CONFINEMENT–DRIVEN SPATIAL VARIATIONS IN THE COSMIC RAY FLUX

PAOLO PADOAN$^1$ AND JOHN SCALO$^2$

ABSTRACT

Low–energy cosmic rays (CRs) are confined by self–generated MHD waves in the mostly neutral ISM. We show that the CR transport equation can be expressed as a continuity equation for the CR number density involving an effective convection velocity. Assuming balance between wave growth and ion–neutral damping, this equation gives a steady–state condition $n_{CR} \propto n_i^{1/2}$ up to a critical density for free streaming. This relation naturally accounts for the heretofore unexplained difference in CR ionization rates derived for dense diffuse clouds (McCall et al. 2003) and dark clouds, and predicts large spatial variations in the CR heating rate and pressure.

Subject headings: ISM: cosmic rays

1. INTRODUCTION

Low–energy cosmic rays (CRs) are coupled to a number of important astrophysical processes. Most of the integrated interstellar CR number density, energy density, pressure and ionization rate are contributed by the low–energy part of the spectrum because of its steepness (see compilation in Antoni et al. 2004). 1–10 GeV CRs control the ionization fraction in the Earth’s lower atmosphere, and so cloud formation and lightning production (e.g. Stozhkov 2003). Tropospheric and stratospheric chemistry are affected, even at energies as low as $\sim 10$ MeV (e.g. Crutzen et al. 1973). In protostellar disks, ionization by ambient low–energy CRs may regulate disk turbulence and affect planet formation (e.g. Matsumura & Pudritz 2002). On the scale of the interstellar medium, low–energy CRs of $\sim 100$ MeV (Webber 1998) dominate the ionization fraction and heating of cool neutral gas, especially dark UV–shielded molecular regions (Goldsmith & Langer 1978). On larger scales, CR “pressure”, most of which is contributed by low–energy CRs, may help confine the Galactic disk and drive the Parker instability (see Hanasz & Lesch 2000), which may itself contribute to the formation of large condensations that form stars, and even drive a Galactic dynamo (Parker 1992); the CR pressure may also affect the hot coronal ISM (Schlickeiser & Lerche 1983). For these reasons, any strong spatial variations of the CR number density could lead to important thermal, chemical, and dynamical effects.

In almost all previous work it is assumed that the CR density and spectrum do not vary significantly in space for length scales smaller than the scale of variation in space density of CR sources or ionization and spallation losses (e.g. Hunter et al. 1997; Wolfire et al. 2003). The primary rationale is that a superposition of CRs propagating diffusively to a point in the Galaxy from many stochastic sources (e.g. supernova remnants) gives an rms variation in CR density of order 1% for typical values of parameters (Lee 1979; see Berezinski et al. 1990, sec. III.10). A more detailed calculation including the spatial distribution and discreteness of sources gives variations that are typically less than 10–20% (Busching et al. 2005). This argument is not valid for low energy CRs because it neglects the possibility that self–confinement of CRs (see below) or ionization losses greatly reduce their propagation distance from the sources. The evidence for CR homogeneity inferred from EGRET $\gamma$–ray data (e.g. Digel et al. 2001) applies only to a spatial resolution of approximately 5 degrees and to CR energies larger than of interest here. Furthermore, the $\gamma$–ray emissivity is derived from an integral over long lines of sight, to which individual clouds or cores may be small contributions, especially at low Galactic latitudes (see Aharonian 2001).

The purpose of this Letter is to show that large spatial variations of the CR number density should indeed exist in the ISM, based on the standard CR transport equation (§ 2), if low–energy CRs are confined along flux tubes by scattering from self–generated Alfvén waves (§ 3). Skilling & Strong (1976) examined a model to exclude CRs from molecular clouds based on screening the CRs by ionization losses enhanced by CR self–confinement, but their model assumes that CR self–confinement sets in suddenly at the edges of molecular clouds because of an extremely large assumed cloud column density. The process examined here is completely independent of that model. Multiple magnetic mirrors could also lead to CR variations of a factor of a few, depending on adopted parameters (Cesarsky & Volk 1978), a process that could be more important in the presence of tangled fields. In reality several effects may contribute to variations on roughly the same scale, but the effect found here has not been previously recognized, and is capable of giving CR variations up to two orders of magnitude. We show that, in a steady state, the CR density should locally scale with the square root of the ion density, up to densities above which damping of the waves allows the CRs to stream freely (§ 3), giving a sharp decline at densities around 500 cm$^{-3}$ (for 100 MeV protons), typical of the transition from diffuse gas to dark molecular clouds. These variations may explain the apparent discrepancy between the large ionization rate derived by McCall et al. (2003) for the diffuse region along the line of sight to Persei and the ionization rate estimated in

---

$^1$ Department of Physics, University of California, San Diego, CASS/UCSD 0424, 9500 Gilman Drive, La Jolla, CA 92093-0424; ppadoan@ucsd.edu

$^2$ Department of Astronomy, University of Texas, Austin, TX 78712; parrot@astro.as.utexas.edu
dark molecular clouds (§ 4). They should also result in large variations in CR pressure and heating rate in the ISM and elsewhere.

2. CONTINUITY EQUATION FOR SELF–CONFINED COSMIC RAYS

It is well–known that CRs streaming along field lines at a speed larger than the Alfvén speed generate MHD waves (Lerche [1965], Wentzel [1968]). The CRs in turn interact with these self–generated waves through resonant pitch–angle scattering with waves whose wavenumbers are multiples of the particle gyroradius for a particle of a given energy (see Cesarsky & Kulsrud [1973]). As shown by Skilling [1971] and others, the waves keep the CRs confined to stream at a speed only slightly larger than the Alfvén speed, if the waves are not efficiently damped. Models of Galactic CR propagation that rely on this effect in order to increase the escape time of CRs from the Galaxy are called self–confinement models, and are reviewed in Wentzel [1974] and Cesarsky [1980]. Self–confinement is generally believed to be efficient for CRs of energies less than about 100 GeV in mostly neutral interstellar material (Kulsrud & Cesarsky [1971]).

To demonstrate how self–confinement can lead to large CR density variations, we begin with the CR transport equation. The usual kinetic equation for the phase space distribution, \( f(x,p,t) \), of CRs interacting with a background plasma or self–generated MHD wave field can be derived from the Vlasov equation, coupled with the CR particle equation of motion and Maxwell’s equations. A number of reasonable assumptions allows transformation of this equation into an equation for the distribution function of guiding centers, assuming the CR anisotropy is small. This equation was derived and discussed in various forms by Kulsrud & Pearce [1966, 1975]; Skilling [1971], Earl [1974], and many others. Lu & Zank [2001] and Lu et al. [2002] give a useful collection of references (see Berezinskii et al. 1990, Schlickeiser 2002 for detailed derivations).

Neglecting diffusion perpendicular to the magnetic field, introducing a non–relativistic background medium velocity with a Galilean transformation (\( V \ll c \)) and including continuous momentum losses, the result is

\[
\frac{\partial f}{\partial t} + \nabla \cdot (V f) - \nabla \cdot [k(x,p) \nabla g] + \frac{p}{3} \nabla \cdot \frac{\partial f}{\partial p} = 0
\]

This well–known transport equation, neglecting the terms on the right, was derived phenomenologically by Parker [1965]. It is especially common in studies of heliospheric cosmic ray transport (Ferreira & Potgieter 2004).

On the lhs the second term is the convection of \( f \) by the background plasma bulk motion at velocity \( V \), which can include guiding center drift velocities (important in heliospheric transport but not in most ISM conditions). When this background motion is due to MHD waves, the appropriate velocity depends on the distribution of the directions of wave propagation relative to the streaming of the CRs. If the waves are self–generated, as we assume here, the directions of the waves and CRs are the same, and \( V \) can be replaced by the Alfvén speed, \( V_A \) (Berezinskii et al. 1990, ch. 10). The third term represents the interaction of the CRs with the MHD waves in the diffusion approximation; \( k(x,p) \) is the pitch–angle averaged spatial diffusion coefficient, which derives from the slight anisotropy of the CR distribution. The fourth term, often called the adiabatic or the Compton–Getting term, represent momentum convection. On the rhs the first term represents momentum (or energy) diffusion, the second term continuous momentum losses due to interactions with plasma particles (e.g. spallation, ionization and radiation losses), and \( S(x,p,t) \) is the source distribution function.

We can safely neglect all the terms on the rhs. It is well–known that momentum or energy diffusion is slow compared to pitch–angle diffusion (transformed into the spatial diffusion on the lhs), by a factor of order \( V_A/c \) (see Berezinskii et al. 1990, ch. 10). We neglect ionization losses because they require a column density of order 100 g/cm² for 100 MeV protons, corresponding to large length scales even when tangled fields or self–confinement are taken into account. The sources (e.g. supernova remnants) can be assumed far from the relatively small (~0.01 to 10 pc) volume under consideration. Similarly, we assume the mean magnetic field varies only over scales much larger than the scattering mean free path, so that we neglect mirroring and drifts due to mean field variations (weak focusing limit). Our resulting scale of variations is probably comparable to that of field variations, so a full calculation should include this focusing term.

Neglecting all terms on the rhs, rearranging the advective term with the Compton–Getting term, and multiplying by \( 4p^2 \), equation 1 yields

\[
\frac{\partial g}{\partial t} + \nabla \cdot (V g) - \nabla \cdot [k(x,p) \nabla g] = \frac{1}{3} \nabla \cdot \frac{\partial f}{\partial p} (pg) \quad (2)
\]

where \( g(x,p,t) = 4\pi p^2 f(x,p,t) \) is the differential number density of CR particles. An integration over \( p \) then gives (assuming that \( p g \) vanishes at infinity)

\[
\frac{\partial n_{cr}}{\partial t} + \nabla \cdot [V n_{cr} - \tilde{k}(x) \nabla n_{cr}] = 0 \quad (3)
\]

where \( n_{cr}(x) = \int_0^\infty g(x,p,t) dp \) is the total number density of CR particles, and \( \tilde{k} \) is the momentum–averaged spatial diffusion coefficient, \( \tilde{k}(x) = \int_0^\infty k(x,p) g(x,p,t) dp/n_{cr}(x,t) \).

Equation 3 can be cast in the form of a continuity equation by noticing that the diffusion along the field is the divergence of a flux that can be represented by the product of a diffusive streaming speed, \( V_{\text{diff}} \), and the CR number density, \( n_{cr} \). We can then define an effective CR streaming speed, \( V_{st} \), as the sum of the convection velocity, \( V \), and the diffusive streaming speed, \( V_{\text{diff}} \). In that case, an equation expressing a steady state for the CR number density in an Eulerian frame is

\[
\nabla \cdot (V_{st} n_{cr}) = 0 \quad (4)
\]

A similar connection between CR transport and a continuity equation has been noted, in different contexts, by Skilling (1971, 1975a), Earl (1974), Schlickeiser & Lerche (1985), Beeck & Wibberenz (1986) and Bieber (1987). Integrating equation 4 over a small volume, using the divergence theorem with a closed surface that corresponds to a segment of a magnetic flux tube, and using the inverse proportionality between the flux tube
cross section and the magnetic field strength, \( B \), we obtain
\[
\frac{\nu_{st} n_{cr}}{B} = \text{const} \quad (5)
\]
As significant spatial variations of both \( B \) and \( \nu_{st} \) are certainly present in the ISM, we should also expect significant variations in the CR number density, \( n_{cr} \), CR pressure, \( P_{cr} \), and CR ionization rate, \( \zeta_{cr} \), as discussed in § 4.

3. COSMIC RAY STREAMING VELOCITY

We can show how \( n_{cr} \) and associated quantities should vary with ISM parameters by deriving an expression for \( \nu_{st} \). If the CR scattering is primarily due to resonant scattering off magnetic waves generated by the CRs themselves, \( \nu_{st} \) can be computed by requiring that the wave growth rate is balanced by the wave damping rate (see Wentzel 1974 sec. 2.3.2.5). Considering only protons, the growth rate of waves propagating in the direction of the magnetic field, \( \mathbf{B} \), as a function of the CR streaming velocity along \( \mathbf{B} \), \( \nu_{st} \) is (eg Kulsrud & Cesarsky 1971):
\[
\Gamma(k_z) = \frac{\pi (\gamma - 3) \Omega_0 n_{cr}(k_z) m_H}{4(\gamma - 2) n_i m_i} \left( \frac{3 \nu_{st}}{\gamma V_A} - 1 \right) \quad (6)
\]
where \( \Omega_0 = eB/m_Hc \) is the non-relativistic cyclotron frequency of a proton of mass \( m_H \), the CR spectrum has been assumed to be a power law in momentum with exponent \( \gamma \) (empirically \( \gamma = 4.7 \) for \( E \geq 10 \) GeV), \( n_i \) is the ion mass, and \( n_{cr}(k_z) \) is the number density of protons with momentum \( p > eB/k_zc \), corresponding to the resonant condition.

In the mostly neutral ISM damping is primarily due to collisions of ions with neutral particles (Kulsrud & Pearce 1969), because neutrals do not take part in the wave motion, as the wave frequency is larger than the ion-neutral collision frequency at the scales of interest.\(^3\) The ion–neutral damping rate is then:
\[
\Gamma_{in} = \frac{1}{2} n_i \langle \sigma v \rangle_{in} m_n m_i \quad (7)
\]
We take the collision rate \( \langle \sigma v \rangle_{in} = 2.1 \times 10^{-9} \text{ cm}^3\text{s}^{-1} \) in diffuse regions and \( \langle \sigma v \rangle_{in} = 1.6 \times 10^{-9} \text{ cm}^3\text{s}^{-1} \) in molecular regions (Osterbrock 1961), assuming \( n(\text{He})/[n(\text{H}) + 2n(\text{H}_2)] = 0.14 \).

Assuming a balance of wave growth and damping, \( \Gamma = \Gamma_{in} \) (such balance is reached in less than a year for typical parameters in the absence of nonlinear wave cascades; see Kulsrud & Pearce 1969), equations (6) and (7) give an expression for the CR streaming velocity (Wentzel 1969):
\[
\nu_{st}(p) = \frac{\gamma V_A}{3} \left[ 1 + \frac{2(\gamma - 2)}{\pi (\gamma - 3) \Omega_0} \frac{n_i}{n_{cr}(p) m_H} \right] \quad (8)
\]
The importance of the second term on the r.h.s. of equation (8) depends on the gas density and fractional ionization. For mostly neutral ISM conditions and for ionizing protons of energy \( E \lesssim 100 \) MeV, the second term is \( \ll 1 \) up to a gas density \( n_{H2} \approx 500(B_0/10 \mu G) \text{ cm}^{-3} \), and low energy CRs are well confined, streaming at a velocity of order the Alfvén speed. In this case, equation (8) yields:
\[
n_{cr} \propto n_i^{1/2} \quad (9)
\]
Notice that the magnetic field strength has dropped out for this regime; it does affect the value of \( P_{H2,fs} \) above which eq. (9) no longer holds. As the gas density is increased above \( n_{H2,fs} \), the second term on the r.h.s. of equation (8) becomes important and the CR streaming velocity approaches the particle velocity, causing a drop in the value of \( n_{cr} \), based on equation (9).

4. RESULTS AND CONCLUSIONS

Figure 1 shows the result, for 100 MeV protons, of iteratively solving equations (8) and (9) for the dependence of total CR number density on total gas density, taking \( n_{cr} = n_{cr} \). We have assumed \( n_i = 1.4 \times 10^{-4} n \) (Cardelli et al. 1996) and \( B = B_0 \) in diffuse clouds, and \( n_i = 10^{-4}(n/10^{-4} \text{ cm}^{-3})^{1/2}(\zeta_{cr}/2 \times 10^{-17} s^{-1})^{1/2} \) (eg Elmegreen 1979) and \( B = B_0(n/200 \text{ cm}^{-3})^{1/2} \) (Crutcher 1999; Bourke et al. 2001) in molecular clouds, with \( B_0 = 10 \mu G \). These are crude approximations due to the large observed scatter. The vertical normalization assumes a CR energy density of approximately \( 1 \text{ eV/cm}^3 \) at \( n = 0.1 \text{ cm}^{-3} \), corresponding to demodulation of the energy spectrum near the Sun (Webber 1998).

Corresponding variations are expected also for the CR pressure, \( P_{cr} = \frac{1}{2} c \int_0^{\infty} \beta(p) g(x,p,t) dp \) (\( \beta \approx 1 \) for relativistic CRs, and \( \beta \approx p/(mc) \) for non-relativistic CRs) and for the CR ionization rate \( \zeta_{cr} = C_\zeta \int_0^{\infty} \sigma_i(p) dp \) where \( \sigma_i(p) \) is the cross section for the ionization of a hydrogen atom by a CR particle of momentum \( p \), and the factor \( C_\zeta \) accounts for heavy CR particles and for secondary electrons (eg Spitzer & Tomasko 1965). These integrals will be computed elsewhere. Here we only stress that \( P_{cr} \) and \( \zeta_{cr} \) clearly increase with \( n_{cr} \). If they are nearly proportional to \( n_{cr} \), we then expect them to be nearly proportional to \( n_i^{1/2} \) as well, based on equation (9). The

---

\(^3\) Nonlinear cascade damping of hydromagnetic waves (e.g. Skilling 1972; Farmer & Goldreich 2003) is slow compared to collisional damping in the cool mostly neutral ISM.
values of $P_{cr}$ and $\zeta_{cr}$ should also drop when the density is large enough to cause the free-streaming of the CRs (when the second term on the r.h.s. of equation (8) is large). Such drop should occur at a gas density above $\approx 500(B_0/10 \, \mu G) \, \text{cm}^{-3}$ in mostly neutral diffuse regions, and above $\approx 10^5(B_0/10 \, \mu G) \, \text{cm}^{-3}$ in dark molecular cores, where $n_i$ is much smaller than in diffuse regions. The precise value of this critical density for free-streaming depends on the magnetic field strength and on the normalization of the CR spectrum. Furthermore, $P_{cr}$ and $\zeta_{cr}$ should decrease by a factor of 10–50 in the transition from diffuse regions to dark molecular cores, also due to the reduced value of $n_i$.

Our result accounts for the heretofore unexplained enhancement of the CR ionization rate in $\zeta$ Persei diffuse gas (McCall et al. 2003; see also Le Petit et al. 2004) compared to dense molecular clouds, where it is estimated using molecular abundance ratios (Williams et al. 1998, Doty et al. 2002, Padoan et al. 2004, and references therein) or from the HI/H$_2$ ratio (Goldsmith & Li 2005, see Liszt (2003) for a discussion of other evidence concerning the CR ionization rate). Density estimates for the $\zeta$ Persei gas using a variety of techniques are in the range 100–400 cm$^{-3}$, putting that gas near the upper limit of our predicted CR density for diffuse gas, just before the free-streaming regime. In molecular regions, shielding from UV radiation allows carbon to remain in the form of atomic or molecular carbon, resulting in much smaller ion density than in diffuse regions, despite the large total gas density. This ion density is low enough that self-confinement is effective (up to a density of approximately $10^5$ cm$^{-3}$), but with a larger streaming speed (larger ion Alfvén speed) than in diffuse regions. As a result, the CR density can be 10–50 times lower than in diffuse regions, due to the constant-flux constraint expressed by equation (4) (see Figure 1). Figure 1 should not be interpreted as predicting a one-to-one or universal relation. For example, we expect a large dispersion in the relation between magnetic field strength and gas density and the CR flux normalization may vary with position relative to nearby CR sources.

We have not discussed ionization due to electrons because the electron spectrum is very uncertain (see Casadei & Bindi 2004), especially below 100 MeV, where it is sensitive to the models used to demodulate the CR electron flux at the Earth (e.g., Webber 1998) or used to disentangle the $\gamma$-ray bremsstrahlung, inverse compton and unresolved point source emission at MeV energies (e.g., Strong 2001). However, CR electrons, like CR protons, are confined to magnetic waves and should therefore follow the CR protons (Melrose & Wentzel 1970), and have the same density variations, although the critical densities for free-streaming may be different.

This work points out the possibility of significant spatial variations in the CR pressure, ionization and heating rates. However, here we have computed explicitly only variations of $n_{cr}$ and not of $P_{cr}$ and $\zeta_{cr}$. An explicit derivation of the corresponding variations in $P_{cr}$ and $\zeta_{cr}$, the inclusion of electrons, and a detailed discussion of the observed and predicted CR ionization rates will be given elsewhere. The variations we find should also be important for the ionization fraction in planetary atmospheres and protoplanetary disk ionization and chemistry, as well as the heating rate and CR-initiated ion–molecule chemistry in molecular clouds.

We are grateful to Steve Federman, Michail Malkov, Ben McCall, Andy Strong, and Donald York for useful discussions. We thank the referee for many helpful comments and suggestions. J.S. acknowledges support from NASA grants NAG5-13280 and NNG04GK43G and the NASA Astrobiology Institute Virtual Planetary Lead Team.
—. 1974, ARAA, 12, 71