1996ApJS..103..255P

THE ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES, 103:255–267, 1996 March
© 1996. The American Astronomical Society. All rights reserved. Printed in U.S.A.

FOKKER-PLANCK EQUATIONS OF STOCHASTIC ACCELERATION: A STUDY OF NUMERICAL METHODS

BRIAN T. PARK[1] AND VAHÉ PETROSIAN[1,2]
Center for Space Science and Astrophysics, Stanford University, Stanford, CA 94305
Received 1995 April 5; accepted 1995 August 15

ABSTRACT

Stochastic wave-particle acceleration may be responsible for producing suprathermal particles in many astrophysical situations. The process can be described as a diffusion process through the Fokker-Planck equation. If the acceleration region is homogeneous and the scattering mean free path is much smaller than both the energy change mean free path and the size of the acceleration region, then the Fokker-Planck equation reduces to a simple form involving only the time and energy variables. In an earlier paper (Park & Petrosian 1995, hereafter Paper 1), we studied the analytic properties of the Fokker-Planck equation and found analytic solutions for some simple cases. In this paper, we study the numerical methods which must be used to solve more general forms of the equation. Two classes of numerical methods are finite difference methods and Monte Carlo simulations. We examine six finite difference methods, three fully implicit and three semi-implicit, and a stochastic simulation method which uses the exact correspondence between the Fokker-Planck equation and the Itô stochastic differential equation. As discussed in Paper I, Fokker-Planck equations derived under the above approximations are singular, causing problems with boundary conditions and numerical overflow and underflow. We evaluate each method using three sample equations to test its stability, accuracy, efficiency, and robustness for both time-dependent and steady state solutions. We conclude that the most robust finite difference method is the fully implicit Chang-Cooper method, with minor extensions to account for the escape and injection terms. Other methods suffer from stability and accuracy problems when dealing with some Fokker-Planck equations. The stochastic simulation method, although simple to implement, is susceptible to Poisson noise when insufficient test particles are used and is computationally very expensive compared to the finite difference method.

Subject headings: acceleration of particles — diffusion — methods: numerical

1. INTRODUCTION

Ever since its introduction by Fermi (1949, 1954) and Davis (1956), stochastic (or second-order Fermi) acceleration has been advanced as a mechanism for accelerating electrons and ions to suprathermal energies. One commonly assumed agent of this acceleration is plasma wave turbulence, which is expected to be present in nonequilibrium conditions of highly magnetized plasmas. Charged particles, spirialing along magnetic field lines, are then accelerated through resonant interactions with plasma waves. This problem is often treated in the quasi-linear approximation (see, e.g., Schlickeiser 1989) and leads to the Fokker-Planck equation with a diffusion coefficient whose magnitude and form depends on the power spectrum and other characteristics of the plasma turbulence. In general, the resultant equation is complicated, and one must resort to some simplifying approximations. Two commonly used simplifications are the following.

First, if the rate of pitch angle scattering is much larger than the rate of energy change and other relevant rates (e.g., rate of particle escape), then the distribution of particles can be assumed to be isotropic. We can integrate out the pitch angle variable in the Fokker-Planck equation by averaging over a timescale longer than the timescale for pitch angle diffusion, but shorter than the timescale for energy diffusion. Second, the dependence on the spatial variable can be eliminated by using a volume-integrated distribution, leading to a considerable simplification if the magnetic field, the particle density, and the turbulence energy density are nearly constant throughout the acceleration region. For efficient acceleration, the scattering mean free path of particles must be much smaller than the size of turbulent region, in which case the spatial convection of particles can be approximated by spatial diffusion. Eventually, particles will leave the acceleration region. The loss of these particles can be modeled by adding an energy-dependent escape term to the Fokker-Planck equation. Under these assumptions, the Fokker-Planck equation becomes a function of only time t and energy x. In spite of these simplifications, analytic solutions can be found only for limited and cases; numerical methods must be used for more general cases.

In Park & Petrosian (1995, hereafter Paper I), we examined the analytical properties of Fokker-Planck equations having this simplified form. Previous treatments of these equations suffered from incorrect or ambiguous boundary conditions because they did not use singular boundary conditions to account for the singularity of the Fokker-Planck equation at the boundary points $x = 0$ and $x = \infty$. We described an extension of the familiar Sturm-Liouville eigenfunction expansion theory which can deal with these special problems. Using this technique, we solved the steady state and the time-dependent Green's functions of three specific cases to study the dependence of the solution on the coefficients of the Fokker-Planck equation. In general, we found that the solutions have a power-law or an exponential energy dependence. The forms of these solutions can be estimated from the energy dependences of the

[1] Department of Applied Physics.
[2] Department of Physics.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

diffusion, advection, and escape timescales of the Fokker-Planck equation.

In this paper, we study the numerical properties of the Fokker-Planck equation and address two problems mentioned in Paper I. The first problem is determining the numerical boundary conditions. Analytically, the entire energy range from 0 to ∞ can be examined. Numerically, only a finite range can be studied because a singular Fokker-Planck equation can diverge at the boundaries. In addition, singular boundary conditions cannot be implemented numerically. A finite range implies that the boundary points are regular so we must determine what regular boundary conditions must be applied. The second problem is the possibility of numerical underflow or overflow caused by the divergence of the solutions near the boundaries. Over the energy range of interest, the solution may vary by many orders of magnitude and numerical truncation errors or catastrophic cancelation errors can occur. We must find a method which is not susceptible to these errors.

Numerical methods for solving Fokker-Planck equations fall into two broad classes, Monte Carlo simulations and finite difference schemes. Bai (1982) used a Monte Carlo method to study electron transport under the effects of Coulomb scattering. Miller & Ramaty (1989) extended the Monte Carlo study to ultrarelativistic electrons under the influence of synchrotron losses, magnetic convergence, pitch angle scattering from plasma turbulence, and other processes. More recently, MacKinnon & Craig (1991) advocated a Monte Carlo simulation technique using the exact equivalence between the Fokker-Planck equation and the Itô stochastic differential equation (van Kampen 1992, chap. 9), and demonstrated the study of a particle transport problem. Achterberg & Krülls (1992) and Krülls & Achterberg (1994) used this method to study the stochastic acceleration of particles. Although Monte Carlo simulations are relatively easy to implement, they are computationally very expensive. This limits the number of test particles which can be simulated, making the results susceptible to Poisson noise. For the same reason, it is difficult to perform searches through the parameter space of the Fokker-Planck equation.

Finite difference schemes are computationally very efficient and produce accurate solutions. They are, however, more difficult to implement because they have various stability constraints. Miller, Guessoum, & Ramaty (1990) used a fully implicit finite difference scheme to solve the time-dependent and steady state solutions of the "hard-sphere" (see, e.g., Ramaty 1979) Fokker-Planck equation. Unfortunately, they do not discuss the exact method used and the boundary conditions imposed. Some equations examined in this paper cannot be solved using some fully implicit methods because the methods become unstable. Hamilton, Lu, & Petrosian (1990) examined the time-dependent solutions of a multidimensional Fokker-Planck equation by using the technique of operator splitting to build composite methods from simpler finite difference schemes. Hamilton (1990) and Hamilton & Petrosian (1992) used these techniques to solve the time-dependent one-dimensional Fokker-Planck equation for electrons undergoing wave-particle acceleration in the presence of Coulomb losses (Steinacker, Dröge, & Schlickeiser 1988).

In § 2, we give a description of the Fokker-Planck equation and the boundary conditions. In § 3, we examine the general properties of finite difference schemes for Fokker-Planck equations. In § 4, we give detailed evaluations of six finite difference methods. The first is the simple fully implicit difference scheme described, for example, by Press et al. (1992, chap. 19). The second is a method proposed by Chang & Cooper (1970, hereafter CC70). The third is a method proposed by Larsen et al. (1985, hereafter LLPS85). The next three are semi-implicit versions of these three methods. We evaluate the accuracy and robustness of each method by solving test equations. We also discuss some peculiar boundary effects caused by the singular nature of the Fokker-Planck equations. In § 5, we evaluate the Monte Carlo stochastic simulation method on the same test equations and compare it to the finite difference methods. In § 6, we give the conclusions of this paper.

2. THE MODEL

2.1. *The Equation*

The Fokker-Planck equation described in Paper I (eq. [1]) contains the time variable t and the normalized energy or momentum variable x which extends over the interval $0 < x < \infty$. Following the notation used by CC70 and others, we rewrite this equation as

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

where $u(x, t)A(x)dx$ is the number of particles in the interval x and $x + dx$ at time t. This equation is identical to Paper I (eq. [1]) except for minor renaming of variables and the addition of the phase factor $A(x)$, which is 1 if x represents energy but $4\pi x^2$ if x represents momentum. Although we can always redefine our variables [$u(x, t) \rightarrow A(x)u(x, t)$] to eliminate the phase factor, it does not complicate the derivation of the finite difference scheme so we retain it for compatibility with previous studies. We use $B(x)$, $C(x)$, $T(x)$, and $Q(x)$ for the advective, diffusive, escape, and source terms, respectively. These coefficients are determined by the physical conditions of the background plasma which is responsible for the acceleration process. For pure wave-particle stochastic acceleration, these can be approximated by power-law forms over a wide range of energies. Numerical methods do not require simplified coefficients, but we consider these cases for comparison with known analytic solutions. These coefficients are considered independent of time because we assume that they vary over a timescale longer than the timescale of interest. The numerical techniques discussed in this paper can be readily extended to time-varying coefficients, but the properties of Fokker-Planck equations with time-varying coefficients are not well-understood, and it becomes impossible to verify the numerical solutions with known analytic ones.

Over the energy interval $0 < x < \infty$, we assume that these coefficients have the following characteristics. First, the phase space must have positive volume, so $A(x) > 0$. Second, a mathematically well-posed diffusion equation must have $C(x) > 0$. If $C(x) < 0$, then we obtain an "inverse" diffusion problem that is unstable under small perturbations in the initial condition. This situation does not correspond to the physics that we are studying here. Third, $B(x)$ corresponds to the advective

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

term, describing a systematic tendency for upward or downward drift of particles; this can have any value between $-\infty$ and ∞. Fourth, the escape time must be non-negative $T(x) \geq 0$, otherwise the zero solution, $u(x, t) \equiv 0$, becomes unstable under perturbations. Finally, $Q(x) \geq 0$ because it is the rate of particle injection.

2.2. *Boundary Conditions*

We found in Paper I that equation (1) is singular at $x = 0$ and $x = \infty$, causing the solutions to diverge at these points. We used singular boundary conditions (Paper I, eqs. [17] and [18]) to determine the analytic solutions, but they cannot be implemented numerically. To avoid these problems, we must evaluate the equation over the finite interval $0 < x_0 < x < x_M < \infty$, where at x_0 and x_M, we must impose regular boundary conditions.

In Paper I, we explained that any regular boundary condition of the form $\alpha u + \beta u' = 0$ can be used if the boundaries extend sufficiently beyond the energy range of interest. Two conditions come to mind. The first is the *no-particle* condition, which sets

$$u(x_0) = u(x_M) = 0 , \tag{2}$$

forcing the number density to vanish at the boundaries. The second is the *no-flux* condition (see Paper I, eq. [3]), which requires that

$$F(x_0, t) = F(x_M, t) = 0 , \tag{3}$$

where the particle flux in x-space is given by

$$F(x, t) = C(x)\frac{\partial u}{\partial x} + B(x)u . \tag{4}$$

Here, we follow the definition used by CC70 which contains an overall sign opposite to the customary definition of the flux (Paper I, eq. [2]) to reduce the propagation of repeated minus signs in later derivations.

For the following four reasons, we use the no-flux condition instead of the no-particle condition. First, the analytical study of Paper I suggests that most solutions do not satisfy the no-particle condition. It does suggest that many equations are consistent with the no-flux condition, even if singular boundary conditions must be used to solve them analytically.

Second, as discussed in Paper I, the no-flux condition provides a good approximation for the effects of additional physical processes not directly incorporated into equation (1). Particles cannot gain energy indefinitely because the source of energy which drives the acceleration process is finite. Similarly, particles cannot lose energy below the mean thermal energy of the background plasma because of collisions with these background particles. While the no-particle condition makes some sense at high energies, it is not justified at low energies.

Third, the no-flux condition provides a conservation law for the Fokker-Planck equation. In the absence of sinks and sources [$T(x) = \infty$, $Q(x) = 0$], the total number of particles in the system

$$\mathcal{N}(t) = \int_0^\infty A(x)\, dx\, u(x, t) \tag{5}$$

satisfies the property $d\mathcal{N}/dt = F(x_M, t) - F(x_0, t) = 0$ and is conserved. The no-particle condition gives no such conservation law.

Last, CC70 pointed out that the no-flux condition, when used with some numerical methods, guarantees positive solutions. Positivity is an inherent property of equation (1), and any numerical method should preserve this property. Negative solutions can cause numerical instability in the calculation of the particle distribution. They can also cause subsequent calculations of electron transport effects and the photon production to become unstable.

All numerical methods, regardless of the boundary condition, will produce boundary effects because they must use a finite energy interval instead of an infinite one. If the boundaries are sufficiently far away, then the no-flux boundary condition will produce an accurate solution in the interior region. Near the boundaries, the numerical solution will be accurate if the true solution is consistent with the no-flux condition. However, some equations do not satisfy the no-flux boundary condition. In these cases, we can expect sharp transients near the boundary. Examples are shown in § 4.

3. GENERAL PROPERTIES OF FINITE DIFFERENCE SCHEMES

3.1. *Notational Conventions*

The discrete time steps are indicated by t_n. We can omit the explicit reference to n in defining

$$\Delta t = t_{n+1} - t_n \tag{6}$$

(even though Δt will not be constant) because we will only be considering "one-step" finite difference schemes whose solutions at each step depend only on the immediately previous time step. Multistep finite difference schemes (Strikwerda 1989, chap. 4; Press et al. 1992, chap. 19) are usually higher order, hence potentially more accurate, but they are more difficult to implement and their properties are less well understood. They do not appear to be better suited for our problem, and we do not consider them in this paper.

The continuous variable x over the range x_0 to x_M is divided into $M + 1$ discrete points indicated by x_m, with the integer index m ranging from 0 to M. The midpoint between two mesh points is defined by the simple arithmetic mean

$$x_{m+1/2} = (x_{m+1} + x_m)/2 . \tag{7}$$

Although other definitions (e.g., geometric mean or harmonic mean) may be better suited in special cases, the difference will be minimal if the mesh size is small ($\Delta x_m/x_m \ll 1$). The midpoint difference, defined by

$$\Delta x_{m+1/2} = x_{m+1} - x_m, \tag{8}$$

gives us $\Delta x_m = (x_{m+1} - x_{m-1})/2$, after using equation (7).

Using the above definitions, we can write the Fokker-Planck coefficients at each mesh point as

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

$$A_m = A(x_m), \tag{9}$$

and similarly for B_m, C_m, T_m, and Q_m. At midpoints, however, we define

$$A_{m+1/2} = (A_m + A_{m+1})/2 \tag{10}$$

to avoid the evaluation of $A(x_{m+1/2})$. This is a good approximation if $\Delta x_{m+1/2}$ is sufficiently small because $A_{m+1/2} = A(x_{m+1/2}) + O[(\Delta x_{m+1/2}/x_{m+1/2})^2]$, where we have used the familiar "big-oh" notation $O(x)$ (see, e.g., Ames 1977, chap. 1). Similar definitions hold for $B_{m+1/2}$, $C_{m+1/2}$, and so on. For the time-dependent variables $u(x, t)$, we define

$$u_m^n = u(x_m, t_n), \tag{11}$$

with the superscript n denoting the time step. Combining this with definition (10) allows us to write $u_m^{n+1/2} = (u_m^n + u_m^{n+1})/2$.

3.2. *Mesh Generation*

To obtain the solution at t_{i+1} from the previous solution at t_i, we use the time steps generated by

$$\Delta t \simeq \min (t_{i+1} - t_i, \tau_{\text{natural}})/N, \tag{12}$$

where N is the number of steps between successive solutions, and τ_{natural} is the natural timescale for the specific Fokker-Planck equation. For the equations considered later in § 4, we set $N \simeq 20$ because this gives reasonably accurate solutions, and $\tau_{\text{natural}} \simeq 1$ because the time variable can be renormalized to make the dominant processes have timescales of order unity.

The energy interval of interest in this paper is approximately from $x_0 \simeq 10^{-3}$ to $x_M \simeq 10^3$. If x represents the energy of an electron in units of the rest mass, then this corresponds to an interval from 500 eV to 500 MeV. Ideally, the mesh size Δx_m should be smaller than the characteristic scale over which the solution varies at x_m,

$$\Delta x_m \lesssim \left| \frac{u}{du/dx} \right|_m . \tag{13}$$

If the solution $u(x)$ is exponential or quasi-periodic over a wide energy range, then the right-hand side of equation (13) is nearly constant. The optimal mesh layout is the uniform mesh given by Δx_m = constant. If the solution is a scale-free power law distribution, $u \propto x^\delta$ for some slowly varying index δ, then condition (13) can be satisfied by the familiar logarithmic mesh $\Delta x_m/x_m$ = constant $\lesssim 1/\delta$. In practice, we require also that $\Delta x_m/x_m \lesssim 1$ to minimize errors due to the nonuniform mesh.

Hamilton (1990) examined more general mesh layouts, generated by

$$\frac{\Delta x_{m+1}}{\Delta x_m} = r, \tag{14}$$

where the adjustable parameter r is a constant for all m. The uniform mesh and the logarithmic mesh are special cases with $r = 1$ and $r = (x_M/x_0)^{1/M}$, respectively. The parameter r can be used to shift the density of mesh points between the lower energies and the higher energies depending on the location of the shortest energy scales. Unfortunately, it is difficult to know where this occurs without some a priori knowledge of the solution.

The analytic solutions given in Paper I show that the Fokker-Planck equations often produce scale-free power-law distributions over a wide range of energy. Therefore, we use the logarithmic mesh layout with $M = 100$ which gives $\Delta x_m/x_m \simeq r - 1 \simeq 0.15$ over 6 orders of magnitude. We pick $M = 100$ because it gives accurate solutions without computational costs.

3.3. *Flux Conservative Finite Difference Schemes*

CC70 proposed a flux conservative difference scheme for an equation similar to equation (1) without the source and escape terms. We discretize equation (1) as

$$\frac{u_m^{n+1} - u_m^n}{\Delta t} = \frac{1}{A_m} \frac{F_{m+1/2}^{n+1} - F_{m-1/2}^{n+1}}{\Delta x_m} - \frac{u_m^{n+1}}{T_m} + Q_m \tag{15}$$

for $m = 0, \ldots, M$, where we have used equation (4) for the flux F_m^{n+1}. We show below that the essential properties of the CC70 scheme, developed for $T_m = \infty$ and $Q_m = 0$, can be retained. Following CC70, we implement the no-flux condition (3) as

$$F_{M+1/2}^{n+1} = F_{-1/2}^{n+1} = 0, \tag{16}$$

because, in the absence of sinks or sources, it conserves the numerical particle number

$$\mathcal{N} = \sum_{m=0}^{M} u_m A_m \Delta x_m, \tag{17}$$

analogous to equation (5), regardless of the exact form for F_m^{n+1} or the mesh layout Δx_m.

CC70 proposed one particular choice of F_m^{n+1}; this and five others are evaluated in § 4. In each case, substituting it into equation (15) results in a tridiagonal system of linear equations, which can be written as

$$\begin{cases} -a_m u_{m-1}^{n+1} + b_m u_m^{n+1} - c_m u_{m+1}^{n+1} = r_m, \\ a_0 = c_M = 0, \end{cases} \tag{18}$$

where r_m is a function of u_m^n. Tridiagonal systems can be solved using a very efficient Gaussian elimination with back substitution routine (Press et al. 1992, chap. 2) whose computational time is of $O(M)$ instead of $O(M^3)$.

An important result from CC70 (also Richtmyer & Morton 1967, chap. 8) states that tridiagonal systems satisfying the positivity condition

$$\begin{cases} |b_m| \geq |a_m| + |c_m|, \\ a_m, b_m, c_m \geq 0, \\ r_m \geq 0, \end{cases} \tag{19}$$

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

will guarantee $u_m^{n+1} \geq 0$. The true solution to equation (1) is always positive, so the ideal numerical method should also give positive solutions. Note that condition (19) is the sufficient but not the necessary condition to guarantee positive solutions.

CC70 derived an expression for F_m^{n+1} which, when coupled with the numerical no-flux condition (16), satisfies condition (19). We verify in § 4.3 that this condition continues to hold even with the addition of the sink and source terms. LLPS85 found another expression for F_m^{n+1} which also satisfies the positivity property. We evaluate this method in § 4.4.

3.4. *Explicit and Semi-implicit Methods*

All time indices on the right-hand side of equation (15) are $n + 1$ which makes this discretization *fully implicit.* If we use instead n, we obtain an *explicit* scheme, which is simpler because it does not require the solution of a tridiagonal system of equations. Unfortunately, explicit schemes have a stability criterion $\Delta t/\Delta x_m^2 \lesssim A_m/C_m$ (Strikwerda 1989, chap. 6; Press et al. 1992 chap. 19), which forces the time step to be too small for practical use.

Fully implicit routines, however, are unconditionally stable for all Δx and Δt and therefore are extremely useful for diffusion or Fokker-Planck type problems. One disadvantage of the fully implicit scheme is that it is only first-order accurate in time (Strikwerda 1989, chap. 6), leading to an examination of higher order schemes.

A second-order method can be obtained by replacing all occurrences of $n + 1$ on the right-hand side of equation (15) with $n + \frac{1}{2}$. The resulting system of equations remains tridiagonal. A well-known example of the semi-implicit method is the Crank-Nicholson scheme (Strikwerda 1989, chap. 6; Press et al. 1992, chap. 19). It is unconditionally stable, which makes it useable for diffusion type problems. Unfortunately, it is not dissipative (Strikwerda 1989, chap. 5), which prevents short-wavelength noise from decaying away. Hence, it can be less accurate than a lower order dissipative routine (Strikwerda 1989, chap. 10) such as the fully implicit method. Another major drawback is that the tridiagonal matrices from semi-implicit methods do not guarantee positive solutions because the condition $r_m \geq 0$ in equation (19) can be violated. We show some evidence of this in § 4.5.

All of the preceding discussion relies on the von Neumann stability analysis (Strikwerda 1989, chap. 2) which assumes a uniform mesh and constant coefficients, and neglects the effect of the boundary conditions. While more sophisticated methods are available (see, e.g., Strikwerda 1989, chaps. 9, 11) that do not suffer from these drawbacks, they are much more difficult to apply. Empirical evidence (Press et al. 1992, chap. 19) suggests that the von Neumann analysis is valid even for the nonuniform mesh and variable coefficients used in this paper.

3.5. *Operator Splitting Method*

The technique of *fractional steps* or *operator splitting* (Richtmyer & Morton 1967, chap. 8; Press et al. 1992, chap. 19) is not a distinct method for solving the Fokker-Planck equation but a way of reducing a large problem into a series of smaller ones. It solves a partial differential equation with J *differential* operators ($\mathscr{L}_j$)

$$\frac{\partial u}{\partial t} = \mathscr{L}_1 u + \mathscr{L}_2 u + \cdots + \mathscr{L}_{J-1} u + \mathscr{L}_J u \,, \tag{20}$$

using a sequence of J finite *difference* operators (L_j) to get

$$u^{n+1} = L_J L_{J-1} \cdots L_2 L_1 u^n \,. \tag{21}$$

Each finite difference operator L_j solves the differential equation $\partial u/\partial t = \mathscr{L}_j u$ by advancing the solution Δt in time from u^n to u^{n+1} using $u^{n+1} = L_j u^n$. Press et al. (1992, chap. 19) states that as a rule of thumb, the composite finite difference solution is stable if the operator with the highest number of derivatives is stable, even if the rest are unstable.

Traditionally, this method has been used successfully for multidimensional problems, where the operators are grouped according to their independent variables. For example, Hamilton et al. (1990) used this technique to solve the time-dependent Fokker-Planck equation describing the evolution of electrons in three variables, energy, pitch angle, and spatial distance. Hamilton (1990) and Hamilton & Petrosian (1992) applied this technique for the one-dimensional Fokker-Planck equation (1) with good results. We have discovered that for best results, the differential operator should not be split within a single independent variable for the following reasons.

First, if the operators are splitting haphazardly, the boundary conditions may become ambiguous. For example, suppose one operator contains the second-order diffusive term that requires two boundary conditions, while the other operator contains the first-order advective term that requires only one. The overall effect of different operators solving parts of the whole problem with different boundary conditions cannot be easily predicted.

The second problem is that the finite difference operators do not generally commute, that is to say, $L_i L_j \neq L_j L_i$ for $i \neq j$. On the other hand, the differential operators $\mathscr{L}_j$ do commute under the addition operator. Consequently, there exists an ambiguity on the ordering of the difference operators L_j. The solution obtained from one particular ordering can differ from another. For example, if the timescale of one operator $\tau_1 \simeq u/|\mathscr{L}_1 u|$ differs from the timescale of another $\tau_2 \simeq u/|\mathscr{L}_2 u|$, then the solution $u^{n+1} = L_2 L_1 u^n$ may be significantly different from $u^{n+1} = L_1 L_2 u^n$ if either τ_1 or τ_2 is much less than Δt. In § 4.6, we discuss this problem further by studying a concrete example.

4. EVALUATION OF SPECIFIC METHODS

In this section, we evaluate six finite difference schemes resulting from three different expressions for the flux F_m^{n+1}. Because numerical methods are normally used when no analytic solutions can be found, it is important that the numerical method be robust. A method which works well for one equation can fail for another. Therefore, we test each method under three equations (given below) whose exact analytic solutions are known.

4.1. *Test Equations*

The following three equations were solved numerically over an interval from $x_0 = 10^{-3}$ to $x_M = 10^3$ using a logarithmically spaced mesh with $M = 100$:

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(x^2\frac{\partial u}{\partial x} - xu - u\right) - u + \delta(x - x_{\rm inj})\Theta(t) \quad (22)$$

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(x^2\frac{\partial u}{\partial x} - xu\right) - u/x + \delta(x - x_{\rm inj})\Theta(t), \quad (23)$$

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(x^3\frac{\partial u}{\partial x} - x^2u\right) - u + \delta(x - x_{\rm inj})\delta(t). \quad (24)$$

In all three cases, the injection function is a monoenergetic distribution at $x_{\rm inj} = 0.1$, which we simulate as a narrow Gaussian distribution whose width is smaller than Δx at $x_{\rm inj}$. For the first two equations, monoenergetic particles are injected at $x_{\rm inj}$ beginning at $t = 0$ at the rate of one particle per unit time. In the third equation, the impulsive source term at $t = 0$ is implemented as an initial value for $u(x, t = 0)$. There is no further injection of particles, so we set $Q(x) = 0$.

The first equation is the familiar hard-sphere equation (see, e.g., Ramaty 1979) with the addition of a term corresponding to relativistic Coulomb losses (Paper 1, eq. [62]; Hamilton & Petrosian 1992; Steinacker et al. 1988). We can define the timescales for diffusion, advection, and escape (see Paper I, eqs. [46], [47], [48], and [75]) as $\tau_C = x^2/C(x)$, $\tau_B = x/|B(x)|$, and $\tau_T = T(x)$, respectively. For this equation, $\tau_C = \tau_T = 1$, but $\tau_B = x/(1 + x)$, which can vary by 3 orders of magnitude from x_0 to x_M. This will test the ability of the numerical method to resolve a widely varying advective timescale. The exact analytic solution of equation (22) at *steady state* ($\partial u/\partial t = 0$) is given by Paper I (eq. [71]; see also Steinacker et al. 1988). The second equation also looks like the familiar hard-sphere equation but the escape time, $\tau_T = x$, is energy-dependent and varies by 6 orders of magnitude over the energy interval. The analytic *steady state* solution of this equation is given by Paper I (eq. [57]; see also Dröge & Schlickeiser 1986). The third equation tests the *time-dependent* properties of the numerical methods by comparing them to the analytic solution given by Paper I (eq. [59]). Also, both the diffusive and advective timescales ($\tau_C = \tau_B = 1/x$) vary by 6 orders of magnitude over the energy interval so this forms another test of the numerical method.

4.2. *Simple Fully Implicit Method*

The simplest finite difference method (Press et al. 1992, chap. 19) uses the midpoint difference for both the advection and the diffusion terms to write the flux (4) as

$$F^{n+1}_{m+1/2} = B_{m+1/2}u^{n+1}_{m+1/2} + C_{m+1/2}\frac{u^{n+1}_{m+1} - u^{n+1}_m}{\Delta x_{m+1/2}}, \quad (25)$$

$$= \frac{C_{m+1/2}}{\Delta x_{m+1/2}}[(1 + w_{m+1/2}/2)u^{n+1}_{m+1} - (1 - w_{m+1/2}/2)u^{n+1}_m], \quad (26)$$

where

$$w_{m+1/2} = \frac{B_{m+1/2}}{C_{m+1/2}}\Delta x_{m+1/2}. \quad (27)$$

The second equality (26) is useful for comparing this method to other methods described in §§ 4.3 and 4.4.

Substituting equation (25) into equation (15) produces the tridiagonal system of equations:

$$\begin{cases} a_m = \dfrac{\Delta t}{A_m\Delta x_m}\dfrac{C_{m-1/2}}{\Delta x_{m-1/2}}(1 - w_{m-1/2}/2), \\ c_m = \dfrac{\Delta t}{A_m\Delta x_m}\dfrac{C_{m+1/2}}{\Delta x_{m+1/2}}(1 + w_{m+1/2}/2), \\ b_m = 1 + \dfrac{\Delta t}{A_m\Delta x_m}\left[\dfrac{C_{m-1/2}}{\Delta x_{m-1/2}}(1 + w_{m-1/2}/2) + \dfrac{C_{m+1/2}}{\Delta x_{m+1/2}}(1 - w_{m+1/2}/2)\right] + \Delta t/T_m, \\ r_m = \Delta t Q_m + u^n_m, \end{cases} \quad (28)$$

which is valid for $m = 1, \ldots, M - 1$. The no-flux boundary condition (3) should be added to obtain the coefficients for $m = 0$ and $m = M$. If the mesh were uniform (Δx_m = constant), then this method would be first-order accurate in time and second-order accurate in space. For a nonuniform mesh, it degrades to first-order accurate in space.

Figure 1 shows that this method produces unstable negative solutions when solving the steady state solution of equation (22). This verifies the analysis of CC70. Figure 2 shows very

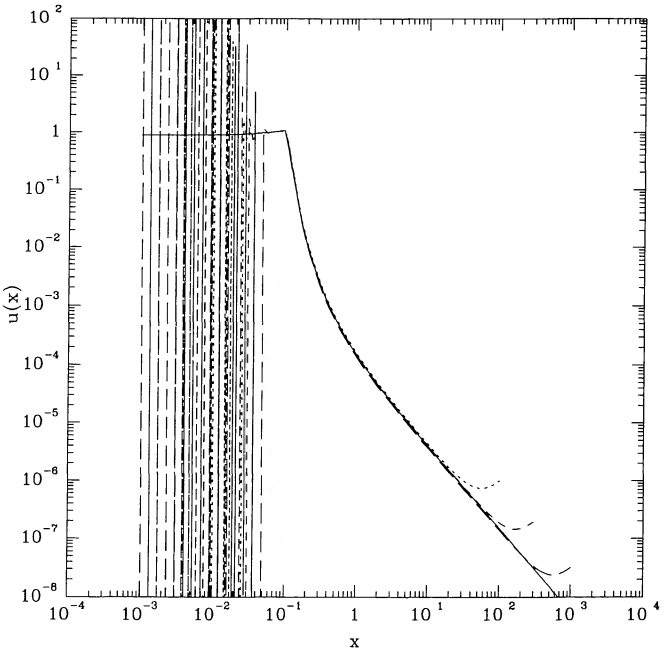

FIG. 1.—Numerical solutions (*dashed lines*) of eq. (22) using the simple fully implicit method (28) compared to the exact analytic solution (*solid line*) from Paper I. Three different numerical boundaries are shown to illustrate the effect of varying the locations of the boundary points: (*short-dashed lines*) $x_0 = 10^{-2}$, $x_M = 10^2$; (*medium-dashed lines*) $x_0 = 10^{-2.5}$, $x_M = 10^{2.5}$; (*long-dashed lines*) $x_0 = 10^{-3}$, $x_M = 10^3$. The steady state numerical solutions were obtained at $t = 10$ (normalized units). The numerical solution produces unstable oscillatory solutions at low energy for all three boundaries.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

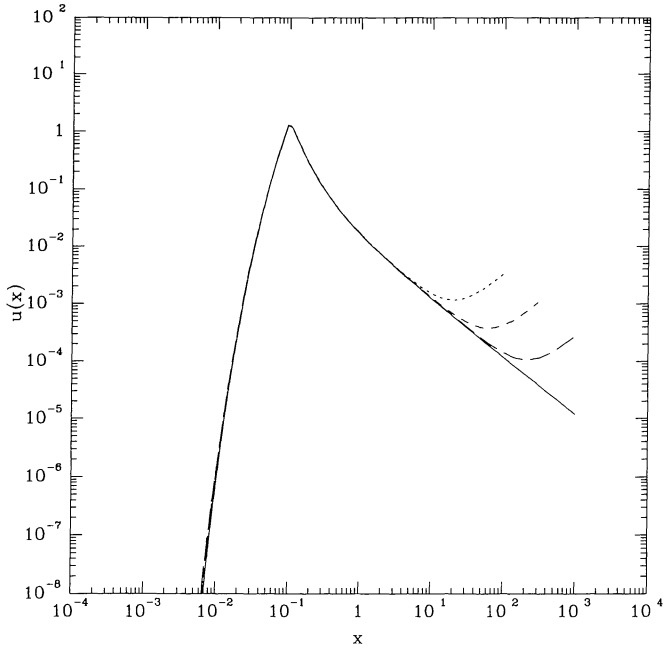

FIG. 2.—Same as Fig. 1 but for eq. (23). Note the good fit to the analytic solution over many orders of magnitude except near the right boundary. The cause of this boundary effect is discussed § 4.7.

good correspondence between the numerical and analytic solutions of equation (23), except near the right boundary. In Figure 3, the time-dependent solutions of equation (24) are compared for three different times. The discrepancy at $t = 30$ is noticeable, but at this point, the particle number in the system has become negligible (note that this solution has been scaled by a factor of 10^{11}). The numerical method corresponds with the exact solutions very well near $x_{\rm inj}$, but rapidly becomes less accurate closer to either boundaries. For $x \lesssim x_{\rm inj}$, the errors are caused by large gradients in the particle distribution which corresponds to a diffusion timescale $\ll \Delta t$. For $x \gtrsim x_{\rm inj}$, the errors are caused by the finite boundary. Although this method appears to be successful for a large number of Fokker-Planck equations, it produces oscillatory negative solutions for some equations, so it is not recommended.

4.3. *Chang-Cooper Method*

A second expression for F_m^{n+1} comes from CC70 (eqs. [16] and [18]), which uses the centered difference on the diffusive term, but a weighted difference on the advective term. The flux is written as

$$F_{m+1/2}^{n+1} = (1 - \delta_{m+1/2}) B_{m+1/2} u_{m+1}^{n+1} + \delta_{m+1/2} B_{m+1/2} u_m^{n+1} + C_{m+1/2} \frac{u_{m+1}^{n+1} - u_m^{n+1}}{\Delta x_{m+1/2}}, \quad (29)$$

$$= \frac{C_{m+1/2}}{\Delta x_{m+1/2}} [W_{m+1/2}^{+} u_{m+1}^{n+1} - W_{m+1/2}^{-} u_m^{n+1}], \quad (30)$$

where δ_m and $W_m^{\pm}$ are defined by

$$\delta_m = \frac{1}{w_m} - \frac{1}{\exp(w_m) - 1}, \quad (31)$$

$$W_m^{\pm} = W_m \exp\left(\pm \frac{w_m}{2}\right), \quad (32)$$

$$W_m = \frac{w_m}{2} \Big/ \sinh \frac{w_m}{2}, \quad (33)$$

and w_m is given by equation (27). Our definitions of w_m and W_m are slightly different from CC70 so that W_m is a symmetric function of w_m and $W_m^{+} - W_m^{-} = w_m$. For $w_{m+1/2} \ll 1$, equation (30) reduces to the simple fully implicit method given by equation (26).

Substituting equation (30) into equation (15), we obtain a tridiagonal system of equations whose coefficients are

$$\begin{cases} a_m = \dfrac{\Delta t}{A_m \Delta x_m} \dfrac{C_{m-1/2}}{\Delta x_{m-1/2}} W_{m-1/2}^{-}, \\ c_m = \dfrac{\Delta t}{A_m \Delta x_m} \dfrac{C_{m+1/2}}{\Delta x_{m+1/2}} W_{m+1/2}^{+}, \\ b_m = 1 + \dfrac{\Delta t}{A_m \Delta x_m} \left[\dfrac{C_{m-1/2}}{\Delta x_{m-1/2}} W_{m-1/2}^{+} + \dfrac{C_{m+1/2}}{\Delta x_{m+1/2}} W_{m+1/2}^{-} \right] + \Delta t / T_m, \\ r_m = \Delta t Q_m + u_m^n. \end{cases} \quad (34)$$

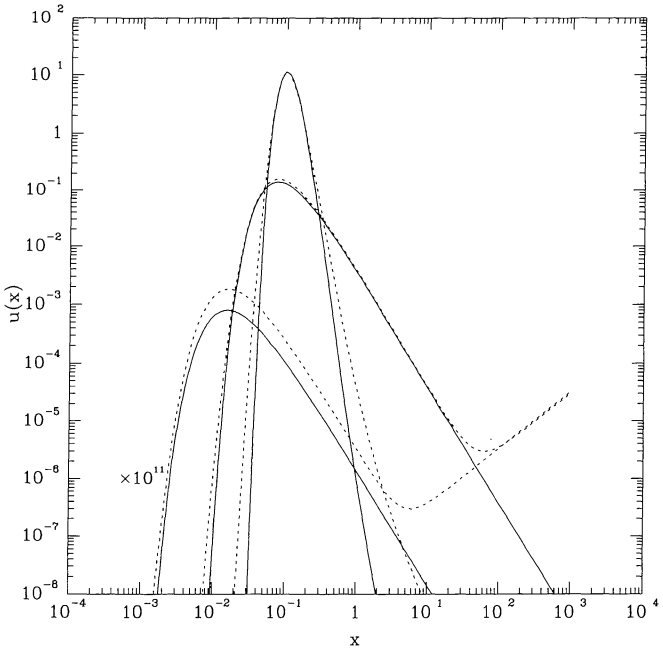

FIG. 3.—Time-dependent solutions (*dashed lines*) of eq. (24) using the simple fully implicit method compared to the analytic solutions (*solid lines*), for $t = 0.3$, 3, and 30 (normalized units) in decreasing height. Solution at $t = 30$ is scaled up by a factor of 10^{11}. Numerical boundary is given by $x_0 = 10^{-3}$ and $x_M = 10^3$. Large deviations occur only when the number of particles has dropped by several decades.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

As explained in CC70, the parameter $\delta_{m+1/2}$ is determined so that the stationary solution to the numerical scheme is identical to the analytic stationary solution without source or escape. CC70 required $B(x)$ to be strictly positive. A more careful examination shows that any value of $B(x)$ is allowed, so $w_{m+1/2}$ can take on values from $-\infty$ to $+\infty$ and the parameter $\delta_{m+1/2}$ then varies from 1 to 0, respectively. This scheme adjusts the weight $\delta_{m+1/2}$ so that the differencing on the advective term is always "upwind." In other words, if $B(x) \gg 0$, forcing particles to flow to low energy, the difference scheme puts more weight on the higher energy mesh points when determining the flux. Conversely, if $B(x) \ll 0$ so that particles flow to high energy, then more weight is given to the lower energy mesh points. This property makes the Chang-Cooper method first-order accurate in both space and time, even over a uniform mesh.

Although expression (33) is useful for analytical manipulations, it may cause numerical overflow or underflow problems in practice because $|w_m|$ may become extremely small or large. For computational purposes, we use

$$W_m = \begin{cases} \left[1 + \dfrac{w_m^2}{24} + \dfrac{w_m^4}{1920}\right]^{-1} & (|w_m| < 0.1), \\ \dfrac{|w_m| \exp(-|w_m|/2)}{1 - \exp(-|w_m|)} & (|w_m| \geq 0.1), \end{cases} \quad (35)$$

which is valid when using typical "single precision" floating point numbers. The corresponding expression for "double precision" numbers can be easily derived. Notice from equation (32) that both W_m^+ and W_m^- increase in magnitude only as a linear power of $|w_m|$ for large $|w_m|$. This is important because it prevents the coefficients of the tridiagonal matrix from overflowing.

The original method considered by CC70 did not contain the escape or source terms. For that case, they showed that the method gives positive solutions for some values of Δt and Δx. LLPS85 (eq. [34]) showed that it guarantees positivity for any values of Δt and Δx. We can easily show that equation (34), which contains the additional escape and source terms, also gives positive solutions. If we make the substitution $\bar{u}_m^n = u_m^n / u_m^\infty$ and $u_m^\infty = \exp(-\sum_{j=0}^{m-1} w_{j+1/2})$ into equations (15) and (29), it can be verified that the tridiagonal matrix for $\bar{u}_m^{n+1}$ satisfies the positivity condition (19) for $\bar{u}_m^{n+1}$. We conclude that u_m^{n+1} must also be positive.

Figure 4 uses the Chang-Cooper method to solve the steady state solution of equation (22). The oscillatory solutions produced by the previous method is not present, and the numerical solution shows very good agreement with the analytic solution. The boundary effect on the left occurs over one mesh interval and does not go away as the boundaries are extended further apart. There is some evidence that this method is not as accurate as the fully implicit method shown in Figure 1 for $x \gtrsim 1$. However, the difference is slight, and the elimination of instability at low energy more than compensates for this loss of accuracy. When this method is used to solve equations (23) and (24), the results look identical to Figures 2 and 3.

4.4. *Larsen-Levermore-Pomraning-Sanderson Method*

LLPS85 studied finite difference schemes for both linear and nonlinear Fokker-Planck equations. Two expressions are relevant for the linear problem. One of them (LLPS85, eq. [23]) becomes identical to the Chang-Cooper method for linear Fokker-Planck equations so we do not have to consider it any further. The other (LLPS85, eq. [9]) is the third method considered in this paper, and it writes the flux F_m^{n+1} as

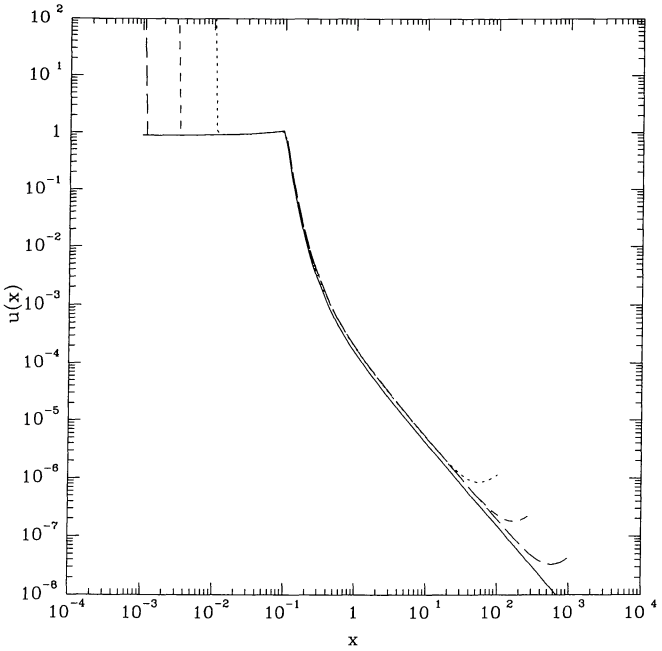

FIG. 4.—Same as Fig. 1 but using the Chang-Cooper method (34). The fit is very good even at low energies where no stability problems are apparent. This method is the most robust among the ones tested in the paper. The sharp boundary transient on the left-hand side occurs over a single mesh point. It is not unique to this method as discussed further in § 4.7.

$$F_{m+1/2}^{n+1} = \frac{C_{m+1/2}}{\Delta x_{m+1/2}} [e^{w_{m+1/2}/2} u_{m+1}^{n+1} - e^{-w_{m+1/2}/2} u_m^{n+1}], \quad (36)$$

where w_m is given by equation (27). Like the Chang-Cooper method, this expression reduces to expression (26) in the limit of $w_m \ll 1$,

Substituting the flux (36) into equation (15), the resulting tridiagonal system can be written exactly as the Chang-Cooper method (34) with the replacement of

$$W_m^\pm = \exp(\pm w_m/2) \quad (37)$$

for equation (32). This method can be shown to guarantee positive solutions like the Chang-Cooper method. Upon closer examination, we discover that problems can arise when applied to equations considered in this paper. Notice that W^+ and W^- can now increase *exponentially* for large $|w_m|$. For some equations, $|w_m|$ becomes so large that the coefficients overflow the floating point number system. Even if an overflow error does not occur, this scheme becomes inaccurate as $|w_m| \gtrsim 1$ because equation (36) no longer approximates the flux accurately.

In Figure 5, we solve equation (22) using the LLPS method. The numerical solution deviates from the true analytic solution at low energies. The deviation starts where $|w_m| \gtrsim 1$, at which equation (36) no longer approximates the flux correctly

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

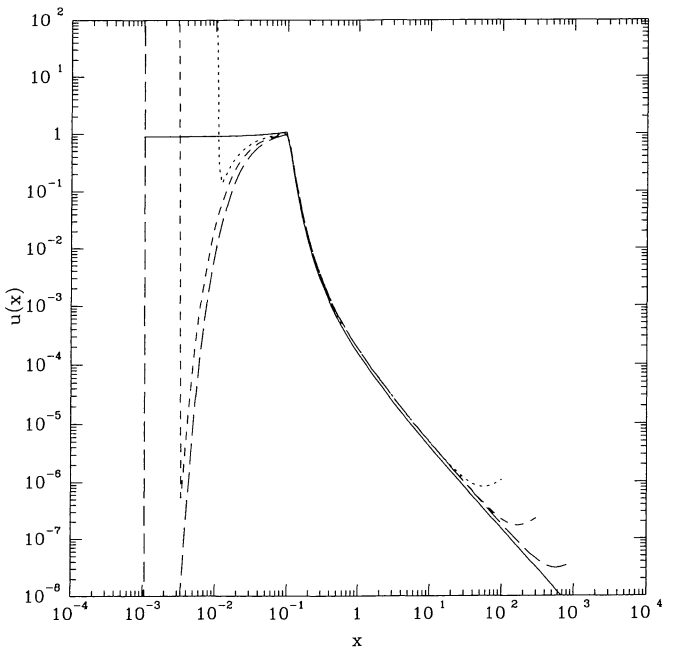

FIG. 5.—Same as Fig. 1 but using the LLPS method, which can be expressed identically as eq. (34) with the substitution of eq. (37). Significant deviation occurs at low energy, making this method unsuitable for many equations.

because of its exponential dependence on $|w_m|$. When this method is used to solve equations (23) and (24), the results are identical to Figures 2 and 3, respectively. For these equations, $|w_m| \lesssim 1$ over the energy interval shown, so this method is able to solve them accurately.

4.5. *Semi-Implicit Methods*

For each of the three methods presented in §§ 4.2, 4.3, and 4.4, the semi-implicit counterpart can be obtained by replacing $F_{m+1/2}^{n+1}$ with $F_{m+1/2}^{n+1/2}$, as explained in § 3.4. For definiteness, we call the semi-implicit version of the simple fully implicit method the "Crank-Nicholson" method (see, Press et al. 1992, chap. 19); the semi-implicit version of the Chang-Cooper method is called the "semi-implicit Chang-Cooper" method; and the semi-implicit version of the LLPS method is called the "semi-implicit LLPS" method. Inspecting the resulting expressions, we notice that the tridiagonal system of equations for all the semi-implicit methods can be written as

$$\begin{cases} a'_m = a_m/2\,, \\ c'_m = c_m/2\,, \\ b'_m = (b_m - 1)/2 + 1\,, \\ r'_m = r_m + u_m^n - a'_m u_{m-1}^n - b'_m u_m^n - c'_m u_{m+1}^n\,, \end{cases} \tag{38}$$

where a_m, b_m, c_m, and r_m are the coefficients of the fully implicit methods as defined by equations (28) and (34). As noted in § 4.4, the tridiagonal system of equations for the LLPS method is identical to the Chang-Cooper method, except for the substitution of equation (37). These semi-implicit methods do not satisfy condition (19) so they cannot be shown to guarantee positive solutions.

For each of the three semi-implicit methods, we can perform the same tests as we did with the three fully implicit methods. For the test equations (22) and (23), there was no discernible difference between the fully implicit and the semi-implicit routines. The solutions to equation (22) looked identical to Figures 1, 4, and 5. That is to say, that the Crank-Nicholson method showed oscillatory negative solutions like the simple fully implicit method; the semi-implicit LLPS method showed exponential divergence from the true solution just like the fully implicit LLPS method. The semi-implicit Chang-Cooper method accurately solved both equations just like its fully implicit counterpart. The solution for equation (23) for all three semi-implicit methods looked identical to Figure 2.

When these methods were tested with equation (24), differences appeared between the semi-implicit methods and the fully implicit methods. As Figure 6 shows, all the semi-implicit methods exhibit unstable oscillatory solutions at $t = 30$. They result from either an inherent property of the semi-implicit methods (recall that the positivity condition [19] is sufficient but not necessary), or by numerical round-off errors from the extra matrix multiplication in the calculation of r'_m in equation (38).

There is evidence in Figure 6 that the semi-implicit methods produce more accurate time-dependent solutions for $x \lesssim 1$. This is consistent with their second-order accuracy in time which should give better results at a given mesh size. However, they appear to be more unstable for $x \gtrsim 1$ than the fully implicit methods, where Δx_m becomes increasingly larger. If high time-dependent accuracy is desired, then the semi-implicit Chang-Cooper method may be used with care, but the results should

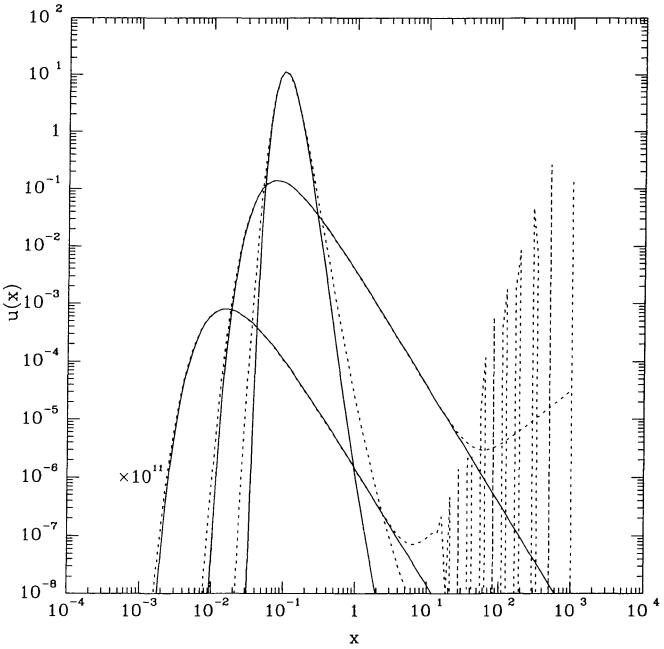

FIG. 6.—Same as Fig. 3 but using semi-implicit schemes. All three semi-implicit versions of the simple fully implicit, Chang-Cooper, and LLPS methods, produce solutions essentially identical to this figure. These methods are more accurate than their corresponding fully implicit methods at the same mesh sizes. However, they can produce unstable oscillatory solutions, as shown for $t = 30$, so they should be used with caution.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

always be checked against the more robust fully implicit Chang-Cooper method.

4.6. *Operator Splitting*

In § 3.5, we stated that the operator splitting method can produce different results, depending on how the operator is split into its smaller parts. For example, consider equation (23) which may be split into two suboperators, $\mathscr{L}_1 u = (x^2 u' - xu)'$ and $\mathscr{L}_2 u = -u/x + \delta(x - x_{\rm inj})\Theta(t)$. One solution can be written as $u^{n+1} = L_1 L_2 u^n$ and the other can be written as $u^{n+1} = L_2 L_1 u^n$, where L_1 and L_2 are the respective finite difference methods for each operator. For illustrative purposes, we used the fully implicit Chang-Cooper method (without the source and escape terms) for the operator L_1. The operator $\mathscr{L}_2$ has an exact analytic solution which was converted to a finite difference form for L_2.

Figure 7 shows the numerical solutions of equation (23) using the $L_1 L_2$ combination (*long dashed line*) and the $L_2 L_1$ combination (*short dashed line*) compared with the exact analytic solution (*solid line*). The boundary effects exhibited by both numerical solutions near the right boundary are explained in § 4.7, and we ignore them for now. Near the left boundary, we notice that the two numerical solutions differ noticeably. The short-dashed line approximates the true solution acceptably, while the long-dashed line does not. However, the short dashed line has difficulty around the injection points $x_{\rm inj}$, while the other handles this region better. Without possessing the exact analytic solution, it would be difficult to decide which was the correct solution.

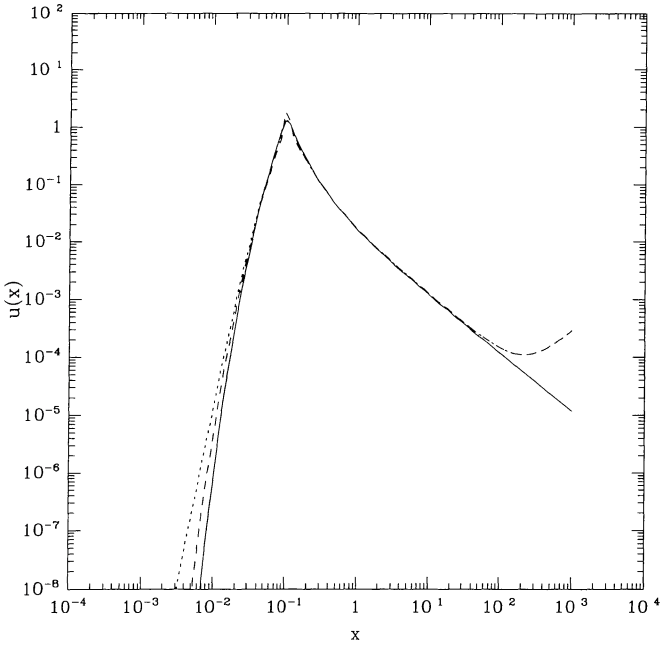

FIG. 7.—Two numerical solutions of eq. (23) using two different operator splitting methods (*long-dashed and short-dashed lines*) and the exact analytic solution (*solid line*). Depending on the ordering of the difference operators, two different results can be obtained at low energies for this equation.

The difference between the two methods can be understood in terms of intrinsic timescales of the individual difference operators L_1 and L_2. As discussed in § 4.1, the timescale for L_1 is equal to the diffusion and advection timescales $\tau_C = \tau_B \simeq 1$, but the timescale for L_2 is the escape timescale $\tau_T = x$, which varies between 10^{-3} to 10^3. The time step used in the numerical integration was $\Delta t \simeq 0.05 \gg \tau_T$ at low energies. As a result, the numerical method cannot resolve the rapid variations in the solution from successive applications of L_1 and L_2.

4.7. *Boundary Effects*

In general, boundary effects are caused by solving the Fokker-Planck equation over the finite interval $x_0 < x < x_M$, instead of the infinite interval $0 < x < \infty$. This implies that boundary effects are generic features of any numerical methods. Normally, we would expect the boundary effects to diminish as the numerical boundaries are moved further apart. Some Fokker-Planck equations have this behavior because their true steady state solutions are consistent with the no-flux boundary condition at $x = 0$ and $x = \infty$.

For other equations, the boundary effect does not diminish as the numerical interval is increased. In all numerical solutions equation (22) regardless of the method, for example Figure 4, a sharp pileup of particles appears at the left boundary. This transient effect occurs over a single mesh point. We also notice that the pileup occurs for all three energy intervals and does not diminish as we make the interval larger. To understand this effect, we examine the flux of particles of this equation at steady state, which is $F \propto$ constant as $x \to 0$ (Paper I, Table 4). No matter how far away we push the left boundary, the numerical no-flux boundary condition can never be satisfied and an unavoidable pileup of particles results. The mathematical reason for this is related to the singular nature of this Fokker-Planck equation and was discussed in Paper I. The physical reason comes from extending the relativistic Coulomb loss term of equation (22) ($\dot{E} \propto u$) into the low-energy regime where the relativistic approximation breaks down.

At the right boundaries of Figures 1, 4, and 5, the boundary effects are much less severe because the analytic flux of particles for equation (22) is $F \to 0$ as $x \to \infty$ (Paper I, Table 4). The right boundary of Figure 2, however, shows significant boundary effects. Here, the flux of particles for equation (23) is given by $F \propto$ constant as $x \to \infty$ (Paper I, Table 2), which causes the pileup of particles. This is an inherent property of a singular Fokker-Planck equation, and no numerical technique will be able to circumvent its boundary effects completely.

5. STOCHASTIC SIMULATION METHOD

A method which arrives at a solution through statistical averages is typically called a "Monte Carlo" method. In the context of stochastic acceleration of particles, a Monte Carlo method tries to simulate the microscopic scatterings of particles. Unfortunately, the microscopic timescales of wave-particle interactions are considerable smaller than the macroscopic timescales of interest, making Monte Carlo simulations computationally very expensive. One class of Monte Carlo methods which suffers less from this problem is the technique of stochastic simulation. It relies on the exact equivalence of the Fokker-Planck equation to the Itô stochastic differential equa-

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

tion (van Kampen 1992, chap. 9) so that, in the absence of source or escape terms, equation (1) can be written as (Krülls & Achterberg 1994; Achterberg & Krülls 1992; MacKinnon & Craig 1991)

$$dx = [C'(x) - B(x)]dt + [2C(x)]^{1/2} r(t) dt . \quad (39)$$

Here, the first term on the right-hand side comes from the advective and diffusive drift terms, the prime denotes differentiation with respect to x, and the second term models the fluctuations responsible for the diffusion term, where $r(t)$ is a Gaussian random noise satisfying $\langle r(t) \rangle = 0$ and $\langle r(t) r(s) \rangle = \delta(t - s)$. The ensemble-averaged phase space density of many particles, each evolving according to equation (39), is equivalent to the solution of equation (1).

Equation (39) immediately leads to a numerical simulation by evolving test particles at discrete time steps according to

$$\Delta x = [C'(x) - B(x)]\Delta t + [2C(x)]^{1/2} \Delta r , \quad (40)$$

where $\Delta r = \int_0^{\Delta t} r(t) dt$ is a Gaussian random variable. The properties of $r(t)$ imply that $\langle \Delta r \rangle = 0$ and $\langle \Delta r_i \Delta r_j \rangle = \Delta t \delta_{ij}$ for two discrete time intervals i and j, where δ_{ij} is the Kronecker delta function. It has been implicitly assumed that the time step satisfies $dt \ll \Delta t \lesssim \tau_{\text{natural}}$, where dt is the microscopic scattering timescale and τ_{natural} is a characteristic timescale of the equation.

To deal with the escape term in equation (1), we calculate the probability of escape,

$$P_{\text{escape}}(\Delta t) = 1 - e^{-\Delta t / T(x)} , \quad (41)$$

at each time step and remove each escaping particle as determined by a uniform random number generator. This is more efficient than the method suggested by Achterberg & Krülls (1992), which assigns a commulative escape probability weight to each particle. Equation (41) works for an energy-dependent escape time as well.

In the simulation of an initial value problem where no new particles are injected after the initial injection, the number of particles in the system at any future time cannot be greater than the initial number. The memory requirements and the execution time of the computer simulation is fixed and known beforehand. When calculating the steady state distribution, new particles are injected into the system continually, and the steady state is reached when the rate of particle injection is equal to the rate of particle escape. When the escape time is constant, the final particle number can be calculated to be $\mathcal{N} = T \int dx Q(x)$. When the escape time is energy-dependent, the final particle number cannot be calculated a priori, so we cannot constrain the memory requirements and the excution time of the simulation.

A modification which avoids this problem is the following. Instead of simply removing the escaping particles from the system, we reinject them back at x_{inj}, thereby forcing a balance in the rate of escape and the rate of injection. We note that an injection function other than the monoenergetic distribution can be implemented by reinjecting the particle with a probability proportional to $Q(x)$. If we run this simulation to $t_{\text{steady state}} \lesssim \tau_{\text{natural}}$, we will obtain the steady state distribution up to some undetermined normalization constant. Unfortunately, solutions at intermediate times $t < t_{\text{steady state}}$ are unphysical because we have artificially enforced a balanced between the rate of particle injection and the rate of escape. If time-dependent solutions are required, then we must keep the two rates independent of each other, as done by Achterberg and Krülls (1992). Despite these problems, this particular modification is useful for obtaining the steady state solutions because it guarantees a fixed upper limit to the total number of particles at steady state.

Although the stochastic simulation method is less sensitive to the effects of boundary conditions than the finite difference method, the implementation of the boundary conditions requires careful consideration. In particular, particles can overshoot and cross the $x = 0$ boundary at finite t because the time step size Δt is discrete and the fluctuation term can be randomly large. Additionally, particles may reach $x = \infty$ in a finite time for certain equations. In equation (23), ignoring the diffusion term (which makes some particles reach $x = \infty$ even faster), the equation governing these particles is $dx \propto x^2 dt$. A particle injected at $x_{\text{inj}} = 0.1$ reaches $x = \infty$ after $t \sim 1/x_{\text{inj}} = 10 \simeq t_{\text{steady state}}$. If particles are removed from the system when they cross the boundary, then we obtain the no-particle condition (2). If the particles are simply pegged to the boundary, then this corresponds to the no-flux condition (3). For some equations, losing too many particles through the no-particle condition increases the Poisson noise to an unacceptable level. To avoid this possibility and to facilitate comparisons with the finite difference methods discussed in the previous sections, we use the no-flux condition, with the lower and upper boundaries set to 10^{-3} and 10^3. In practice, these boundaries should be moved much further apart to reduce the boundary effects.

Figures 8–10 show the solutions of equations (22) to (24), respectively, using the stochastic simulation method. The steady state solutions in Figures 8 and 9 were vertically rescaled by eye to match the analytic solutions; Figure 10 required no normalization because it solved the time-dependent problem. All three plots show good agreement with the analytic solution where there are sufficient number of particles in the bins. As the number of particles in the bin decreases, the Poisson noise increases. For Figure 10, the solution at $t = 30$ was not plotted because no particles were left in the system. Just like the finite difference methods, boundary effects are apparent.

In terms of computational efficiency, the stochastic simulation method, for the parameters used here, takes about 50 to 100 times longer than the corresponding finite difference method. The stochastic simulation method may offer advantages for multidimensional equations, but these advantages may be canceled by the larger number of test particles required for the larger phase space volume. A more detailed study is required to address this issue. For single-dimensional Fokker-Planck equations, however, the finite difference method offers better accuracy at lower computational cost.

6. CONCLUSIONS

The best finite difference method for solving the Fokker-Planck equation obtained in the study of stochastic acceleration is essentially the Chang-Cooper method given by CC70.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

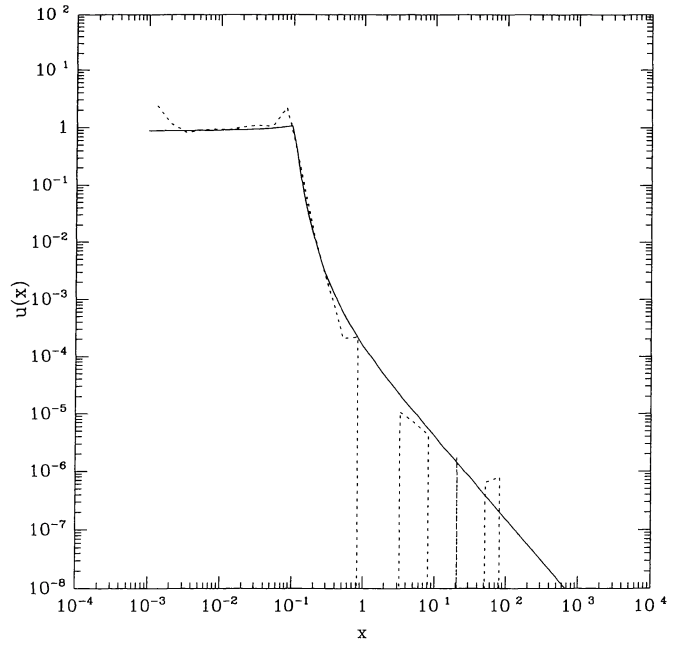

FIG. 8.—Same as Fig. 1 but dashed line solved using the stochastic simulation method (40). The simulation used 10^4 particles in 30 logarithmicly spaced bins, a time step of $\Delta t = 0.05$, and boundaries at 10^{-3} and 10^3. Although the solution matches the analytic solution well when sufficient particles fall within the bins, the Poisson noise becomes dominant when the particle number density decreases by just a few decades.

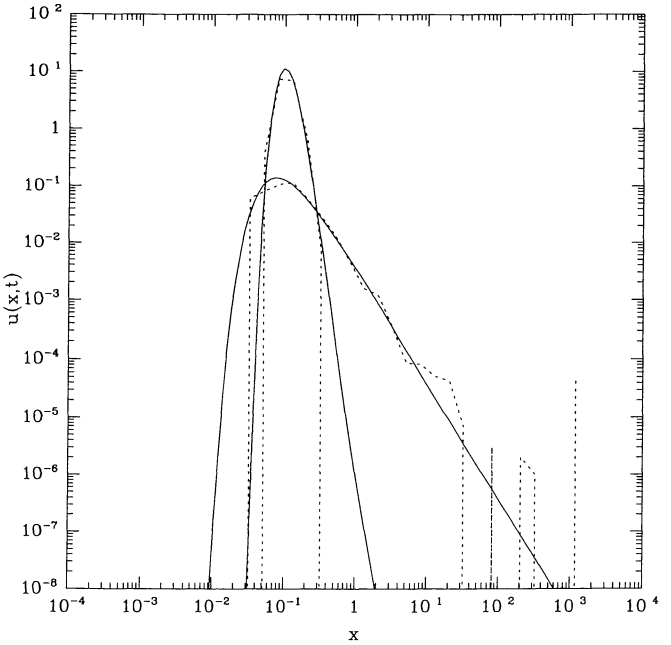

FIG. 10.—Same as Fig. 3 but dashed lines solved using the stochastic simulation method. The solutions shown for $t = 0.3$ and 3, in decreasing height; the solution at $t = 30$ was not plotted because no particles were remaining in the system. Other parameters are same as Fig. 8. Numerical results do not agree well with analytic results. This method suffers from Poisson noise when the number of particles in a bin becomes small. Because a fixed number of test particles must be used, the dynamic range of Monte Carlo methods, in general, is small.

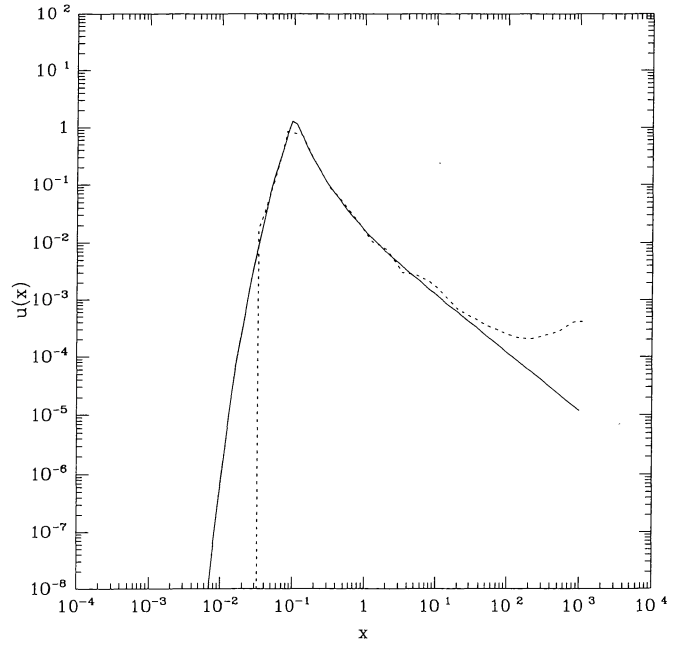

FIG. 9.—Same as Fig. 2 but dashed line solved using the stochastic simulation method. Parameters are same as Fig. 8. The numerical results do not match the analytic solution very well.

Our implementation (34) makes a small extension to include the escape and source terms while maintaining the guarantee of positive solutions.

The semi-implicit Chang-Cooper method (§ 4.5) seems to be the most robust among semi-implicit methods studied in this paper. It sometimes suffers from unstable oscillatory negative solutions. It is accurate to second-order in time, producing slightly more accurate solutions for time-dependent problems. When high accuracy is a requirement, then this method should be considered.

The most useful numerical boundary condition is the no-flux condition given by equation (3). The combination of this condition and the Chang-Cooper method guarantees positive solutions. The boundary effect produced by the application of the no-flux condition on Fokker-Planck equations which do not admit a no-flux solution is caused by the singular properties of the Fokker-Planck equation. Solutions in the region of interest will be accurate if the numerical boundaries are sufficiently far away. Boundary effects are inherent features of any numerical solutions of singular Fokker-Planck equations.

The problem of numerical overflow and underflow when solving equations over many orders of magnitudes is adequately handled by using the Chang-Cooper method. The tridiagonal matrix elements produced by this method are not as susceptible to numerical underflow or overflow as the elements of the LLPS method.

© American Astronomical Society • Provided by the NASA Astrophysics Data System

1996ApJS..103..255P

The operator splitting method, which is extremely useful for multidimensional problems, should not be used for one-dimensional Fokker-Planck equations. Different results can be obtained for different operator orderings. In any case, there is no compelling reason to use the operator splitting method because the one-dimensional differential operator can easily be evaluated using the Chang-Cooper method in its entirety.

The stochastic simulation method is not recommended for one-dimensional Fokker-Planck equations. It is computationally expensive and susceptible to Poisson noise. The same equation can be solved faster and more accurately using finite difference methods. For multidimensional problems, however, it may offer some advantages, but further study is required.

From the solution of the Fokker-Planck equation, we can calculate the spectra of photons produced by these accelerated particles. Comparing them with observational data from solar flares (Park, Petrosian, & Schwartz 1996), for example, gives us constraints on the Fokker-Planck coefficients, which in turn reveals the nature of the stochastic acceleration process.

B.T.P. gratefully acknowledges the fellowship support from the Natural Sciences and Engineering Research Council (NSERC) of Canada. We thank useful discussions with Russ Hamilton and Marcos Montes. This research has been funded by NSF ATM 90-11528, NASA NAGW 1976, and NASA NAGW 2290.

REFERENCES

Achterberg, A., & Krülls, W. M. 1992, A&A, 265, L13
Ames, W. F. 1977, Numerical Methods for Partial Differential Equations, (2d ed.; New York: Academic Press)
Bai, T. 1982, ApJ, 259, 341
Chang, J. S., & Cooper, G. 1970, J. Comp. Phys., 6, 1 (CC70)
Davis, L. 1956, Phys. Rev., 101, 351
Dröge, W., & Schlickeiser, R. 1986, ApJ, 305, 909
Fermi, E. 1949, Phys. Rev., 75, 1169
———. 1954, ApJ, 119, 1
Hamilton, R. J. 1990, Ph.D. thesis, Stanford Univ.
Hamilton, R. J., Lu, E. T., & Petrosian, V. 1990, ApJ, 354, 726
Hamilton, R. J., & Petrosian, V. 1992, ApJ, 398, 350
Krülls, W. M., & Achterberg, A. 1994, A&A, 286, 314
Larsen, E. W., Levermore, C. D., Pomraning, G. C., & Sanderson, J. G. 1985, J. Comp. Phys., 61, 359 (LLPS85)
MacKinnon, A. L., & Craig, I. J. D. 1991, A&A, 251, 693
Miller, J. A., Guessoum, N., & Ramaty, R. 1990, ApJ, 361, 701
Miller, J. A., & Ramaty, R. 1989, ApJ, 344, 973
Park, B. T., & Petrosian, V. 1995, ApJ, 446, 699 (Paper I)
Park, B. T., Petrosian, V., & Schwartz, R. A. 1996, in preparation
Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1992, Numerical Recipes in C: The Art of Scientific Computing, 2nd ed. (New York: Cambridge Univ. Press)
Ramaty, R. 1979, in AIP Conf. Proc. 56, Particle Acceleration Mechanisms in Astrophysics, ed. J. Arons, C. Max, & C. McKee (New York: AIP), 135
Richtmyer, R. D., & Morton, K. W. 1967, Difference Methods for Initial-Value Problems, 2d ed. (New York: Wiley)
Schlickeiser, R. 1989, ApJ, 336, 243
Steinacker, J., Dröge, W., & Schlickeiser, R. 1988, Solar Phys., 115, 313
Strikwerda, J. C. 1989, Finite Difference Schemes and Partial Differential Equations (Belmont: Wadsworth)
van Kampen, N. G. 1992, Stochastic Processes in Physics and Chemistry, 2d ed. (Amsterdam: North-Holland)

© American Astronomical Society • Provided by the NASA Astrophysics Data System