The Astrophysical Journal, 642:244–259, 2006 May 1

NONLINEAR SHOCK ACCELERATION BEYOND THE BOHM LIMIT

M. A. Malkov and P. H. Diamond
University of California at San Diego, La Jolla, CA 92093-0319; mmalkov@ucsd.edu, phd@physics.ucsd.edu
Received 2005 September 7; accepted 2005 December 12

ABSTRACT

We suggest a physical mechanism whereby the acceleration time of cosmic rays (CRs) by shock waves can be significantly reduced. This creates the possibility of particle acceleration beyond the knee energy at $\sim 10^{15}$ eV. The acceleration results from a nonlinear modification of the flow ahead of the shock supported by particles already accelerated to the knee momentum at $p \sim p_*$. The particles gain energy by bouncing off converging magnetic irregularities frozen into the flow *in the shock precursor* and not so much by recrossing the *shock itself*. The acceleration rate is thus determined by the gradient of the flow velocity and turns out to be formally independent of the particle mean free path (mfp). The velocity gradient is, in turn, set by the knee particles at $p \sim p_*$ as having the dominant contribution to the CR pressure. Since it is independent of the mfp, the acceleration rate of particles above the knee does not decrease with energy, unlike in the linear acceleration regime. The reason for the knee formation at $p \sim p_*$ is that particles with $p > p_*$ are effectively confined to the shock precursor only while they are within limited domains in the momentum space, while other particles fall into "loss islands," similar to the "loss cone" of magnetic traps. This structure of the momentum space is due to the character of the scattering magnetic irregularities. They are formed by a train of shock waves that naturally emerge either from unstably growing and steepening magnetosonic waves or as a result of acoustic instability of the CR precursor (CRP). These losses steepen the spectrum above the knee, which also prevents the shock width from increasing with the maximum particle energy.

Subject headings: acceleration of particles — cosmic rays — shock waves — supernova remnants — turbulence

Online material: color figures

1. INTRODUCTION

The first-order Fermi acceleration, also known as diffusive shock acceleration (DSA), is regarded as the principal mechanism whereby Galactic CRs are produced. The physics of the energy gain is very simple: particles get kicked sequentially by scattering centers that are on the opposite side of the shock front and so approach each other. Its modern version has been developed by Krymsky (1977), Axford et al. (1977), Bell (1978), Blandford & Ostriker (1978), and others. Soon after that, however, it was realized that the mechanism can only marginally (if at all; Lagage & Cesarsky 1983) account for the acceleration of the Galactic CRs up to the knee energy at 10^{15} eV (see also Jones et al. 1998; Kirk & Dendy 2001, for more recent good discussions). Up to the knee, the spectrum is nearly a perfect power law, so it is most probably produced by a single mechanism. At the same time, the low particle scattering rate makes the mechanism somewhat slow and the particle confinement to the shock insufficient to accomplish this task over the active lifetime of a typical Galactic supernova remnant (SNR) shock, which is considered to be the most probable site of CR acceleration.

Not surprisingly, there have been many suggestions as to how to improve the performance of the DSA. In one way or another they targeted the most important and, at the same time, the most uncertain parameter that determines the acceleration time, the particle diffusivity κ. The acceleration rate can be written roughly as $\tau_{\rm acc}^{-1} = U_{\rm sh}^2/\kappa(p)$, where $U_{\rm sh}$ is the shock velocity. Therefore, the acceleration rate *decreases* with particle momentum p, since the particle diffusion coefficient κ usually *grows* with it. In the Bohm limit, which is reasonably optimistic, one can represent κ as $\kappa = cr_g/3$, where r_g is the particle gyroradius (substituted in the last formula as a particle mfp λ) and c is the velocity of light. The condition $\lambda \sim r_g$ is justified when the magnetic field perturbations that scatter accelerated particles (and are driven by them at the same time) reach the level of the ambient magnetic field, $\delta B \sim B_0$. Such a high fluctuation level has been long considered as a firm upper limit (McKenzie & Völk 1982), and there have been calculations indicating that the turbulence may saturate at a somewhat lower level due to nonlinear processes (e.g., Achterberg & Blandford 1986).

It should be noted, however, that in cases of efficient particle acceleration, i.e., when a significant part of the shock ram pressure $\rho U_{\rm sh}^2$ is converted to the pressure of accelerated particles P_c, there is indeed sufficient free energy of accelerated particles that can potentially be transformed to fluctuations with $\delta B \gg B_0$. This would decrease λ significantly below the particle gyroradius (calculated by the unperturbed field B_0), so the acceleration time would also decrease. From simple energetics principles, one can estimate the maximum fluctuation energy as (McKenzie & Völk 1982)

$$\frac{\delta B^2}{B_0^2} \sim M_{\rm A}\frac{P_c}{\rho U_{\rm sh}^2}, \tag{1}$$

where $M_{\rm A} = U_{\rm sh}/V_{\rm A}(B_0)$ is the shock Alfvén Mach number. Bell & Lucek (2001) suggested that this upper bound may indeed be reached, corroborating this idea by the numerical simulation (Lucek & Bell 2000). The supporting simulation was rather limited in the system size and particle energy range, which in turn truncates the wave spectrum. Therefore, the saturation effects predicted earlier by McKenzie & Völk (1982), Achterberg & Blandford (1986), and Blandford & Eichler (1987) have not been observed in the simulations by Lucek & Bell (2000). In any case, equation (1) provides only an upper bound to the actual wave amplitude, which may be much lower due to the saturation effects, not included in its derivation. On the other hand, Bell (2004) suggested that the magnetic field can reach an even higher level if the CR current that drives the instability is fixed. According to this suggestion, the instability saturation mechanism is due to magnetic

tension. Under these circumstances the estimate given by equation (1) acquires an additional factor $\sim M_A(U_{sh}/c)$, which may be quite large in the case of young SNRs, making the turbulence level sufficient to increase the maximum particle momentum by a factor of 10^3.

A different approach to the same goal of increasing the turbulent component beyond the ambient field level has been pursued in the recent papers by Ptuskin & Zirakashvili (2003, 2005). These authors studied the DSA in the presence of a Kolmogorov-type magnetohydrodynamic (MHD) cascade initiated by the strong MHD instability in the CRP. In addition, recently Bell (2004) took up from the Bell & Lucek (2000) and Lucek & Bell (2000) approaches and replaced the time-consuming particle simulations by providing an estimate of the growth rate caused by the CR current at the shock. With the current kept fixed, the waves indeed grow to a level sufficient to increase the maximum particle momentum by a factor of 10^3. It remains unclear, however, whether the CR trapping by the self-generated turbulence and other wave-particle interactions (e.g., Völk et al. 1984; Achterberg & Blandford 1986) saturates the instability at a lower level. Besides, the alleged growth of δB beyond B_0 clearly invalidates treatments of the linear instability that are based on the $\delta B \ll B_0$ assumption.

Diamond & Malkov (2004) suggested an alternative picture in which a strong δB may be generated as a result of a type of inverse cascade in k-space. Strong field perturbations driven resonantly by already accelerated particles are nonlinearly coupled to longer scales, not only facilitating the acceleration of higher energy particles but also providing an ambient field for them as the longest scale field perturbation. Such a cascade can be driven by scattering of the Alfvén waves on acoustic waves driven, in turn, by the Drury instability of the CRP. Another related route to large-scale magnetic fluctuations is modulational instability of the Alfvén waves themselves. The condensation of magnetic energy at long scales allows us to marginally preserve the weak turbulence $\delta B < B_0$ requirement, where the role of B_0 is taken by the longest scale part of the magnetic energy spectrum (longer than the particle gyroradius). This could possibly weaken the saturation factors setting in at $\delta B \sim B_0$, which are noted above.

The above-mentioned works, while aimed at the explanation of acceleration beyond the knee, do not specifically address mechanisms that could be responsible for the formation of the knee itself. An exception is a paper by Drury et al. (2003), in which the authors suggested that the knee can appear as a result of an abrupt slowing down of the acceleration at the sweep-up phase in combination with the Bell hypothesis about the generation of the strong magnetic field. Some further interesting ideas about the knee origin can be found in a recent paper by Sveshnikova (2003).

In this paper we propose a different scenario of faster than Bohm acceleration that is also intimately related to the knee phenomenon. Odd as it sounds, this mechanism does not require super-Bohm magnetic field fluctuations. The reason for that is that its rate does not depend on the particle scattering mfp λ. In § 2 we describe the mechanism by comparing it with both the test particle and nonlinear DSA regimes under the Bohm diffusivity. We analyze the conditions under which the proposed acceleration regime is faster than the latter two. This sets the stage for §§ 3 and 4, in which a preliminary study of particle dynamics and transport in the CRP is presented. Section 5 deals with the estimates of the maximum energy that can be reached in excess of the knee energy. In § 6 an acceleration model that allows us to calculate the slope of the spectrum between the knee and the maximum energy is presented. We conclude with a summary and brief discussion.

2. ACCELERATION MECHANISM: A PRIMER

Perhaps the easiest way to understand why and when the suggested version of the DSA becomes faster than the standard one is to consider why the latter is slow. For that, we turn to the individual particle treatment due to Bell (1978). Upon completing one acceleration cycle, i.e., crossing and recrossing the discontinuity, a particle gains momentum

$$\frac{\Delta p}{p} \sim \frac{\Delta U}{c}, \tag{2}$$

where ΔU is the relative velocity between the upstream and downstream scattering centers $\Delta U = U_1 - U_2 \sim U_1$. Thus, over the acceleration time (when the momentum gain $\Delta p \sim p$), the number of cycles that the particle needs to make is $N_{cycl} \sim c/U_1 \gg 1$. Apart from numerical factors and the differences between the upstream and downstream residence time contributions, the particle acceleration time can be estimated as

$$\tau_{acc} \simeq \frac{\kappa(p)}{U_1^2} \sim \lambda c/U_1^2 \sim \tau_{col} c^2/U_1^2, \tag{3}$$

where λ and τ_{col} are the particle mfp and collision time, respectively. It is also useful to note that the acceleration time is of order the time needed for the fluid element to cross the diffusion zone ahead of the shock, $\tau_{acc} \sim L_{dif}(p)/U_1$, where $L_{dif} = \kappa(p)/U_1$ is the particle diffusion length. Using equations (2) and (3), one can write the following relation between the acceleration time, the time needed to complete one cycle, and the collision time:

$$\tau_{acc} : \tau_{cycl} : \tau_{col} \sim \frac{c^2}{U_1^2} : \frac{c}{U_1} : 1.$$

Therefore, out of the c^2/U_1^2 wave-particle collisions needed to gain a momentum $\Delta p \sim p$, only c/U_1 are productive in terms of the energy gain. Most of the collisions are wasted.

The situation changes fundamentally when the number and energy of accelerated particles increase. First of all, the shock gets shielded with a cloud of accelerated particles (CRs) diffusing ahead of it, and the plasma flow becomes significantly modified by their back-reaction. In addition to an abrupt velocity jump (which can be significantly reduced by this back-reaction), the flow develops an extended CRP of length $L_p \sim \kappa(p_*)/U_1 = L_{dif}(p_*)$, where p_* is the particle momentum corresponding to the maximum contribution to the pressure of accelerated particles. The flow in the CRP gradually slows down toward the main shock (subshock), and the spectrum is flatter than p^{-4} at high momenta, so that $p_* \simeq p_{max}$ and thus $L_p \simeq L_{dif}(p_*)$. Let us assume that a particle undergoes subsequent scattering by two scattering centers, approaching each other at a speed $\delta U \ll c$ (Fig. 1). Suppose that immediately after scattering off the right center, the particle has (in that center reference frame) the momentum p and the cosine of the pitch angle with respect to the plasma flow (moving to the left), $\mu < 0$. After scattering off the left center, the particle momentum becomes p', and the cosine of the pitch angle $\mu' > 0$, i.e., the particle moves back toward the first scattering center. Since scattering is elastic, from particle kinematics we obtain

$$\frac{p'}{p} \simeq \left(1 - \frac{\delta U}{c}\mu\right)\left(1 + \frac{\delta U}{c}\mu'\right). \tag{4}$$

This result is obviously identical to the momentum gain in the conventional DSA theory (Bell 1978). The left scattering center

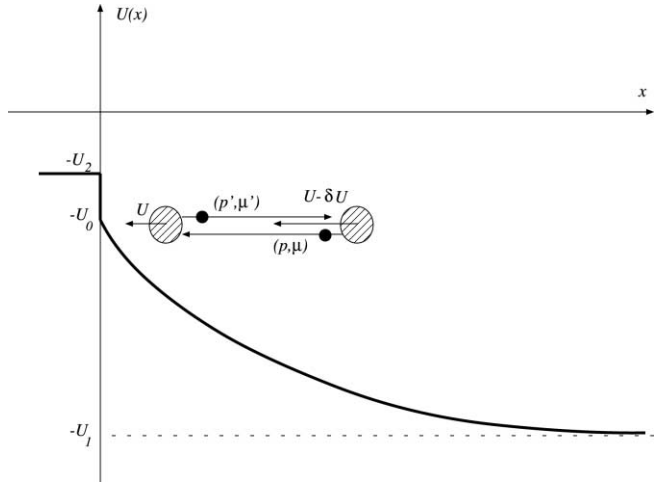

FIG. 1.—Flow velocity profile in a CR dressed shock (*solid curve*). Hatched circles represent two scattering centers moving with the flow and approaching each other at a speed δU. The filled circles depict a CR particle after the first and second collision (see text).

can be identified with one of the downstream, and the right one with one of the upstream, scattering centers. Considering μ and μ' as independent stochastic variables evenly distributed over the intervals $(-1, 0)$ and $(0, 1)$, respectively, one can calculate an average momentum gain,

$$\frac{\langle \delta p \rangle}{p} \simeq 2\frac{\delta U}{c}\langle \mu \rangle = \frac{4}{3}\frac{\delta U}{c}, \tag{5}$$

where

$$\langle \cdot \rangle = \frac{\int_0^1 (\cdot)\mu\, d\mu}{\int_0^1 \mu\, d\mu}.$$

Note that the last result also holds in the case in which there is an intervening scattering that does not reverse the particle direction. This can be seen from equation (4) by noting that in such a case, both μ-values would be of the same sign and cancel on averaging to first order in $\delta U/c$. Just as in the conventional (linear) acceleration theory, a pair of scattering centers does not need (and is unlikely) to be the same at the next acceleration cycle. However, this is quite possible in the scattering environment considered later in this paper. If so, then the momentum gain per cycle δp can be obtained from the conservation of the adiabatic invariant,

$$J_\parallel = \oint p_\parallel\, dl_\parallel = \text{constant},$$

where the integral runs between the two scattering centers. If the interaction time of a particle with a scattering center is much shorter than the flight time between collisions, then $J_\parallel \simeq p_\parallel l =$ constant, where $l(t)$ is the (decreasing) distance between the two centers. Formula (5) immediately follows from the last relation.

The average time needed to complete the cycle is

$$\langle \delta t \rangle = 2\left\langle \frac{\lambda}{c\mu} \right\rangle = 4\frac{\lambda}{c},$$

where we substituted λ for an averaged distance between scattering centers. Now we can write

$$\frac{dp}{dt} \equiv \dot{p} \simeq \frac{\langle \delta p \rangle}{\langle \delta t \rangle} = \frac{1}{3}p\frac{\delta U}{\lambda}.$$

As long as $\lambda \ll L_p$, one can approximate δU as

$$\delta U \simeq -\lambda\frac{\partial U}{\partial z},$$

so that the local acceleration rate is

$$\dot{p} \simeq -\frac{1}{3}p\frac{\partial U}{\partial z}. \tag{6}$$

We observe that the particle mfp dropped out of the acceleration rate in the smooth part of the shock structure, in contrast to the linear acceleration regime that occurs at the shock discontinuity. The last formula is, of course, consistent with the standard diffusion convection equation,

$$\frac{\partial f}{\partial t} + U\frac{\partial f}{\partial z} - \frac{\partial}{\partial z}\kappa\frac{\partial f}{\partial z} = \frac{1}{3}\frac{\partial U}{\partial z}p\frac{\partial f}{\partial p}, \tag{7}$$

if one considers its characteristics, ignoring diffusion effects.[1] Then in the vicinity of the (abrupt) momentum cutoff $p_{\max}$, the main terms in equation (7) are the first one on the left-hand side and the acceleration term on the right-hand side. This clearly shows that a (front) solution propagates in momentum space at the speed given by equation (6), basically independent of z and κ. It is important to emphasize, however, that the independence of the acceleration rate of the particle diffusion coefficient (or mfp, as shown above) does not mean that the acceleration is faster than in the linear case where such dependence is the main factor making the acceleration relatively slow. The reason is very simple. The derivative $\partial U/\partial z$ in equations (6) and (7) itself depends on κ. Indeed, estimating $\partial U/\partial z$ as

$$\frac{\partial U}{\partial z} \simeq \frac{U_1}{L_p} \simeq \frac{U_1^2}{\kappa(p_*)}, \tag{8}$$

one sees that the acceleration time is $\tau_{\rm acc} \sim \kappa(p_*)/U_1^2$; i.e., particles with momenta $p \sim p_*$ are accelerated approximately at the same rate as in the linear theory (apart from a numerical factor $O(1)$; see Malkov & Drury [2001]). Clearly, since the local (i.e., inside of the CRP) acceleration rate is independent of momentum, particles with $p < p_*$ accelerate more slowly than the linear theory acceleration rate, while particles with $p > p_*$ accelerate faster rate than in linear theory and may in principle be accelerated faster than the Bohm rate. The overall acceleration process is illustrated in Figure 2, depending on whether the pressure-dominant momentum p_* is stopped from growing after some $t = t_*$. If it is, particle momentum grows exponentially; otherwise, it continues to grow linearly in time.

There are two problems with realization of this possibility. First, because in the nonlinear acceleration regime the spectrum at the upper cutoff is flatter than p^{-4} (it actually flattens to $p^{-3.25}$ at $p_{\max}$ [Malkov 1997] in the limit of high Mach number and

[1] As was shown by Malkov & Drury (2001), there is a reason to do so in certain situations. In particular, under a strong nonlinear shock modification and Bohm diffusivity ($\kappa \propto p$), the velocity profile is linear, i.e., $\partial U/\partial z \simeq$ constant in a significant part of the CRP.

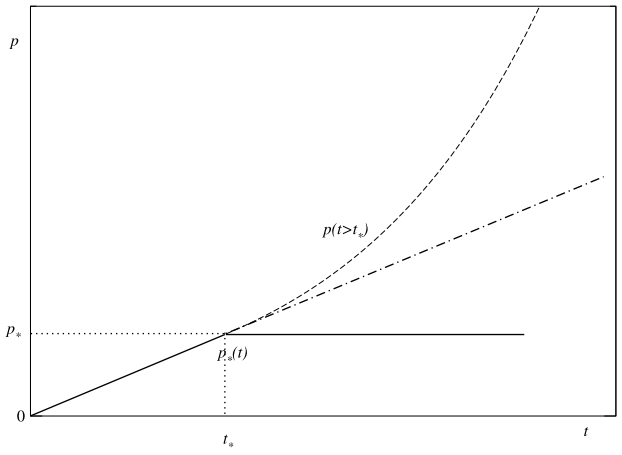

FIG. 2.—Schematic representation of acceleration process. Up to the time t_* the acceleration proceeds at the standard rate, which is similar during both the linear and nonlinear phases (*solid line*). The particle pressure-dominant momentum $p_*(t)$ grows as the maximum momentum of the entire spectrum. If for $t > t_*$ the pressure-dominant momentum p_* remains flat (*solid line*), particles with $p > p_*$ maintain the same acceleration rate that they had when their momentum was equal to p_* (*dashed line*). The dash-dotted line shows for comparison how the maximum momentum $p_{\max} = p_*$ would continue to grow after $t = t_*$, had $p_*(t)$ not stopped growing at $t = t_*$. [*See the electronic edition of the Journal for a color version of this figure*.]

$p_{\max} \gg mc$), p_* is clearly identifiable with $p_{\max}$. Therefore, unless the spectrum slope changes at $p \sim p_*$, it is difficult to accelerate particles beyond p_*. This is why the acceleration rates in both linear and nonlinear regimes are similarly slow, although for rather different reasons. In the linear regime, the acceleration rate slows down with time since the acceleration cycle duration, $\tau_{\rm cycl}$, grows with momentum (as the gyroperiod does), so the particle momentum grows only linearly with time. In the nonlinear regime, the width of the shock grows linearly with p_*, so that the acceleration rate (8) decreases with p_* (not with p) and p_* also grows linearly with time. Clearly, the first requirement would be to stop p_* from growing at some point or at least to slow down its growth considerably. If $p_* =$ constant, then particles with $p > p_*$ increase their momentum exponentially (eqs. [6] and [8]).

One simple reason for the saturation of $p_*(t)$ is a geometric one. The thickness of the CR cloud ahead of the subshock, $L_p(p_*)$, cannot exceed the accelerator size, i.e., some fraction of the shock radius R_s, in a SNR, for example, (e.g., Drury 1983; Berezhko 1996). Thus, we may identify p_* with the maximum momentum achievable in an SNR but limited by the geometric constraints, not by the lifetime of SNRs. What is also required here is that, along with or, independent of this geometric limit to p_*, the character of *particle confinement* to the shock *changes* at p_* when p_* reaches some critical value. This should lead to a break in the spectrum at $p \sim p_*$, such that its slope at $p > p_*$ is steeper than p^{-4} and flatter at $p < p_*$. Then the main contribution to the particle pressure would come from particles *around* p_*.

For example, the spectral break may be initiated by a strong shock modification by accelerated particles (Malkov et al. 2002). Namely, Alfvén wave compression in the CRP shifts the waves to the short-wavelength end of the spectrum, leaving particles with $p > p_*$without the resonant waves. In fact, some of those particles still remain resonant with the waves with $k \geq r_g^{-1}(p_*) \equiv k_*$ as long as their pitch angles satisfy the condition $p\mu < m\omega_c/k_*$, where μ is the cosine of the pitch angle and ω_c is the gyrofrequency. At the same time, the number of these resonant particles rapidly drops with momentum since their pitch-angle distribution is limited by $|\mu| < m\omega_c/k_*p$, and they can hardly form a good pool for the acceleration. Rather, the momentum spectrum can be shown to decay exponentially at $p > p_*$.

However, the latter example is based on a weak-turbulence picture in which a sharp resonance of particles with waves of random phases determines particle dynamics. As we demonstrate below, strong magnetic field perturbations result in a quite different character of particle confinement. In § 3 we consider a scattering environment that (although still simplified) retains some important characteristics of a more realistic strongly unstable turbulent CRP. As discussed earlier in this section, this turbulence should satisfy the following two conditions: (1) particles with momenta $p > p_*$ should not suffer catastrophic losses, and (2) their spectrum must nevertheless be steeper than p^{-4}. As we demonstrate in §§ 3 and 4, a gas of relatively weak shocks, which can emerge in a number of ways, provides the required scattering.

3. PARTICLE DYNAMICS

The conventional paradigm of particle transport in shock environment is the diffusive pitch-angle scattering on self-generated, random phase Alfvén waves. The resulting spatial transport is also diffusive with a coefficient that scales with the wave amplitude as $\kappa(p) \propto \delta B_k^{-2}$, where the fluctuation wavenumber k must be set in the resonance relation with the particle momentum, $kr_g(p) \simeq 1$. While this is the most consistent plasma physics approach that can be derived from first principles using the quasi-linear theory, provided that $\delta B \ll B_0$, there is no guarantee that it holds when the waves go nonlinear (e.g., Völk et al. 1984; Achterberg & Blandford 1986). The latter is the rule rather than an exception in shock environments, which has been documented via observations, theory, and numerical studies. So the Earth's bow shock, as well as interplanetary and cometary shocks, reveals a variety of coherent nonlinear magnetic structures. Usually, they originate upstream as a result of nonlinearly developed instability of the distribution of ions reflected from the shock in the case of oblique propagation. If the shock is quasi-parallel, the unstable ion distribution forms by thermal leakage from the downstream plasma. In cometary plasma, the pick-up ions drive the instability. Very often the magnetic structures are observed as an ensemble of discontinuities referred to as shock trains or shocklets (Tsurutani et al. 1987). There is no shortage of theoretical models describing these features, and we briefly return to them below. Standing somewhat outside of these models, but perhaps more specific to the nonlinear CR shocks, is the acoustic instability driven by the gradient of the CR pressure in the CRP. It was first studied by Drury (see, e.g., Drury & Falle 1986; Zank et al. 1990; Kang et al. 1992, and references therein), and it also evolves into a gas of moderately strong shock waves propagating toward the main shock (subshock) in the CRP. In what follows in this section, we assume that the shock precursor is filled with the gas of such shocks separated by some distance L.

Let us consider such a shock train propagating in the CRP. Obviously the relative speed between the individual shocks and the speed of the bulk flow is small compared to the particle speed, $U(z) \ll c$. Therefore, we can consider the problem of particle scattering by scattering centers (shocks) and particle acceleration caused by the relative motion of the scattering centers, separately. Note that the acceleration part of the problem has been preliminarily considered in § 2, and we return to it in § 6. Turning here to the scattering part of the problem, we assume for simplicity that the shocks are one-dimensional and propagate at the same speed along the shock normal. Furthermore, we transform to their reference frame, neglecting compression by the main flow $U(z)$ and their relative motion. Even under these simplifications, the shock train magnetic structures can be still rather complicated. For

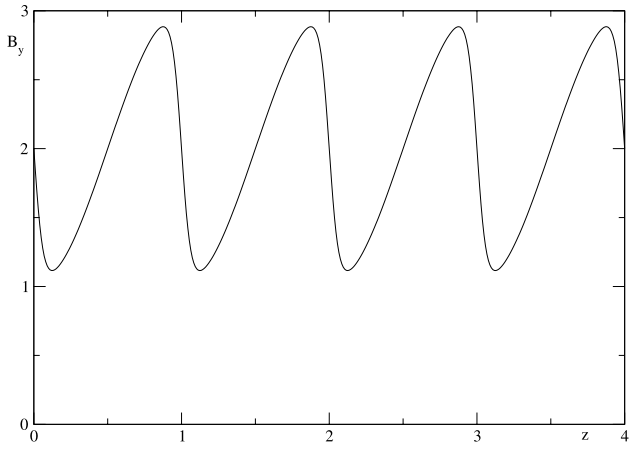

FIG. 3.—The y-component of the magnetic field in the shock train given by eq. (9) for $\nu = 0.2$, $\tilde{B}_y = 1$, and $\overline{B}_y = 2$.

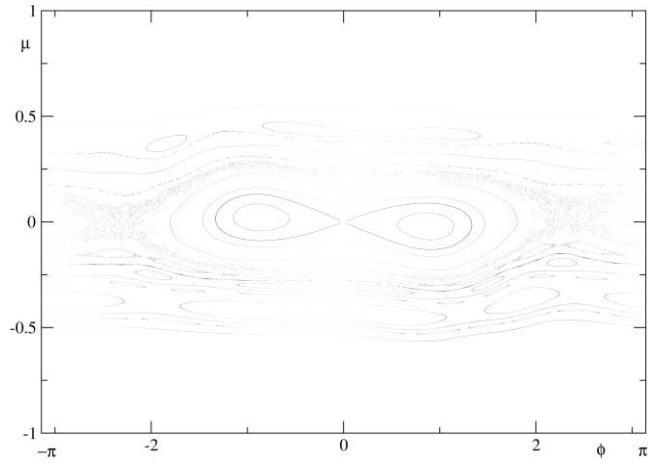

FIG. 4.—Particle orbits shown for $\nu = 0.5$, $\overline{B}_y = 2$, and $\tilde{B}_y = 0.3$ in eq. (9) and $p = 1$ on the (ϕ, μ) plane. Each dot corresponds to the intersection of particle trajectories with one of the planes located at an integer z and coinciding with the shock locations (Poincaré section). Several particle orbits form invariant curves (Kolmogorov-Arnold-Moser tori), as well as stochastic layers around some of the separatrices (see text).

example, they have been extensively studied in the frame of the so-called derivative nonlinear Schrödinger equation and its modifications (Medvedev & Diamond 1996). Alfvén–ion acoustic wave coupling results in wave steepening and gives rise to a shock train. A kinetic effect included in this model is particle trapping, which was shown to be important for the formation of the shock train and which has been studied numerically (Hada et al. 1990; Medvedev et al. 1997). In another model, the shock train forms from a balance of a quasi-periodic driver and the nonlinearity of the magnetic field perturbations. The driver may originate from one or a few unstable harmonics providing the smooth parts of the solution that steepen into shocks. These shocks also have a complicated spatial structure, typically characterized by a fast rotation of the magnetic field vector, caused by dispersion (see Malkov 1996, and references therein). However, since the gyroradius of high-energy particles, even if reduced by possible magnetic field amplification, is still much larger than both the width of the shock transition and the dispersive scale $c/\omega_{\rm pi}$, we replace each shock by a discontinuity of a coplanar magnetic field. A different possibility that results in a similar behavior of the magnetic field occurs when the strong Drury instability discussed earlier evolves into an ensemble of moderately strong shocks.

Both cases are covered by the following representation of the magnetic field, which corresponds to a periodic sequence of shocks, in which only one component of the field varies with the coordinate in a comoving reference frame, $\boldsymbol{B} = \left(0,\, B_y,\, 1\right)$, where

$$B_y(z) = \overline{B}_y + \tilde{B}_y \sin[\pi(z - 1/2)]\, \tanh\left\{\frac{1}{\nu}\cos[\pi(z - 1/2)]\right\}. \tag{9}$$

Here $\overline{B}_y$ is the constant component of the transverse to the shock normal component of the magnetic field, and $\tilde{B}_y$ characterizes the strength of the shocks in the shock train (both normalized to the constant z-component, B_0; Fig. 3). We normalized the coordinate z to the distance between shocks L. Of course, in reality, different shocks in the shock train do not have the same strength so that the coefficient $\tilde{B}_y$ should be replaced by a stochastic variable with a probability density function (pdf), inferred from the shock dynamics. We return to this discussion in § 4, but in this simplified study we assume all the shocks in the shock train have the same strength.

Let us consider particle dynamics in the magnetic field given by equation (9). It is convenient to write the equations of motion using dimensionless variables in which time is normalized to L/c. Since we are primarily concerned with the dynamics of particles having momenta $p \gtrsim p_* \gg mc$, we normalize their speed to the speed of light and set the absolute value of particle velocity to unity, $V \approx 1$. The remaining variable p is normalized naturally so that the corresponding gyroradius $r_g(p)$ is measured in the units of L; that is, the particle momentum p is scaled to eB_0L/c. We also introduce the cosine of the pitch angle $\mu = p_z/p$. The equations of motion read

$$\dot{\boldsymbol{p}} = \boldsymbol{n}\times\boldsymbol{B}, \tag{10}$$

$$\dot{z} = \mu, \tag{11}$$

where $\boldsymbol{n} = \boldsymbol{p}/p$ and $p =$ constant.

Figure 4 shows particle trajectories as a Poincaré section that represents orbit crossings of the shocks, i.e., the planes $z = j$ (where j is an integer) on the gyrophase–pitch-angle plane, (ϕ, μ). To make a connection with the drift approximation, we have transformed the particle momentum to the coordinate system rotated around the x-axis by the angle $\vartheta = \tan^{-1}\overline{B}_y$. In this coordinate system, particles spiral around the z-axis with $\mu =$ constant if $\tilde{B}_y = 0$. The gyroradius is taken to be equal to the distance between the shocks ($p = 1$, resonant case in terms of the linear theory), but the shock transitions are broad so that no distinct shocks with the steep magnetic field gradient are present and the particle dynamics remains largely regular. For smaller p the system becomes almost integrable since it can be reduced to a two-dimensional one, because the particle magnetic moment is conserved with good accuracy. In the case shown in Figure 4, however, some of the separatrices are clearly destroyed. Nevertheless, there is a clear separation of trapped (magnetic mirroring) and passing particles. Based on the conservation of the first adiabatic invariant $I \propto (1 - \mu^2)/B$ (magnetic moment), the trapped particles must be confined in momentum space to $|\mu| < \mu_{\rm crit} = [(B_{\rm max} - B_{\rm min})/B_{\rm max}]^{1/2} \approx 0.27$, which is close to what can be seen from Figure 4. It is important to emphasize here that the trapping area does not shrink with p, as was the case in the linear resonance situation, discussed at the end of § 2. Clearly the spatial transport of these particles is nondiffusive. The trapped particles are convected with the shock train. The untrapped particles ($|\mu| > \mu_{\rm crit}$) propagate ballisitically and escape from the CR cloud. Thus, even

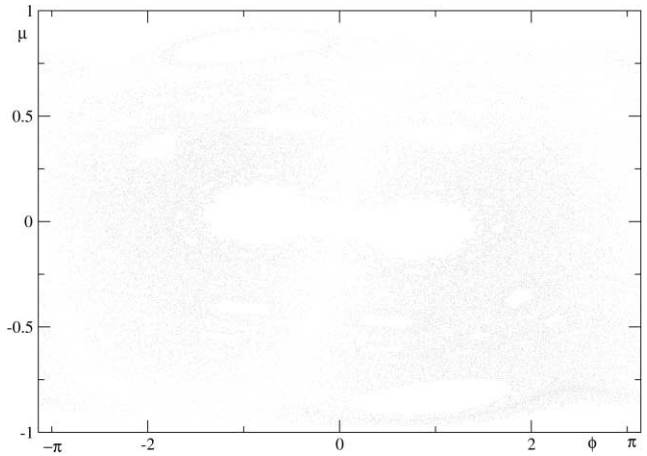

FIG. 5.—Same as Fig. 4, but for the narrow shock transitions, $\nu \approx 0$. In contrast to Fig. 4, however, the entire phase portrait is formed by a single particle orbit that does not visit the "holes" in the phase space.

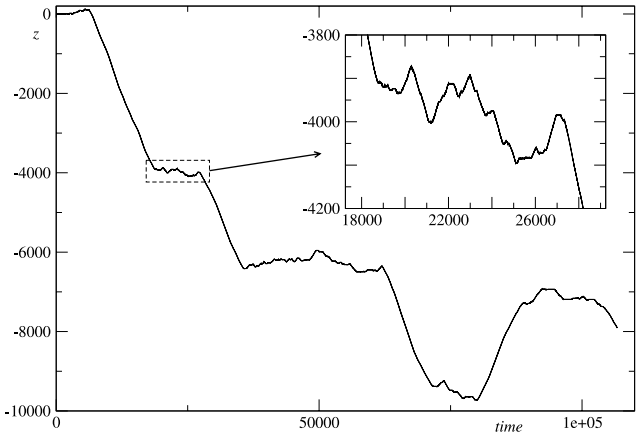

FIG. 6.—Particle trajectory represented as $z(t)$ that corresponds to the phase space shown in Fig. 5. It starts close to the islands near $\mu = 0$ so that the orbit stays close to the origin ($z = 0$) for a long time. Such long rests will be repeated many times at different locations. Overall, particle transport is organized in clusters (five such clusters are shown) where the transport is strongly suppressed. The clusters are connected with long jumps where the transport is ballistic.

a small-amplitude (coherent) shock train provides a substantially different confinement of particles than the small-amplitude random ensemble of Alfvén waves does. It ensures perfect confinement of trapped particles and no confinement at all of untrapped particles. This is what is required for the acceleration mechanism outlined in § 2. When this confinement regime takes over above some $p = p_*$, since there are no resonant waves of the length larger than $r_g(p_*)$, the energy spectrum must become steeper for $p > p_*$. We come back to this point in § 6.

The dynamics changes even more dramatically when the thickness of the shocks in the shock train becomes small compared to the particle gyroradius. Taking the same parameters in equation (9) as those used in Figure 4 except for $\nu \approx 0$, we show in Figure 5 the Poincaré section created by a single trajectory. It should be noted, however, that the actual amplitude of the shocks also increases somewhat in this case, even though the amplitude parameter $\tilde{B}_y$ is fixed. The phase space is now a "stochastic sea" with embedded "islands" of quasi-regular motion, quite typical for the deterministically chaotic systems (see, e.g., Lichtenberg & Lieberman 1983). The dynamics inside the islands is weakly stochastic, so that the particle trajectories there are closer to the nearly integrable case, shown in Figure 4, than in the rest of the phase plane. Therefore, within those islands where the averaged pitch angle $\overline{\mu} \neq 0$, particles propagate ballistically in the z-direction, similar to the untrapped particles in the previous example. There are also islands where $\overline{\mu} \approx 0$, and particle propagation is substantially suppressed there. The islands are stochastic attractors, and the mapping exhibits cycling around them apart from the quasi-periodic motion inside islands, not shown in Figure 5. Note that only a single trajectory is shown. The first type of island ($\overline{\mu} \neq 0$) is responsible for Levy flights (ballistic mode of particle propagation), whereas the second type of island ($\overline{\mu} = 0$) represents long rests or traps. The rest of the phase plane is covered by the region of a global stochasticity where particle propagation appears to resemble that around the islands but with shorter rests and Levy flights. Note that we use the term Levy flights for the above transport events even though the probability distribution function of the jump lengths might have a finite first moment (mean jump length). We investigate particle propagation in the z-direction in more detail in § 4.

4. PARTICLE TRANSPORT

It is useful at this point to recall that in the standard diffusive shock acceleration the pitch-angle diffusion determines the spatial transport, which is also diffusive. This picture results from the scattering of particles by Alfvén waves of small amplitude and random phases. By contrast, in § 3 we have considered a coherent magnetic structure of the finite amplitude shocks, as an alternative particle scattering agent at momenta higher than the pressure-containing momentum scale p_*. Despite the coherence, the particle dynamics, as we have seen, remains strongly chaotic, and the onset of stochasticity depends on both the amplitude of the shocks and their thickness. We thus appeal here to the intrinsic stochasticity caused by unpredictable particle motion in a field that can be perfectly regular. The fact that in a real situation the field is also random is often assumed to be of secondary importance.

The resulting spatial transport is not a "classic" diffusion process. The connection between the particle dynamics, represented by the Poincaré section shown in Figure 5 and the particle spatial transport can be most easily understood by considering equation (11), $\dot{z} = \mu$, as a stochastic differential equation, in which μ is a random process. Some general characteristics of this process can be inferred from the Poincaré section shown in Figure 5. As we mentioned above, there are traps represented by islands that translate into long flights or long rests depending on whether the averaged value of μ, i.e., $\overline{\mu}$ is nonzero or zero, respectively. Here the averaging is to be taken over an island attractor that is a layer of enhanced phase density around the island.

The particle transport can be conveniently treated as a random walk on a lattice of shocks located at the integer values of z. However, in contrast to the classical random walk, this is a non-Markovian process since it is characterized by long rests (Zaslavsky 2002). Particles interact with the same shock repeatedly while they are trapped near an island with $\overline{\mu} = 0$. Similarly, they perform a long jump in one direction and cross many shocks in a row while they are trapped in the phase space around an island with $\overline{\mu} \neq 0$.

The particle trajectory that corresponds to the Poincaré section shown in Figure 5 is demonstrated in Figure 6. The trajectory is indeed nondiffusive. It consists of long "Levy flights"' that connect areas (clusters) of a rather slow particle propagation. Such clustered propagation regime is quite typical for nonlocal (fractional) diffusion models (see, e.g., Metzler & Klafter 2000). The propagation within a typical cluster is magnified and shown in the inset of Figure 6. A certain similarity with the dynamics at

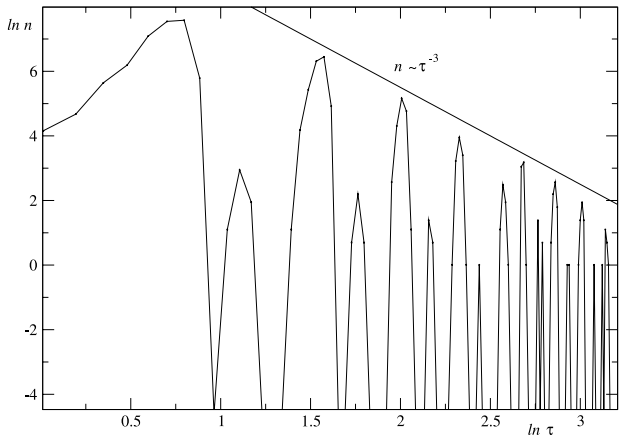

FIG. 7.—Waiting time distribution in a log-log format. Here n is the number of times the orbit has spent time τ at any shock in the shock train. A somewhat discrete character of distribution is related to the gyromotion of particles near a shock so that when they interact with the shock repeatedly, the interaction time is commensurate with the gyroperiod. [*See the electronic edition of the Journal for a color version of this figure.*]

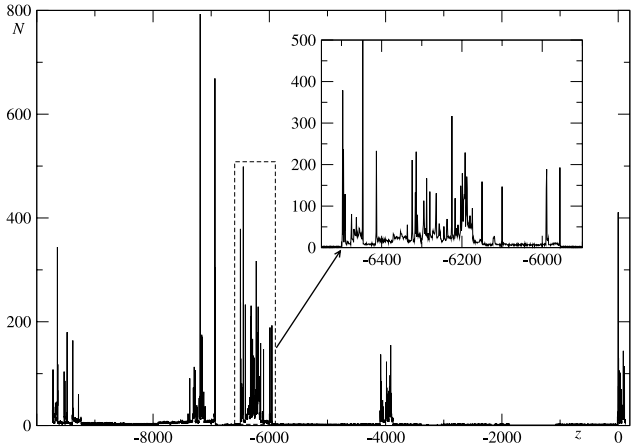

FIG. 9.—Distribution of number of visits of different shocks in the shock train. The magnified box shows the distribution inside of one of the clusters (see text).

large can be noted, which is also not unusual for the deterministically chaotic dynamical systems. However, the intracluster Levy flights are significantly shorter than the intercluster ones, an effect that is shown to be much stronger for stronger shocks. Besides, long rests at different shocks can be seen inside the clusters. Note that it is this intracluster particle dynamics that ensures their confinement to a localized region of the flow convected across the CRP as described in § 2. As we argued there, every such passage through the CRP approximately doubles the particle momentum. The Levy flights, on the other hand, can result in the escape of a particle from the shock precursor. We quantitatively describe and compare both possibilities in § 6.2.

Let us come back to the overall dynamics. There are two important characteristics that describe the random walk process. These are the waiting time and the jump length pdf. We have obtained these quantities from the trajectory shown in Figure 6. The waiting time distribution at different shocks is shown in Figure 7. Due to the gyromotion of particles near the shocks, the pdf is not smooth, but it has a smooth envelope that decays approximately as τ^{-3}, where τ is the waiting time. The distribution of the lengths of the Levy flights is presented in Figure 8. Because of the asymmetry of the shock train (Fig. 3), there is a distinct, directional asymmetry of the jumps. Otherwise, they exhibit an approximate power-law distribution up to the jumps having lengths limited by about 10. The pdf describes both the intracluster and intercluster particle jumps, but only the jumps of the first type are suitable for the continuous acceleration. The longer jumps, as we estimate in § 5, generally result in the loss of particles. Another way of representing the trajectory clustering is illustrated in Figure 9, where the distribution of numbers of visits of different sites in the shock train is shown.

The knowledge of the waiting time and jump distributions is necessary to derive the form of the adequate operators in the transport equation, (e.g., Metzler & Klafter 2000). Due to the non-Markovian character of transport, these operators are integral (also called fractional differential operators, due to the memory in stochastic trajectories) rather than conventional diffusion operators. In § 6 we shall take a somewhat simpler approach, which, however, also requires the flight–rest-time pdf to determine the transport coefficients.

Up to now we have numerically examined the particle transport in shock trains of a period longer or equal to the particles' gyroradii. This corresponds to the resonant wave-particle interaction in the small-amplitude limit. Obviously, even in the case $p \ll 1$, there are resonant Fourier components in the shock train, since its spectrum decays as k^{-2} for $k > 2\pi$ (recall that the shock train period is set to unity). This is illustrated by Figure 10, where

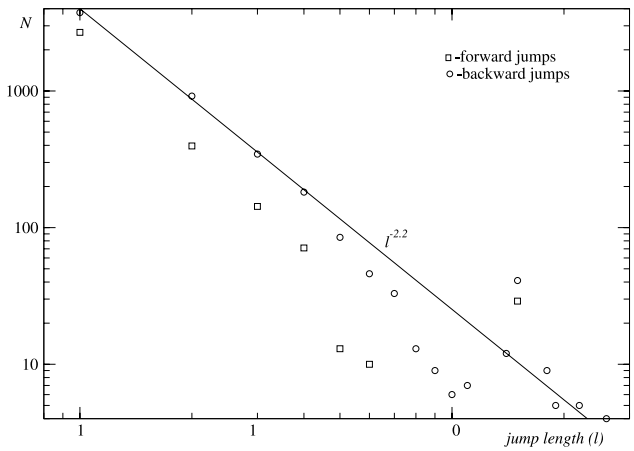

FIG. 8.—Distribution of Levy flights shown for forward and backward jumps. [*See the electronic edition of the Journal for a color version of this figure.*]

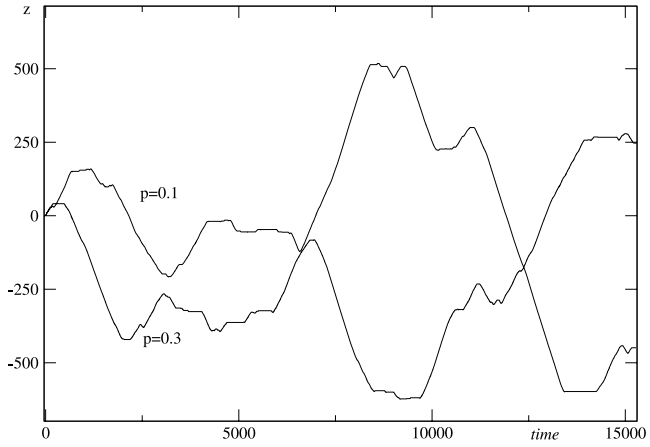

FIG. 10.—Same as Fig. 6, but for smaller particle momenta, $p = 0.1$ and 0.3. [*See the electronic edition of the Journal for a color version of this figure.*]

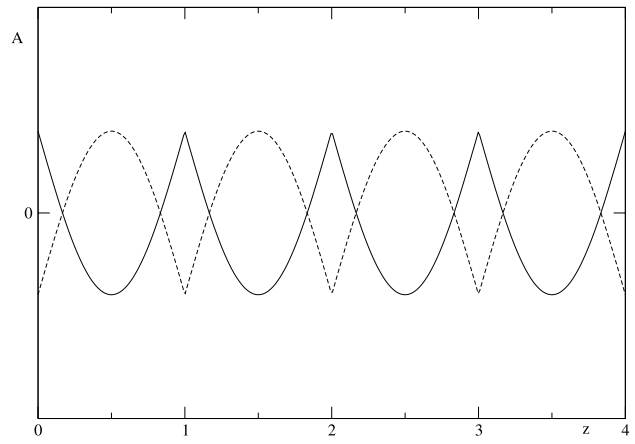

FIG. 11.—Potential $A(z)\cos(\phi)$ in eq. (15), corresponding to the magnetic field $B_y = -\partial A/\partial z$, given by eq. (9) with $\nu = 0$ and $\overline{B}_y = 0$. The function $A(z) = (\tilde{B}_y/\pi)[|\cos \pi(z - 1/2)| - 1/2]$. *Solid line*: $\cos(\phi) = -1$; *dashed line*: $\cos(\phi) = 1$.

particle trajectories with smaller momenta are shown. It should also be noted, however, that particles with smaller gyroradii, such as $p \lesssim p_*$, can be efficiently confined due to the resonance with self-generated waves since, as we argued earlier, their population is significantly more abundant than those with $p > p_*$ because of the steeper spectrum in this momentum range. This result is confirmed later. It is clear that, no matter how long the period of the shock train is, we need to estimate the confinement of particles with a gyroradius larger than that period, i.e., those with $p > 1$. One can expect that their confinement will progressively deteriorate due to the fact that there is only a high-frequency force acting on these particles. To estimate the fraction of $p \gg 1$ particles that can be confined, we write equation (10) in the form of the following system of equations:

$$\dot{\phi} = -\frac{1}{p} + \frac{\tilde{B}_y(z)\mu}{p\sqrt{1-\mu^2}}\sin(\phi), \tag{12}$$

$$\dot{\mu} = \frac{\tilde{B}_y(z)}{p}\sqrt{1-\mu^2}\cos(\phi), \tag{13}$$

$$p = \text{constant},$$

where we have assumed for simplicity that $\overline{B}_y = 0$. Because $\tilde{B}_y(z)$ has zero average, ϕ and μ are the particle gyrophase and the pitch angle, respectively, with respect to the shock normal. Assuming $\mu \ll 1$ (which we verify below), from equations (11) and (13), we then obtain

$$\ddot{z} - \frac{1}{p}\tilde{B}_y(z)\cos(\phi) = 0. \tag{14}$$

Considering ϕ as slowly varying compared to z and μ, we can integrate the last equation once:

$$\frac{p}{2}\dot{z}^2 + A(z)\cos(\phi) = \text{constant} + O(p^{-1}), \tag{15}$$

where we have introduced a periodic function $A(z)$ according to $\tilde{B}_y(z) = -\partial A/\partial z$. Since ϕ is a slowly varying variable (it varies on a timescale $p \gg 1$, while z and μ vary on a timescale $\sqrt{p}$), equation (15) describes an oscillator lattice with a slowly varying potential as shown in Figure 11. The factor $\cos(\phi)$ in equation (15) slowly inverts the lattice potential. However, trapped particles do not detrap completely but rather get trapped in one

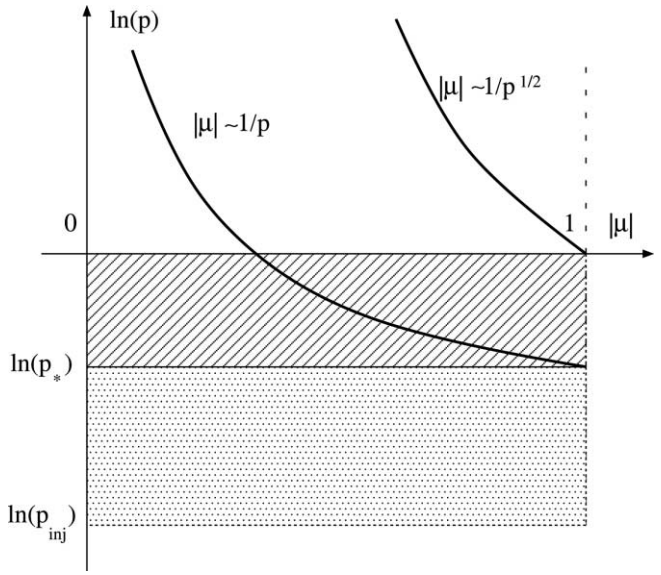

FIG. 12.—Phase space of particles in pitch-angle–momentum representation. The shaded area ($p_{\rm inj} < p < p_*$, $|\mu| < 1$) corresponds to the conventional particle confinement via randomly phased Alfvén waves. For higher momenta, $p > p_*$ [if there were only weakly turbulent waves present with minimum wavenumber, $k = k_{\rm min} = 1/r_g(p_*)$], this type of particle confinement would be limited to $|\mu| < \mu_c \propto 1/p$ because of the resonance condition $kr_g(p)\mu = 1$. A shock train with the spacing $r_g(p_*) < L < r_g(1)$ confines particles similar to the mirroring-type confinement $|\mu| < \mu_c = (\Delta B/B)^{1/2}$ (*hatched area*), except for the phase-space fragmentation (see Fig. 5 and text). Beyond $p = 1$ the particle confinement deteriorates to $|\mu| < \mu_s \propto 1/\sqrt{p}$ (see text).

of the nearby potential wells. The width of the trapping zone in $\dot{z}$ and thus in μ is given by $|\mu| < \mu_s$, where

$$\mu_s = \sqrt{\frac{2}{p}(A_{\rm max} - A_{\rm min})} \ll 1.$$

For $\tilde{B}_y(z)$ given by equation (9), we obtain $A_{\rm min} = -A_{\rm max}$ and $A_{\rm max} = \tilde{B}_y/2\pi$. Particles with $|\mu| > \mu_s$ are not confined to the shock train and propagate ballistically.

To conclude this section, we emphasize that when the momentum of particles grows and their Larmor radius increases beyond the distance between the shocks, the fraction of particles that *do not* escape ballistically shrinks with momentum as $\mu_s \propto 1/\sqrt{p}$ (Fig. 12). This is slower than in the standard, weakly turbulent picture, in which the critical value of μ decays as $1/p$. On the other hand, the latter dependence is based on the linear cyclotron resonance without broadening, while the resonance broadening also improves the confinement (Achterberg 1981). In any case, one has to expect the spectral cutoff at $p \sim 1$ in units used in this section (the Larmor radius is of order the shock spacing). Therefore, the characteristic distance between shocks determines the maximum momentum of accelerated particles. Of course, in a real situation there is a continuous strength/distance distribution of shocks that should produce a smoother spectrum decay at the cutoff momentum. Here we merely estimate the maximum distance that is equivalent to the minimum wavenumber in the standard acceleration picture. In the framework of the acceleration mechanism considered in this paper, the minimum wavenumber of the randomly phased Alfvén waves should be of the order of $r_g^{-1}(p_*)$, while particles with the momenta $p_* < p < 1$ are confined via the mechanism of interaction with the shock train. In § 5 we estimate the distance L between the shocks and thus $p_{\rm max}$, since $r_g(p_{\rm max}) \simeq L$, again, similarly to the standard estimate $r_g(p_{\rm max}) \simeq k_{\rm min}^{-1}$.

5. ESTIMATE OF THE MAXIMUM MOMENTUM

In this section we estimate the maximum distance between the shocks in the shock train, which, according to § 4, is directly related to the maximum momentum. In what follows we assume that the shocks are formed due to the development of an acoustic instability driven by the CR pressure gradient in the CRP. The latter has the scale height L_p as being created by particles accelerated in a standard DSA manner and having maximum momentum p_*. Specifically, $L_p \simeq \kappa(p_*)/U_1$. As was demonstrated by Drury & Falle (1986), the linear growth rate of this instability can be written as

$$\gamma_{\rm D}^{\pm} = -\frac{\gamma_c P_c}{\rho\kappa} \pm \frac{P_c}{C_s\rho}\left(1 + \frac{\partial \ln \kappa}{\partial \ln \rho}\right). \tag{16}$$

Here P_c and P_c are the CR pressure and its gradient (which is antiparallel to the shock normal), respectively; γ_c is their adiabatic index; ρ is the plasma mass density; and C_s is the sound speed. For efficiently accelerating shocks, one can assume $\gamma_c \approx 4/3$. The first term in equation (16) represents the wave damping caused by CR diffusion, calculated earlier by Ptuskin (1981). The second term is positive for the waves propagating along the pressure gradient, while the oppositely propagating waves are damped. The factor in the parentheses can reduce or even completely eliminate the instability if $\partial \ln \kappa/\partial\rho = -1$. However, there are no physical grounds on which this particular selection should be made (Drury & Falle 1986). It should also be noted here that particle diffusivity κ refers here to particles with $p \lesssim p_*$, while the transport of higher energy particles is nondiffusive.

Within the approximation leading to the growth rate given by equation (16), there is no dependence on the wavenumber. A more thorough investigation can be found in Kang et al. (1992), which shows that the wavenumber dependence is indeed not strong. Under these circumstances it is reasonable to assume that the largest seed waves are the most important ones. Furthermore, we assume that the latter are related to the cyclotron instability of the Alfvén and magnetosonic waves (which can provide the compressional seed component) that have excited by the accelerated particles with momenta $p \lesssim p_*$, since particles of higher momenta have a steeper spectrum and do not contribute significantly to the growth rate of the cyclotron instability. We should focus primarily on the farthest part of the precursor, since perturbations starting to grow there have the best chances to develop fully while convected with the flow toward the subshock. This part of the precursor is accessible only to the particles near the maximum momentum achievable within the standard acceleration scenario, i.e., p_*. Hence, we assume that the typical wavelength of the Drury instability is $\lambda_{\rm D} \sim r_g(p_*)$.

While these perturbations grow to a nonlinear level and propagate with the flow toward the subshock, they also steepen into the shocks. This behavior is seen in both two-fluid simulations of Drury & Falle (1986) and, what is particularly relevant to our study, in the kinetic simulations of Kang et al. (1992). Due to the significant difficulties in the numerical realization of the kinetic model, this study is limited in the maximum particle momentum, and thus in the length of the CRP, L_p. Besides that, a monochromatic wave has been chosen as a seed for the instability. As a result, no significant interaction between the shocks has been observed. In a more realistic situation with a much longer precursor and shocks of varying amplitudes, strong interactions between shocks are to be expected (e.g., Gurbatov et al. 1991). In particular, stronger shocks overtake weaker ones and absorb them, so that the interaction has a character of inelastic collisions. Time asymptotically, such decaying shock turbulence is characterized by decreasing spatial density of the shocks.

In the context of the present study, we must consider driven, rather than decaying, turbulence. A simple version of the driven Burgers model that reveals shocks merging and creation of new shocks has been considered by Malkov et al. (1995). In the case of driven turbulence, shocks not only merge, but also new shocks are formed in between them. The new shocks merge with their neighbors and so on. In a steady state, some statistically stationary ensemble of shocks can be assumed with a certain average distance L between the shocks and a shock strength characterized by an average Mach number M. For a systematic study of particle propagation in such a gas of shocks, the pdfs of Mach numbers and distances between shocks are clearly required (Bykov & Toptygin 1981; Webb et al. 2003). As we mentioned above, we make here only rough estimates of these quantities. In any event, for the generation of a statistically stationary shock ensemble, the instability should act faster than the convection of a fluid element across the precursor does, which requires $\gamma_{\rm D} L_p/U_1 \gg 1$. The latter condition can be transformed to the following,

$$\frac{P_c}{\rho U_1 C_s} \gg 1. \tag{17}$$

Note that the last condition does not contain the precursor scale L_p and can also be represented as $\epsilon M \gg 1$, where $\epsilon = P_c/\rho_1 U_1^2$. With P_c referring here to its subshock value, ϵ is the acceleration efficiency, and $M = U(z)/C_s$ is an acoustic Mach number of the flow. Not surprisingly, the criterion $\epsilon M \gg 1$ differs from the condition $\delta B^2/B_0^2 \gg 1$ only by a factor of $M/M_{\rm A} = 1/\sqrt{\beta}$, which follows from equation (1). Indeed, both instabilities share the same source of free energy, the CR pressure gradient.

Now that we can assume that the Drury instability has sufficient time to fully develop to a nonlinear level, we need to examine whether the nonlinear stage of the wave steepening lasts longer than the linear one. For the wave breaking time being shorter than the linear growth, it is sufficient to fulfill $\gamma_{\rm D} < \Delta v/L$, where Δv is the characteristic magnitude of the velocity jump across the shocks and L is the characteristic distance between them. The last condition can be rewritten as

$$\frac{L}{L_p} < \frac{\Delta v}{U_1}\frac{1}{\epsilon M}. \tag{18}$$

Under these circumstances, in a steady state shock merging must be equilibrated with their growth due to the instability, i.e., $\gamma_{\rm D} \sim \delta v/L$, where δv is the characteristic variation of the shock strength, i.e., the rms difference between neighboring values of Δv. The latter condition can be written as

$$\frac{L}{L_p} \sim \frac{\delta v}{U_1}\frac{1}{\epsilon M}. \tag{19}$$

Among the three parameters entering the last expression, $\delta v/U_1$, ϵ, and M, only the local Mach number $M(z)$ needs special attention. Indeed, the far upstream M is considered to be prescribed and may be very large, but because of the multiple shock formation in the CRP, strong heating should occur and drive this parameter down considerably.

Let us estimate the local Mach number M. There are two major factors determining this parameter. After a fluid element passes through a shock in the shock train, it gets heated, but it cools adiabatically before it crosses the next shock. For this rough estimate we can assume that all shocks are of the same compression ratio

and that the overall flow pattern is L-periodic. For this order-of-magnitude type of estimate, we neglect the gradual variation of the flow at the scale $L_p > L$. We thus assume that the flow speed in front of each shock is v_1 and is equal to $v_2 = v_1/r$ behind it, where r is the shock compression ratio. Similarly, the densities are related through $\rho_2 = r\rho_1$. We estimate the shock heating from the standard Rankine-Hugoniot relations. Let us denote the gas pressure in front of a shock in the shock train as P_{g1}. Then, the gas pressure behind it is

$$P_{g2} = \frac{2\gamma M^2 - \gamma + 1}{\gamma + 1} P_{g1}.$$

The subsequent decompression phase decreases the pressure so that it becomes equal to $P'_{g1} = r^{-\gamma} P_{g2}$ in front of the next shock. Combining this with the above equation, we obtain

$$\frac{P'_{g1}}{P_{g1}} = \left(\frac{\gamma - 1 + 2/M^2}{\gamma + 1}\right)^{\gamma} \frac{2\gamma M^2 - \gamma + 1}{\gamma + 1}.$$

Therefore, the total pressure change after passing trough the shock and the smooth flow behind it is zero only for $M = 1$. However, the pressure does not really change for a broad range of Mach numbers around $M = 1$ (where the right-hand side of the last expression is close to unity). This result means that M cannot be large, since it would lead to strong plasma heating driving its temperature to a state with $M \simeq 1$. However, in this situation the calculation of heating using the Rankine-Hugoniot relations is not correct, since shocks are likely to be subcritical. Note that in our estimates, we also ignored the increase of magnetic energy behind the shocks. Since the above estimates show that shocks cannot be highly supercritical also, we conclude that a reasonable estimate for M to substitute in equation (19) is a critical value of the Mach number, i.e., $M \simeq M_*$, which is of order a few ($M_* \simeq 3$, for example), depending on plasma β and ϑ_{nB}, the angle between the shock normal and the magnetic field (see Sagdeev 1966; Papadopoulos 1985, for a review of collisionless shock physics).

The efficiency ϵ may then be calculated once the heating rate is known (see, e.g., Malkov & Drury 2001). We note that the effect of the precursor heating on the shock structure and thus on ϵ is very sharp for a certain range of the heating rate parameter. The latter is difficult to quantify, so that we assume here that it is below the critical value at which a strong effect on the shock structure is expected. As for the parameter $\delta v/U_1$, the first obvious constraint is $\delta v < \Delta v$, which obviously verifies the consistency of equations (19) and (18). In fact, based on the studies of the driven Burgers turbulence (Cheklov & Yakhot 1995; Gotoh & Kraichnan 1998), we can assume that the shock strength variance δv and its mean Δv are related by $\delta v \lesssim \Delta v$.[2] At the same time Δv may be assumed to be not so much weaker than the subshock itself (Drury & Falle 1986; Kang et al. 1992). The subshock strength can be calculated from nonlinear acceleration theory (Malkov 1997; Blasi 2002; Blasi et al. 2005), along with the acceleration efficiency. Given the uncertainty of the heating rate, one can estimate $\epsilon \sim \frac{1}{5}$–$\frac{1}{3}$, so that the estimate $L/L_p \sim \frac{1}{5}$ does not seem to be totally unreasonable.

As follows from § 4, the maximum particle momentum can be estimated from the relation $r_g(p_{\max}) \sim L$. Bearing in mind that $L_p \simeq r_g(p_*)c/U_1$, the maximum momentum $p_{\max}$ can be estimated as

$$\frac{p_{\max}}{p_*} \simeq \frac{c}{U_1}\frac{L}{L_p}.$$

[2] This rough estimate should be taken with caution. Both quantities are really distributions, often with a power-law pdf, and may not even posses finite means.

Since L/L_p, as we argued, can be not very small and c/U_1 is a large parameter, the suggested mechanism can produce a significant additional acceleration beyond the break momentum at $p = p_*$. This stage of acceleration lasts for approximately $\tau_{\rm acc}(p_{\max}) \sim \left[\kappa(p_*)/U_1^2\right] \ln(p_{\max}/p_*)$.

6. PARTICLE SPECTRUM BETWEEN THE BREAK AND MAXIMUM MOMENTUM

Up to now we have calculated the particle energy gain by their scattering on unspecified scattering centers carried by the converging flow in the CRP. Independently of that, we then considered the scattering of particles on an ensemble of shocks. The latter process did not lead to energy gain, since no relative motion between the scatterers was assumed. Recall that we considered the simple shock ensemble as a magnetic structure traveling at a constant speed. Nevertheless, each shock in the shock train may possess its own flow structure with upstream and downstream regions and thus can, in principle, accelerate particles via the standard DSA mechanism. However, as we pointed out earlier, these shocks do not have sufficient resonant turbulence upstream and downstream with $k < 1/r_g(p_*)$. Therefore, there is no sufficient coupling of particles with momenta $p > p_*$ to the converging flow around each such shock in order to gain momentum.

In the situation we consider further in this section the momentum gain results from the combination of particle scattering off the scattering centers (shocks, localized magnetic structures) and gradual flow compression, leading to the convergence of the centers. In fact this is similar to what we have considered in § 2 at an elementary level. An essential difference is that the scattering in momentum space is not homogeneous as the results of § 3 suggest. Therefore, we relax the assumption, also made in § 2, that particle transport in momentum space is diffusion that evenly covers isoenergetic surfaces. This affects the formula for the momentum gain.

6.1. *Particle Momentum Gain*

To calculate the rate at which particles gain momentum, let us write down the equations of particle motion in the subshock reference frame and apply them in the area ahead of the subshock where the velocity of the flow is $\boldsymbol{U}(z, t) = U(z, t)\boldsymbol{e}_z$ (with the z-axis pointing in the shock normal direction, $\boldsymbol{e}_z$),

$$\frac{d\boldsymbol{p}}{dt} = \frac{e}{c}(\boldsymbol{v} - \boldsymbol{U}) \times \boldsymbol{B}. \tag{20}$$

In this approximation, we have replaced the electric field by $\boldsymbol{E} = -c^{-1}\boldsymbol{U} \times \boldsymbol{B}$, assuming that due to the perfect conductivity the electric field vanishes in the local plasma frame. In contrast to § 3, we do not specify the distance between the scattering centers here, so it is convenient to normalize the length to $c/\omega_{\rm ci}$, time to $\omega_{\rm ci}^{-1}$, p to mc, and v and $U(z)$ to c.

Introducing the cosine of the particle pitch angle to the shock normal, μ, as well as the particle gyrophase, ϕ, as in § 4, it is useful to rewrite the equations of motion given by equation (20) as

$$\dot{p} = \sqrt{1 - \mu^2}\, U \,\mathrm{Im}\left(be^{-i\phi}\right), \tag{21}$$

$$\dot{\mu} = \frac{1 - \mu U}{pU}\dot{p}, \tag{22}$$

$$\dot{\phi} = -\frac{1}{p} + \frac{\mu - U}{\sqrt{1 - \mu^2}}\mathrm{Re}\left(be^{-i\phi}\right), \tag{23}$$

$$\dot{z} = \mu. \tag{24}$$

We have introduced the following complex variable instead of the transverse component of the magnetic field by $b = (B_x + iB_y)/B_0$. The above equations are written in the shock frame. Now we need to transform particle momentum p to the local plasma frame, i.e., the frame moving with respect to the shock with the velocity $U(z, t)$, since it is this momentum the convection-diffusion equation [eq. (7)] refers to. This transformation can be written in the approximation $U \ll 1$ as $p' = (1 - U\mu)p$. Now we differentiate the both sides of the last formula with respect to time, taking into account that to this order of approximation $dU/dt \approx \dot{z}\partial U/\partial z$. Making also use of equations (22) and (24), for the acceleration rate in the local plasma frame we obtain

$$\dot{p}' = -\mu^2 \frac{\partial U}{\partial z} p'. \tag{25}$$

Recall that the precursor scale height L_p that determines $\partial U/\partial z$ should be calculated using the spectral break momentum at p_*, i.e., $L_p = \kappa(p_*)/U_1$. We also note that the acceleration rate in equation (25) does not contain the dynamical variable ϕ and is actually almost independent of z. This is because U, to a good approximation, can be considered as a linear function of z in a significant part of the shock precursor (in the case of the Bohm diffusion). As we mentioned earlier, due to the large gyroradius, we can neglect any small-scale variation of U. This includes the structure of individual shocks in the shock train, so that only the gradual variation of U across the CRP is important. In addition, we are really interested in the averaged value of the acceleration rate $\dot{p}'/p'$. For example, if the particle transport in pitch angle is a small step diffusion, μ can be regarded as a random variable evenly distributed over the interval $(-1, 1)$. In this simple case, the average $\overline{\mu^2} = \frac{1}{3}$, and we recover the standard acceleration rate given by equation (6). In the case of structured phase space considered earlier in § 3, one can assume that during most of an acceleration cycle, a particle is bound to the island having $\overline{\mu} = 0$. If $\overline{\mu} \neq 0$, the particle performs the Levy flights, which we discuss below. In the case of small islands (i.e., with the width $\mu_0 \ll 1$), we can replace $\overline{\mu^2}$ by $\overline{\mu}^2$ (for $\overline{\mu} \neq 0$) and by $\mu_0^2/3$ (for $\overline{\mu} = 0$).

The situation becomes somewhat more complicated in the case of an oblique shock, for which the phase space shown, for example, in Figure 3, implies that the cosine of the pitch angle $\tilde{\mu}$ is to be measured with respect to the average magnetic field and not with respect to the shock normal as in equation (25). The transformation to the angles $\tilde{\mu}$ and $\tilde{\phi}$, i.e., to the reference frame having the z-axis aligned with the averaged magnetic field, is given by

$$\mu = \tilde{\mu}\cos\vartheta - \sqrt{1 - \tilde{\mu}^2}\sin\vartheta\sin\tilde{\phi}, \tag{26}$$

where $\vartheta = \tan^{-1}\overline{B}_y$ (see § 3). Therefore, the acceleration rate given by equation (25) does include the gyrophase $\tilde{\phi}$. However, the dynamics on the $(\tilde{\phi}, \tilde{\mu})$ plane in the case of trapping we are interested in here, and that is exemplified in Figure 5, is mostly rotation around the origin, so that one can expect that $\tilde{\mu}\sin\tilde{\phi} \approx 0$. Thus, the acceleration rate can be represented as

$$\frac{\dot{p}}{p} \approx -K\frac{\partial U}{\partial z}, \tag{27}$$

where

$$K = \left[\frac{1}{3}\mu_0^2\cos^2\vartheta + \frac{1}{2}\left(1 - \frac{1}{3}\mu_0^2\right)\sin^2\vartheta\right] \tag{28}$$

and μ_0 denotes the width of the island in which the particle orbit is trapped. We also omitted the primes in the momentum p.

Using this formula it is easy to calculate particle momentum gain assuming that the particle is trapped by the flow at the distance z from the subshock in the CRP where the local flow speed is $U(z)$ and is convected with the flow to the subshock where the flow speed is $-U_0$. Suppose that at the moment t and coordinate z the particle has the momentum p_0. From equation (27) we have for the particle momentum at $z = 0$ where $U = -U_0$,

$$p = p_0\exp\left[-K\int_t^{t_0}\frac{\partial U}{\partial z}\,dt\right] = p_0\exp\left[K\ln\frac{U(z)}{-U_0}\right] = p_0\left(\frac{U}{-U_0}\right)^K. \tag{29}$$

Note that in the conventional diffusive acceleration, in which particles ergodically cover the entire isoenergetic surface, i.e., $\mu_0 = 1$, the acceleration constant $K = \frac{1}{3}$ (independent of ϑ), and the last result signifies the effect of adiabatic compression of the CR gas by the converging flow. On the other hand, if the particle dynamics is such that $\mu_0 \approx 0$ and $\vartheta \approx \pi/2$, which corresponds to two-dimensional compression across the magnetic field, one obtains $K \approx \frac{1}{2}$. In strong CR-modified shocks, the shock precompression $R = U_1/U_0$ scales as $M^{3/4}$ (Kazanas & Ellison 1986; Berezhko et al. 1996). As was shown by Malkov (1997), the latter scaling is only valid for $M < M_* \equiv \left(\nu_{\rm inj}\, p_*/p_{\rm inj}\right)^{4/3}$, and the shock precompression R saturates at $R \sim \nu_{\rm inj}\, p_*/p_{\rm inj}$ when $M > M_*$ (see also Blasi et al. 2005). Here the injection parameter $\nu_{\rm inj} \simeq \left(cp_{\rm inj}/mU_1^2\right)n_c/n_1$ and $p_{\rm inj}$ is the injection momentum, while n_c and n_1 are the number density of accelerated particles and that of the background plasma far upstream, respectively. In both Mach number ranges the shock precompression can be quite significant unless the gas heating in the CRP is strong. At the same time, strong reduction of the shock precompression by plasma heating would diminish the heating itself by weakening the acoustic (Drury) instability of the CRP. The situation may settle at some critical level at which a moderate, but still quite significant, precompression is accompanied by the CRP heating caused by the acoustic instability (Malkov et al. 2000; Malkov & Drury 2001). Further interesting analysis of the nonlinear shock acceleration and its bifurcation has been published recently by Blasi et al. (2005).

6.2. *Acceleration Model: Details*

The picture that emerges from the above study of particle dynamics in the modified shock precursor can be described as follows. There are three groups of particles. One group is made up by particles that are locally trapped in the plasma flow by a structure in the shock train. They are convected with the flow and either do not propagate at all or propagate very slowly with respect to the flow. They are clearly seen in, e.g., Figure 10, as the quasi-horizontal portions of the stochastic trajectory of a single particle. The remaining two groups of particles are particles performing Levy flights, i.e., those propagating ballistically in positive and negative directions at the speeds U_+ and U_-, respectively. These are characterized by the steep portions of the trajectory in Figure 10, with positive and negative slopes, respectively. Let us denote the phase-space density of the above three groups of particles by $F_0(t, z, p)$ and $F_\pm(t, z, p)$. Furthermore, we introduce the probabilities of transition in units of time by denoting by $\alpha_\pm$ the probability of trapping of particles performing Levy flights in positive and negative directions, respectively. We denote the rates of the reverse transitions by $\beta_\pm$. Now we can write the balance of

particles in a phase-space cell between z_1, p_1 and $z_2 = z_1 + dz$, $p_2 = p_1 + dp$ as follows:

$$\frac{\partial}{\partial t}F_0(t,z,p)dzdp = \left(F_0U|_{z_1} - F_0U|_{z_2}\right)dp + \left(\dot{p}F_0|_{p_1} - \dot{p}F_0|_{p_2}\right)dz - (\beta_+ + \beta_-)F_0\,dzdp + (\alpha_+F_+ + \alpha_-F_-)\,dzdp, \tag{30}$$

or after retaining only the linear terms in dz and dp and using equation (27), we obtain the following equation for the phase-space density of trapped particles F_0,

$$\frac{\partial F_0}{\partial t} + \frac{\partial}{\partial z}UF_0 - K\frac{\partial U}{\partial z}\frac{\partial}{\partial p}pF_0 = -(\beta_+ + \beta_-)F_0 + \alpha_+F_+ + \alpha_-F_-. \tag{31}$$

Here the constant K is given by equation (28). Proceeding in a similar manner, we obtain two equations for $F_\pm$, i.e., for the phase-space density of the particles in the Levy flight state,

$$\frac{\partial F_\pm}{\partial t} + \frac{\partial}{\partial z}U_\pm F_\pm - K_\pm\frac{\partial U}{\partial z}\frac{\partial}{\partial p}pF_\pm = -\alpha_\pm F_\pm + \beta_\pm F_0. \tag{32}$$

In contrast to equation (31), the acceleration constants $K_\pm$ are defined by

$$K_\pm = \left[\mu_\pm^2\cos^2\vartheta + \frac{1}{2}\left(1-\mu_\pm^2\right)\sin^2\vartheta\right], \tag{33}$$

where $\mu_\pm$ are the averaged values of the cosines of the particles' pitch angles ($\mu_\pm = \overline{\tilde{\mu}}$) performing Levy flights in the positive and negative direction, respectively, while $U_\pm = c\mu_\pm\cos\vartheta \gg U$ are their speed components along the shock normal (cf. eq. [28]). Note that we ignore direct transitions between the opposite Levy flights as comparatively rare events. This can be inferred from Figure 10, for example.

Equations (31) and (32) form a closed system that must be supplemented with the boundary conditions. Far upstream from the shock, at $z = \infty$, it is natural to impose the following boundary conditions:

$$F_0(\infty) = F_-(\infty) = 0, \tag{34}$$

which simply mean that accelerated particles can only leave the system in that direction, but do not enter from it. Note that $F_+(\infty) \neq 0$ and should be determined from equations (30) and (32). The second boundary condition, that at the subshock location $z = 0$, is somewhat more ambiguous. In the standard DSA scheme it is fixed by the flux of particles convected with the flow downstream. It can be expressed through the particle phase-space density and the downstream flow speed since the particle distribution is isotropic. In the case of high momenta ($p > p_*$) considered here, particles are not bound to the flow so strongly as to be able to make an isotropic distribution. Rather, according to equations (30) and (32), particles convected with the flow at the speed $U(z)$ are being converted into particles propagating at high speed, $U_\pm \sim c$, and can either leave the system upstream (unless they are retrapped) or can penetrate deeply downstream and reach the contact discontinuity. The latter possibility for the high-energy particles has been already discussed in the literature (Berezhko 1996; Blondin & Ellison 2001). The fact that the contact discontinuity and the forward shock may indeed be closer to each other due to the back-reaction of accelerated particles on the shock dynamics, as clearly follows from the nonlinear DSA, seems to find impressive observational confirmation in Warren et al. (2005). The strong magnetic field at the contact discontinuity may very well reflect particles with $p > p_*$, and they can return to the shock. In a simple form this requirement can be obviously formulated by mathematically setting the total particle flux through $z = 0$ equal to zero:

$$-F_0(0)U_0 + F_-(0)U_- + F_+(0)U_+ = 0. \tag{35}$$

In other words, the negative part of the total particle flux, represented by the first two terms ($U_- < 0$), is reflected by the strong magnetic turbulence downstream and ultimately reappears at the shock in the form of the untrapped particles propagating upstream at the speed $U_+ > 0$. The strong turbulence downstream may be formed by the magnetic perturbation convected and compressively amplified from the upstream media and by the Raleigh-Taylor instability of the contact discontinuity (Jun et al. 1996).

The way particle momentum enters equations (31) and (32), as well as the homogeneity of the boundary conditions given by equations (34) and (35), suggest the power-law solution:

$$F_{0,\pm} \propto p^{-q}. \tag{36}$$

Assuming for simplicity $U_\pm = \text{constant}$ (which is close to the truth, as, e.g., Fig. 10 suggests), we can rewrite these equations as follows:

$$\frac{d}{dz}UF_0 = \gamma_0F_0 + \alpha_+F_+ + \alpha_-F_-, \tag{37}$$

$$U_\pm\frac{dF_\pm}{dz} = \gamma_\pm F_\pm + \beta_\pm F_0. \tag{38}$$

Here we introduce the following notation:

$$\gamma_0 = -K(q-1)\frac{\partial U}{\partial z} - \beta,$$

$$\gamma_\pm = -K_\pm(q-1)\frac{\partial U}{\partial z} - \alpha_\pm,$$

where

$$\beta = \beta_+ + \beta_-. \tag{39}$$

From equations (37) and (35), we have

$$F_0(z) = -\frac{1}{U}\int_z^\infty \exp\left(-\int_z^{z'}\frac{\gamma_0}{U}dz''\right)[\alpha_+F_+(z') + \alpha_-F_-(z')]\,dz'. \tag{40}$$

To zeroth-order approximation in $\beta L/U \ll 1$, $U/U_\pm \ll 1$, from equation (38) we have $F_+ \approx$ constant and $F_- \approx$ constant, while from the boundary condition given by equations (34) and (35), we obtain

$$F_+ = \frac{U_0}{U_+}F_0(0)$$

and $F_- \approx 0$. Upon substitution of $F_\pm$ into equation (40) and taking it at $z = 0$, we finally obtain the following equation for the spectral index q:

$$\int_0^\infty \left[\frac{U(z)}{-U_0}\right]^{K(q-1)} \exp\left(\int_0^z \frac{\beta}{U}\, dz'\right)\alpha_+(z)\frac{dz}{U_+} = 1. \quad (41)$$

In fact, the last formula allows a very simple physical interpretation, making it essentially equivalent to the well-known general Fermi (1949) result. Indeed, Fermi showed that if a group of particles undergoes continuous acceleration at the rate $\tau_{\rm acc}^{-1}$ and have certain probability to escape that can be characterized by the escape time $\tau_{\rm esc}$, then the energy [momentum distribution normalized above as $F_0(p)$] has the following spectral index:

$$q_F = 1 + \frac{\tau_{\rm acc}}{\tau_{\rm esc}}. \quad (42)$$

At least for some short period of time, by definition of these timescales, one can state that the momentum gain $p(t)/p(0) \equiv p/p_0 = \exp(t/\tau_{\rm acc})$. The probability of staying in the accelerator if escape is a Poisson process is $\mathcal{P}_{\rm conf} = \exp(-t/\tau_{\rm esc})$. Then equation (42) can be recast as (cf. Bell 1978; Blandford & Eichler 1987; Achterberg et al. 2001):

$$q_F = 1 + \frac{\ln\left(1/\mathcal{P}_{\rm conf}\right)}{\ln(p/p_0)}. \quad (43)$$

Note that t can be regarded here as the duration of an acceleration cycle and it does not enter the spectral index explicitly. It should also be clear that determination of $\mathcal{P}_{\rm conf}$ and the momentum gain (or in other words, $\tau_{\rm esc}$ and $\tau_{\rm acc}$) requires in our model more information about particle transport, in particular, the particle trapping and untrapping rates α and β.

To demonstrate that the index q in our formula (41) must be interpreted in the same way as in the classical Fermi problem, let us analyze various factors in the integrand in equation (41) separately. Assume that a particle enters an acceleration cycle at $z = 0$ as a Levy flight in a positive direction with a speed U_+ against the plasma flow of speed $U(z)$ in the shock precursor, located in a half-space $z > 0$. During its flight, the particle has a possibility to be trapped in the flow that is characterized by the rate α_+. In other words, during time dt the probability to be trapped between z and $z + dz$ equals $\alpha_+(t)dt = \alpha_+(z)dz/U_+$. We see this probability density in the integral given by equation (41). Assume that the trapping event actually occurred at some z. Then the particle is convected (with the flow) back to the subshock at the speed $U(z)$. During this convection it has a certain probability to escape, which is characterized by the rate β. This means that if the particle has a probability of being in the flow $P(t)$, then the probability to remain there at $t + dt$ is $\mathcal{P}(t + dt) = \mathcal{P}(t)(1 - \beta dt)$, or if the particle is certainly in the flow at $t = 0$ (just trapped there, for example), then $\mathcal{P}(t) = \exp\left(-\int_0^t \beta\, dt\right)$. Noting that $dt = dz/U$, we can write the probability of returning with the flow from z to $z = 0$ as

$$\exp\left(-\int_z^0 \beta\, dz/U\right).$$

Combining with the above probability of being trapped between z and $z + dz$, we can express the probability of completing an acceleration cycle of the length between z and $z + dz$ as[3]

$$d\mathcal{P}_{\rm cycl}(z) = \exp\left(\int_0^z \frac{\beta}{U}\, dz'\right)\alpha_+(z)\frac{dz}{U_+}.$$

The total probability of completing one such cycle of any possible length $0 < z < \infty$ is

$$\mathcal{P}_{\rm cycl} = \int_0^\infty \frac{d\mathcal{P}_{\rm cycl}}{dz}\, dz.$$

We can now rewrite equation (41) as follows:

$$\int_0^\infty \left[\frac{U(z)}{-U_0}\right]^{K(q-1)} \frac{d\mathcal{P}_{\rm cycl}}{dz}\, dz = 1. \quad (44)$$

Here the quantity $G(z) = |U(z)/U_0|^K$ is equal to a particle momentum gain p_f/p_i after it is convected with the flow from z to the origin (eq. [29]). Using the mean value theorem, equation (44) can be represented as

$$\left\langle \frac{p_f}{p_i} \right\rangle^{q-1} \mathcal{P}_{\rm cycl} = 1,$$

where $\langle p_f/p_i \rangle = G(\bar{z})$ is the mean momentum gain per cycle and $\mathcal{P}_{\rm cycl}$ is, as stated above, the probability of completing this acceleration cycle. From the last formula we obviously have

$$q = 1 + \frac{\ln(1/\mathcal{P}_{\rm cycl})}{\ln\langle p_f/p_i\rangle}, \quad (45)$$

which is formally equivalent to the classical Fermi result.

The quantities α_+ and β, which the index q depends on, can be inferred from the particle stochastic trajectories shown, e.g., in Figures 6 and 10. Indeed the probability for a particle to be in the trapped state can be written as $\tau_{\rm tr}/\left(\tau_{\rm tr} + \tau_{\rm L}^+ + \tau_{\rm L}^-\right)$, where $\tau_{\rm tr}$ and $\tau_{\rm L}$ are the average trapping and Levy flight times, respectively, that can be calculated from the particle trajectory. Then the particle trapping rate α_+ can be written as

$$\alpha_+ \approx \frac{\tau_{\rm tr}}{(\tau_{\rm tr} + \tau_{\rm L}^+ + \tau_{\rm L}^-)}\frac{1}{\tau_{\rm L}^+};$$

similarly,

$$\beta_\pm \approx \frac{\tau_{\rm L}^\pm}{(\tau_{\rm tr} + \tau_{\rm L}^+ + \tau_{\rm L}^-)}\frac{1}{\tau_{\rm tr}}.$$

For the case of sufficiently strong shocks in the shock precursor, the trapping process is quite efficient, so we can assume that

$$\int_0^{L_p} \frac{\beta}{U}\, dz \ll 1. \quad (46)$$

[3] Particle behavior during a Levy flight is non-Markovian, since it bears elements of deterministic motion (this is also true for the trapped particles). However, the joint probability of two such events is multiplicative, since they evolve independent of each other.

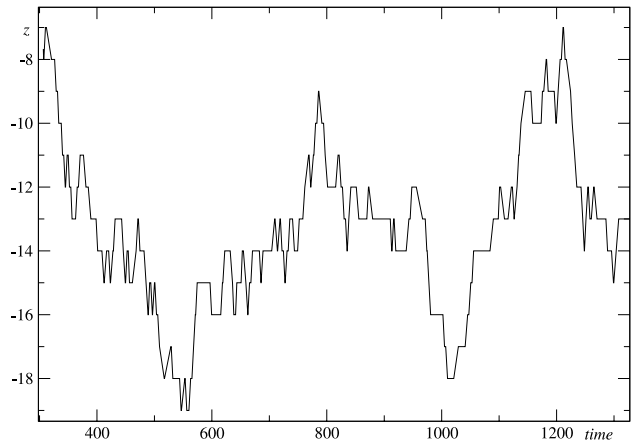

FIG. 13.—Stochastic particle trajectory represented in the same format as in Fig. 10, but for $p = 0.3$, $\tilde{B}_y = 1$, and $\tilde{B}_y = 2$.

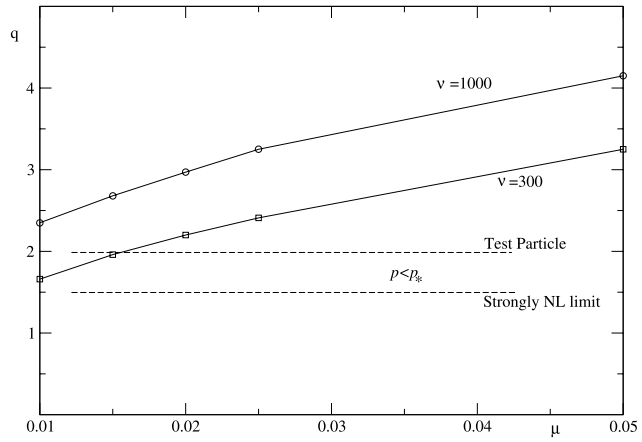

FIG. 14.—Particle power-law index q for $p > p_*$ as a function of μ for different values of ν and $K = \frac{1}{3}$ (see text). The test particle result ($q = 2$) and an index corresponding to a strongly modified shock ($q = 1.5$) are also shown.

Here the shock precursor length L_p can be related to the trapping probability as

$$\frac{L_p \sim \int_0^\infty \alpha_+(z) z\, dz}{\int_0^\infty \alpha_+(z)\, dz}.$$

The requirement for the inequality in equation (46) is then

$$\frac{\tau_p \tau_{\rm L}}{\tau_{\rm tr}^2} \ll 1, \tag{47}$$

where the precursor crossing time is

$$\tau_p = \int_0^{L_p} \frac{dz}{|U|}.$$

An example of a particle trajectory with such long trapping times is shown in Figure 13. Note that when a particle is bouncing between two to three neighboring shocks it can also be considered as "trapped" in the flow, as is the case for a particle stuck to a single shock. These kind of events are clearly seen in Figure 13. Using equation (41), we present in § 6.3 a simplified calculation of the power-law index q.

6.3. *Estimate of the Power-Law Index between the Knee and Cutoff*

The power-law index q in the momentum range beyond the knee ($p_* < p < p_{\rm max}$) essentially depends on the flow profile $U(z)$ (eq. [41]). For the Bohm-type diffusivity $\kappa(p)$ in the region $p < p_*$ [i.e., the pressure-dominated part of the spectrum, thus important for determination of $U(z)$] the flow profile has been calculated by Malkov (1997). Assuming here that the integral in equation (41) cuts off by the exponential factor rather than by $\alpha_+(z)$ at such values of z [where $U(z)$ is still not very different from its subshock value $-U_0$, i.e., $|U + U_0| \ll U_1$], one can represent $U(z)$ as follows:

$$-U(z)/U_0 \approx 1 + \sigma\sqrt{z}.$$

Here

$$\sigma \approx 2\eta \frac{U_0 - U_2}{U_0} \sqrt{p_0 p_*} \sqrt{\pi |U_0|/\kappa_0},$$

where η is the particle injection rate and p_0 and p_* are the injection and the knee momenta, respectively, in units of mc and the particle diffusivity $\kappa_0 \equiv \kappa(p_0)$ (see also Fig. 2). In the last expression for σ we omitted an additional factor that is close to unity. Equation (41) for q can be rewritten as follows:

$$2\frac{\alpha_+ \mu^{-\lambda-\mu-2}}{U_+ \sigma^2} \int_0^\infty (x+\mu)^{\lambda+\mu} e^{-x} x\, dx = 1, \tag{48}$$

where instead of the power-law index q we use

$$\lambda = K(q-1)$$

and where

$$\mu \equiv \frac{\beta \kappa_0}{2\pi \eta^2 p_0 p_* (U_0 - U_2)^2}.$$

The latter parameter contains the important quantity $\eta^2 p_0 p_*$, which was shown to be the control parameter of nonlinear shock modification (Malkov 1997). Namely, this parameter should be large in the case of strong shock modification by accelerated particles. The remaining parameter $\beta\kappa_0/(U_0 - U_2)^2$ represents the detrapping probability of a high-energy particle ($p > p_*$) over the time of crossing the diffusion length of a freshly injected particle ($p \sim p_0$). It is implied here that $U_0 - U_2 \sim U_0$. Therefore, we can assume that $\mu \ll 1$. In general, equation (48) can be represented in terms of an incomplete gamma function, but since $\mu \ll 1$, it takes the following simple form:

$$\Gamma(\lambda + 2) = \nu \mu^{\lambda+1}, \tag{49}$$

where Γ is the gamma function. Here we introduced a new parameter ν according to

$$\nu = \frac{U_+ \beta}{U_0 \alpha_+} \simeq \frac{U_+}{U_0} \frac{\tau_{\rm L} \tau_{\rm L}^+}{\tau_{\rm tr}^2} \sim \frac{c}{U_0} \gg 1.$$

Since $\mu \ll 1$, equation (49) has a solution at $\lambda \lesssim 1$. It depends on the parameters ν and μ, but taking the limit of large ν, one can apply the Stirling formula for the gamma function, which leads to the following estimate of the power-law index $q = 1 + \lambda/K$:

$$q = 1 + \frac{1}{K} \frac{\ln(\nu\mu)}{\ln(1/\mu)}.$$

The ratio of logarithms is typically below unity, while the constant K is in the range of $\frac{1}{3}$–$\frac{1}{2}$. Thus, the power-law index is (almost) definitely well above 2 (the particle pressure clearly converges), but it can hardly be significantly higher than 3 for strongly nonlinear fast shocks (see Fig. 14 for a numerical solution of eq. [49]). For a more accurate determination of the power-law index in the $p > p_*$ range, further study of particle scattering in the shock precursor is needed. Fortunately, the index q only weakly (logarithmically) depends on the most uncertain parameters of the present model, namely, on μ and ν or, equivalently, on the trapping and untrapping rates α and β.

7. SUMMARY AND CONCLUSIONS

The principal conclusion of this paper is that particle acceleration inside the CR shock precursor may very well be faster and proceed to higher energies during SNR shock evolution than the conventional DSA theory suggests. The requirement for such enhanced acceleration is some limitation on the momentum of particles that contribute the most to the CR pressure in order to prevent further inflation of the shock precursor. This can be achieved by spatial limitations of momentum growth, as discussed in detail by, e.g., Berezhko (1996), by the wave compression and blueshifting of the wave spectrum in the CRP (Malkov et al. 2002), and by the change of acceleration regime after the end of the free expansion phase (Drury et al. 2003). As a result of these limitations, a break on the particle spectrum is formed and maintained at $p = p_*$ beyond which particles do not contribute to the pressure significantly but are still accelerated at the same rate as particles with momenta $p < p_*$.

The fundamental acceleration mechanism is essentially identical to the nonlinear DSA. *The difference from the linear DSA is that frame switching that results in the energy gain occurs predominately between the scatterers convected with the gradually converging upstream flow in the precursor and not between the upstream and downstream scattering centers, as usually assumed*. The maximum momentum is estimated to be

$$p_{\max} \sim \frac{c}{U_1}\frac{L}{L_p}p_*, \tag{50}$$

with the maximum value of $L/L_p \lesssim \frac{1}{5}$. Therefore, this mechanism provides up to $\sim 0.2(c/U_1)$ enhancement to the maximum momentum, based on the standard spatial constraints (e.g., (Berezhko 1996). The spectral index between p_* and $p_{\max}$ depends on both the ratio of the particle trapping to flight time $\tau_{\rm tr}/\tau_{\rm L}$ in the upstream turbulent medium and on the shock precompression R. This is in deep contrast with the spectrum below p_*, which tends to have a form quasi-independent of R (e.g., Malkov & Drury 2001). The acceleration time is

$$\tau_{\rm acc}(p_{\max}) \sim \tau_{\rm NL}(p_*)\ln\frac{p_{\max}}{p_*}, \tag{51}$$

where $\tau_{\rm NL}(p) \simeq 4\kappa(p)/U_1^2$ is the nonlinear acceleration time, which only slightly differs from the upstream contribution to the standard linear acceleration time (Malkov & Drury 2001).

The overall temporal development of the acceleration process can be described as follows. It starts as the standard DSA mechanism from some slightly suprathermal momentum $p_{\rm inj}$ and proceeds at the rate $\tau_{\rm acc}^{-1}(p_{\max}) \sim U_1^2/\kappa(p_{\max})$, where $\kappa(p) \sim cr_g(p)$ and $r_g(p_{\max})k_{\min} \sim 1$. The minimum wavenumber $k_{\min}$ decreases with growing $p_{\max}(t)$, since the MHD waves confining particles with momenta $p_{\rm inj} < p < p_{\max}$ to the shock are resonantly generated by the accelerated particles themselves. After $p_{\max}$ reaches a certain value (that can be calculated analytically, e.g., Malkov & Drury 2001) and is clearly seen in numerical time-dependent solutions (e.g., Berezhko et al. 1996), a CRP forms, and the acceleration then proceeds mostly in the CRP. The acceleration rate, however, behaves with time approximately as it does at the previous (linear) stage, since dU/dz decreases with the growing width of the CRP, $L_p \sim \kappa(p_{\max})/U_1$. Note that dU/dz sets the acceleration rate. The next acceleration-boosting transition occurs when L_p reaches one of its above-mentioned "natural" limits, such as the accelerator size [a significant fraction of an SNR radius, for example, i.e., $\kappa(p_{\max})/U_1 \sim R_{\rm SNR}$]. Assume $p_{\max} \simeq p_*$ at this point. Then particles with momenta $p > p_*$ cannot be confined to the shock diffusively, since their diffusion length $L_{\rm dif}(p) = \kappa(p)/U_1 > L_p \simeq \kappa(p_*)/\mathrm{U}_1$. Fortunately, by this time instabilities of the CRP produce an ensemble of shocks or similar nonlinear magnetic structures. These ensembles of structures turn out to be capable of confining particles with $p > p_*$, since the mean distance between the structures (which is the scale that replaces the wavelength of particle-confining turbulence in the quasi-linear picture) is much longer than the particle gyroradius, $L \gg r_g(p_*)$. Particle confinement is, however, selective in terms of particle location in phase space. This is similar to, but more general than, particle confinement in magnetic traps, in which only particles having pitch angles satisfying $|\mu| < (\Delta B/B)^{1/2}$ are trapped while others leave the system freely. The price to be paid for this is a spectrum at $p > p_*$ that is even steeper than p^{-4}. This is, in fact, more a blessing than a curse since these particles do not contribute to the CR pressure significantly and L_p does not grow, thus maintaining the acceleration rate at the same level $U_1^2/\kappa(p_*)$ for all particles with momenta $p_* < p < p_{\max}$.

The total acceleration time up to the maximum momentum $p_{\max}$ (given by eq. [51]) is thus only logarithmically larger than the acceleration time to reach the spectral break (knee) at $p \simeq p_*$, and the maximum momentum itself is also pinned to the break momentum p_* through equation (50). The break momentum is simply the maximum momentum in the standard DSA scheme, whatever physical process stops it from growing further. The realization of the above scenario, which obviously produces a significant acceleration enhancement in terms of both the maximum energy and acceleration time, depends crucially on the presence of the ensemble of shocks or similar scattering magnetic structures in the CRP. Their existence in various (even CR-unmodified) collisionless shocks is perhaps beyond any reasonable doubt, as we discussed in § 3, but the real challenge is to calculate the statistical characteristics of their spacing L and amplitude parameter Δv. A simple approach that we pursued in § 5 leads to the conservative estimate of $L/L_p \sim \frac{1}{10}$, so that taking $U_1/c \sim 10^3$, the maximum momentum exceeds its standard value by a factor of 100 (eq. [50]).

At this point it is worthwhile to recapitulate the major differences between our approach and other recent models discussed in § 1 that are also aimed at the enhancement of acceleration efficiency. They primarily seek to increase the turbulent magnetic energy by tapping the free energy of (already) accelerated particles. This would result in decreasing the particle dwelling time (more frequent shock crossing and thus faster acceleration) and, concomitantly, the particle diffusion length (better confinement, higher maximum energy). Our approach is based on the appreciation of the following:

1. In a nonlinear regime, the acceleration time is reduced because of a steeper velocity profile in the smooth part of the shock transition (where acceleration mostly occurs). The steep gradient is maintained due to the formation of a break on the particle

spectrum. Without the break, the gradient would flatten with growing particle energy slowing down the acceleration. Apart from this gradient, the acceleration rate does not depend on the mfp explicitly, as long as the latter is smaller than the shock transition.

2. Particle confinement (maximum energy) is enhanced by increasing the wavelength (shock spacing in a shock train inside the CRP, for example) of turbulence that is in a distinctively strong regime. Departure from the weak turbulence is marked by modification of particle confinement in the CRP. This steepens the momentum spectrum beyond the spectral break and ensures its existence, as required in (1).

It remains unclear, whether the realization of the first (i.e., strong magnetic field amplification) hypothesis would preclude or otherwise strongly influence the acceleration scenario suggested in this paper. It would almost certainly not, if the magnetic energy increases predominantly at longest scales as suggested by Diamond & Malkov (2004) . This should merely scale the estimates provided in this paper by renormalizing such basic quantities as L_p and $\tau_{\rm NL}$, thus making acceleration improvement even stronger. This seems to be necessary for the successful explanation of the Galactic CR spectrum between the knee (at $\approx 4 \times 10^{15}$ eV) and the second knee (at $\approx 10^{18}$ eV), e.g., Hörandel et al. (2005). The knee position can thus be interpreted as p_* corresponding to the interstellar magnetic field moderately amplified (by a factor of 10 or so), whereas the spectrum between the knees is formed by the mechanism described in §§ 6.2 and 6.3. It is also worthwhile to recall in this regard that due to the scale invariance of the equations of motion of the accelerated particles (in which particle momentum and charge enter in combination p/Z, where Z is the charge number), the high-energy spectra obtained for the protons remain valid for other particles with the same value of p/Z. This is what observations actually suggest (e.g., Ellison et al. 1997; Hörandel et al. 2005).

One disadvantage of acceleration enhancement with significant field amplification, however, is that the increased magnetic field inevitably leads to increased radiative losses with a reduction of the particle maximum energy as an implication. This important acceleration constraint has been studied in detail by Aharonian et al. (2002), including applications to many popular acceleration sites. Naturally, the most prominent impact is expected on the acceleration of the extremely high energy CRs (EHECR; $\gtrsim 10^{20}$ eV). Detailed models of such acceleration have been proposed by Miniati et al. (2001), Kang & Jones (2005), and others. Clearly, the acceleration improvement without significant increase of the magnetic field potentially provides a significant edge on the field amplification scenario in the general context of the problem of EHECR acceleration.

Support by NASA grant ATP 03-0059-0034 and under US Department of Energy grant FG 03-88ER53275 is gratefully acknowledged.

REFERENCES

Achterberg, A. 1981, A&A, 98, 161
Achterberg, A., & Blandford, R. D. 1986, MNRAS, 218, 551
Achterberg, A., Gallant, Y. A., Kirk, J. G., & Guthmann, A. W. 2001, MNRAS, 328, 393
Aharonian, F. A., Belyanin, A. A., Derishev, E. V., Kocharovsky, V. V., & Kocharovsky, V. V. 2002, Phys. Rev. D, 66, 023005
Axford, W. I., Leer, E., & Skadron, G. 1977, Proc. 15th Int. Cosmic Ray Conf. (Plovdiv), 11, 132
Bell, A. R. 1978, MNRAS, 182, 147
———. 2004, MNRAS, 353, 550
Bell, A. R., & Lucek, S. G. 2001, MNRAS, 321, 433
Berezhko, E. G. 1996, Astropart. Phys., 5, 367
Berezhko, E. G., Yelshin, V., & Ksenofontov, L. 1996, Soviet Phys.-JETP, 82, 1
Blandford, R. D., & Eichler, D. 1987, Phys. Rep., 154, 1
Blandford, R. D., & Ostriker, J. P. 1978, ApJ, 221, L29
Blasi, P. 2002, Astropart. Phys., 16, 429
Blasi, P., Gabici, S., & Vannoni, G. 2005, MNRAS, 361, 907
Blondin, J. M., & Ellison, D. C. 2001, ApJ, 560, 244
Bykov, A. M., & Toptygin, I. N. 1981, Izv. Ser. Fizika, 45, 474
Chekhlov, A., & Yakhot, V. 1995, Phys. Rev. E, 51, R2739
Diamond, P. H., & Malkov, M. A. 2004, J. Plasma Fusion Res., 6, 28
Drury, L. O'C. 1983, Rep. Progr. Phys., 46, 973
Drury, L. O'C., & Falle, S. A. E. G. 1986, MNRAS, 223, 353
Drury, L. O'C., van der Swaluw, E., & Carroll, O. 2003, preprint (astro-ph/0309820)
Ellison, D. C., Drury, L. O'C., & Meyer, J. P. 1997, ApJ, 487, 197
Fermi, E. 1949, Phys. Rev., 75, 1169
Gotoh, T., & Kraichnan, R. H. 1998, Phys. Fluids, 10, 2859
Gurbatov, S. N., Malakhov, A. N., & Saichev, A. I. 1991, Nonlinear Random Waves and Turbulence in Nondispersive Media: Waves, Rays, Particles (Manchester: Manchester Univ. Press)
Hada, T., Kennel, C. F., Buti, B., & Mjolhus, E. J. 1990, Phys. Fluids, 2, 2581
Hörandel, J. R., Kalmykov, N. N., & Timokhin, A. V. 2005, preprint (astro-ph/0508015)
Jones, T. W., et al. 1998, PASP, 110, 125
Jun, B. I., Jones, T. W., & Norman, M. L. 1996, ApJ, 468, L59
Kang, H., & Jones, T. W. 2005, ApJ, 620, 44
Kang, H., Jones, T. W., & Ryu, D. 1992, ApJ, 385, 193
Kazanas, D., & Ellison, D. C. 1986, ApJ, 304, 178
Kirk, J. G., & Dendy, R. O. 2001, J. Phys. G., 27, 1589
Krymsky, G. F. 1977, Dokl. Akad. Nauk SSSR, 234, 1306 (Engl. transl. in Sov. Phys.-Dokl. 23, 327 [1977])
Lagage, P. O., & Cesarsky, C. J. 1983, A&A, 125, 249
Lichtenberg, A. J., & Lieberman, M. A. 1983, Regular and Stochastic Motion (New York: Springer)
Lucek, S. G., & Bell, A. R. 2000, MNRAS, 314, 65
Malkov, M. A. 1996, Physica D, 95, 62
———. 1997, ApJ, 485, 638
Malkov, M. A., Diamond, P. H., & Jones, T. W. 2002, ApJ, 571, 856
Malkov, M. A., Diamond, P. H., & Völk, H. J. 2000, ApJ, 533, L171
Malkov, M. A., & Drury, L. O'C. 2001, Rep. Progr. Phys., 64, 429
Malkov, M. A., Kotelnikov, A. D., & Kennel, C. F. 1995, Physica D, 86, 480
McKenzie, J. E., & Völk, H. J. 1982, A&A, 116, 191
Medvedev, M. V., & Diamond, P. H. 1996, Phys. Plasmas, 3, 863
Medvedev, M. V., Diamond, P. H., Shevchenko, V. I., & Galinsky, V. L. 1997, Phys. Rev. Lett., 78, 4934
Metzler, R., & Klafter, J. 2000, Phys. Rep., 339, 1
Miniati, F., Ryu, D., Kang, H., & Jones, T. W. 2001, ApJ, 559, 59
Papadopoulos, K. 1985, in Collisionless Shocks in the Heliosphere, ed. B. T. Tsurutani & R. G. Stone (Geophys. Monogr. 35; Washington: AGU), 59
Ptuskin, V. S. 1981, Ap&SS, 76, 265
Ptuskin, V. S., & Zirakashvili, V. N. 2003, A&A, 403, 1
———. 2005, A&A, 429, 755
Sagdeev, R. Z. 1966, Rev. Plasma Phys., 4, 23
Sveshnikova, I. G. 2003, A&A, 409, 799
Tsurutani, B. T., Thorne, R. M., Smith, E. J., Gosling, J. T., & Matsumoto, H. 1987, J. Geophys. Res., 92, 11074
Völk, H. J., Drury, L. O'C., & McKenzie, J. F. 1984, A&A, 130, 19
Warren, J. S., et al. 2005, ApJ, 634, 376
Webb, G. M., Ko, C. M., Zank, G. P., & Jokipii, J. R. 2003, ApJ, 595, 195
Zank, G. P., Axford, W. I., & McKenzie, J. F. 1990, A&A, 233, 275
Zaslavsky, G. M. 2002, Phys. Rep., 371, 461