) to be solved which would be washed out by noise in a particle simulation.

The sample code, with documentation, is available from the author. It is designed to be run on the National Magnetic Fusion Energy Computer Center CRAY computers, but should be adaptable to any high speed computer.

ACKNOWLEDGMENTS

I thank Professor C. K. Birdsall and Dr. T. L. Crystal for their valuable advice and support. This work was supported by U. S. Department of Energy contract DE-AT03-76ET53064, and the computing was performed on the CRAY computers at the National Magnetic Fusion Energy Computer Center.

REFERENCES

1. S. ROUSSET, "Time-Dependent Child-Langmuir Diode Simulation," ERL Report UCB/ERL M83/39, 11 July 1983.
2. P. H. SAKANAKA *et al.*, *Phys. Fluids* **14** (1971), 611.
3. K. ESTABROOK *et al.*, *Phys. Fluids* **14** (1971), 1792.
4. J. DENAVIT, *J. Comput. Phys.* **9** (1972), 75.
5. F. H. HARLOW AND J. E. WELCH, *Phys. Fluids* **8** (1965), 2182.
6. H. L. BERK AND K. V. ROBERTS, *Phys. Fluids* **10** (1967), 1595; K. V. ROBERTS AND H. L. BERK, *Phys. Rev. Lett.* **19** (1967), 297.
7. C. Z. CHENG AND GEORG KNORR, *J. Comput. Phys.* **22** (1976), 330.
8. P. K. TIEN AND J. MOSHMAN, *J. Appl. Phys.* **27** (1956), 1067.
9. C. K. BIRDSALL AND W. B. BRIDGES, "Electron Dynamics of Diode Regions," Chap. 6, Academic Press, New York, 1966.
10. I. LANGMUIR, *Phys. Rev.* **21** (1923), 419.
11. C. K. BIRDSALL AND A. B. LANGDON, "Plasma Physics Via Computer Simulation," McGraw–Hill, New York, 1985.