JOURNAL OF COMPUTATIONAL PHYSICS 6, 1–16 (1970)

A Practical Difference Scheme for Fokker–Planck Equations*

J. S. CHANG AND G. COOPER

Lawrence Radiation Laboratory, University of California, Livermore, California 94550

Received July 28, 1969

A practical finite difference scheme for initial value problems of Fokker–Planck equations has been studied. In addition to satisfying the conditions of convergence and unconditional stability, this scheme provides numerical solutions which preserve some of the more important intrinsic properties of the original partial differential equation. In particular, the solutions are non-negative, particle conserving in the absence of external sources or sinks, and exact representations of the analytic solution upon equilibration. Furthermore, coupled with variable mesh size, this scheme actually significantly reduces the number of mesh points required with no loss of accuracy.

INTRODUCTION

In the study of an infinite, isotropic, fully ionized plasma, one often requires numerical solutions of certain Fokker–Planck equations [1, 2, 3, 4] of the form

$$\frac{\partial u}{\partial t} = \frac{1}{A(x)} \frac{\partial}{\partial x} \left[B(x, t)u + C(x, t) \frac{\partial u}{\partial x} \right] \tag{1}$$

in the domain $0 \leqslant t \leqslant t_0$, $0 \leqslant x < \infty$, with A, B, and C all positive functions of their arguments.[1] The quantity x usually represents a velocity (or momentum) variable and $u(x, t)$ is the single particle distribution function in that space. Depending on the particular problem, B and C may or may not be functions of u. If it is desired to find the solution to the above equation for a selected class of initial conditions and for $t_0 \leqslant t_{eq}$ (t_{eq} is some characteristic equilibration time), one of the many standard finite difference schemes in existence may be employed [5]; by choosing sufficiently small time and velocity intervals, a stable, convergent solution to Eq. (1) will be obtained [1, 3]. If, however, the Fokker–Planck equation

* Work performed under the auspices of the U. S. Atomic Energy Commission.

[1] Physically the usual meaning of these functions are $4\pi A(x)\, dx = d^3x$ is the spherically averaged volume element, $B(x, t)$ is the dynamic friction coefficient and $C(x, t)$ is the dispersion coefficient.

581/6/1-1

is only part of a larger calculation (e.g., a coupled many species problem) or one is limited by the size and speed of available computers, the standard numerical techniques may prove unsuitable. In this paper, we will elucidate some of the shortcomings of these methods and present a practical scheme for their correction, i.e., a scheme that will perit larger Δt and Δx with satisfactory accuracy and significant decrease in real machine time.

The underlying objectives of any practical numerical scheme are the following: (1) large time and velocity steps, (2) accuracy and stability, and (3) preservation of any intrinsic properties implied by the Fokker–Planck equation. Clearly, one may not be able to accomplish all three simultaneously and completely; instead, we must strive to satisfy these requirements as best as compatibility allows. The criterion for judgement must be the sensitivity or degree of dependence of the physics on particular errors introduced by the numerical scheme.

I. Intrinsic Properties

The structure of Eq. (1) along with certain properties of its coefficients dictates various properties which any exact solution must possess. These are all interpretable in terms of physical quantities and are consequently properties which we must consider when setting up our numerical scheme.

We first show that the positiveness of the coefficient functions A, B, and C is sufficient to guarantee $u(x, t) \geqslant 0$ as long as $u(x, 0) \geqslant 0$. This property is, of course, mandatory for the interpretation of u as a probability density. To establish the nonnegative character of u we consider the following theorem: for Eq. (1), if for any t_0 there is a point x_0 such that $u(x_0, t_0) = 0$ and $u(x, t_0) \geqslant 0$ for all other x, then $\partial u/\partial t \,|_{x=x_0} \geqslant 0$. This is proven by noting that u is a local minimum at x_0, so that

$$\left.\frac{\partial u}{\partial x}\right|_{x=x_0} = 0 \qquad \text{and} \qquad \left.\frac{\partial^2 u}{\partial x^2}\right|_{x=x_0} \geqslant 0.$$

Then looking at Eq. (1) we see that since A and C are positive,

$$\left.\frac{\partial u}{\partial t}\right|_{x=x_0} = \frac{C(x, t)}{A(x)} \frac{\partial^2 u}{\partial x^2} \geqslant 0 \tag{2}$$

which is the desired relation. Using this theorem plus the fact that since $\partial u/\partial t$ exists, u must be continuous in time, we see immadiately that

$$u(x, 0) \geqslant 0 \Rightarrow u(x, t) \geqslant 0,$$

i.e., u must go through zero to become negative, but $u(x, t_0) = 0 \Rightarrow u$ cannot decrease.

Another property we wish to demonstrate is the existence of a conservation law. From Eq. (1) we see that the quantity

$$N = \int_0^\infty u(x, t)\, A(x)\, dx \tag{3}$$

satisfies the equation

$$\frac{\partial N}{\partial t} = \left[B(x, t)u + C(x, t)\frac{\partial u}{\partial x}\right]\Bigg|_0^\infty \tag{4}$$

which we may interpret physically as follows: the time rate of change of N is dependent only on the generalized flux, $[Bu + C(\partial u/\partial x)]$, at the boundaries. If, as is usually the case, the boundary conditions on u and $\partial u/\partial x$ are such that to make the right side of Eq. (4) vanish, then we have a truely conserved quantity. In most cases, $4\pi A(x)\, dx = d^3x$, and N turns out to be the total number of particles in the system.

A third property of importance concerns the quasiequilibrium solutions (or equilibrium solutions). By quasiequilibrium solutions we mean a time dependent function $\tilde{u}(x, t)$ such that

$$\frac{\partial}{\partial x}\left[B(x, t)\tilde{u} + C(x, t)\frac{\partial \tilde{u}}{\partial x}\right] \equiv 0 \tag{5}$$

and if $\tilde{u}(x, t) = \tilde{u}(x)$, then it is the equilibrium solution. If Eq. (1) is linear, then the quasiequilibrium solution is

$$u(x, t) = \alpha \exp\left[-\int^x \frac{B(y, t)}{C(y, t)}\, dy\right], \tag{6}$$

where α is a constant of integration.

In many problems the quasiequilibrium solution may be function of a slowly varying time dependent parameter, say $\theta(t)$, i.e., $u(x, t) = u(x, \theta(t))$, and $\partial\theta/\partial t \approx 0$. Consequently (6) will be a fairly accurate first approximation to the solution of Eq. (1). For a complex problem,[2] the various parts of the system may have widely differing time constants so that the solution of Eq. (1) may reach a quasi-equilibrium state and then develops adiabatically as the parameter $\theta(t)$ proceeds to equilibrium. This is usually reflected in the fact that u is described by an equilibrium distribution with a time dependent temperature. For such problems, in order to study the energy transfer among different parts of the system[3] some of

[2] For example, a system of mutually interacting particles which gives rise to a system of coupled equations each of which is of form of Eq. (1).

[3] This could be among different species of the plasma, or physical regions in coordinate space.

which may be in quasiequilibrium states, it is mandatory that a numerical scheme represent this quasi-equilibrium solution accurately. Otherwise, small energy exchange may take place due to numerical error, and experience has shown that these exchanges tend to be cumulative so that the entire character of the problem may change drastically.

II. Basic Difference Scheme and Particle Conservation

To develop a numerical solution for Eq. (1), we represent $u(x, t)$ by the discrete set $u_j^n = u(j\,\Delta x, n\,\Delta t)$ where Δx and Δt define the velocity and time mesh sizes. For convenience these are taken to be constant, though they may in general be variable, Δx_j and Δt_n. In an effort to avoid all nonessential complexity and without loss of generality (at least in as much as what we want to discuss) we shall base our discussion on a commonly used difference scheme. The time derivative will be forward differenced, i.e., $\partial u/\partial t = 1/\Delta t(u^{n+1} - u^n)$ and all the velocity (momentum) derivatives will be center differenced, e.g.,

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_j = \frac{1}{(\Delta x)^2}(u_{j+1} - 2u_j + u_{j-1}).$$

The resulting difference equation will be solved implicitly, i.e., the finite difference representation of the right side of Eq. (1) will contain u^{n+1}. The implicit method is preferred over the explicit because of its inherent advantage in stability.

The standard way to set up a difference scheme for Eq. (1) is to rewrite it as

$$\begin{aligned}\frac{\partial u}{\partial t} &= \frac{1}{A}\left[\frac{\partial B}{\partial x}u + \left(B + \frac{\partial C}{\partial x}\right)\frac{\partial u}{\partial x} + C\frac{\partial^2 u}{\partial x^2}\right] \\ &= \frac{1}{A}\left[\alpha u + \beta\frac{\partial u}{\partial x} + \gamma\frac{\partial^2 u}{\partial x^2}\right] \qquad (7)\end{aligned}$$

and then do the differencing. This has the apparent advantage that the derivatives of B and C may be evaluated exactly in any difference scheme thus hopefully decreasing the truncation error for a given Δx. Unfortunately, as we shall show, this type of scheme does not guarantee particle conservation so is unsuitable for our practical purposes.

To represent the fact that we are using an implicit scheme, we will write the u's appearing in the right side of Eq. (7) as $\tilde{u}_j$. These will be some linear combinations of u_j^n and u_j^{n+1} depending on the particular implicit differencing scheme. We also

will use values of α, β, and γ evaluated at time step n, though other options are available such as using time extrapolated values. Equation (7) is represented by

$$\frac{1}{\Delta t}(u_j^{n+1} - u_j^n) = \frac{1}{A_j}\left[\alpha_j^n \tilde{u}_j + \beta_j^n \frac{(\tilde{u}_{j+1} - \tilde{u}_{j-1})}{2\Delta x} + \gamma_j^n \frac{(\tilde{u}_{j+1} - 2\tilde{u}_j + \tilde{u}_{j-1})}{(\Delta x)^2}\right]$$

$$= \frac{1}{A_j}\left[\left(\frac{1}{(\Delta x)^2}\gamma_j^n - \frac{1}{2\Delta x}\beta_j^n\right)\tilde{u}_{j-1} + \left(\alpha_j^n - \frac{2}{(\Delta x)^2}\gamma_j^n\right)\tilde{u}_j \right.$$

$$\left. + \left(\frac{1}{(\Delta x)^2}\gamma_j^n + \frac{1}{2\Delta x}\beta_j^n\right)\tilde{u}_{j+1}\right]. \tag{8}$$

This is defined for a set of j values from 0 to J, where $J\,\Delta x$ is the maximum velocity which is chosen large enough so that all $u_{j>J}$ may be taken as zero.

For the present purposes, we take the simple definition

$$N^n \equiv \sum_{j=0}^{J} A_j u_j^n \,\Delta x \tag{9}$$

for the number of particles. A more accurate representation of the total particle number integral (Eq. (3)) will not in general change the results. For strict conservation we require $N^{n+1} = N^n$ for all n. Looking at Eq. (8), we find

$$\frac{1}{\Delta t\,\Delta x}(N^{n+1} - N^n) = \sum_{j=0}^{J}\left\{\left(\frac{1}{(\Delta x)^2}\gamma_j^n - \frac{1}{2\Delta x}\beta_j^n\right)\tilde{u}_{j-1} + \left(\alpha_j^n - \frac{2}{(\Delta x)^2}\gamma_j^n\right)\tilde{u}_j \right.$$

$$\left. + \left(\frac{1}{(\Delta x)^2}\gamma_j^n + \frac{1}{2\Delta x}\beta_j^n\right)\tilde{u}_{j+1}\right\}$$

$$= \sum_{j=1}^{J-1}\left[\frac{1}{(\Delta x)^2}(\gamma_{j+1}^n - 2\gamma_j^n + \gamma_{j-1}^n) + \frac{1}{2\Delta x}(\beta_{j-1}^n - \beta_{j+1}^n) + \alpha_j^n\right]\tilde{u}_j$$

$$+ \left(\frac{1}{(\Delta x)^2}\gamma_J^n + \frac{1}{2\Delta x}\beta_J^n\right)\tilde{u}_{J+1}$$

$$+ \left[\frac{1}{(\Delta x)^2}(\gamma_{J-1}^n - \gamma_J^n) + \frac{1}{2\Delta x}\beta_{J-1}^n + \alpha_J^0\right]\tilde{u}_J$$

$$+ \left[\frac{1}{(\Delta x)^2}(\gamma_1^n - 2\gamma_0^n) - \frac{1}{2\Delta x}\beta_1^n + \alpha_0^n\right]\tilde{u}_0$$

$$+ \left(\frac{1}{(\Delta x)^2}\gamma_0^n - \frac{1}{2\Delta x}\beta_0^n\right)\tilde{u}_{-1}\,. \tag{10}$$

Since α, β, and γ are usually nonconstant functions of x and t, the right side of Eq. (10) in general will not vanish. There are two kinds of terms in this expression, the interior terms ($j = 1, \ldots, J-1$) and boundary related terms ($j = -1, 0, J, J+1$).

On the other hand, the conservation equation (Eq. (3)) has only boundary terms. Consequently, even if approximate particle conserving boundary conditions were applied the present difference scheme still will not have particle conservation due to the interior terms which in effect are sources or sinks for particles. By expanding α, β, and γ in Taylor series we can show that the change in particle number per time step is proportional to $(\Delta x)^2$. Assuming the boundary terms vanish, we will have the relation

$$N^{n+1} - N^n = \Delta t \frac{(\Delta x)^2}{12} A(\beta, \gamma, \tilde{u})$$

where

$$A(\beta, \gamma, \tilde{u}) = \sum_{j=1}^{J} \left[\frac{\partial^3}{\partial x^3} \frac{\partial \gamma}{\partial x} - 2\beta \right]_{x=j\Delta x} \tilde{u}_j \Delta x \simeq \int_0^\infty \left[\frac{\partial^3}{\partial x^3} \left(\frac{\partial \gamma}{\partial x} - 2\beta \right) \right] \tilde{u}\, dx$$

or

$$N^{n+1} - N^n \simeq \Delta t \frac{(\Delta x)^2}{12} I(\beta, \gamma, \tilde{u}), \tag{10'}$$

$I(\beta, \gamma, \tilde{u})$ being the above integral. In a practical problem, when Δx is very small (say $\Delta x / x_{\max} \ll 10^{-2}$) the error indeed may be small and even change sign. But after reaching quasiequilibrium the error becomes monotonic, because then u changes very slowly. The accumulation of this unidirectional steady error will become significant.[3] For academic problems this actually may be used as a practical guide to terminate the problem. But if this equation (Eq. (1)) is only part of a bigger system, this error will best be removed, due to the abundance of quasi-equilibrium regions.

To achieve strict conservation without regard to Δx or number of time steps, we will difference Eq. (1) as it stands. This is slightly less accurate in the sense that some derivatives which could be done exactly are approximated by a finite difference representation. However, a brief consideration of why the scheme of Eq. (8) did not conserve particles will show that the inconsistancy of evaluating the derivatives of B and C exactly while doing those of u approximately was the real culprit. Thus, the apparent increase of accuracy is somewhat dubious, at least when judged on the basis of this fundamental moment of u. Of course, as Δx is made smaller and smaller, the answers obtained from Eq. (8) will tend to stay closer to the exact results. Reasonable values of Δx can be obtained, though, which when used with the difference scheme to be presented give sufficient convergence, strict particle conservation, and other advantages.

To difference Eq. (1) we define a generalized current $F = Bu + C(\partial u/\partial x)$. Then using a centered difference in velocity, we obtain

$$\frac{1}{\Delta t}(u_j^{n+1} - u_j^n) = \frac{1}{A_j \Delta x}(\tilde{F}_{j+1/2} - \tilde{F}_{j-1/2}) \tag{11}$$

for our representation of the differential equation. It should be remarked that here we require $\tilde{F}_{j\pm1/2}$ which means knowing $\tilde{u}_{j\pm1/2}$. Since the $\tilde{u}$'s are defined only at the j's, we have a certain amount of ambiguity in the definition of $\tilde{F}_{j\pm1/2}$. In the next section, we will exploit this fact to achieve some desired intrinsic properties for u. For the present, we just assume that the $\tilde{F}$'s are defined consistently so that $\tilde{F}_{j-1/2} = \tilde{F}_{k+1/2}\,|_{k=j-1}$.

From the difference equation above, we see that

$$\sum_{j=0}^{J} A_j \frac{\Delta x}{\Delta t}(u_j^{n+1} - u_j^n) = \sum_{j=0}^{J}(F_{j+1/2} - F_{j-1/2}) = F_{J+1/2} - F_{-1/2} \tag{12}$$

and this is the exact numerical equivalent of Eq. (4). Number conservation is obtained by imposing the boundary conditions $F_{J+1/2} = 0 = F_{-1/2}$.[4] To the order of accuracy of the difference scheme, these are consistent with the conditions on the differential equation, i.e., as $\Delta x \to 0$, these reduce to

$$[Bu + C(\partial u/\partial x)]_{x=0} = [Bu + C(\partial u/\partial x)]_{x=x_{\max}} = 0.$$

The physical interpretation of the particle conserving boundary conditions is quite evident, i.e., no net current flow through the boundaries. Also, the fact that strict conservation holds regardless of the size of Δx can be seen and is the main reason this type of scheme is so useful in a practical calculation.[5]

III. Nonnegative Spectrum and Equilibration

The usual difference methods for Fokker–Planck type equations, be they particle conserving or not, may produce negative u's for values of Δx used in a practical problem. This behavior appears in the large x portion of the spectrum and may be traced to the fact that the convergence of the difference scheme to the differential equation is determined by $x\,\Delta x$ instead of Δx. In order to treat high velocity source terms, as many practical problems require, we clearly need an alternative scheme. The most obvious is to change variables to x^2 since $x\,dx = \frac{1}{2}d(x^2)$; however, a consideration of the number of $\Delta(x^2)$ mesh points needed to represent the entire problem accurately shows that this is an undesirable alternative. We

[4] From now on we have implicitly assumed that the current vanishes at the boundary. This assumption simplifies some of our discussions (and is more physical) which can always be extended to the more general cases.

[5] An analogous exploitation of differencing the divergence has been applied to an anisotropic plasma [8].

choose, instead, to modify the difference equations in such a way as to eliminate this nonuniformity without having to use many mesh points. To obtain the desired modification, we look also at the equilibrium spectrum predicted by the numerical scheme since this has the same convergence problems. It turns out that by assuring proper equilibration we can prove that the spectrum remains nonnegative for all times as required.

We base our arguments on the particle conserving scheme, Eq. (11), using a fully implicit representation, i.e., $\tilde{u}_j = u_j^{n+1}$. In the generalized current, $F_{j+1/2} = [Bu + C(\partial u/\partial x)]_{j+1/2}$, we use a centered difference for the x derivative to get

$$\frac{\partial u}{\partial x}\Big|_{j+1/2}^{n+1} \simeq \frac{1}{\Delta x}(u_{j+1}^{n+1} - u_j^{n+1}).$$

This is expressed in terms of implicitly known values of u so no ambiguity exists. However, we also will require $u_{j+1/2}^{n+1}$ which is not known. The *usual* choice for this quantity is $u_{j+1/2}^{n+1} \simeq \frac{1}{2}(u_{j+1}^{n+1} + u_j^{n+1})$ which can be shown to give rise to a second order accurate scheme in Δx when used in Eq. (11). We now show that this choice gives rise to negative u_j's if Δx is not small enough.

Inserting the above definitions for $u_{j+1/2}$ and $\partial u/\partial x\,|_{j+1/2}$ into the expression for the current gives

$$F_{j+1/2} = \left(\frac{1}{2}B_{j+1/2}^n + \frac{1}{\Delta x}C_{j+1/2}^n\right)u_{j+1}^{n+1} - \left(\frac{1}{\Delta x}C_{j+1/2}^n - \frac{1}{2}B_{j+1/2}^n\right)u_j^{n+1}. \quad (13)$$

At equilibrium (or quasiequilibrium), this goes to zero which means the spectrum satisfies

$$\frac{u_{j+1}^{n+1}}{u_j^{n+1}} = \frac{\frac{1}{\Delta x}C_{j+1/2}^n - \frac{1}{2}B_{j+1/2}^n}{\frac{1}{\Delta x}C_{j+1/2}^n + \frac{1}{2}B_{j+1/2}^n}. \quad (14)$$

Now if for some j we have

$$\Delta x > \frac{1}{2}\frac{C_{j+1/2}^n}{B_{j+1/2}^n},$$

then it is clear that u_{j+1}^{n+1} and u_j^{n+1} have different signs (B and C are positive) so one must in general be negative. This means that even though all of the u_j's are initially nonnegative, some of them will eventually change sign, certainly near equilibrium if not sooner. A precise statement as to when this occurs of course depends on the initial shape of u. We find that a localized, delta-function type input will immediately give negative u's; this being probably the most pathological

case. However, even *smooth* functions (e.g., a flat spectrum) develop negative elements quite early. Thus, unless we are willing to use

$$\Delta x < \frac{1}{2}\frac{C^n_{j+1/2}}{B^n_{j+1/2}} \qquad \text{for all } j,$$

and this is much too strong a condition since B and C may be time dependent, a different scheme must be employed.

To see how to eliminate the negative spectra, we compare the numerical estimate of the equilibrium with the exact result. From Eq. (6), we find that the latter is given by[6]

$$\frac{u^{n+1}_{j+1}}{u^{n+1}_{j}} = \exp\left[-\int_{j\Delta x}^{(j+1)\Delta x} \frac{B^n(y)}{C^n y)}\right] dy$$

$$\simeq \exp\left[-\frac{B^n_{j+1/2}}{C^n_{j+1/2}}\Delta x\right] \tag{15}$$

which agrees with Eq. (14) through order $(\Delta x)^2$ as long as $\Delta x \ll C^n_{j+1/2}/B^n_{j+1/2}$. It is exactly the breakdown of this condition which causes negative equilibrium spectra; the numerical scheme overestimates the relative fall off of the spectrum to such an extent as to make it go negative. We, thus, must find a way to weaken this estimate.

The approach we take is to note that it is consistent with the difference scheme to define $u^{n+1}_{j+1/2} = (1 - \delta_j)\, u^{n+1}_{j+1} + \delta_j u^{n+1}_j$ for any $0 \leqslant \delta_j \leqslant 1/2$, i.e., the only thing we know is that the $(j + 1/2)$ valuc is somewhere between the j and $j + 1$ values. We now show that this parameter may be exploited to give proper equilibration and nonnegative spectra. To see how the former may be attained, we note that the current is now given by

$$F_{j+1/2} = \left[(1 - \delta_j)\, B^n_{j+1/2} + \frac{1}{\Delta x} C^n_{j+1/2}\right] u^{n+1}_{j+1} - \left(\frac{1}{\Delta x} C^n_{j+1/2} - \delta_j B^n_{j+1/2}\right) u^{n+1}_j \tag{16}$$

so that the equilibrium spectrum satisfies

$$\frac{u^{n+1}_{j+1}}{u^{n+1}_{j}} = \frac{\dfrac{1}{\Delta x} C^n_{j+1/2} - \delta_j B^n_{j+1/2}}{\dfrac{1}{\Delta x} C^n_{j+1/2} + (1 - \delta_j)\, B^n_{j+1/2}}. \tag{17}$$

[6] This is under the assumption of vanishing boundary currents.

We then chooses δ_j such that this ratio is $\exp(-\Delta x\, B^n_{j+1/2}/C^n_{j+1/2})$, the exact result. Defining $w_j = \Delta x\, B^n_{j+1/2}/C^n_{j+1/2}$, we find

$$\delta_j = \frac{1}{w_j} - \frac{1}{\exp(w_j) - 1} \tag{18}$$

which can easily be shown to be monotonically decreasing from 1/2 to 0 as w_j goes from 0 to ∞. It may also be remarked that when this choice for $u^{n+1}_{j+1/2}$ is used in a first derivative, the scheme continuously shifts from a centered difference ($\delta_j \equiv 1/2$) to a forward difference ($\delta_j \equiv 0$). It may be asked, "Why not use the forward difference ($\delta_j \equiv 0$) right from the beginning, since this will guarantee the positiveness of u_j ?" The simple answer is given by formulas (15) and (17). Simple forward difference will not give convergence unless $\Delta x \ll C^n_{j+1/2}/B^n_{j+1/2}$ for all j and this is precisely what we want to avoid.

Putting the expression for δ_j into the current and using this in Eq. (11), gives

$$\frac{1}{\Delta t}(u_j^{n+1} - u_j^n) = \frac{1}{A_j(\Delta x)^2}\,[C^n_{j+1/2}W_j \exp w_j\, u^{n+1}_{j+1} - (C^n_{j+1/2}W_j + C^n_{j-1/2}W_{j-1}\exp w_{j-1})\, u_j^{n+1} + C^n_{j-1/2}W_{j-1}u^{n+1}_{j-1}], \tag{19}$$

where $W_j = w_j/(\exp w_j - 1)$. Now the nonnegative aspect of the solution may be proven directly from the structure of Eq. (19).[7]

The above equation may be written as

$$-\tilde{A}_j u^{n+1}_{j+1} + \tilde{B}_j u^{n+1}_j - \tilde{C}_j u^{n+1}_{j-1} = u_j^n, \qquad j = 1,\ldots, J \tag{20}$$

with $\tilde{A}_j$, $\tilde{B}_j$, and $\tilde{C}_j$ all $\geqslant 0$. The boundary conditions are $\tilde{C}_0 = 0 = \tilde{A}_J$. Applying the "back substitution algorithm", assuming $u_j^{n+1} = e_j^{n+1}u^{n+1}_{j+1} + f_j^{n+1}$, we obtain

$$e_j^{n+1} = \frac{\tilde{A}_j}{\tilde{B}_j - \tilde{C}_j e^{n+1}_{j-1}} \qquad \text{and} \qquad f_j^{n+1} = \frac{u_j^n + C_j f^{n+1}_{j-1}}{\tilde{B}_j - \tilde{C}_j e^{n+1}_{j-1}}. \tag{21}$$

Following [5] there exists $a = \Delta t/(\Delta x)^2$ such that $0 \leqslant e_j^{n+1} \leqslant 1$ for all j. Considering the family of e_j^{n+1}'s as functions of a and using the recursion relations (21), we can show inductively that $e_j^{n+1} > 0$ and bounded for all j and a.

Then if all $u_j^n \geqslant 0$, we readily see that $f_j^{n+1} \geqslant 0$ for all j since

$$f_0^{n+1} = u_0^n e_0^{n+1}/A_0 \geqslant 0.$$

[7] The following proof is a generalization of reference [5], p. 119.

Thus we find that

$$u_j^{n+1} = e_j^{n+1} u_{j+1}^{n+1} + f_j^{n+1} \geqslant 0.$$

remembering that $\tilde{A}_J = 0 \Rightarrow u_J^{n+1} = f_J \geqslant 0$. This completes the demonstration that the present difference scheme gives nonnegative spectra.

In summary, the preceeding discussions give us the following procedure for a numerical solution of the Fokker–Planck Eq. (1).

1. Obtain the quasiequilibrium solution $\bar{u}$ of the ordinary differential equation ($t = (n + 1)\,\Delta t$ is considered to be a parameter)

$$\frac{1}{A(x)} \frac{\partial}{\partial x} \left[B(x, t)\, \bar{u} + C(x, t) \frac{\partial \bar{u}}{\partial x} \right] = 0.$$

Usually this will yield $\bar{u}$ in a closed form. But at the worst we can simply calculate formula (15) numerically, since we only require $\bar{u}_{j+1}/\bar{u}_j$ in step 2.

2. Calculate δ_j^{n+1}'s from the current condition

$$\left[(1 - \delta_j^{n+1})\, B_{j+1/2}^n + \frac{1}{\Delta x} C_{j+1/2}^n \right] \bar{u}_{j+1}^{n+1} - \left(\frac{1}{\Delta x} C_{j+1/2}^n - \delta_j^{n+1} B_{j+1/2}^n \right) \bar{u}_j^{n+1} = 0.$$

3. Solve the following implicit difference equation by the *back substitution* algorithm.

$$\begin{aligned} \frac{1}{\Delta t} (u_j^{n+1} - u_j^n) = \frac{1}{A_j \Delta x} \Big\{ & \left[(1 - \delta_j^{n+1})\, B_{j+1/2}^n + \frac{1}{\Delta x} C_{j+1/2}^n \right] u_{j+1}^{n+1} \\ & - \left[\frac{1}{\Delta x} (C_{j+1/2}^n + C_{j-1/2}^n) + (1 - \delta_{j-1}^{n+1}) B_{j-1/2}^n - \delta_j^{n+1} B_{j+1/2}^n \right] u_j^{n+1} \\ & + \left[\frac{1}{\Delta x} C_{j-1/2}^n - \delta_{j-1}^{n+1} B_{j-1/2}^n \right] u_{j-1}^{n+1} \Big\}. \end{aligned}$$

Here we see that the extra calculations involved in getting δ_j's is insignificant as compared with the rest of the program. In some cases, the use of δ_j's actually results in a net saving in the amount of calculation required to obtain the difference coefficients (for example see Section IV). Consequently, with practically no increase in real machine time, we have succeeded in removing the strongest conditions on Δx and Δt as required by conservation of particles, positiveness of solution, and accuracy of equilibrium solution. The subsequent freedom in choosing Δx and Δt as directed by accuracy alone will significantly decrease the real computation time. We have illustrated these points with two examples.

IV. Thermalization of Fast Ions in a Plasma

As a simple application of the proposed difference scheme, we consider the slowing down of a charged test particle in a Maxwell–Boltzmann plasma. This will be described by linearizing the Fokker–Planck equation of Rosenbluth, MacDonald and Judd [2] assuming a singly ionized background plasma of density n, and temperature T. Replacing the velocity and time variables by the dimensionless ones

$$x = v\left(\frac{M}{kT}\right)^{1/2} \qquad \text{and} \qquad \tau = \frac{4\pi n e^4 \ln \Lambda}{M^{1/2}(kT)^{3/2}}\, t,$$

where M is the test particle mass (assumed equal to the mass of the plasma ions) and $\ln \Lambda$ is the usual Coulomb logarithm, the relevant equation [1] is

$$\frac{\partial u}{\partial \tau} = \frac{1}{x^2}\frac{\partial}{\partial x}\left(xGu + G\frac{\partial u}{\partial x}\right), \qquad (0 < \tau < \infty, 0 \leqslant x < \infty). \tag{23}$$

In this expression u is the distribution function of the test particle, assumed to be spherically symmetric, and G is given by

$$G(x) = \left(\frac{2}{\pi}\right)^{1/2}\left[R(x) + \rho R\left(\frac{x}{\rho}\right)\right]$$

with ρ^2 the ion-electron mass ratio and

$$R(x) = \frac{1}{x}\int_0^x \exp(-y^2/2)\, dy - \exp(-x^2/2).$$

The initial condition for u is taken to be a delta function at some x_0 corresponding to the introduction of the test particle into the plasma at a well-defined velocity. Then, the above differential equation describes the slowing down and subsequent thermalization of the particle to a distribution proportional to $\exp(-x^2/2)$.

To set up the difference scheme, we must determine the parameters δ_j defined by Eq. (18). In the present case, $B = xG$ and $C = G$ so that

$$w_j = \Delta x \frac{B^n_{j+1/2}}{C^n_{j+1/2}} = x_{j+1/2}\,\Delta x$$

and thus

$$\delta_j = \frac{1}{w_j} - \frac{1}{\exp(w_j) - 1}.$$

The difference equation then becomes

$$u_j^{n+1} - u_j^n = \frac{\Delta\tau}{x_j^2(\Delta x)^2}\left[F_{j+1/2}^{n+1} - F_{j-1/2}^n\right]$$

with

$$F_{j+1/2}^{n+1} = G_{j+1/2}\frac{w_j}{[\exp(w_j) - 1]}\left[\exp(w_j)\, u_{j+1}^{n+1} - u_j^{n+1}\right]$$

and the boundary conditions

$$F_{1/2} = 0 = F_{j_{\max}+1/2}\,.$$

With $10 \leqslant x_0 \leqslant 30$, it is found that $\Delta x = 1.5$ and $\Delta\tau = 1.0$ work sufficiently well when compared with any finer mesh or other method of calculation (e.g., the results of reference [1]). For larger Δx, the distribution function becomes somewhat distorted; the three constraints of number conservation, nonnegative u, and proper equilibration will of course remain true no matter what Δx is chose.

For comparison, Eq. (23) was differenced according to Eq. (8) which represents the *typical*, unconstrained scheme. Using $x_0 = 25$, we found $\Delta x \leqslant .05$ was required to eliminate negative values of the distribution function. With the threshold value $\Delta x = .05$, the equilibrium temperature was .3 % too low, while the total number of particles at the equilibrium time $\tau \simeq 500$ was .2 % low and showing a systematic loss rate of $(1/N)(\Delta N/\Delta\tau) \sim 10^{-5}$. The time evolution and spectral characteristics of this calculation were found to be in agreement with the practical scheme (using $\Delta x = .15$) to within 1 %. Thus, as long as this Fokker–Planck equation is used in calculations where no times greater than about $10\tau_{eq}$ are required, Eq. (8) would represent the solution with sufficient accuracy. However, using the practical scheme, we see that a factor of 3 in mesh size and computation time may be gained. Moreover, a wider class of problems may be handled since no restriction is placed on the length of time which the calculation runs. Finally, using procedures similar to those presented in the next section, a more general difference scheme using variable mesh sizes may be employed since now the large x behavior of the system is handled accurately. Thus, an even greater saving in time and number of mesh points is possible.

V. Compton Cooling of Hot Electrons

Although the above example is a linear partial differential equation, the proposed scheme applies equally well to nonlinear equations. Consider the problem of the nonequilibrium time evolution of a photon distribution due to Compton scattering

with an infinite, homogeneous Maxwellian electron plasma. Assuming isotropic scattering and in the nonrelativistic limit this process may be described by the nonlinear Fokker–Planck equation [6, 7]

$$\frac{\partial f}{\partial t} = \frac{\sigma_0 \rho_e}{k^2} \frac{\partial}{\partial k} \left\{ k^4 \left[\theta_e(t) \frac{\partial f}{\partial k} + f + f^2 \right] \right\}, \qquad (0 < t < \infty, 0 \leqslant k < \infty), \tag{24}$$

where $\sigma_0 = .67$ barns, $f = (hc)^3/2\ f_p(k, t)$ and $f_p(k, t)$ is the nonequilibrium photon distribution; $k = h\nu$, $\theta_e(t)$ and ρ_e are the electron temperature and density, respectively. The total photon number $N(t)$ and total photon energy $E(t)$ can be found by

$$N(t) = 4\pi \int_0^\infty k^2 f(k, t)\, dk$$

and

$$E(t) = 4\pi \int_0^\infty k^3 f(k, t)\, dk. \tag{25}$$

For the numerical solution of Eq. (24) the upper limit for k is approximated by $k_{\max}$ which is sufficiently large so as to minimize the distortion of the spectrum. It was found that for the usual finite difference schemes [5], in order to obtain an accurate convergent and stable numerical solution $k_{\max}$ must be very large and Δk must be small, consequently, a large number of mesh points was required. However, with the proposed practical scheme we may use nonuniformly spaced mesh points to great advantage (by saving as much as 20 times or more the number of mesh points with essentially no loss in accuracy). It should be noted that even with nonuniform mesh Δk_j the usual scheme will still have to satisfy the convergence condition

$$\Delta k_j \ll \frac{1}{2} \frac{C^n_{j+1/2}}{B^n_{j+1/2}}$$

and the resultant restriction on the smallness of Δk_j.

If we let $f_j^n = (k_j, n\,\Delta t)$, $\Delta k_{j+1/2} = k_{j+1} - k_j$, and

$$\Delta k_j = \frac{\Delta k_{j+1/2} + \Delta k_{j-1/2}}{2},$$

we obtain the implicit difference equation,

$$\frac{f_j^{n+1} - f_j^n}{\Delta t} = \frac{\alpha}{k_j^2} \frac{1}{\Delta k_j} \left[k^4_{j+1/2} F^{n+1}_{j+1/2} - k^4_{j-1/2} F^{n+1}_{j-1/2} \right]$$

and

$$F_{j+1/2}^{n+1} = \frac{\theta_e^{n+1/2}}{\Delta k_{j+1/2}} (f_{j+1}^{n+1} - f_j^{n+1}) + (1 + f)_{j+1/2}^{n} f_{j+1/2}^{n+1} \tag{25}$$

where $\alpha = \sigma_0 \rho_e = \text{const}$. The energy group positions k_j are input parameters and nonuniformly spaced. The half point values $f_{j+1/2}^{n}$ and $f_{j+1/2}^{n+1}$ will be calculated as follows

$$f_{j+1/2}^{n} = (1 - \delta_j^{n+1}) f_{j+1}^{n} + \delta_j^{n+1} f_j^{n}$$

and

$$f_{j+1/2}^{n+1} = (1 - \delta_j^{n+1}) f_{j+1}^{n+1} + \delta_j^{n+1} f_j^{n+1},$$

where δ_j^{n+1}'s are to be computed at each time step from the quasiequilibrium solution f^e.

At any time t, the quasiequilibrium solution f^e is

$$f^e = \frac{1}{C \exp[k/\theta_e(t)] - 1} \tag{26}$$

where C is some constant to be determined. For Compton scattering the total photon number N is a conserved quantity and for distribution (26) N is related to θ_e and C by

$$N = 8\pi \theta_e^{\,3} \sum_{n=1}^{\infty} \frac{C^{-n}}{n^3}.$$

Hence, knowing N and $\theta_e(t)$ at any instant in time we can compute C iteratively. Numerically, quasiequilibrium is achieved if $F_{j+1/2}^{n+1} = 0$, for all j, i.e.,

$$\frac{\theta_e^{n+1/2}}{\Delta k_{j+1/2}} (f_{j+1}^{e} - f_j^{e}) + [1 + (1 - \delta_j^{n+1}) f_{j+1}^{e} + \delta_j^{n+1} f_j^{e}][(1 - \delta_j^{n+1}) f_{j+1}^{e} + \delta_j^{n+1} f_j^{e}] = 0,$$

and this quadratic equation in δ_j^{n+1} can be solved easily. As it turns out for this particular problem there is always one and only one positive real root satisfying $0 \leqslant \delta_j \leqslant 1/2$, and this is the desired result.

In the test problems, we have taken $\rho_e = 6 \times 10^{22}$ particles/cm^3, $\Delta t = 5 \times 10^{-11}$ sec and $\theta_e = 10,25,50$ keV, while the initial photon spectra is composed of low energy lines of different intensity. The total energy of the system is held fixed so that we may observe the cooling of the initially hotter electrons. For the above values of Δt and ρ_e the hottest problem takes about 1200 cycles to reach equilibrium; using 20 mesh points we obtained convergence (the usual method would

take at least 400–800 mesh points in order to obtain the same accuracy). The time dependent energy exchange rates, when compared with the converged results for the usual method, are accurate to more than .5 % at all times and the equilibrium spectrum can be said to be exactly duplicated. Energy conservation is of course exact, since the energy gain by the photons is used to determine the energy loss of the electrons.

REFERENCES

1. H. C. KRANZER, *Phys. Fluids* **4** (1961), 214.
2. M. N. ROSENBLUTH, W. M. MACDONALD, AND D. L. JUDD, *Phys. Rev.* **109** (1957), 1
3. J. KILLEEN AND K. D. MARX, "Methods in Computational Physics," Vol. 9, Academic Press, New York, (1969).
4. D. C. MONTGOMERY AND D. A. TIDMAN, "Plasma Kinetic Theory," McGraw–Hill, New York, 1964.
5. R. D. RICHTMYER AND K. W. MORTON, "Difference Methods for Initial-Value Problems," 2nd ed., Wiley (Interscience), New York, 1967.
6. A. R. FRASER, AWRE Report No. 0–82/65, 1965.
7. G. C. POMRANING, Los Alamos Report No. LA–4006–MS, 1968.
8. J. WHITNEY, Los Alamos Report No. LA–3990, 1968; Columbia University Plasma Laboratory Report No. 45, 1969.