JOURNAL OF THE KOREAN ASTRONOMICAL SOCIETY
34: S231 ~ S235, 2001

Numerical Simulation of Cosmic-Ray Acceleration

T. W. JONES (1)
(1) Department of Astronomy, University of Minnesota, Minneapolis, MN 55455, USA
(Received Nov. 15, 2001; Accepted Nov. 15, 2001)

ABSTRACT

Cosmic-ray acceleration, although physically important in many astrophysical contexts, is difficult to incorporate into numerical models, because it involves microphysics that is generally far from thermodynamic equilibrium, and also because the length and time scales for that physics typically range over many orders of magnitude, reflecting the huge range of particle rigidities that must be represented. The most common accelerator models are stochastic in nature and involve nonequilibrium plasma properties that are also often poorly understood. Still, nature clearly finds a way to produce simple, robust and almost scale-free energy distributions for the cosmic-rays. Their importance has inspired a number of approaches to examining the production and transport of cosmic-ray particles in numerical simulations. I offer here a brief comparison of some of the methods that have been introduced.

Key Words : cosmic-rays—particle acceleration

I. Introduction

In a companion talk I discussed the roles that high energy particles, or "cosmic-rays" (CRs), play in astrophysics (Jones 2001) and how they are thought to be accelerated. Such particles, which generally refer to ions and electrons with energies well above the local "thermal" energies, are ubiquitous. Ordinarily the CRs appear to be distributed in momentum approximately as power laws, so that they carry little or no information about scales with them. The spectra can extend over many orders of magnitude. CR existence can be understood to be a natural consequence of the simple facts that tenuous plasmas are common and that they rarely come to a genuine thermodynamic equilibrium. Binary, Coulomb collision rates are very slow compared to other timescales, especially as one considers higher energies. Most of us are already well acquainted with this lack of "LTE" through the properties of atomic and molecular emission lines. Partly because of this situation, dynamical couplings typically depend on "collective" plasma processes rather than Coulomb collisions.

In our context, the point is most easily seen through comparisons of some of the basic interaction lengths of collisionless plasmas. Roughly, we have for the Coulomb scattering length for singly charged particles,

$$l_c \sim \frac{mv^2}{n\pi e^4} \sim \frac{(3kT)^2}{n\pi e^4} \sim 10^{18}\frac{T_6^2}{n}\ \mathrm{cm}, \qquad (1)$$

where the last form expresses the temperature in units of 10^6K and n is the plasma number density. Especially in hot, rarefied environments such as supernova remnants, or intergalactic media this length becomes obviously too large to be locally relevant, and certainly too long to produce fluid-like behaviors. By contrast the "collective" lengths associated with electrostatic (e.g., two-stream) and low frequency magneto-ionic instabilities are the ion inertial length or Debye length,

$$l_i = v/\omega_{pi} \approx \sqrt{\frac{3kT}{4\pi ne^2}} \approx 7\times 10^3\sqrt{\frac{T_6}{n}}\ \mathrm{cm}, \qquad (2)$$

and the gyro radius,

$$r_g = \frac{pc}{eB} \approx 1.6\times 10^9\frac{\sqrt{T_6}}{B_{-6}}\mathrm{cm}, \qquad (3)$$

where ω_{pi} is the proton plasma frequency and B_{-6} expresses the magnetic field strength in μGauss. The first form of r_g is general, whereas the last form applies to thermal nonrelativistic protons. Associated timescales for all three examples are found simply through division by $v_t = \sqrt{3kT/m} \approx 1.6\times 10^7\sqrt{T_6}$ cm/sec. Extensions to greater speeds than v_t are straightforward. These scales actually represent lower bounds for randomization of particle motions, especially for l_i and r_g at CR energies, since the latter represent resonant scattering interactions with waves, so typically involves multiple scattering events.

Generally, "thermalization" through collective effects is good enough that the bulk population has dynamical behaviors that we associate with fluids. This includes the formation of shocks where flow kinetic energy is dissipated inside relatively thin layers. Detailed statistical balances, however, can be quite different from LTE and often continue to evolve over times long compared to those associated with bulk dissipation. These differences from LTE can be much more than cosmetic, since substantial energy can be transferred to relatively small numbers of particles, especially once they become relativistic. Weak coupling at high energies allows those CRs to diffuse and to stream against the bulk flows, generating plasma turbulence, transporting energy and and pressure, thus, modifying the flows themselves. Those effects, in turn, alter the transport properties of the CRs. An important objective is, then, to include these effects in numerical

simulations. That is quite a challenge, as it turns out, both on account of the complexity and uncertainties in the microphysics and also because the length and time scales of that physics. Those are typically definable in terms of $r_g \propto p$, which can range over many orders of magnitude, extending upwards from close to the dissipation scales.

Although CRs are occasionally modeled simply as ordinary fluids with a softer equation of state than ordinary, nonrelativistic gas, that really misses many of their most important features, and in particular their weak coupling over a range of scales to the bulk plasma. More realistic properties are difficult and/or computationally expensive to model, but they are so important that a number of approaches have been developed to include them. In the following sections I will briefly outline some of the methods. The focus here is particle acceleration. As emphasized in Jones 2001, collisionless shocks are widely thought to be the most effective sites for CR acceleration through the "diffusive shock acceleration process" (DSA). Consequently, the following discussion stresses treatments of CR transport around those structures, although all of the methods can be and are applied elsewhere. Through my own work I am most familiar with approaches based on a Fokker-Planck, Boltzmann equation, so will offer more detail there.

II. Numerical Approaches

We can naturally divide the different approaches to calculating CR acceleration at shocks into two conceptual groups. One kind of approach follows individual particle motions and then constructs statistical measures of the population, such as the flux at a given velocity or energy. The other approach works with the probability distribution of the particles using either the Boltzmann equation in some form or a mathematically equivalent formulation of the problem. The former class generally makes no formal distinction between CRs and the rest of the particle population. The latter, on the other hand, generally assumes the fluid limit for the bulk of the population, then considers that higher energy particles propagate with respect to the bulk population in a diffusive manner. All methods must deal with the likelihood that characteristic scattering lengths will increase in some proportion to r_g, at least once the particles are superthermal. Each approach has its physical and numerical advantages as well as some significant limitations. To keep this introductory discussion simple, I ignore many important issues, such as obliquity of the ambient magnetic field. The cited literature can be used as a guide to explore these other matters.

(a) Particle Methods

i) Hybrid Plasma Simulations

Hybrid plasma simulations solve explicit equations of motion for individual ions in the presence of locally defined electric and magnetic fields. Electrons, whose individual motions are much quicker, so harder to follow individually, are treated as a light, charged fluid. Ions respond only to relatively slow variations in the E & M fields, so that is expected to be a very reasonable approximation. The electric and magnetic fields are evolved in the presence of these charge distributions directly through Maxwell's equations. The result is a relatively complete, fully nonlinear and self-consistent picture of the state of the plasma, including detailed distributions of the ions in space, time and momentum, plus properties of the plasma waves that are used to describe their behaviors. To simulate shocks these calculations typically inject a population of particles to reflect off a boundary and then stream back towards their injection point. That behavior leads to plasma instabilities and dissipation of the organized kinetic energy. Except for a limited number of *in situ* measurements of heliospheric shocks (e.g., Terasawa et al. 2001) hybrid simulations provide the most realistic picture we have of the structures of collisionless shocks. Designed more to study plasma dynamics than particle acceleration, these methods have nonetheless been applied to the latter and especially to issues associated with the so-called "injection problem". The very important issue there is how and how many particles manage to reach high enough energies at shocks so that they behave diffusively, as CRs are expected to do (e.g., Quest 1987; Giacalone et al 1997).

Hybrid simulations, because they are so detailed, are computationally very expensive. Generally they are intended to model behaviors over relatively short times and within only a relatively modest number of nominal shock thicknesses (typically of order several l_i or r_g depending on shock geometry). Most commonly hybrid simulations are carried out with at least one ignorable coordinate (that is, in either 2D or 1D) in order to reduce costs. Recent studies, however, have emphasized the importance of fully 3D particle orbits, especially when there is an oblique large scale magnetic field (e.g., Giacalone & Ellison 2000). Prohibitive cost makes hybrid simulations impractical tools for examining the acceleration of CRs to high energy, since that would require very long simulations and very large spatial domains. They are similarly not suitable for modeling time evolution of astrophysically large structures.

ii) Monte Carlo Simulations

This method streamlines the plasma simulation approach by assuming a predefined, universal scattering law for the ions; i.e., by assuming the spectral form and intensity of scattering waves. Only large scale magnetic fields are included explicitly. Fluctuating fields

are not modeled, so effectively the plasma wave field is assumed. Steady properties of the plasma are determined after a large set of numerical experiments for individual particles that establish the statistical properties of a population incident on a barrier, analogous to full plasma simulations. The scattering centers are physically supposed to be "attached" to the bulk flow or at least to possess motions definable with respect to the bulk flow. Thus, a self-consistent solution requires an iterative sequence of experimental "configurations" to define that frame of reference at each point (e.g., Ellison & Eichler 1984). The simplifications speed up the calculations over plasma simulations. Consequently, they are able to track particles over a wide range of momenta. Electrons remain difficult to track at low momenta, but once electron momenta (rigidities) are comparable to the ions it is straightforward to include them as well. Monte Carlo methods are readily applied to multiple charge species, and have even been used to explore the acceleration of charged dust grains as a source of heavy ion CRs (e.g., Ellison et al. 1997)

Because of the iterative approach, these methods are hard to apply to time dependent situations, and in common with plasma simulations, they have so far only been applied directly to single, plane shock structures. Finite time and geometry-related consequences are effectively included, however, in order to achieve convergence, as well as to model real physics. In particular, either finite acceleration times or finite shock sizes will limit the maximum momentum to which CRs can be accelerated. Similarly, to keep the CR pressure finite in strong, highly modified shocks, there must be an effective upper momentum cutoff to the CR spectrum. Monte Carlo simulations achieve this property either by a defined maximum momentum, or, usually, by applying a so-called "free escape boundary" placed some distance upstream of the shock. Since scattering lengths generally increase with momentum, this effectively limits the maximum momentum achievable. Comparisons between suitably chosen plane Monte Carlo shocks and spherical shocks computed by a Boltzmann equation method have shown good agreement (Ellison et al. 2000).

Like plasma simulations, Monte Carlo schemes make no distinction between thermal and nonthermal particles. In that sense they model CR injection naturally as "thermal leakage". On the other hand, they generally apply a simple scattering law over the entire momentum range (typically as a power law $l_s \propto p^\beta$), independent of time or position. Nonlinear analytic models of collisionless shock formation suggest that resonant particle trapping in postshock wave turbulence can be very effective for particle speeds up to several times the postshock flow speed, then rather abruptly diminish for higher speeds (Malkov 1998). Such differences will influence significantly the net injection rate, and especially its behavior as the shock evolves or in the event that the large-scale ambient magnetic field is oblique.

(b) Boltzmann Equation-based Methods

The alternate conceptual approach usually works with the distribution function, $f(p, x, t)$, for each particle species, usually beginning from the collisionless Boltzmann, or Vlasov equation. A full treatment of all particle momenta in this way would be impractical for all the reasons suggested above regarding other methods. So, here, one generally separates the particles of a given species into low momentum, "bulk" particles and high momentum CRs. The bulk particles are presumed to be thermalized and strongly coupled, so that they can be treated by conventional continuum gas dynamics or magnetogasdynamics. The CRs are presumed to be sufficiently strongly scattered by local wave turbulence that they are approximately isotropic with respect to the wave motions, but weakly enough coupled that they diffuse with respect to the bulk flow in response to density gradients in f. The resulting formalism for CRs is a Fokker-Planck equation (equation 6 in Jones 2001) that was first derived heuristically (e.g., Parker 1965) and subsequently more formally (e.g., Skilling 1975). This equation is of a form commonly called a "diffusion-convection equation" or "DCE" for short. There are several variants of the approach, as outlined below.

As with the particle methods, Boltzmann equation methods have both strengths and weaknesses. One of the principal strengths of this approach is that it can be applied pretty generally, at least in concept, since the formalism lends itself to any number of physical dimensions as well as time dependent treatments. The formalism is also adaptable to both analytic and numerical methods. Powerful and robust methods are available for both. On the other hand, these methods are still computationally expensive, especially when more than one spatial dimension is included, so numerical applications have been restricted to date, and analytic calculations, once they include nonlinear effects, quickly become very sophisticated. The separation of particles into bulk and CR populations is not clean, and treatments of injection intended to bridge the gap are still in their formative stages (see, e.g., Gieseler et al. 2000; Kang 2001). Similarly, while these methods allow self-consistent treatment of the scattering wave turbulence, and its evolution in space, wave-number and time, those behaviors are rarely included, since key properties are not well understood, and they add significant computational complexity and cost. In practice, then, the diffusion coefficient has generally been assumed *a priori* (but, see, e.g., Bell 1978, Jones 1993; Malkov et al. 2001).

i) Application of the DCE to DSA

The most straightforward approach in this paradigm is direct solution of the DCE. At shocks, the momentum diffusion term is usually ignored as being small. The spatial diffusion coefficient, κ, really is a tensor with respect to the magnetic field direction; in plane

shocks, it can be treated as a scalar corresponding to diffusion along the shock normal.

In numerical treatments several details are important to understand in order to obtain reliable solutions. First, is that the DCE applies only *outside* the dissipative shock structure. In fact, the formalism describes DSA physics on account of the fact that CRs cross shocks uninhibited. Correct solutions require one to match solutions across the shock, in particular to make the distribution f continuous there. Analytically, that is straightforward and has been even applied in semi-analytic computational schemes (e.g., Berezhko et al. 1995). However, this is a problem on a discrete grid where the shock has a finite thickness, Δx_s, typically several zones across. The DCE is solved everywhere, including the interior of the (unphysical) numerical shock. The key is to minimize the error in $\partial f/\partial t$ caused by this fact. The dominant error comes from adiabatic compression across the numerical shock structure, which gives a contribution to df/dt as

$$\frac{f}{3}\frac{\partial \ln f}{\partial \ln p}\nabla \cdot u = -\frac{qf}{3}\nabla \cdot u. \quad (4)$$

Physically, f should not change across the shock, so this should integrate to $(-qf/3)\Delta u$. But, owing to the finite numerical shock thickness and an upstream gradient $\partial f/\partial x \approx f/x_d$, where $x_d = \kappa/u_s$, the distribution function change across the shock structure will be of order $\Delta f \sim f\Delta x_s/x_d$. This adds an extra term giving a fractional error,

$$\left.\frac{\Delta f}{f}\right|_{err} \approx \frac{q}{3}\frac{\Delta u}{u_s}\frac{\Delta x_s}{x_d}. \quad (5)$$

An accurate solution requires that $|\Delta f/f|_{err} << 1$, and that, in turn requires a numerical grid fine enough that $\Delta x_s/x_d << 1$.

This constraint applies at all CR momenta, and, hence all r_g. Since $x_d \propto \kappa \propto \eta r_g$, (see, e.g., Jones 2001), this can be a serious problem. In particular, κ can cover many orders of magnitude when one tries to model CR momenta ranging from just superthermal to ultrarelativistic.

There have been at least two serious attempts to deal with the enormous span of length and time scales just described. The most direct is the semi-analytical method of Berezhko and collaborators (Berezhko et al. 1995). They rescale the physical lengths for each momentum in terms of the diffusion length, x_d, then integrate the DCE on both sides of a shock, applying the necessary matching conditions at the shock itself. Using a simplified gasdynamics they have effectively used this approach to study CR acceleration at spherical blast waves. The method does require *a priori* models for the diffusion, but otherwise is fully time dependent. The alternative approach is adaptive mesh refinement (AMR). Our group has been developing an AMR finite difference code that zooms on multiple levels around shocks in order to handle the same issues (Kang et al. 2001; Kang 2001). So far we applied it only to individual plane shocks, but we intend to apply it to more complex systems, such as spherical SNRs with multiple shocks.

ii) Numerical Schemes

The most straightforward Fokker-Planck based simulations are based on finite difference codes. Computationally the bulk flow is generally treated by one of the standard schemes for ideal gasdynamics or magnetogasdynamics based on their hyperbolic form. The DCE is not hyperbolic, however, since it contains a diffusive or parabolic term, so cannot be solved with those methods alone. The parabolic term can be handled effectively by explicit or implicit schemes. The constraint in equation 5, combined with the standard CFL condition on solution of hyperbolic equations means that an explicit diffusion timestep, $\Delta t_c << \Delta t_g$, where Δt_g is the characteristic gasdynamical time step. Thus, explicit approaches to this problem require multiple diffusion time steps in each gasdynamical step; i.e., "subcycling" (e.g., Jun et al. 1994). An alternative semi-implicit method is the Crank-Nicholson scheme. It avoids the use of subcycles, and gives comparable solutions to the explicit method (e.g., Falle & Giddings 1987; Kang & Jones 1991).

The need to follow evolution of $f(p,x,t)$ in an extra dimension (momentum) often makes the DCE several times more computationally expensive as gasdynamics or magnetogasdynamics in a given calculation. One must compute advection of particles in this dimension. The corresponding adiabatic compression term (equation 4) is as important to DSA as the diffusion term in the DCE. It is the only means by which particles can increase their energy when there is no momentum diffusion. The term depends on $q = -\partial \ln f/\partial \ln p$, so an accurate solution to the DCE at a given momentum depends on good numerical estimates for this slope. In a finite difference approach this generally means that one must use at least moderate resolution in momentum space, making this a major consideration. Two approaches have been introduced to cut this computational cost factor. Both attempt to reduce the required resolution in momentum by providing q or equivalent information.

The first and most commonly applied method to economize on the DCE was the so-called "two fluid" method (e.g., Drury & Völk 1981; Jones et al. 1994). That method replaces the DCE with its energy moment equation. One still needs an effective q, which, in this case is supplied by the closure relationship $P_c = (\gamma_c - 1)E_c$, where P_c and E_c are the CR pressure and energy densities respectively, while γ_c is the adiabatic index for the CR 'gas'. If $f(p)$ is a pure powerlaw, one can easily show that $\gamma_c = q/3$. Two fluid methods have been applied to both analytic and numerical treatments of CR acceleration and have provided a number of important insights. Their big weakness is

that q and γ_c are solutions to the problem, so generally cannot be assumed *a priori*. This is a particular problem when $q \to 4$ ($\gamma_c = 4/3$), since solutions can result with unrealistically large CR pressures. The most controversial example may be a steady state analytical solution for a CR modified shock in which $P_c = 0$ upstream, but $P_c \neq 0$ downstream even in the absence of injection (Drury & Völk 1981). That is an obviously pathological solution that could not be reached in any time dependent treatment, but it illustrates the issue.

A second approach we introduced divides the momentum space into broad, but finite bins, then integrates the DCE or its moments across those bins (Jun & Jones 1999; Jones et al. 1999; Miniati 2001). By assuming some simple sub-bin model for $f(p)$, most conveniently a piecewise powerlaw, fluxes between momentum bins can be computed, just as for finite volume gasdynamic methods. By requiring continuity and following multiple moments of the DCE one can obtain self-consistent solutions for $q(p)$ while still reducing the needed resolution in momentum by a large factor over traditional methods. This approach has good potential as a method that can be applied in complex, 2D and 3D flows. So far this method has been applied only to test-particle simulations of that kind, but there is no fundamental reason it cannot be extended to simulations in which CR dynamical feedback is included. The issue there comes back to obtaining fine spatial and time resolution required to simulate shock modification properly.

Finally I mention briefly an alternate approach that is conceptually related to the Boltzmann equation methods, but is operationally more like some particle methods. This is the use of stochastic differential equations to follow the evolution of the distribution function, f. From Liouville's theorem we know we can evolve f simply by following the trajectory of an element of phase space over time. When the evolution is deterministic one computes that trajectory with ordinary differential equations of motion. In the presence of diffusion, that motion takes on a stochastic behavior that can be modeled by conducting numerical experiments for the trajectories analogous to Monte Carlo particle methods. Formally this is equivalent to the Boltzmann equation method and its applicability to CR acceleration has been demonstrated (Achterberg & Krülls; Marcowith & Kirk 1999). It is yet to be applied to practical astrophysical contexts, but we can assume that is forthcoming very soon.

III. Conclusion

Simulating CR acceleration and transport is challenging for both technical and physical reasons. It is also very important to manage, since CRs are ubiquitous and very probably play central roles in the dynamics and energy budgets of many environments. CRs also provide unique signatures of the physics in those environments, if we can only learn to interpret them. Substantial progress has taken place in accomplishing that task, although we still lack numerical methods that are robust and economical enough to apply broadly. That time may come soon. however.

ACKNOWLEDGEMENTS

I am grateful to the organizers of this excellent workshop for their invitation to participate and to their generous hospitality during our visit to Pusan. This work was supported in part by NASA grant NAG5-10774, by NSF grant AST00-71167, and by the University of Minnesota Supercomputing Institute.

REFERENCES

Achterberg, A. & Krülls, W. M. 1992, A&A, 265, L13

Bell, A. R. 1978, MNRAS, 182, 147

Berezhko E. G., Ksenofontov, L. T. & Yelshin, V. K. 1995, Nuclear Phys. B., 39A, 171

Drury, L. O'C. 1983, Rep. Prog. Phys., 46, 973

Drury, L. O'C. & Völk 1981, ApJ, 248, 344

Ellison, D. C., Berezhko, E. G. & Baring, M. G. 2000, ApJ, 540, 292

Ellison, D. C., Drury, L. O'C. & Meyer, J.-P. 1997, ApJ., 487, 197

Ellison, D. C. & Eichler, D. 1984, ApJ., 286, 691

Falle, S. A. E. G. & Giddings, J. R. 1987, MNRAS, 225, 399

Gieseler, U. D. J., Jones, T. W. & Kang, H. 2000, A&A, 364, 911

Jones, T. W., 1993, ApJS, 90, 969

Jones, T. W., Ryu, D. & Engel, A. 1999, ApJ, 512, 105

Jones, T. W. 2001 (these proceedings)

Jun, B.-I., Clarke, D. A. & Norman, M. L. 1994, ApJ, 429, 748

Jun, B.-I. & Jones, T. W. 1999, ApJ, 511, 774

Kang, H. 2001 (these proceedings)

Kang, H. & Jones, T. W. 1991, MNRAS, 249, 439

Kang, H., Jones, T. W., LeVeque, R. J. & Shyue, K. M. 2001, ApJ, 550, 737

Kirk, J. G., Guthmann, A. W., Gallant, Y. A. & Achterberg, A. 2000, ApJ, 542, 235

Malkov, M. A., 1997a, ApJ., 485, 638

Malkov, M. A., 1997b, ApJ., 491, 584

Malkov, M. A. 1998, Phys Rev. E., 58, 4911

Malkov, M. A., Diamond, P. H. & Vólk, H. J. 2000, ApJ, 533

Marcowith A. & Kirk, J. G. 1999, A&A, 347, 391

Miniati, F. 2001, Comm. Comp. Phys, (in press)

Parker, E. N. 1965, Planet. Space Sci., 13, 9

Quest, K. B. 1988, J. Geophys. Res., 93, 9649

Skilling, 1975, MNRAS, 172, 557

Terasawa, T. et al. 2001, 26th ICRC (Hamburg), 6, 528.