Mon. Not. R. Astron. Soc. **321**, 433–438 (2001)

Cosmic ray acceleration to very high energy through the non-linear amplification by cosmic rays of the seed magnetic field

A. R. Bell★ and S. G. Lucek
Blackett Laboratory, Imperial College, London SW7 2BZ

Accepted 2000 September 8. Received 2000 September 7; in original form 1999 August 9

ABSTRACT

The maximum energy for cosmic ray acceleration at supernova shock fronts is usually thought to be limited to around 10^{14}–10^{15} eV by the size of the shock and the time for which it propagates at high velocity. We show that the magnetic field can be amplified non-linearly by the cosmic rays to many times the pre-shock value, thus increasing the acceleration rate and facilitating acceleration to energies well above 10^{15} eV. A supernova remnant expanding into a uniform circumstellar medium may accelerate protons to 10^{17} eV and heavy ions, with charge Ze, to $Z \times 10^{17}$ eV. Expansion into a pre-existing stellar wind may increase the maximum cosmic ray energy by a further factor of 10.

Key words: acceleration of particles – magnetic fields – plasmas – shock waves – turbulence – cosmic rays.

1 INTRODUCTION

The theory of diffusive shock acceleration of cosmic rays (Axford, Leer & Skadron 1977; Bell 1978; Blandford & Ostriker 1978; Krymsky 1977; reviewed by Drury 1983) is well established, but it cannot easily account for the probably Galactic origin of cosmic rays (CR) between the spectral knee at 10^{14}–10^{15} eV and energies of 10^{18}–10^{19} eV, above which the origin is probably extragalactic (Axford 1994). Lagage & Cesarsky (1983) showed that the characteristic time for acceleration to a momentum p is $\tau = 4D(p)/u^2$ where u is the shock velocity and $D(p)$ is the spatial diffusion coefficient of cosmic rays with momentum p. In the time τ the shock travels a distance $4D(p)/u$. The shock eventually slows by interaction with the surrounding medium and this places a limit on the energy to which CR can be accelerated. The minimum realistic coefficient for diffusion that is not perpendicular to the magnetic field is $D(p) = r_g c/3 = pc/3ZeB$, where r_g is the CR gyroradius, Z is the charge on the CR in units of e and B is the magnetic field in front of the shock. Protons accelerated to 10^{15} eV in a magnetic field of 3×10^{-6} G by a shock travelling at $10^4\,\mathrm{km\,s^{-1}}$ require the shock to propagate at high velocity for a distance $4.4 \times 10^{17}\,\mathrm{m} = 14\,\mathrm{pc}$, taking a time $\tau = 1400\,\mathrm{yr}$. $D(p)$ can be less for perpendicular shocks when CR diffusion is across the magnetic field, but this is a special case. Acceleration to 10^{15} eV by supernova remnants (SNR) is marginal and acceleration to higher energies improbable unless the magnetic field is larger than typical interstellar values. McKenzie & Volk (1982) point out that CR might indeed be able to amplify the magnetic field above that in the undisturbed medium ahead of the shock. This would increase the rate of acceleration and facilitate acceleration to higher energies. Here we further investigate this possibility.

2 MAGNETIC FIELD GENERATION

CR in the vicinity of the shock are scattered by irregularities in the magnetic field. In the linear theory, these irregularities consist of Alfven waves generated by the CR themselves as they diffusively drift through the plasma upstream of the shock. If U_a is the energy density of the Alfven waves and P_{cr} is the CR pressure, then the growth and advection of U_a in the shock rest frame (the inertial frame in which the shock is at rest) obeys the equation

$$\frac{\partial U_a}{\partial t} + u\frac{\partial U_a}{\partial x} = v_a \frac{\partial P_{cr}}{\partial x}, \qquad (1)$$

where the plasma flows with positive velocity u into the shock from the direction of $x = -\infty$, the Alfven speed v_a is assumed to be much less than u, and Alfven wave damping is neglected. The waves propagate in the direction opposite to the plasma flow. In a steady state $\partial U_a/\partial t = 0$. If the background magnetic field is uniform then v_a is independent of x, and the equation can be integrated to give $U_a = v_a P_{cr}/u$. If P_{cr0} is the CR pressure at the shock, ρ is the upstream mass density, B is the background magnetic field and ΔB is the fluctuating magnetic field ($U_a = \Delta B^2/\mu_0$), then

$$\left(\frac{\Delta B}{B}\right)^2 = M_a \frac{P_{cr0}}{\rho u^2}, \qquad (2)$$

where $M_a = u/v_a$ is the Alfven Mach number of the shock. CR acceleration can be very efficient (Volk, Drury & McKenzie 1984;

★ E-mail: t.bell@ic.ac.uk

Bell 1987; Falle & Giddings 1987), $P_{\mathrm{cr}0} \sim \rho u^2$, and M_{a} is necessarily greater than 1, implying that the magnetic field is amplified non-linearly. Naive application of this relation gives $\Delta B/B \gg 1$ for parameters typical of young supernova remnants. A shock moving at $10^4\,\mathrm{km\,s^{-1}}$ into a medium containing 1 proton $\mathrm{cm^{-3}}$ and a magnetic field of 3×10^{-6} G has an Alfven Mach number of $M_{\mathrm{a}} = 1500$. A naive application of equation (2) suggests an increase in the magnetic field energy by a factor of 1000. The conclusion is first that the linear theory breaks down and secondly that it is at least possible that the magnetic energy density close to the shock is much greater than that in the undisturbed upstream plasma.

As the Alfven disturbance grows and $\Delta B/B$ approaches 1, it can no longer be considered as a linear Alfven wave and energy can be transferred by damping mechanisms and wave–wave coupling to the thermal plasma, to kinetic motions, to non-Alfvenic waves or to Alfven waves at other wavelengths. A full treatment of CR-excited Alfven turbulence is beyond the scope of this paper, and we propose the following approximate model.

In the linear case it was assumed that all Alfven disturbances propagate (relative to the plasma) towards $x = -\infty$ at the Alfven velocity. In the non-linear case, Alfven turbulence will consist of structures that move in all directions. We assume that the tensioning and relaxation of field lines isotropizes the motion of magnetic structures on a time-scale τ_{a}. Since the characteristic scale of the structures moving at velocity v_{a} is a CR gyroradius r_{g}, we assume that the characteristic winding and unwinding time of field lines is $r_{\mathrm{g}}/v_{\mathrm{a}}$, so we set $\tau_{\mathrm{a}} = r_{\mathrm{g}}/v_{\mathrm{a}}$. We notionally separate the Alfven turbulence into a component with energy density U_+ propagating at speed v_{a} in the positive x direction and a component with energy density U_- propagating at the same speed in the negative x direction. We define a mean local velocity of the Alfven turbulence as

$$v_{\mathrm{av}} = v_{\mathrm{a}} \frac{U_+ - U_-}{U_+ + U_-}, \tag{3}$$

and the rate at which CR transfer energy to Alfven turbulence (force acting through a distance) is

$$u \frac{\partial}{\partial x}(U_+ + U_-) = -v_{\mathrm{av}} \frac{\partial P_{\mathrm{cr}}}{\partial x}. \tag{4}$$

The CR pressure gradient $\partial P_{\mathrm{cr}}/\partial x > 0$ increases the energy density U_- of the reverse travelling component of the Alfven turbulence and correspondingly decreases U_+. We assume that the rate of turbulent energy transfer from U_- to U_+ is proportional to U_-, and that the rate of energy transfer from U_+ to U_- is similarly proportional to U_+. This allows us to propose the following equations, including both excitation and isotropization, for the evolution of each component:

$$u \frac{\partial U_-}{\partial x} = \frac{U_-}{U_+ + U_-} v_{\mathrm{a}} \frac{\partial P_{\mathrm{cr}}}{\partial x} - \frac{v_{\mathrm{a}}}{r_{\mathrm{g}}}(U_- - U_+); \tag{5}$$

$$u \frac{\partial U_+}{\partial x} = -\frac{U_+}{U_+ + U_-} v_{\mathrm{a}} \frac{\partial P_{\mathrm{cr}}}{\partial x} - \frac{v_{\mathrm{a}}}{r_{\mathrm{g}}}(U_+ - U_-). \tag{6}$$

The first term on the right-hand side of each equation represents wave excitation, and the second term represents wave isotropization. The energy given by CR to each component of the wave is allocated in proportion to its amplitude. In writing these equations we made the approximations that u is constant and that the advection of each component is at velocity u rather than $u \pm v_{\mathrm{a}}$, which would allow for the motion of components at velocity $\pm v_{\mathrm{a}}$ relative to the plasma. These approximations should be adequate for the high-energy CR, which in themselves have insufficient energy density to modify the shock strongly or increase the magnetic energy density to a value comparable with the plasma kinetic energy density. Low-energy cosmic rays, having smaller gyroradii, are confined more closely to the shock and have no effect on the regions upstream occupied by the highest energy cosmic rays.

Equations (5) and (6) determine the growth of the magnetic field. We also need an equation for the CR spatial distribution in response to scattering by the turbulence. CR are most strongly scattered by magnetic structures with scalelengths comparable to their gyroradius. For a complete treatment we would need to make U_+ and U_- functions of scalelength or wavenumber k, and we would have to make P_{cr} a function of energy or CR gyroradius. To simplify the analysis we confine it to CR with a characteristic gyroradius r_{g} and Alfven turbulence with a scalelength equal to r_{g}. We take U_+ and U_- to be the energy density of structures with wavenumber $k = 1/r_{\mathrm{g}}$ over waveband $\Delta k = k$, and assume that these structures interact with CR over a momentum range $\Delta p = p$, where p is the momentum of CR with a gyroradius r_{g}. We will then extend the argument to a larger range of k and E.

We assume that the CR interact diffusively with the Alfven turbulence with a diffusion coefficient $D = cr_{\mathrm{g}}/3$, where $r_{\mathrm{g}} = p/eB$ and $B = \sqrt{\mu_0(U_+ + U_-)}$. This corresponds to CR interacting with that part of the magnetic field constituting the Alfven turbulence with the required scalelength. The CR distribution is then determined by a balance between advection and diffusion:

$$\frac{\partial P_{\mathrm{cr}}}{\partial x} = \frac{uP_{\mathrm{cr}}}{D} = \frac{3euP_{\mathrm{cr}}}{pc}[\mu_0(U_+ + U_-)]^{1/2}. \tag{7}$$

The set of equations for U_+, U_- and P_{cr} is completed by an equation for the Alfven velocity v_{a}. In the region upstream of the shock occupied by the highest energy cosmic rays, the main contribution to the magnetic field comes from the Alfven turbulence generated by these CR. Hence we can write

$$v_{\mathrm{a}} = \left[\frac{U_+ + U_-}{\rho}\right]^{1/2}, \tag{8}$$

where ρ is the upstream plasma mass density, which is uniform because u is uniform.

Writing $W = U_- - U_+$, equations (5)–(8) can be manipulated to give the set of equations

$$2\rho^2 v_{\mathrm{a}}^2 u \frac{\partial v_{\mathrm{a}}}{\partial P_{\mathrm{cr}}} = W \tag{9}$$

$$u \frac{\partial W}{\partial P_{\mathrm{cr}}} = v_{\mathrm{a}}\left(1 - \theta \frac{W}{P_{\mathrm{cr}}}\right); \qquad \theta = \frac{2D}{ur_{\mathrm{g}}}, \tag{10}$$

$$\frac{\partial P_{\mathrm{cr}}}{\partial x} = \frac{u}{D} P_{\mathrm{cr}}. \tag{11}$$

Since P_{cr} increases monotonically with x, equations (9) and (10) can be viewed as coupled equations for W and v_{a}, which are both functions of P_{cr}. In this model, θ is independent of x. Since $D = r_{\mathrm{g}}c/3$, $\theta = 2c/3u$ and $\theta \gg 1$ for cases considered here. These equations can be put in dimensionless form:

$$\frac{\mathrm{d}\mu}{\mathrm{d}\eta} = \phi\left(1 - \frac{\mu}{\eta}\right); \qquad \frac{\mathrm{d}(\phi)^3}{\mathrm{d}\eta} = \frac{3}{2}\mu; \qquad \frac{\mathrm{d}\eta}{\mathrm{d}\lambda} = 2\eta\phi,$$

where

$$\mu = \theta^2 \frac{W}{\rho u^2}; \qquad \phi = \theta \frac{v_a}{u}; \qquad \eta = \theta \frac{P_{cr}}{\rho u^2};$$

$$\lambda = \frac{xue\sqrt{\mu_0 \rho}}{\theta pc} = \frac{xu}{\theta r_{g0} v_{a0}}. \tag{12}$$

r_{g0} and v_{a0} are the CR gyroradius and Alfven velocity in some reference magnetic field such as a typical interstellar field of 3 μG. In essence, for a fixed density ρ and flow velocity u, ϕ is a measure of magnetic field, η a measure of CR pressure, λ a measure of distance and μ a measure of the anisotropy in the Alfven turbulence. Note the importance of the parameter θ in scaling the equations. The first two equations of (12) can be solved, without reference to the third, for the magnetic field as a function of the CR pressure. The third equation then gives the magnetic field and CR pressure spatial profiles as functions of x.

The solution separates into two regimes:

(i) a large u/c, relatively low CR pressure, small η and $\mu \ll \eta$ regime (regime A) in which isotropization of the Alfven turbulence can be neglected and the first of equations (12) reduces to $d\mu/d\eta = \phi$;

(ii) a small u/c, high CR pressure, large η, regime (regime B) in which isotropization dominates and the advective term (term on the left-hand side) in the first of equations (12) can be neglected, giving $\mu = \eta$. Regime B, if it applies at all, applies close to the shock at late times when u/c is small.

A solution in regime A is $\mu = \eta^2/4$, $\phi = \eta/2$. This solution assumes that the magnetic field ($\propto\phi$) tends to zero as the CR pressure ($\propto\eta$) tends to zero. It neglects the interstellar magnetic field far upstream where the CR pressure is very small, but this neglect is acceptable because we are mainly interested in cases where the field is amplified by a large factor. The solution in regime B is $\mu = \eta$, $\phi = (3/4)^{1/3}(\eta^2 - \eta_0^2)^{1/3}$, where η_0 is a constant of integration. The solutions for μ cross at $\eta = 4$, and the solutions for ϕ match at this point if $\eta_0^2 = 16/3$. Making this approximation that the asymptotic solutions for regimes A and B hold for all $\eta < 4$ and all $\eta > 4$ respectively gives the overall solution

$$\phi = \begin{cases} \eta/2; & \mu = \eta^2/4 \quad \text{if } \eta < 4, \\ (3\eta^2/4 - 4)^{1/3}; & \mu = \eta \quad \text{if } \eta > 4. \end{cases} \tag{13}$$

This approximate use of asymptotic solutions in place of the correct solution of equations (12) causes a maximum error, close to $\eta = 4$, of 20 per cent in ϕ. This is a small error in comparison with the uncertainties of the model.

Reverting to dimensional quantities, the solution for the Alfven velocity can be written in terms of the ratio of the magnetic pressure to the kinetic energy density

$$\frac{B^2/2\mu_0}{\rho u^2/2} = \begin{cases} \frac{1}{4}\left(\frac{P_{cr}}{\rho u^2}\right)^2 & \text{if } \frac{P_{cr}}{\rho u^2} < \frac{6u}{c} (\text{Regime A}), \\ \left(\frac{9uP_{cr}^2}{8c\rho^2 u^4} - \frac{27u^3}{2c^3}\right)^{2/3} & \text{if } \frac{P_{cr}}{\rho u^2} > \frac{6u}{c} (\text{Regime B}). \end{cases} \tag{14}$$

It is likely that an SNR shock enters regime B late in the life of the SNR when $u \ll c$. Hence magnetic field amplification will be weak when the SNR is well into the Sedov phase. During the early free expansion phase, when the shock velocity is around or above 10 000 km s^{-1}, it is likely that the solution for Regime A applies throughout the upstream plasma since P_{cr} is the pressure of only the high-energy CR. In this case, the magnetic field is

$$B = \left(\frac{u}{10^4\,\text{km s}^{-1}}\right)\left(\frac{n_e}{\text{cm}^{-3}}\right)^{1/2}\left(\frac{P_{cr}}{0.1\rho u^2}\right) \times 2.3 \times 10^{-4}\,\text{G}. \tag{15}$$

The magnitude of the magnetic field depends upon the uncertain ratio of the CR pressure P_{cr} to the plasma momentum flux ρu^2. The total (thermal plus CR) pressure at a strong shock is $3\rho u^2/4$, so a CR modulated shock may have a CR pressure as large as $\rho u^2/2$. In our model, P_{cr} represents the pressure of CR at the highest energy over a momentum range $\Delta p = p$. In the standard linear theory of shock acceleration, the CR pressure and energy density are equally spread over every decade of the spectrum. If the spectrum spreads evenly between 10^9 and 10^{17} eV, then P_{cr} is equal to the total CR pressure over the whole spectrum divided by $\ln(10^{17}/10^9) = 18.4$, in which case, for a CR modulated shock, $P_{cr} \approx 0.03\rho u^2$. However, Bell (1987) showed that strong CR modulation flattens the CR spectrum, and much of the energy is given to the highest energy CR. In these conditions, P_{cr} might easily reach values as large as $0.1\rho u^2$, implying magnetic fields approaching a mG for a rapidly expanding SNR.

3 COMPARISON WITH SIMULATIONS

In a companion paper (Lucek & Bell 2000) we have simulated the CR excitation of Alfven turbulence. CR, modelled as particles, are initialized with a streaming velocity u, equivalent to the shock velocity. The magnetic field is initialized as a uniform field with a low-amplitude Alfvenic perturbation. The perturbation is amplified by the streaming CR to produce a tangled magnetic field with an amplitude many times that of the original nearly uniform field. The CR streaming is isotropized by the magnetic field. The effect is largest when the wavelength of the initial perturbation is about 1/5 of the initial CR gyroradius. As the magnetic field is amplified the CR gyroradius decreases until it matches the scalelength of the magnetic field structure as expected for strong interaction between waves and particles. The streaming is annulled in about one CR gyroperiod, increasing the magnetic energy density to $(B^2/2\mu_0)/P_{cr} \sim (u/c)^2$, which differs from $(B^2/2\mu_0)/P_{cr} = P_{cr}/8\rho u^2$ found for regime A in equation (14). For reasonable SNR parameters, the two estimates give similar absolute magnitudes but the processes are different because the simulated turbulence is not excited by a stationary pressure gradient. The turbulence is excited more rapidly in the simulation but this may be a result of the initial conditions, and the simulated turbulence is seen to grow less rapidly after one gyroperiod. It is encouraging that the simulation (i) demonstrates that a magnetic field can be amplified well above its seed value, (ii) shows that CR isotropization occurs on a time-scale of a gyroperiod, implying a CR mean path of a gyroradius and the validity of Bohm diffusion, and (iii) generates a magnetic field even more rapidly than we assume in this paper.

4 THE MAXIMUM COSMIC RAY ENERGY

Previous analysis (Lagage & Cesarsky 1983; Axford 1994) has shown that a SNR expanding into a typical interstellar magnetic field of a few μG is capable of accelerating protons to energies around the spectral knee at 10^{14}–10^{15} eV. Non-linear amplification of the magnetic field to 100 times this value reduces the CR gyroradius, the upstream scalelength of the CR distribution, and the acceleration time by a corresponding factor of 100. Since the maximum energy, as calculated by Lagage & Cesarsky, is determined by the acceleration rate, it appears that magnetic field amplification might offer an explanation for the continuation of the CR spectrum beyond the knee. As shown by Hillas (1984), the maximum CR energy depends upon the value of LvB, where L, v and B are the characteristic distance, velocity and magnetic field in the accelerating region. Across a wide range of possible acceleration sites, the value of this parameter limits the maximum energy to that of the knee. Magnetic field amplification offers a means to increase LvB and accelerate CR to higher energies.

The following argument is an adaptation for our circumstances of that by Lagage & Cesarsky. The acceleration rate is proportional to B/p. Acceleration is rapid at low momentum and slow at high momentum. The dependence of the acceleration rate on B means that acceleration is relatively more rapid at momenta at which a large number of CR have amplified the magnetic field. Because of these two considerations, a front in the CR density will develop in momentum space at some momentum p_{front}. At momenta below p_{front}, the CR density, B, and the acceleration rate are all large. Above the front, beyond p_{front}, there are few CR, B has not been amplified, and acceleration is slow. The rate at which p_{front} increases can be estimated from the equation for the CR phase space distribution function:

$$\frac{\partial f}{\partial t}+\frac{\partial (uf)}{\partial x}-\frac{1}{3}\frac{\partial u}{\partial x}\frac{1}{p^2}\frac{\partial (fp^3)}{\partial p}-\frac{\partial}{\partial x}\left(D\frac{\partial f}{\partial x}\right)=0, \tag{16}$$

where $f(x,p,t)$ is the isotropic part of the distribution function which is a function of position x, magnitude of momentum p and time. Integration in x, in the approximation that f is constant across the region in which u changes, gives

$$\frac{\partial}{\partial t}\int_{-\infty}^{\infty} f\,\mathrm{d}x=-u_2F-\frac{u_1-u_2}{3}\frac{1}{p^2}\frac{\partial (Fp^3)}{\partial p}, \tag{17}$$

where u_1 and u_2 are the upstream and downstream fluid velocities, respectively, in the shock frame, and F is the value of f at the shock. At the front in momentum space, the first term on the right-hand side of equation (17) is much smaller than the second term, so integration across the front from just below p_{front} to just above p_{front} gives

$$\begin{aligned}\frac{\mathrm{d}p_{\mathrm{front}}}{\mathrm{d}t}&=\frac{u_1-u_2}{3}p_{\mathrm{front}}F(p_{\mathrm{front}})\left[\int_{-\infty}^{\infty} f(p_{\mathrm{front}})\,\mathrm{d}x\right]^{-1}\\ &\approx\frac{u_1-u_2}{6L_{\mathrm{cr}}}p_{\mathrm{front}}\approx\frac{u^2eB}{2c},\end{aligned} \tag{18}$$

where we have made the approximation $\int_{-\infty}^{\infty} f(p_{\mathrm{front}})\,\mathrm{d}x=2L_{\mathrm{cr}}F(p_{\mathrm{front}})$ and used $L_{\mathrm{cr}}=D/u=(c/u)(r_{\mathrm{g}}/3)$. The factor of 2 in the approximation for the spatial integral allows for the CR spending equal times in the downstream and upstream regions. The upstream and downstream dwelling times are the same if both the plasma flow velocity and the CR diffusion coefficient decrease by the compression ratio at the shock as expected (Lagage & Cesarsky 1983). Equation (18) can be integrated over the expansion of the SNR to give the maximum CR energy E_{front} (in eV) after the SNR has expanded to a radius R:

$$E_{\mathrm{front}}=\frac{cp_{\mathrm{front}}(R)}{e}=\frac{1}{2}\left(\frac{P_{\mathrm{cr0}}}{\rho u^2}\right)\mu_0^{1/2}\int_{R_0}^{R} u^2\rho^{1/2}\,\mathrm{d}r, \tag{19}$$

where the magnetic field at any time is taken to be that immediately upstream of the shock $B^2/\mu_0=(P_{\mathrm{cr0}}/\rho u^2)^2\rho u^2$, as derived earlier, and P_{cr0} is the value of P_{cr} at the shock. R_0 is the shock radius when acceleration begins. This may be set equal to the radius of the progenitor star.

The SNR enters the self-similar Sedov phase when it has swept up more than its own mass of circumstellar material. During the Sedov phase, $u\propto R^{-3/2}$, so E_{front} increases little once this phase has been entered. Indeed, CR acceleration is restricted by the condition that the CR scaleheight upstream of the shock must be smaller than the radius of the SNR. Because magnetic field amplification weakens or ceases in the Sedov phase, thus increasing the CR diffusivity, acceleration cannot then take place beyond the usual Lagage & Cesarsky limit and CR previously accelerated above the limit may be able to escape the remnant.

The maximum CR energy $E_{\max}$ attained during the lifetime of the SNR can be estimated by integrating through the preceding free expansion, $u=$ constant, until $R=R_{\mathrm{free}}$, i.e. $E_{\max}=E_{\mathrm{front}}(R_{\mathrm{free}})$ where

$$\int_{R_0}^{R_{\mathrm{free}}} 4\pi\rho r^2\,\mathrm{d}r=M_{\mathrm{SN}} \tag{20}$$

and M_{SN} is the ejected mass.

5 A UNIFORM CIRCUMSTELLAR MEDIUM

We first calculate $E_{\max}$ for a uniformly dense circumstellar medium, giving

$$E_{\max}=\frac{1}{2}(\mu_0\rho)^{1/2}\left(\frac{P_{\mathrm{cr0}}}{\rho u^2}\right)u^2\left(\frac{3M_{\mathrm{SN}}}{4\pi\rho}\right)^{1/3}, \tag{21}$$

$$\begin{aligned}E_{\max}=&\left(\frac{u}{10^4\,\mathrm{km\,s^{-1}}}\right)^2\left(\frac{n_{\mathrm{e}}}{\mathrm{cm^{-3}}}\right)^{1/6}\left(\frac{P_{\mathrm{cr0}}}{0.1\rho u^2}\right)\left(\frac{M_{\mathrm{SN}}}{M_0}\right)^{1/3}\\ &\times 1.5\times10^{16}\,\mathrm{eV}.\end{aligned} \tag{22}$$

This result suggests that expansion into a uniform circumstellar medium can accelerate CR above the knee. Since the SNR expansion velocity u can be as large as $40\,000\,\mathrm{km\,s^{-1}}$, proton acceleration to 10^{17} eV, and acceleration of heavier ions above 10^{18} eV, is feasible.

6 EXPANSION INTO A STELLAR WIND

Volk & Biermann (1988) suggested that CR could be accelerated to higher energies if the SNR expanded into a pre-existing stellar wind from the giant progenitor to a Type II SN. They estimated that the magnetic field in the wind, transported by the wind from the surface of the star, could be many Gauss, resulting in rapid CR acceleration. Their assumed stellar magnetic field may be unreasonably large (Axford 1994), and here we consider the possibility that the field might be amplified by CR streaming. As reference parameters we take those appropriate for SN1993J in which the SN shock expands at a velocity of around $20\,000\,\mathrm{km\,s^{-1}}$ into a stellar wind with a mass loss rate $\dot{M}\cong 5\times10^{-5}M_0\,\mathrm{yr^{-1}}$ at a velocity $v_{\mathrm{wind}}\cong 10\,\mathrm{km\,s^{-1}}$ (values taken from Fransson & Bjornsson 1998). The progenitor was probably a red giant (e.g. Marcaide et al. 1997).

In a steady wind, the circumstellar density is proportional to R^{-2} such that $4\pi R^2\rho v_{\mathrm{wind}}=\dot{M}$. Substituting this density into equation (19), the maximum energy resulting from acceleration during the free expansion phase is

$$E_{\max} = \frac{1}{2}\left(\frac{P_{\rm cr0}}{\rho u^2}\right)u^2\left(\frac{\mu_0 \dot{M}}{4\pi v_{\rm wind}}\right)^{1/2}\ln\left(\frac{R_{\rm free}}{R_0}\right)$$

$$= \left(\frac{u}{2\times10^4\,{\rm km\,s^{-1}}}\right)^2\left(\frac{\dot{M}}{5\times10^{-5}M_0\,{\rm yr^{-1}}}\right)^{1/2}\left(\frac{P_{\rm cr0}}{0.1\rho u^2}\right)$$

$$\times\left(\frac{v_{\rm wind}}{10\,{\rm km\,s^{-1}}}\right)^{-1/2}\left[\frac{\ln(R_{\rm free}/R_0)}{9}\right]\times1.0\times10^{18}\,{\rm eV}, \quad (23)$$

where the value of the logarithmic term is based on a choice of $R_0 = 10^{13}$ m. The maximum CR energy increases equally in each decade of increase in radius, as shown by the logarithmic dependence on $R_{\rm free}$. Hence $E_{\max}$ is not strongly dependent on the initial and final expansion radii (R_0 and $R_{\rm free}$). The increase in $E_{\max}$ over the uniform density case is a result of the increased density and magnetic field (dependent on density as given by equation 15), especially during the early stages of the SN explosion. These are given by

$$n_{\rm e} = \left(\frac{\dot{M}}{5\times10^{-5}M_0\,{\rm yr^{-1}}}\right)\left(\frac{v_{\rm wind}}{10\,{\rm km\,s^{-1}}}\right)^{-1}\left(\frac{R}{10^{13}\,{\rm m}}\right)^{-2}$$

$$\times1.5\times10^8\,{\rm cm^{-3}}, \quad (24)$$

$$B = \left(\frac{u}{2\times10^4\,{\rm km\,s^{-1}}}\right)\left(\frac{\dot{M}}{5\times10^{-5}M_0\,{\rm yr^{-1}}}\right)^{1/2}\left(\frac{P_{\rm cr0}}{0.1\rho u^2}\right)$$

$$\times\left(\frac{v_{\rm wind}}{10\,{\rm km\,s^{-1}}}\right)^{-1/2}\left(\frac{R}{10^{13}\,{\rm m}}\right)^{-1}\times5.6\,{\rm G}. \quad (25)$$

The large magnetic field when R is small yields very rapid acceleration. For comparison, Fransson & Bjornsson deduce from observations of SN 1993J a magnetic field which is an order of magnitude larger than that given by equation (15), $B \cong 64(R/10^{13}\,{\rm m})^{-1}$ G, which they suggest is the result of post-shock turbulent amplification. However, their magnetic field is probably a post-shock value, which will be 2–4 times the pre-shock value, and our value of B is that on the scalelength of the gyroradius of the highest energy CR. When shock compression and the contribution of short scalelength fields is allowed for (e.g. by putting $P_{\rm cr0} \approx 0.5\rho u^2$), then our estimate matches the measurement by Fransson & Bjornsson surprisingly well. It is notable that the dependence on R is the same. In any case, it is encouraging that their results show that our estimated magnetic field is not extravagantly large.

This looks very encouraging, but there is another limit on $E_{\max}$ that we must consider, which in fact restricts $E_{\max}$ to a lower value than that given in equation (23). So far, we have considered a limitation on the time available for acceleration. There is also a limitation on space. The CR scaleheight upstream of the shock, $L_{\rm cr}$, cannot exceed the radius of the SNR:

$$L_{\rm cr} = \frac{1}{3}\frac{c}{u}r_{\rm g} = \frac{1}{3}\frac{c}{u}\frac{p_{\max}}{eB} \leqslant R, \quad (26)$$

where $E_{\max} = cp_{\max}/e$. This spatial limit is close to that obtained by Volk & Biermann by including adiabatic losses, and indeed the two are closely connected. Observations of SN1006 (Tanimori et al. 1998), show that, in at least some cases of CR acceleration, the maximum CR energy does not fall far short of this limit. For constant u and expansion into a stellar wind, $B \propto \rho^{1/2} \propto R^{-1}$, so the spatial limit on $E_{\max}$ is independent of R for expansion into a stellar wind. The spatial limit on the CR energy is

$$E_{\max} = \left(\frac{u}{2\times10^4\,{\rm km\,s^{-1}}}\right)^2\left(\frac{\dot{M}}{5\times10^{-5}M_0\,{\rm yr^{-1}}}\right)^{1/2}$$

$$\times\left(\frac{P_{\rm cr0}}{0.1\rho u^2}\right)\left(\frac{v_{\rm wind}}{10\,{\rm km\,s^{-1}}}\right)^{-1/2}\times3.4\times10^{17}\,{\rm eV}. \quad (27)$$

Comparison with the temporal limit shows that this is tighter by a factor $\ln(R_{\rm free}/R_0)/3$. As the spatial limit dominates, and this limit is independent of the radius of the shock front, acceleration to the limit occurs before the end of the free expansion phase. This spatial limit does not affect the acceleration in a uniform circumstellar medium discussed in Section 5.

Another possible limit on proton acceleration might be nuclear interactions with background matter. A CR proton has a cross-section of $\sigma = 3\times10^{-30}\,{\rm m^2}$ for interaction with other protons (Harwit 1973), giving rise to a loss time $\tau_{\rm loss} = (n_{\rm e}c\sigma)^{-1} = 10^{15}(n_{\rm e}/{\rm cm^{-3}})^{-1}$ s. Equation (24) for $n_{\rm e}$ in a stellar wind, implies, for $\dot{M} \cong 5\times10^{-5}M_0\,{\rm yr^{-1}}$ and $v_{\rm wind} \cong 10\,{\rm km\,s^{-1}}$, a loss time of around 10^7 s at a radius of 10^{13} m, which is an order of magnitude longer than the SNR expansion time and therefore cannot inhibit acceleration. As the SNR radius increases, the nuclear loss time increases (inversely proportional to density) as R^2, whereas the acceleration and expansion times increase only as R. Nuclear losses do not inhibit the CR acceleration in the circumstances considered here, but they might be important in circumstances that are not much more extreme.

Nuclear losses might not inhibit acceleration but they are large enough to suggest substantial gamma-ray emission up to PeV energies through the production and decay of neutral pions. In the conventional theory of shock acceleration, the highest energy CR are produced in the Sedov phase and gamma-ray emission is relatively low during the earlier free expansion phase. In contrast, if CR streaming amplifies the field as suggested here, the highest energy CR are produced during free expansion. If the SN expands into a pre-existing stellar wind, the gamma-ray luminosity is greatest early in the free expansion phase. Although the number of accelerated CR increases with shock radius $r_{\rm s}$ during free expansion into a pre-existing wind, the characteristic density, and hence the rate of gamma-ray production by each CR, decreases as $r_{\rm s}^{-2}$. This suggests that SNR should be most visible at gamma-ray energies soon after the shock initially breaks out of the progenitor atmosphere. If this scenario is correct, the emission of gamma rays at the highest energy may cease during the Sedov phase owing to the cessation of field amplification and the probable escape from the SNR of the highest energy CR. It has been shown (Drury, Aharonian & Volk 1994) that CR-produced gamma rays from SNR should be on the verge of detectability. The current upper limits on TeV gamma-ray emission (Buckley et al. 1998) are beginning to place constraints on the theory of diffusive shock acceleration. Buckley et al. concentrated on observations of SNR in the Sedov phase in the expectation that the SNR should then be most visible at high energies. Our model suggests that SNR might be more visible during free expansion.

7 CONCLUSIONS

A simple model for MHD turbulence suggests that CR might amplify the pre-shock magnetic field to a magnitude much greater than the seed interstellar value. This increases the rate of CR acceleration, facilitating acceleration to energies above the knee in the CR spectrum. SNR expansion into a uniform circumstellar medium should be able to accelerate protons to 10^{17} eV and heavy ions above 10^{18} eV. Expansion into a pre-existing stellar wind should enable acceleration to higher energies. For parameters suitable for SN1993J, protons may be accelerated to 3×10^{17} eV. The maximum energy is proportional to the square of the expansion velocity. Expansion at 40 000 km s^{-1} into the circumstellar environment of SN1993J would accelerate protons to 10^{18} eV and heavy ions to $Z \times 10^{18}$ eV. Hence it appears that CR amplification of magnetic field can provide the explanation for the origin of CR between the spectral knee at 10^{15} eV and the onset of extragalactic acceleration at 10^{18}–10^{19} eV (Wdowczyk & Wolfendale 1989; Axford 1994). It is encouraging that our estimated magnetic fields for expansion into a stellar wind are in line with observational estimates, and that computational simulations suggest, if anything, more rapid growth of magnetic field than we assume here. Our model implies that the highest energy CR are produced in the free expansion phase instead of the Sedov phase, and that the gamma-ray luminosity of a SNR expanding into a pre-existing stellar wind is greatest at the beginning of the free expansion phase.

Because CR both above and below the knee are accelerated by SNR, our theory has the advantage of naturally producing a CR spectrum which connects smoothly the spectrum below the knee to the steeper spectrum above the knee. The spectral steepening at the knee probably arises from source statistics, because (i) magnetic field amplification and acceleration above the knee only takes place in young freely expanding SNR, and (ii) acceleration to the highest energies in the range 10^{15}–10^{18} eV is dependent on the presence of a pre-existing stellar wind.

ACKNOWLEDGMENTS

This work was funded by the UK Particle Physics and Astronomy Research Council (grant GR/L 21129). We thank Peter Meikle of Imperial College for helpful advice and gratefully acknowledge helpful discussions with Luke Drury and other members of the European TMR network in Astro Plasma Physics.

REFERENCES

Axford W. I., 1994, ApJS, 90, 937
Axford W. I., Leer E., Skadron G., 1977, Proc. 15th Int. Cosmic Ray Conf., 11, 132
Bell A. R., 1978, MNRAS, 182, 147
Bell A. R., 1987, MNRAS, 225, 615
Blandford R. D., Ostriker J. P., 1978, ApJ, 221, L29
Buckley J. H. et al., 1998, A&A, 329, 639
Drury L.O'C., 1983, Rep. Prog. Phys., 46, 973
Drury L.O'C., Aharonian F. A., Volk H. J., 1994, A&A, 287, 959
Falle S. A. E. G., Giddings J. R., 1987, MNRAS, 225, 399
Fransson C., Bjornsson C.-I., 1998, ApJ, 509, 861
Harwitt M., 1973, Astrophysical Concepts. John Wiley & Sons, New York
Hillas A. M., 1984, ARA&A, 22, 425
Krymsky G. F., 1977, Sov. Phys. Dokl., 23, 327
Lagage O., Cesarsky C. J., 1983, ApJ, 125, 249
Lucek S. G., Bell A. R., 2000, MNRAS, 314, 65
Marcaide J. M. et al., 1997, ApJ, 486, L31
McKenzie J. F., Volk H. J., 1982, A&A, 116, 191
Tanimori T. et al., 1998, ApJ, 497, L25
Volk H. J., Biermann P. L., 1988, ApJ, 333, L65
Volk H. J., Drury L.O'C., McKenzie J. F., 1984, A&A, 130, 19
Wdowczyk J., Wolfendale A. W., 1989, Ann. Rev. Nucl. Part. Sci., 39, 43

This paper has been typeset from an electronic file prepared by the author.