The Astrophysical Journal, 591:954–961, 2003 July 10
© 2003. The American Astronomical Society. All rights reserved. Printed in U.S.A.

ON PARTICLE ACCELERATION AROUND SHOCKS. I.

Mario Vietri
Scuola Normale Superiore, I-56100 Pisa, Italy
Received 2002 December 16; accepted 2003 March 25

ABSTRACT

We derive a relativistically covariant (although not manifestly so) equation for the distribution function of particles accelerated at shocks, which applies also to extremely relativistic shocks and arbitrarily anisotropic particle distributions. The theory is formulated for arbitrary pitch-angle scattering and reduces to the well-known case for small-angle scatterings via a Fokker-Planck approximation. The boundary conditions for the problem are completely reformulated introducing a physically motivated Green's function; the new formulation allows derivation of the particle spectrum both close and far away from the injection energy in an exact way, while it can be shown to reduce to a power law at large particle energies. The particle spectral index is also recovered in a novel way. Contact is made with the Newtonian treatment.

Subject headings: acceleration of particles — cosmic rays — shock waves

1. INTRODUCTION

The theory of particle acceleration at shocks is currently unsatisfactory. We cannot follow this process from first principles, i.e., the development of a collisionless shock, with the ensuing injection of nonthermal particles, nor do we understand the later shock evolution, which should include the interaction between the population of high-energy, nonthermal particles and the shock structure. Even in the test-particle limit, to which this paper is totally confined, there are at least two major problems: First, the properties of the scattering agents are very poorly known; this problem is not dealt with in this paper. Second, the general foundations of the theory have not been laid down. In particular, there is no general equation for the particle distribution function for arbitrary (i.e., even relativistic) shock speed, and no general treatment is available when the distribution function is anisotropic. Most results available to us either have been derived in the vanishing speed (thus fully isotropic) limit (Bell 1978; Blandford & Ostriker 1978) or descend from numerical (Bednarz & Ostrowski 1998) or seminumerical (Kirk & Schneider 1987) approaches. This is true even in the idealized situation of a plane shock expanding forever in a uniform medium, with particle scattering due only to pitch-angle scattering. It is the aim of this paper to derive a sufficiently general equation describing in a relativistically covariant way this process, to elucidate the conditions leading to the determination of the particle spectrum, and to apply these results to the extremely relativistic limit, at which the distribution function anisotropy is expected to be maximal. In passing, a number of old results are recovered: it is shown that the probability of reaching infinity is independent of the particle momentum, that the spectrum is the superposition of many bumps, each corresponding to a set of particles that have crossed the shock 0, 1, 2, . . ., N, . . . times, and that, for very large particle energy, the spectrum from the superposition of these bumps leads to a power law in the particle impulse: all these results are shown to hold for all shock speeds, Newtonian, relativistic, or intermediate ones. In addition, the explicit dependence on the injection spectrum is presented. In a future paper, this formalism will be applied to the extremely relativistic limit, including the effects of small- or large-angle scatterings and a mean magnetic field.

The plan of this paper is as follows: In § 2 we derive the new equation for the particle distribution function in an explicitly relativistic covariant way. In § 3 we reformulate the boundary conditions to which this problem is subject in a physically motivated way, finding eventually the dependence of the distribution function on the injection spectrum. In § 4 we concentrate on the distribution function at energies large compared to those of injection; here we show that it must be a power law, showing what fixes the power-law index analytically and deriving the relativistic generalization of Bell's law. We also make contact with the Newtonian limit. The important probability distributions P_d and P_u, whose existence is merely postulated in the previous sections, are explicitly written out in terms of eigenfunctions of the angular part of the scattering equation in § 5. In § 6 we summarize our results.

2. AN EQUATION FOR THE DISTRIBUTION FUNCTION

We consider a shock propagating at arbitrary speed in a homogeneous fluid. We place ourselves in the shock reference frame and choose coordinates such that $z = 0$ identifies the shock position. The upstream fluid is located at $z < 0$, so that the fluid speed is always more than 0, both upstream and downstream. Suitable jump conditions at the shock are assumed (Landau & Lifshitz 1987), but they are not necessary in this section.

We assume now that we can break down the effects of the magnetic field into two parts, as is customarily done. The first part is due to a small–coherence-length magnetic field, possibly self-excited by the particles; this provides for an effective scattering, which is included in the collisional term. Another part, however, is due to a long–coherence-length field, which provides instead for a smooth deflection of particles in phase space; this effect is included in the convective term (see below). We are thus using a dichotomic description of the magnetic field; most of the results to be described here, which are wholly unavoidable within this description, ultimately depend on the correctness of this assumption about the magnetic field.

The distribution function is of course subject to the collisionless Vlasov equation, supplemented with suitable

collisional and injection terms. In a somewhat abstract form, we have

$$\frac{df}{ds} = \left(\frac{\partial f}{\partial s}\right)_{\rm coll} + \phi_{\rm inj} \ . \tag{1}$$

The right-hand side is the collisional term, with collisions per unit proper time, plus the rate of particles injected at the shock.

The convective term is

$$\frac{df}{ds} \equiv \frac{dx^\nu}{ds}\frac{\partial f}{\partial x^\nu} + \frac{dp^\nu}{ds}\frac{\partial f}{\partial p^\nu} \ . \tag{2}$$

In the above equation, the term containing the derivative with respect to the energy is not usually explicitly written out; instead, it is imposed ab initio that particles be on their mass shell. Here we write it out (showing later on that it makes no difference) in order to emphasize the covariant character of the equation. In fact, both the first and the second terms on the right-hand side are invariant for Poincaré group transformations. We make use of this by computing each term in different frames of reference and with different orientations for the coordinate axes. In the first term, we place ourselves in the shock frame and orient the z-axis along the shock normal, so that we can make use of $\partial/\partial t = \partial/\partial x = \partial/\partial y = 0$, because we assume the shock to be in a steady state and to have planar symmetry. We also write the z-component of the particle four-speed in terms of quantities in the fluid frame:

$$\frac{dz}{ds} = \gamma_p \gamma (u + \mu) \ , \tag{3}$$

where μ is the cosine of the angle that the particle speed makes with the shock normal z in the fluid frame, u and γ are the shock speed (in units of c) and Lorentz factor speed with respect to the fluid, and γ_p is the particle Lorentz factor also with respect to the fluid. All particles are here assumed relativistic in any reference frame (so that their three-speeds are always $\approx c$), even those that are just injected. Strictly speaking, this is accurate only for relativistic shocks. The generalization to include Newtonian particles, although easy in principle to do, introduces some cumbersome formulae (with radicals) that complicate this work uselessly: it is not considered any further.

The second term on the right-hand side of equation (2) is instead computed in the fluid frame, so that we can neglect any motional electric field. In this case, the particle energy is conserved, and the term including the derivative $\partial f/\partial E$ disappears. The remaining components can be written as (Berezinsky et al. 1990, eq. [9.16])

$$\frac{dp^\nu}{ds}\frac{\partial f}{\partial p^\nu} = \frac{dt}{ds}\frac{dp^\nu}{dt}\frac{\partial f}{\partial p^\nu} = -\gamma_p \omega \frac{\partial f}{\partial \phi} \ , \tag{4}$$

where $\omega = eB/E$ is the particle Larmor frequency in the fluid frame and ϕ is the longitudinal angle around the direction of the magnetic field, in momentum space.

To estimate the collisional term, we first place ourselves in the fluid frame and then remark that, in full generality, such a term has the form

$$\left(\frac{\partial f}{\partial t}\right)_{\rm coll} = -d(\mu) f + \int w(\mu,\ \mu') f(\mu')\, d\mu' \ . \tag{5}$$

Here, we have made use of the assumption of linearity: we assume, that is, that the magnetic field turbulence responsible for the scattering is independent of the particle density. We consider particle scattering when there may be a correlation between the directions before and after the scattering event, so that the scattering coefficient $w(\mu,\ \mu')$ depends on both. In addition, we are not assuming the diffusion approximation: the above equation applies to arbitrarily few, and arbitrarily large-angle, scattering laws. It would reduce to the well-known Ginzburg-Syrovatskii (Ginzburg & Syrovatskii 1964) form in the diffusive approximation. The first term on the right-hand side represents the scattering of particles away from their direction of motion, while the second term represents the total contribution to the distribution function in the direction of motion under consideration by scattering from all other directions. Clearly, d and w are related through

$$d(\mu) = \int_{-1}^{+1} w(\mu',\ \mu)\, d\mu' \ , \tag{6}$$

which simply expresses probability conservation. Here w (and thus also d) may depend on p. Often, one makes the simplifying assumption that the p-dependence in w (and thus in d) factors out; we comment later on why this may come in handy but do not assume this. We also remark that $w, d > 0$.

Since we need $\partial f/\partial s$, while above (eq. [5]) we have evaluated the collisional term in the fluid frame, we correct it by means of $dt/ds = \gamma_p$. In addition, we wish to include the injection term, $\phi_{\rm inj}$. We assume that injection takes place at the shock only, but we leave the energy and direction dependence of the term otherwise arbitrary: $\phi_{\rm inj} = G_{\rm inj}(p,\ \mu)\delta(z)$, where $\delta(x)$ is the Dirac delta function.

Putting together all of the above, we find

$$\gamma(u+\mu)\frac{\partial f}{\partial z} = \left[-d(\mu) f + \int w(\mu,\ \mu') f(\mu')\, d\mu'\right] + \omega\frac{\partial f}{\partial \phi} + G_{\rm inj}(p,\ \mu)\delta(z) \ , \tag{7}$$

which is the equation we were searching for. It is worth remarking that all terms in the above equation are computed in the fluid frame, except for the space dependence and its attaining derivative. In addition, the above equation is valid both upstream and downstream, but now care must be taken because the quantities p, μ, μ', ω, ... are all evaluated in the local fluid frame, and thus quantities bearing the same name in the two distinct reference frames are not identical: they are related instead by a Lorentz transformation, which we give here for future reference. Using subscripts u and d to denote quantities in the upstream and downstream frame, respectively, we have

$$p_u = p_d \gamma_r (1 - u_r \mu_d), \quad \mu_u = \frac{\mu_d - u_r}{1 - u_r \mu_d} \ , \tag{8}$$

where u_r and γ_r are the modulus of the relative speed between the upstream and downstream sections and its associated Lorentz factor.

Whenever the scattering in each individual event is small, it is possible to apply to the above equation the very same treatment that leads to a Fokker-Planck equation. In this

case one obtains (Kirk & Schneider 1987)

$$\gamma(u+\mu)\frac{\partial f}{\partial z}=\frac{\partial}{\partial \mu}\left[D(\mu,\ p)\left(1-\mu^2\right)\frac{\partial f}{\partial \mu}\right] + \omega\frac{\partial f}{\partial \phi}+G_{\rm inj}(p,\ \mu)\delta(z)\ , \qquad (9)$$

where the effective diffusion coefficient is related to the rms angle of diffusion. It is worth remarking that it may occur that the above equation holds for the downstream section, but it is not justified for the upstream section, where the full equation (7) has to be employed. The reason is that the diffusive approximation on which the above equation is based requires each scattering event to lead to an rms deflection $O(1/\gamma)$, where γ is the fluid Lorentz factor with respect to the shock. Thus, when the shock becomes extremely relativistic, the correctness of the diffusive approximation breaks down: it follows that one can use equation (9) for the downstream section but is forced to use equation (7) for the upstream one.

In the following, we drop the term due to the long–coherence-length magnetic field, which will be considered in the next paper in this series.

3. OBTAINING THE PARTICLE SPECTRUM

We begin by remarking that in this problem there is a net mean flux of particles across any surface; if the surface is steady in the shock frame, then the flux is also time-independent. Since the particles' speed is always assumed to be $\approx c$ (≈ 1 in our units), the infinitesimal flux across a unit shock area per unit time is given by

$$dJ=\mu_s p_s^2 f\, d\mu_s\, dp_s=\gamma_u(u+\mu_u)p_u^2 f\, d\mu_u\, dp_u = \gamma_d(u_d+\mu_d)p_d^2 f\, d\mu_d\, dp_d\ , \qquad (10)$$

depending on whether we choose to express f in terms of variables in the shock, upstream, or downstream frames. The representation in terms of downstream variables is especially useful, because in the downstream section the downstream momentum p_d is conserved, so that a net flux across the shock of particles of impulse p_d can be due only to the fact that some fraction of these particles is advected to downstream infinity, with probability given by equation (21).

From now on in this section, we use coordinates in the downstream frame exclusively, which we thus indicate, for ease of notation, without subscripts.

The mathematical treatment of this problem must be *physically* motivated: we envision a system in which particles are injected at low energies at the shock, from which they cannot reach upstream infinity because they would have to swim upstream, but from which they do reach downstream infinity. In order to study how particles generated at the shock reach downstream infinity, we first concentrate on the downstream part of the problem.

Consider now a downstream section of finite length L, beginning at the shock; then place two observers at either end and have them inject particles into the downstream section. Then they can collect outgoing particles at either end. The problem is obviously physically well defined. The speed of a particle (in downstream-frame coordinates) with respect to the shock is

$$v_s=\frac{u_d+\mu}{1+u_d\mu}\ . \qquad (11)$$

Thus, the observer located at the shock injects particles for $\mu > -u_d$ (here u_d is the downstream fluid speed with respect to the shock) and collects the outgoing ones (i.e., those with $\mu < -u_d$). Therefore, reasonable boundary conditions at the shock can be posed for $\mu > -u_d$ only. The observer at the other end of the downstream section disappears as $L \to \infty$, so that there can be no entering flux there; we can only impose the regularity condition that f neither grows to infinity nor goes to zero.

Similarly, for the upstream section we see that we can impose boundary conditions for $\mu < -u_d$ at the shock; at upstream infinity, the only reasonable boundary condition is that $f \to 0$, because we cannot expect our particles to swim against the fluid advection all the way to upstream infinity.

We now see that the boundary conditions for the upstream problem are provided by the outgoing particles of the downstream section, and, vice versa, the outgoing particles of the upstream section provide the boundary conditions of the downstream section, thus setting up an obvious (fixed point) problem. In order to make this explicit, let us introduce the conditional probability $P_d(\mu_{\rm in},\ \mu_{\rm out})d\mu_{\rm out}$ that a particle, given that it crossed the shock toward downstream along a direction $\mu_{\rm in}$, will recross it along the direction $\mu_{\rm out}$. Then, calling $J_{\rm in}=(u_d+\mu_{\rm in})f(\mu_{\rm in})$ the entering flux, the outgoing one will be given by

$$J_{\rm out}(\mu_{\rm out})=\int_{-u_d}^{1} d\mu_{\rm in}\, P_d(\mu_{\rm in},\ \mu_{\rm out})J_{\rm in}(\mu_{\rm in})+G(p,\ \mu_{\rm out}) \equiv P_d * J_{\rm in}+G_{\rm out}\ , \qquad (12)$$

where we have also included in the outgoing flux the part deriving from the injection $G_{\rm out}=G_{\rm inj}(\mu_{\rm out})$ for $\mu < -u_d$. The symbolic notation is helpful in the following.

Before specifying the equivalent relationship for the upstream part of the fluid, we pause to establish the connection between P_d and the solutions of equation (7) (or eq. [9]). Given the definition of P_d, it seems obvious enough that P_d is the particle flux $(u+\mu)f(\mu)$ given by the solution of equation (7) (or of eq. [9]) with the boundary condition $(u_d+\mu)f=\delta(\mu-\mu_{\rm in})$ (unit entering flux) for $\mu > -u_d$ at the shock ($z=0$) and f regular at downstream infinity. In fact, we can see that this solution represents a monochromatic (in energy) flux of particles, all moving initially in the same direction, subject to pitch-angle scattering, which is in turn responsible for kicking them back toward the shock, occasionally, along a direction $\mu_{\rm out}$ with nonuniform probability.

One way of looking at P_d is to realize that it is somewhat related to the Green's function for a diffusive problem. As an analogy, consider the paradigmatic diffusion equation

$$\frac{\partial n}{\partial z}=\frac{\partial^2 n}{\partial \mu^2}\ , \qquad (13)$$

which is solved by convolving the given boundary conditions with its Green's function:

$$2(\pi z)^{-1/2}\exp\left[\frac{-(\mu-\mu_i)^2}{4z}\right]\to\delta(\mu-\mu_i) \qquad (14)$$

as $z \to 0$. Here too, the Dirac delta function appears in the boundary conditions and is smoothed out as the time variable (z) evolves. The major difference is that suitable boundary conditions for the above equation include the whole range in μ, at $z = 0$, while in our problem we can specify boundary conditions only for $\mu > -u_d$. The function P_d is the projection of the solution of the problem on exactly that part of the boundary at $z = 0$ where we cannot specify the boundary conditions, a part that does not exist for equation (13). Still, because of this relationship, we refer to P_d and to P_u, shortly to be determined, as the Green's functions of the problem.

As an aside, we show that, under some physically realistic conditions, P_d does not depend on the particle impulse p. This may be at first surprising because both coefficients $d(\mu)$ and $w(\mu,\ \mu')$ in equation (7) are allowed to depend on p: the proof follows. Rewrite equation (7) outside the shock (which means that we can drop the injection term) without the term depending on the magnetic field, using as a new variable $y \equiv d(\mu)z$; this can always be done because, as noted above, $d > 0$. We obtain

$$\gamma(u + \mu)\frac{\partial f}{\partial y} = -f + \int g(\mu,\ \mu')f(\mu')\,d\mu' \ . \tag{15}$$

Since we assumed that the p-dependence in w and d factors out, we see from equation (6) that g is independent of p. Then, from equation (15) we see that p has disappeared from the problem altogether and that any solution can depend on p only through the term $y = d(\mu)z$. However, P_d is the flux at $z = 0$, the shock position, so that we see that when this is computed ($y = 0$), all dependence on p drops out: P_d does not depend on p. This result was known to Bell (1978), who proved it for Newtonian shocks; the proof given here is, however, valid for all shock speeds.

Upstream, we also have a probability distribution $P_u(\mu_{\rm out},\ \mu_{\rm in})$ that a particle leaving downstream along a direction $\mu_{\rm out}$ will recross the shock along a direction $\mu_{\rm in}$. An argument entirely identical to the one given above shows that, often, P_u does not depend on the particle's impulse. There is, however, a small complication in that particles of impulse p_i in the downstream frame, exiting it along a direction $\mu_{\rm out}$ and reentering it along a direction $\mu_{\rm in}$, emerge with a different impulse p (as measured again in the downstream frame), given by

$$p = \frac{1 - u_r\mu_{\rm out}}{1 - u_r\mu_{\rm in}}p_i \equiv Gp_i \ . \tag{16}$$

Here, G is the energy amplification a particle receives as it cycles once around the shock. Since $-1 \le \mu_{\rm out} \le -u_d$ and $-u_d \le \mu_{\rm in} \le 1$, we have

$$1 \le G \le \frac{1 + u_r}{1 - u_r} \ . \tag{17}$$

Thus, particles leaving the downstream section with impulse p_i provide the reentering flux at a different impulse. A trivial computation yields the entering flux at impulse p as

$$\begin{aligned} J_{\rm in}(p,\ \mu_{\rm in}) &= \int_{-1}^{-u_d} d\mu_{\rm out} P_u(\mu_{\rm out},\ \mu_{\rm in})\left(\frac{1 - u_r\mu_{\rm in}}{1 - u_r\mu_{\rm out}}\right)^3 \\ &\quad \times J_{\rm out}(p_i,\ \mu_{\rm out}) + G_{\rm in}(p,\ \mu_{\rm in}) \\ &\equiv P_u * J'_{\rm out} + G_{\rm in}(p,\ \mu_{\rm in}) \ , \end{aligned} \tag{18}$$

where again we added to the incoming flux the injected one $G_{\rm in} = G_{\rm inj}(\mu_{\rm in})$ for $\mu_{\rm in} > -u_d$, and the prime over $J_{\rm out}$ reminds us that the outgoing flux is to be computed at the particle energy p_i given by equation (16).

A comment on normalization is in order. All particles that enter upstream will in due time be brought backward across the shock by the fluid's advection, so that we expect

$$\int_{-u_d}^{1} d\mu_{\rm in}\, P_u(\mu_{\rm out},\ \mu_{\rm in}) = 1 \ , \tag{19}$$

while the analogous statement for the downstream Green's function P_d does not hold for the same reason: a fraction of all particles will be advected toward downstream infinity, and the returning flux will be smaller than the departing one. Thus, in general, $\int_{-1}^{-u_d} d\mu_{\rm out}\, P_d(\mu_{\rm in},\ \mu_{\rm out}) < 1$. The average probability of coming back to cross the shock is given by averaging P_d over the whole incoming flux:

$$P = \frac{\int_{-1}^{-u_d} d\mu_{\rm out} \int_{-u_d}^{1} d\mu_{\rm in}\, P_d(\mu_{\rm in},\ \mu_{\rm out})J_{\rm in}(\mu_{\rm in})}{\int_{-u_d}^{1} d\mu_{\rm in}\, J_{\rm in}(\mu_{\rm in})} \ . \tag{20}$$

When the contribution of injection can be neglected, we see from the above and from equation (12) that

$$P = \frac{\int_{-1}^{-u_d} d\mu_{\rm out}\, J_{\rm out}(\mu_{\rm out})}{\int_{-u_d}^{1} d\mu_{\rm in}\, J_{\rm in}(\mu_{\rm in})} \ . \tag{21}$$

The reason for this is quite simple: in a steady state, such as that assumed here, there can be no accumulation of particles except at infinity; since there is a larger flux entering the downstream region than leaving it, there must be an accumulation of particles in it, which can thus be realized only at downstream infinity, i.e., by leaving the region of the shock.

Our cardinal equations, equations (12) and (18), can now be simply combined to obtain

$$J_{\rm in} = P_u * P_d * J'_{\rm in} + \left(P_u * G'_{\rm out} + G_{\rm in}\right) \equiv Q * J'_{\rm in} + X \ . \tag{22}$$

This equation can be solved iteratively, using the term in parentheses as the first guess,

$$\begin{aligned} J_{{\rm in},(0)} = X, \quad J_{{\rm in},(1)} &= Q * X' + X \ , \\ J_{{\rm in},(N)} &= Q * J'_{{\rm in},(N-1)} + X \ldots \ , \end{aligned} \tag{23}$$

to arrive at the solution in the form

$$J_{\rm in} = X + Q * X' + Q * Q * X'' + Q * Q * Q * X''' + \ldots \ . \tag{24}$$

The interpretation of this equation is simple: X is the flux of injected particles, including both those that are directly moving in and those that were initially moving out but have been turned back exactly once. The operator Q transforms an ingoing flux into another ingoing flux, made of the fraction of all particles that have completed exactly one loop around the shock. Thus, the terms $Q * Q * X''$ and so on represent the particles that have made two, three, ..., N, ... loops around the shock, after injection. The fact that this theory manages to recover the well-known fact that the spectrum is the superposition of infinite bumps, each corresponding to particles that have looped an integer number of times around the shock, can be considered a useful test that has been successfully cleared. It also has the advantage of

yielding the distribution function at the shock for the first time both very close to the injection energy and very far from it. In addition, it substantiates physical intuition that the solution of the problem must depend only on the injected flux (X) and on the scattering properties of the medium. Finally, the equation concerning $J_{\rm out}$ carries no new information: once $J_{\rm in}$ has been found, it can be plugged into equation (12), which yields a $J_{\rm out}$ that automatically satisfies equation (18).

4. THE LARGE-p LIMIT

We see from equation (16) that after half a cycle (only the upstream section changes the particles' energies), the energy of a particle just injected with energy $p_{\rm inj}$ can span the range ($-1 \le \mu_{\rm out} \le -u_d \le \mu_{\rm in} \le 1$)

$$p_{\rm inj} < p < p_{\rm inj}\frac{1+u_r}{1-u_r} \equiv p_{\rm max} \ . \tag{25}$$

This shows that the evolution of that part of the particle distribution function that describes particles making many loops around the shock must extend toward larger and larger energies; in fact, since the minimum amplification is $G = 1$, the process under discussion does not describe a spreading of the injection distribution around the initial value but a drift toward larger and larger energies (plus a spreading, of course): Fermi acceleration of type I. Therefore, the question naturally arises of what the asymptotic spectrum of particles is, as we consider asymptotically larger energies.

If injection provides particles of greatest energy $p_{\rm inj}$, we can neglect the injection terms in the above equations for energies exceeding $p_{\rm max}$, and we thus consider equation (22) in this case:

$$J_{\rm in}(p,\ \mu) = \int_{-1}^{-u_d} d\mu_{\rm out}\, P_u(\mu_{\rm out},\ \mu)\left(\frac{1-u_r\mu}{1-u_r\mu_{\rm out}}\right)^3 \times \int_{-u_d}^{1} d\xi\, P_d(\xi,\ \mu_{\rm out}) J_{\rm in}\left(\frac{1-u_r\mu_{\rm in}}{1-u_r\mu_{\rm out}}p,\ \xi\right) . \tag{26}$$

We show now that this equation, except for the null solution, has a solution that is a simple power law in p. Let us first consider two situations in which injection occurs at slightly different energies, p_0 and $p_0 + \delta p_0$. At large energies, $p > p_{\rm max}$, the spectra of these two situations differ by very little, $(\partial J_{\rm in}/\partial p_0)\delta p_0$. The difference between two solutions of the inhomogeneous problem (i.e., including injection) must be a solution of the associated homogeneous problem (the one without injection, eq. [26]), so that $\partial J_{\rm in}/\partial p_0$ is a solution of equation (26). For obvious dimensional reasons, $J_{\rm in} = h(p/p_0,\ \mu)/p^3$, so that [using $\dot{h}(z,\ \mu) \equiv \partial h/\partial z$]

$$\frac{\partial J_{\rm in}}{\partial p} = -\frac{3h}{p^4} + \frac{\dot{h}}{p^3 p_0} \ , \tag{27}$$

$$\frac{\partial J_{\rm in}}{\partial p_0} = -\frac{\dot{h}}{p^2 p_0^2} \ , \tag{28}$$

and, on eliminating h,

$$p\frac{\partial J_{\rm in}}{\partial p} = -3J_{\rm in} - \frac{1}{p_0}\frac{\partial J_{\rm in}}{\partial p_0} \ , \tag{29}$$

which shows $p\partial J_{\rm in}/\partial p$ to be the linear combination of two solutions of the homogeneous equation (26) (for $p > p_{\rm max}$), and thus a solution itself. If we now assume equation (26) to have a unique solution, we find that $p\partial J_{\rm in}/\partial p$ must be proportional to $J_{\rm in}$, when injection can be neglected. We thus have

$$\frac{1}{J_{\rm in}(p,\ \mu)}\frac{\partial J_{\rm in}}{\partial p}(p,\ \mu) = -\frac{s}{p} \ , \tag{30}$$

where s is a constant yet to be determined, and the obvious solution,

$$J_{\rm in}(p,\ \mu) = \frac{(u_d+\mu)g(\mu)}{p^s} \ , \tag{31}$$

follows, where we have arbitrarily factored out $u_d + \mu$ to remind us that J is a particle flux; thus, $g(\mu)/p^s$ is the particle distribution function. It must be remarked that this power-law dependence has been derived without assuming that P_d and P_u are independent of the particles' momentum p.

The importance of the above is obvious: it shows that the particle distribution function is a power law in the particles' momentum, for every shock speed. This result had been obtained before in all numerical solutions for arbitrary shock speeds and in analytic form in the Newtonian limit, but in the full relativistic regime it was put in by hand (Kirk & Schneider 1987), and the boundary conditions could have equally well been satisfied by any other function with a free parameter. The present derivation is universal, and it has the rather unexpected result (at least, for this author) that the power-law solution is exact, not asymptotic, as soon as we consider energies exceeding the injection ones, i.e., for $p > p_{\rm max}$ and not just $p \gg p_{\rm max}$.

With the result provided by equation (31), our principal equations, equations (12) and (18), assume the following form, when injection can be neglected:

$$(u_d+\mu)g(\mu) = \int_{-u_d}^{1} d\xi\, P_d(\xi,\ \mu)(u_d+\xi)g(\xi) \ , \tag{32}$$

$$(u_d+\mu)g(\mu) = \int_{-1}^{-u_d} d\mu_{\rm out}\, P_u(\mu_{\rm out},\ \mu)\left(\frac{1-u_r\mu}{1-u_r\mu_{\rm out}}\right)^{3-s} \times (u_d+\mu_{\rm out})g(\mu_{\rm out}) \ . \tag{33}$$

The first one concerns the outgoing flux (with respect to the downstream section) and thus is valid for $-1 \le \mu \le -u_d$, while the second one concerns the entering flux and applies for $-u_d \le \mu \le 1$. In addition, use of equation (31) in equation (21) shows that the probability of reaching downstream infinity does not depend on the particle's momentum p.

Again, they can be combined to obtain

$$(u_d+\mu)g(\mu) = \int_{-u_d}^{1} d\xi\, Q^T(\xi,\ \mu)(u_d+\xi)g(\xi) \ , \tag{34}$$

where

$$Q^T(\xi,\ \mu) \equiv \int_{-1}^{-u_d} d\nu\, P_u(\nu,\ \mu)P_d(\xi,\ \nu)\left(\frac{1-u_r\mu}{1-u_r\nu}\right)^{3-s} \ , \tag{35}$$

which applies to the incoming flux only: $-u_d \le \mu \le 1$. Once again, the outgoing flux carries no new information: if the above is solved, plugging the solution into equation (32) returns an outgoing flux that automatically solves equation

(33). All of the problem's information is in one of the subintervals only, even in the case of the homogeneous equation.

The above equation, for an arbitrary value of s, does *not* have a solution. The equation that *does* have a solution is

$$\lambda(u_d+\mu)g(\mu)=\int_{-u_d}^{1} d\xi\, Q^T(\xi,\ \mu)(u_d+\xi)g(\xi)\ . \qquad (36)$$

In fact, the above equation is a homogeneous equation of the Fredholm type, with a smooth and bounded kernel over a compact interval (Courant & Hilbert 1953), which has at most a finite number of solutions, each belonging to an eigenvalue λ that is to be determined simultaneously with the eigenvector $g(\mu)$. We thus see that what fixes s is the requirement that $\lambda=1$: the physical value of s is that which has a unitary eigenvalue. We are forced by our limited mathematical skills to assume that equation (34) has exactly one solution, neither more nor fewer, but apart from this, the task of determining the particle distribution function is completed: equation (36) with $\lambda=1$ yields the particle distribution function at the shock and the all-important particle spectral index s, even in the case of arbitrary particle anisotropy.

4.1. *The Newtonian Limit*

In order to gain physical insight into the condition that fixes s ($\lambda=1$), we rewrite equation (36) somewhat. First, we write equation (33) as

$$\lambda(u_d+\mu)g(\mu)=\int_{-1}^{-u_d} d\mu_{\rm out}\, P_u(\mu_{\rm out},\ \mu)\left(\frac{1-u_r\mu}{1-u_r\mu_{\rm out}}\right)^{3-s}\times(u_d+\mu_{\rm out})g(\mu_{\rm out})\ . \qquad (37)$$

This can now be integrated over the whole range in μ, $-u_d\le\mu\le 1$, and divided by the whole outgoing flux to obtain

$$\lambda\frac{\int_{-u_d}^{1} d\mu(u_d+\mu)g(\mu)}{\int_{-1}^{-u_d} d\mu_{\rm out}(u_d+\mu_{\rm out})g(\mu_{\rm out})}=\int_{-u_d}^{1} d\mu\int_{-1}^{-u_d} d\mu_{\rm out}\left[P_u(\mu_{\rm out},\ \mu)\times\left(\frac{1-u_r\mu}{1-u_r\mu_{\rm out}}\right)^{3-s}(u_d+\mu_{\rm out})g(\mu_{\rm out})\right]\times\left[\int_{-1}^{-u_d} d\mu_{\rm out}\,(u_d+\mu_{\rm out})g(\mu_{\rm out})\right]^{-1}\ . \qquad (38)$$

The term on the left-hand side can be rewritten by means of equation (21) as λ/P. The term on the right-hand side is clearly the average value of the $(s-3)$ power of the amplification G (eq. [16]) over the whole particle distribution, $\langle G^{s-3}\rangle$. The above is thus

$$\lambda=P\langle G^{s-3}\rangle\ , \qquad (39)$$

and demanding that $\lambda=1$ implies that

$$P\langle G^{s-3}\rangle=1\ . \qquad (40)$$

In the Newtonian limit, in which $G-1\ll 1$, we obviously have $\langle G^{s-3}\rangle\approx\langle G\rangle^{s-3}$, and thus the energy spectral index $k=s-2$ is given by

$$s-2\equiv k=1-\frac{\log P}{\log\langle G\rangle}\ , \qquad (41)$$

which is exactly Bell's equation for the spectral index. We have thus established that the condition $\lambda=1$ is equivalent to $P\langle G^{s-3}\rangle=1$, which appears to be the correct relativistic generalization of Bell's Newtonian results.

In the Newtonian limit, where $g(\mu)\approx$ const, independent of μ (the isotropic limit), it is easy to derive from equation (21) that $P\approx 1-u$ and from equation (38) that $\langle G\rangle\approx 1+u$, from which $s=4$, as per Bell's results for a strong shock.

There are some differences between the fully relativistic approach and the Newtonian one. First, in the relativistic approach, in general $\langle G^{s-3}\rangle$ cannot be approximated by $\langle G\rangle^{s-3}$, contrary to Peacock's (1981) arguments. Second, here G and P are functions of s, so that equation (40) sets up a transcendental equation for s. Third, equation (40) is only of symbolic value, since it requires knowledge of the anisotropic particle distribution function, which is not known a priori, but luckily, the full relativistic approach developed here returns simultaneously the anisotropic distribution function (eq. [34]).

4.2. *Bits and Pieces*

A useful identity is obtained by rewriting equation (18) without the injection term as

$$\frac{J_{\rm in}(p,\ \mu_{\rm in})}{(1-u_r\mu_{\rm in})^3}=\int_{-1}^{-u_d} d\mu_{\rm out}\, P_u(\mu_{\rm out},\ \mu_{\rm in})\frac{J_{\rm out}(p_i,\ \mu_{\rm out})}{(1-u_r\mu_{\rm out})^3}\ , \qquad (42)$$

integrating over $\mu_{\rm in}$, using equation (19), and then inserting equation (31) to obtain

$$\int_{-u_d}^{1} d\mu_{\rm in}\,(1-u_r\mu_{\rm in})^{s-3}(u_d+\mu_{\rm in})g(\mu_{\rm in})=\int_{-1}^{-u_d} d\mu_{\rm out}\,(1-u_r\mu_{\rm out})^{s-3}(u_d+\mu_{\rm out})g(\mu_{\rm out})\ , \qquad (43)$$

which, of course, is not trivial, because the intervals of integration are different.

It is worth noting that, in all of § 3, no reference whatsoever has been made to the explicit form of the right-hand side of equation (7); all that is necessary is to determine the equation's Green's functions P_d and P_u, but otherwise all results apply to all right-hand side members and can then be seen as a property of the left-hand side member $(u+\mu)\partial f/\partial z$. They are thus of very general validity.

Finally, we have assumed that equations (34) and (36) have one, and just one, solution. If they had none, then the steady state problem would have no solution, and the problem would be intrinsically time-dependent. If they had several, then a rather interesting case would arise, whereby shocks moving at the same speed in the same media, but with different injection properties, would have different asymptotic spectral indexes s. The near universality of the index s points of course in the opposite direction, but it is also true that a proof of the existence of a unique solution is beyond our modest abilities.

5. P_u AND P_d

In this subsection we show how to build P_u and P_d. For the sake of definiteness, we concentrate on P_u, but a wholly analogous treatment holds for P_d.

We now wish to find an explicit expression for $P_u(\mu_{\rm out},\ \mu_{\rm in})$. That this quantity exists is physically obvious: it represents the probability that a net flux of particles leaving the downstream section along a direction $\mu_{\rm out}$ (in downstream variables) reenters along a direction $\mu_{\rm in}$. We can imagine an experiment to determine this quantity: an experimenter located next to the shock with a small cannon shooting out particles along the direction $\mu_{\rm out}$ can collect them on a screen, determining how many come back along $\mu_{\rm in}$. It corresponds to a solution of equation (7) or (9) without the injection term, with the boundary condition

$$dJ = \delta(\mu_d - \mu_{\rm out})\, d\mu_d = \delta\left(\mu_u - \mu_{\rm out}^{(u)}\right) d\mu_u\ . \qquad (44)$$

Here, $\mu_{\rm out}^{(u)}$ is the value of $\mu_{\rm out}$ in the upstream-frame variables, and the last identity is a trivial property of the Dirac delta function.

In order to find P_u, we now use separation of variables to solve equation (7) or (9). For instance, writing $f(\mu_u,\ z) = A(z)B(\mu_u)$, we find the solutions of

$$\gamma\frac{1}{A(z)}\frac{dA}{dz} = \lambda_n = \frac{1}{(u+\mu_u)B(\mu_u)} \times \frac{d}{d\mu_u}\left[D(\mu_u,\ p_u)\left(1-\mu_u^2\right)\frac{dB}{d\mu_u}\right]\ . \qquad (45)$$

Calling B_n the angular eigenvectors, we find a generic solution of the above as

$$f = \sum_n a_n \exp\left(\frac{\lambda_n z}{\gamma}\right) B_n(\mu_u)\ , \qquad (46)$$

where a_n are coefficients that we must choose so as to satisfy the initial conditions, equation (44). We must thus have, at $z = 0$,

$$\gamma(u+\mu_u)f\ = \sum_n a_n\left(\mu_{\rm out}^{(u)}\right)(u+\mu_u)\gamma B_n(\mu_u) = \delta\left(\mu_u - \mu_{\rm out}^{(u)}\right)\ . \qquad (47)$$

Obviously, the coefficients a_n depend on the point where the delta function is located, $\mu_{\rm out}^{(u)}$.

Here there is a tricky but important point. The above eigenvector problem is well known to have solutions belonging to both positive and negative eigenvalues (plus the $\lambda = 0$ case). Eigenvectors belonging to negative values of λ_n, of which there are an infinite number, are physically ill behaved at upstream infinity, as can be seen from equation (46), so that we surely have to restrict ourselves in all the sums above to the well-behaved eigenvectors, i.e., those with $\lambda_n > 0$. However, does the above equation then hold? The answer is yes, as we now show.

The fact that an infinite set of functions can be arranged to satisfy the above equation is called by mathematicians "completeness," indicating that any function can be written as the superposition of B_n's with suitable coefficients. If the above equation holds, in fact, multiplying both sides by any function $F(\mu_{\rm out}^{(u)})$ and integrating over the whole range of $\mu_{\rm out}^{(u)}$, we find that

$$F(\mu_u) = \sum_n\left[\int d\mu_{\rm out}^{(u)}\, a_n\left(\mu_{\rm out}^{(u)}\right)F\left(\mu_{\rm out}^{(u)}\right)\right] \times \gamma(u+\mu_u)B_n(\mu_u)\ , \qquad (48)$$

proving that we can write any function as the superposition of functions from the set, with suitable coefficients, and thus that the set is, by definition, complete. Now, the set of all eigenfunctions B_n is well known to be complete, but here is the rub: the full set includes both the eigenvectors that have $\lambda > 0$ (which thus are physically well behaved at upstream infinity) and those that have $\lambda \le 0$, which diverge at upstream infinity and which we thus cannot use.

The half-set of the well-behaved eigenvectors is not complete over the whole range in μ, but here we are helped by an unusual property of the eigenvectors of this equation: it can be shown in fact (Freiling, Yurko, & Vietri 2003) that the half-set of all well-behaved eigenvectors (i.e., those with $\lambda > 0$) is complete in the restricted range $-1 \le \mu_u \le -u$. This property (sometimes called half-range half-completeness) implies that, for $\mu_u \le -u$, we can always write

$$\gamma(u+\mu_u)f = {\sum_n}' a_n(u+\mu_u)\gamma B_n(\mu_u) = \delta\left(\mu_u - \mu_{\rm out}^{(u)}\right)\ , \qquad (49)$$

where, however, now the prime over the sum reminds us that the summation is extended only to well-behaved eigenvectors. For $\mu_u > -u$, the above equation is *not* verified, and the sum gives some function, which we can now identify with our required P_u. Thus, P_u is the continuation to the rest of the interval ($\mu_u > -u$) of the sum above, which inside the restricted range $\mu \le -u$ satisfies the above equality.

One word of caution should be stated here: the property of half-range half-completeness has been proved for the eigenvectors of equation (9) but not for those of equation (7). Still, since they are known to be so similar (both are called the third, or polar, type, one within the realm of integral equations and the other one of differential equations, to emphasize several formal similarities), and since the above physical argument makes it extremely plausible that P_u exists also for this case, we make no further distinction between the two cases and proceed as if this argument went through for both equations (9) and (7).

We now find the coefficients a_n. The well-behaved eigenvectors are not orthonormal over the restricted range $\mu_u \le -u$. They can easily be made so by means of a well-known procedure (Schmidt's diagonalization; Courant & Hilbert 1953), which leaves the lowest order eigenvector unaffected (except for the normalization) and suitably modifies the others in a finite number of steps. Once this new basis P_n has been found, since it is complete (from the half-range half-completeness), we easily find, for $\mu_u \le -u$,

$$(u+\mu_u)\sum_n P_n\left(\mu_{\rm out}^{(u)}\right)P_n(\mu_u) = \delta\left(\mu_u - \mu_{\rm out}^{(u)}\right)\ , \qquad (50)$$

and thus, by working backward to equation (10) (for $\mu_{\rm out} < -u_d, \mu_{\rm in} > -u_d$),

$$P_u(\mu_{\rm out},\ \mu_{\rm in})\, d\mu_{\rm in} = (u+\mu_u)\sum_n P_n\left(\mu_{\rm out}^{(u)}\right)P_n(\mu_u)\, d\mu_u\ . \qquad (51)$$

6. SUMMARY

The major results presented in this paper are the following:

1. We have presented a relativistically covariant equation (eq. [7]) for the particle distribution function, which includes pitch-angle scattering (without assuming that the diffusive approximation holds), the effect of smooth magnetic fields, and particle injection at the shock; this equation applies separately in the upstream and downstream frames.

2. We have described the transport properties of the upstream and downstream media by means of two Green's functions, P_d and P_u, which are independent of the particle distribution functions and which have been explicitly constructed in terms of the eigenfunctions of the angular part of the problem (eq. [51]).

3. We have established two relations (eqs. [12] and [18]) that provide the true boundary conditions of the problem, giving the flux entering the downstream section in terms of the one leaving the upstream section and vice versa, plus the injection terms. This naturally sets up an equation for the particle flux, equation (22).

4. We have solved this equation, equation (24), in a general way, which yields the particle spectrum both close to and far from the injection energies.

5. We have shown that under physically realistic conditions, the probabilities of returning to the shock, P_u and P_d, do not depend on the particles' energy p.

6. In the limit of large particle energies, we have shown that the spectrum is a pure power law in the particles' momentum, even when the probabilities of returning at the shock, P_u and P_d, depend on p (eq. [31]).

7. In this limit, we have then simplified our main equations, arriving at a system of equations (eq. [32], [33], or their combination, eq. [34]) and further showing that the requirement that equation (34) has the eigenvalue $\lambda = 1$ fixes the all-important asymptotic particle spectral index s.

8. We have recovered the low-speed, Newtonian limit of Bell (1978) and Blandford & Ostriker (1978) and provided a relativistic generalization (eq. [40]) that is, however, of little use because it requires knowledge of the particle anisotropic distribution function, which can be obtained only through the full solution of equation (34), which automatically incorporates equation (40).

Although the treatment presented in this paper may appear rather abstract, it is not without practical consequences. For instance, the fact that the asymptotic spectral index s is not asymptotic but applies as soon as $p > p_{\max} = p_{\rm inj}(1 + u_r)/(1 - u_r)$, where u_r is the modulus of the relative speed between the upstream and downstream fluids, implies that numerical simulations trying to compute s need not extend to very large particle energies but can save precious computational time by sticking to energies exceeding $p_{\max}$ by a small factor, dictated exclusively by numerical questions. In addition, the treatment presented in § 3 can be somewhat simplified by a judicious use of Fredholm's formulae, resulting in a considerable simplification for the calculational side of the theory. Finally, our results allow for the treatment of the hyperrelativistic limit, as will be shown in a future paper.

REFERENCES

Bednarz, J., & Ostrowski, M. 1998, Phys. Rev. Lett., 80, 3911
Bell, A. R. 1978, MNRAS, 182, 147
Berezinsky, V. S., Bulanov, S. V., Dogiel, V. A., Ginzburg, V. L., & Ptuskin, V. S. 1990, Astrophysics of Cosmic Rays (Amsterdam: North-Holland)
Blandford, R. D., & Ostriker, J. P. 1978, ApJ, 221, L29
Courant, R., & Hilbert, D. 1953, Methods of Mathematical Physics (New York: Wiley)
Freiling, G., Yurko, S., & Vietri, M. 2003, Lett. Math. Phys., submitted
Ginzburg, V. L., & Syrovatskii, S. I. 1964, The Origin of Cosmic Rays (Oxford: Pergamon)
Kirk, J. G., & Schneider, P. 1987, ApJ, 315, 425
Landau, L. D., & Lifshitz, E. M. 1987, Fluid Mechanics (Oxford: Pergamon)
Peacock, J. A. 1981, MNRAS, 196, 135