Kinetic Processes and Plasma Transport

G.P. Zank
Center for Space Plasma and Aeronomic Research (CSPAR)
Department of Space Science
University of Alabama in Huntsville

Lecture Outline

- Lectures 1 and 2
  1) Derivation of gyrophase-averaged transport equation
  2) Derivation of Axford-Gleeson- Parker transport equation
     - Reference relativistic particle derivation

  3) Magnetic correlation tensor

  4) Quasi-linear scattering tensor

  5) Nonlinear guiding theory (NLGC)

  6) Diffusive shock acceleration

  7) Chapman-Enskog derivation of multi-component plasma/fluid description

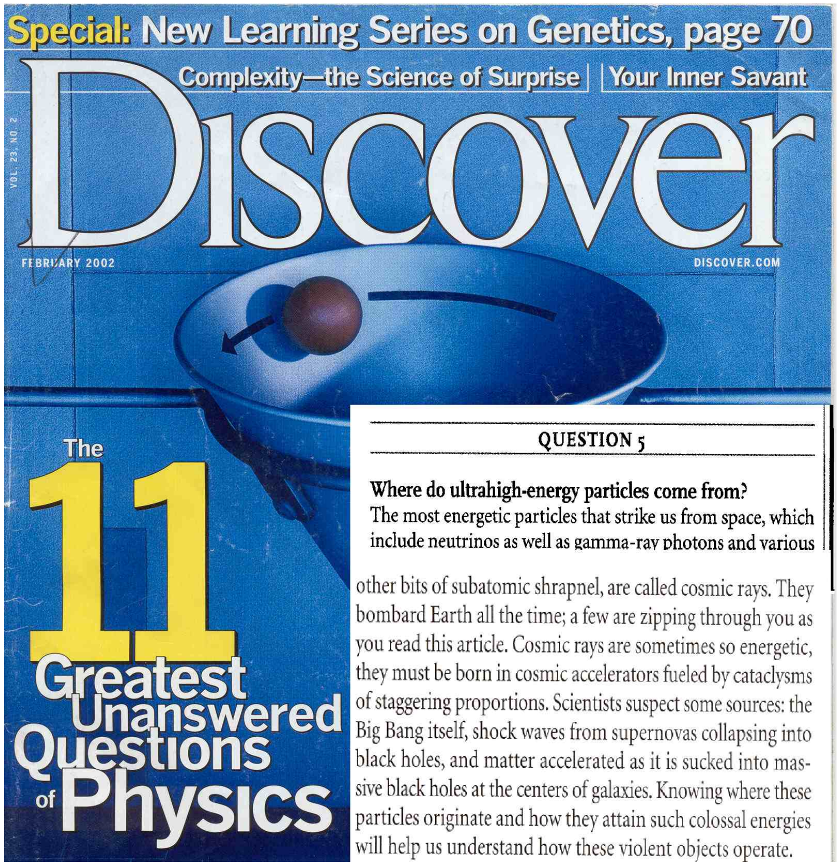
References on which lectures based:

- G.P. Zank, *Transport Processes in Space Physics and Astrophysics*, Springer Lecture Notes in Physics 877, 2014
- G.P. Zank et al., The Astrophysical Journal, 797:87, 2014
- G.P. Zank, Geoscience Letters, 3:22, DOI 10.1186/s40562-016-0055-2, 2016

Brief Motivation

- Energetic particles cornerstone of much of astrophysics and space physics
- Fundamental equation describing transport of energetic particles first derived in late 1960's and has been investigated repeatedly since then
- Lectures provide background I wish I had when entering the field …

Examples



Special: New Learning Series on Genetics, page 70

Complexity—the Science of Surprise | Your Inner Savant

Discover

VOL. 23, NO. 2

FEBRUARY 2002

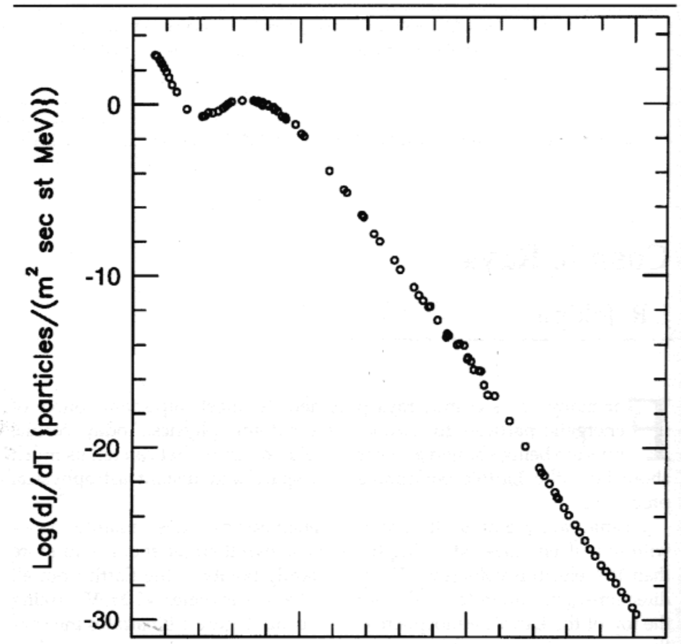
DISCOVER.COM

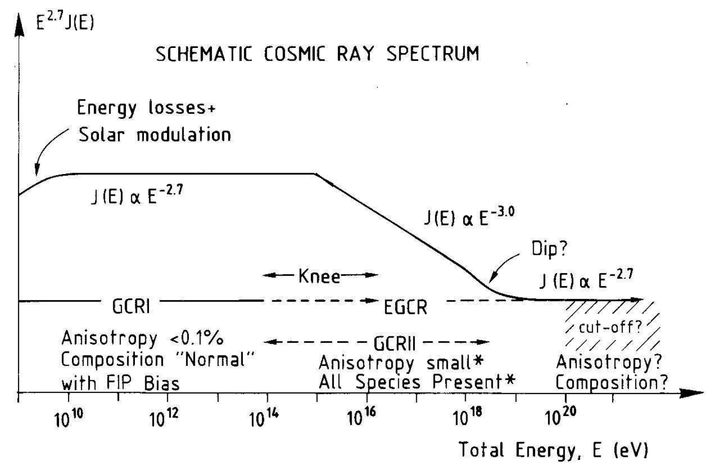
The 11 Greatest Unanswered Questions of Physics

QUESTION 5

Where do ultrahigh-energy particles come from?

The most energetic particles that strike us from space, which include neutrinos as well as gamma-ray photons and various other bits of subatomic shrapnel, are called cosmic rays. They bombard Earth all the time; a few are zipping through you as you read this article. Cosmic rays are sometimes so energetic, they must be born in cosmic accelerators fueled by cataclysms of staggering proportions. Scientists suspect some sources: the Big Bang itself, shock waves from supernovas collapsing into black holes, and matter accelerated as it is sucked into massive black holes at the centers of galaxies. Knowing where these particles originate and how they attain such colossal energies will help us understand how these violent objects operate.

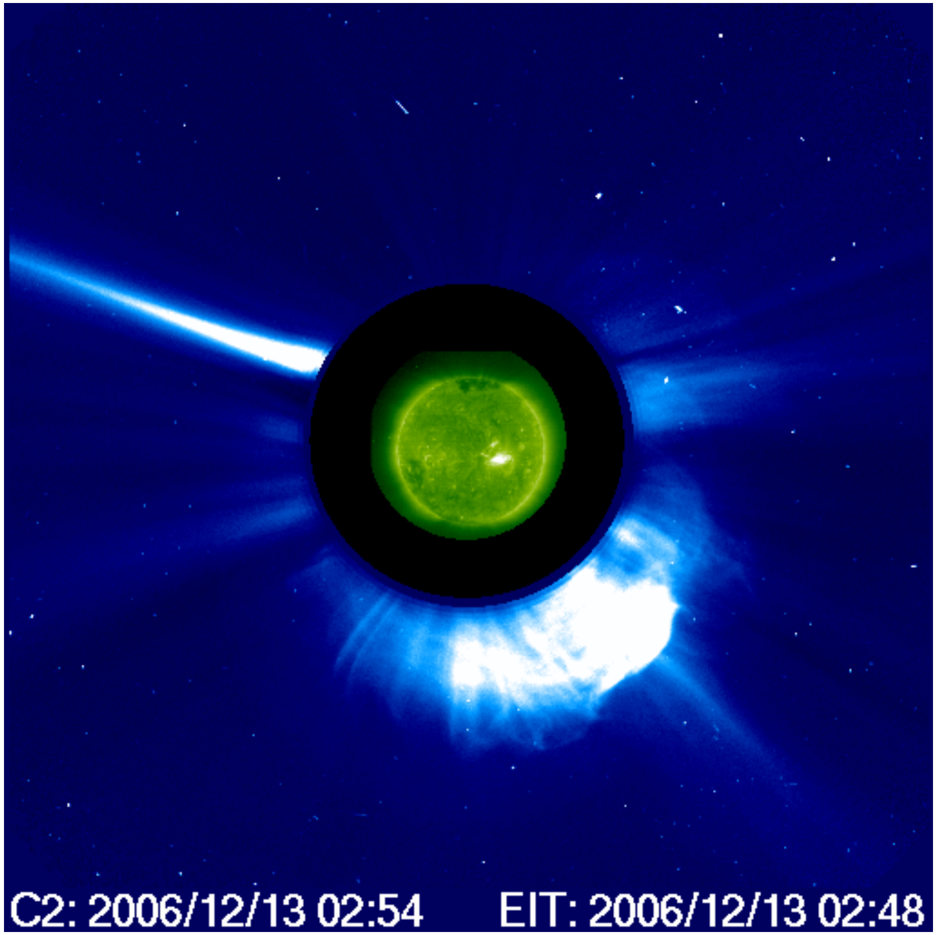




Explain the origin of energetic particle power law spectra (e.g., the universal galactic cosmic ray spectrum).

Examples

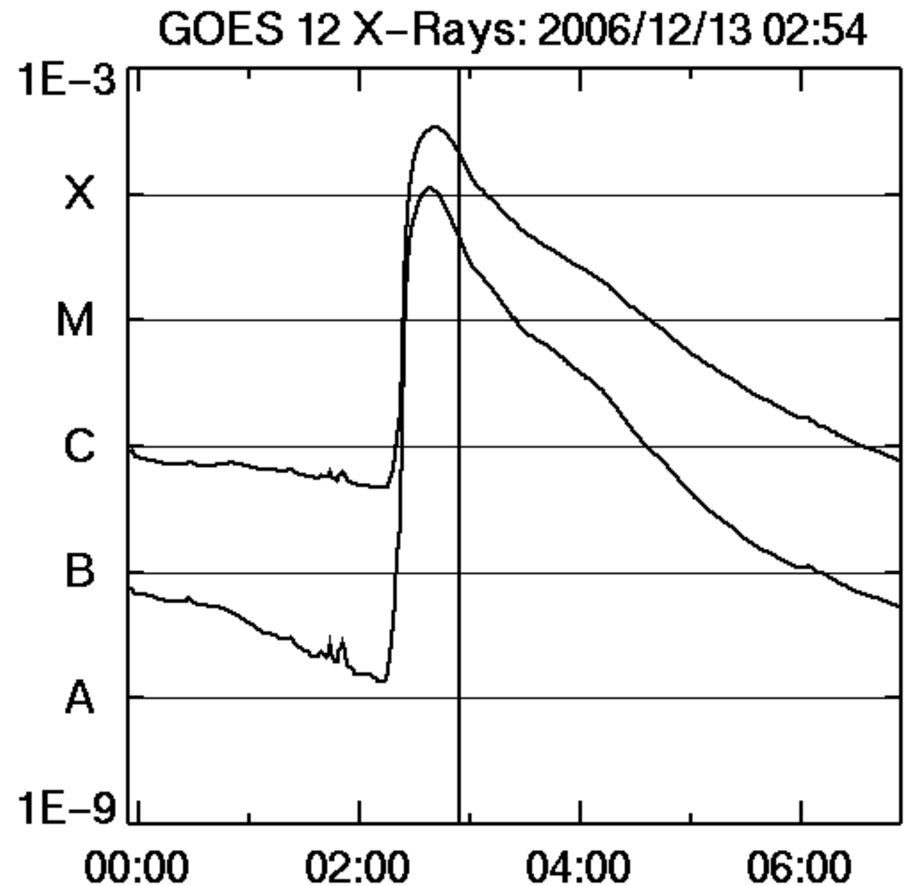
December 13, 2006 event



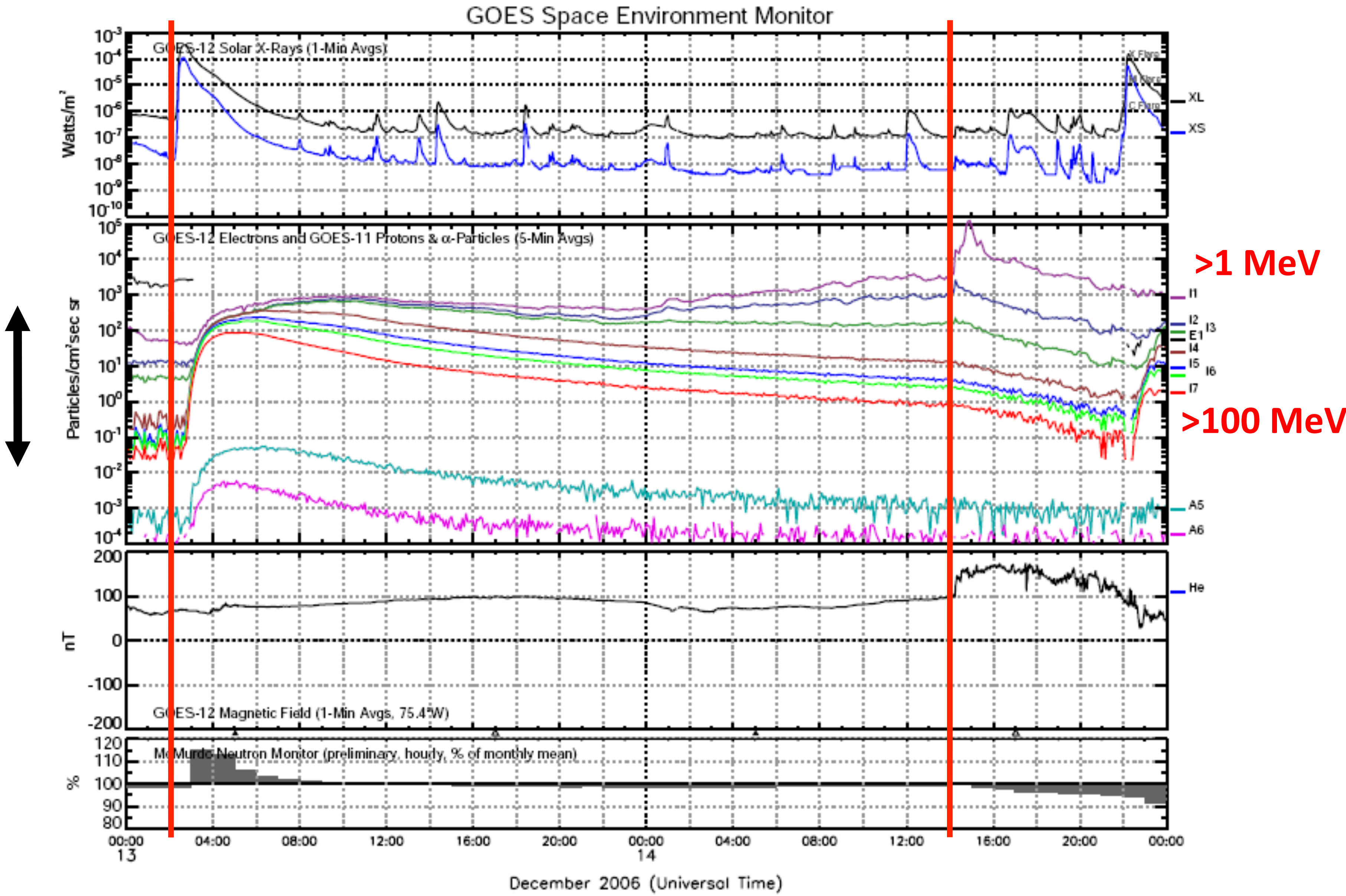
Halo CME: Dec. 13, 2006: 02:54:04

linear speed: 1774 km/sec;
speed at 20 R: 1573 km/sec;

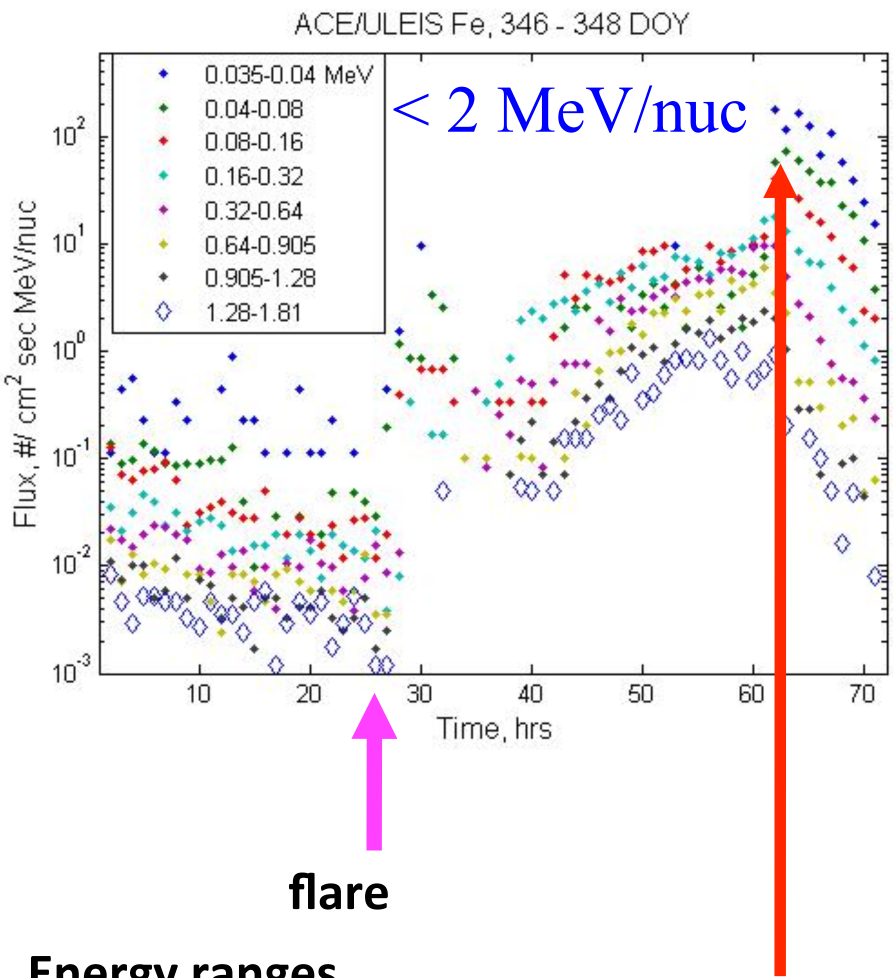
(from the SOHO/LASCO CME Catalog, courtesy of the CDAW Data Center, GSFC).

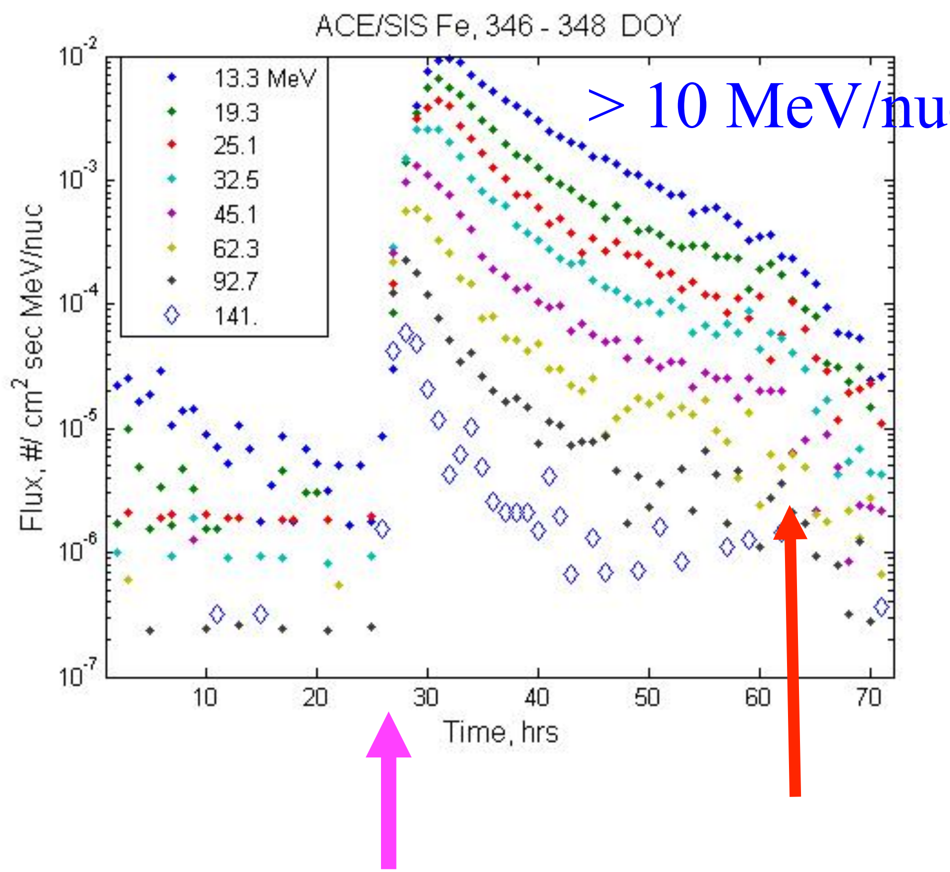


ESP Event (shock arrival at ACE: Dec. 14, ~1400 UT)


GOES Space Environment Monitor
GOES-12 Solar X-Rays (1-Min Avgs)
Watts/m²
XL
XS
GOES-12 Electrons and GOES-11 Protons & α-Particles (5-Min Avgs)
Particles/cm²sec sr
>1 MeV
>100 MeV
I1
I2
I3
E1
I4
I5
I6
I7
A5
A6
GOES-12 Magnetic Field (1-Min Avgs, 75.4°W)
nT
He
McMurdo Neutron Monitor (preliminary, hourly, % of monthly mean)
%
December 2006 (Universal Time)

ACE observations: Fe ions





flare

Energy ranges are in MeV/nuc

Approximate shock arrival time.

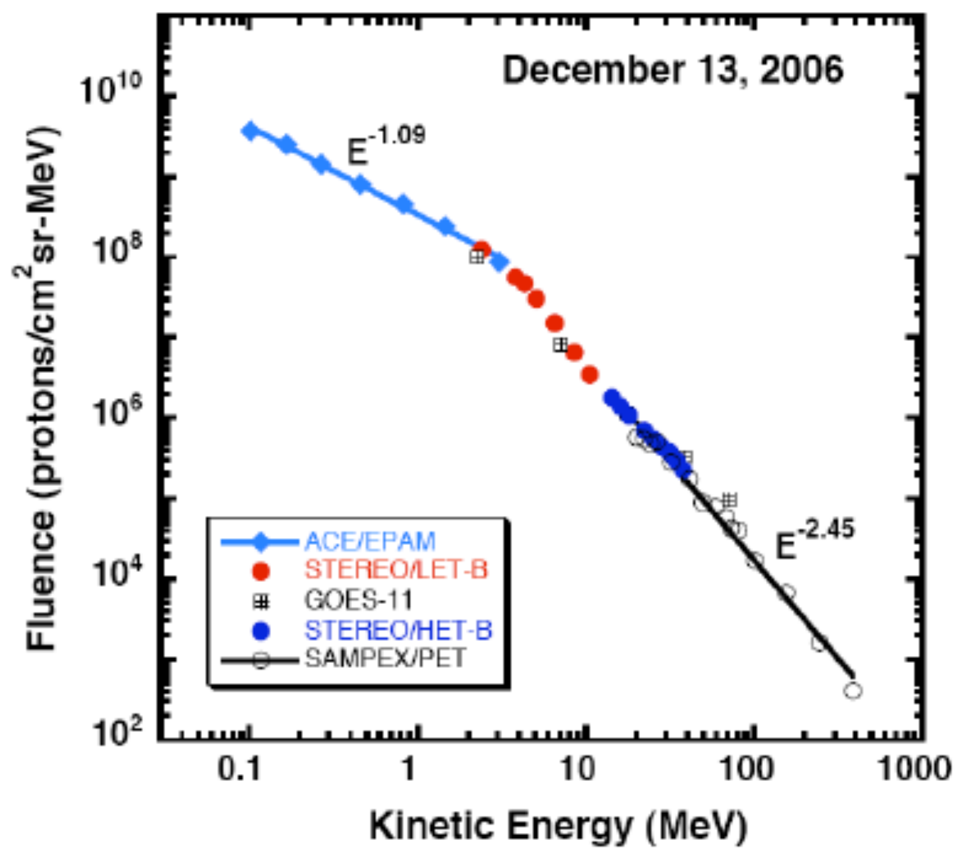


Figure 3. Event-integrated energetic proton spectra obtained with ACE, STEREO, GOES-11 and SAMPEX for the December 13, 2006 SEP event.

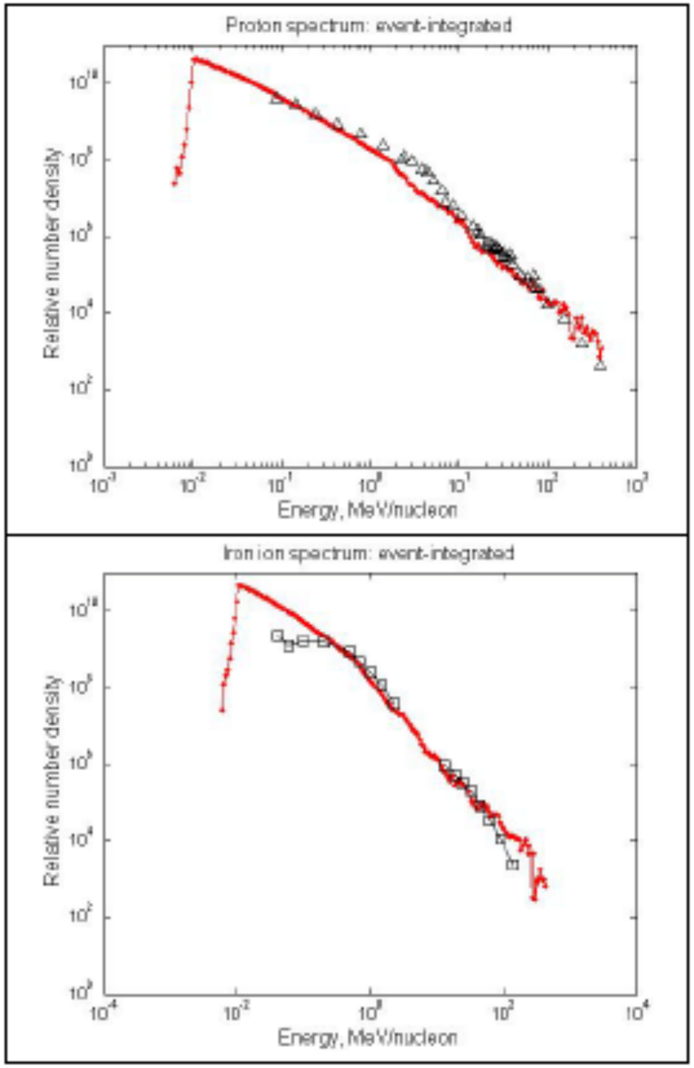


Figure 8. a) Event-integrated proton spectrum. PATH code results are shown by the red line. Fluences obtained with ACE, STEREO, GOES-11 and SAMPEX are shown by triangles. b) Event-integrated iron ion spectrum. PATH code results are shown by the red line. Fluences derived from ACE measurements are shown by squares.

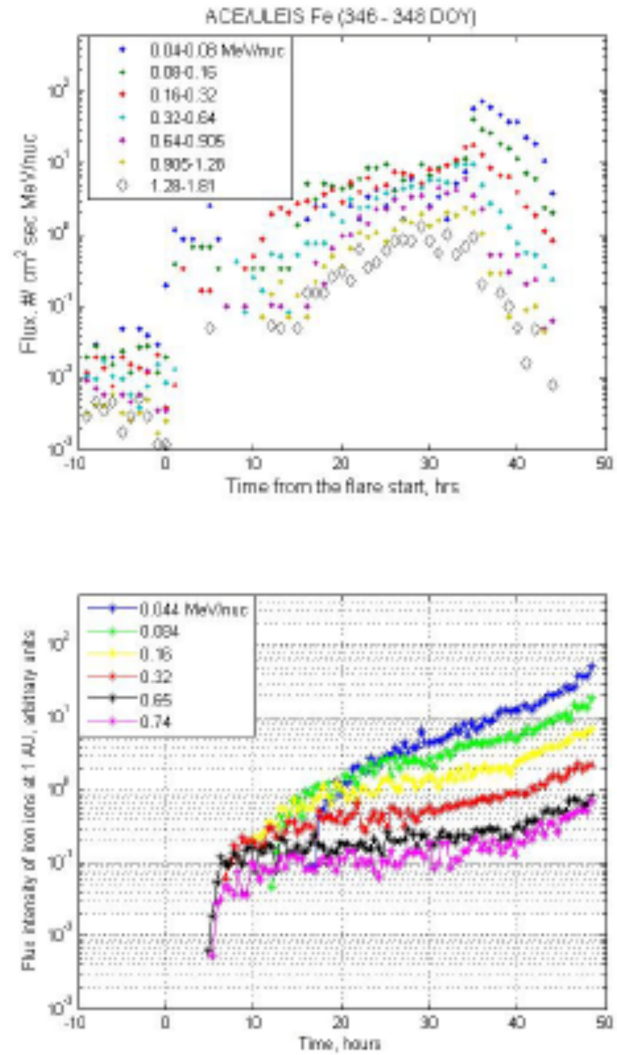


Figure 7. Iron ion fluxes derived from ACE/ULEIS measurements in a low-energy range of ~ 35 keV/nuc – 2 MeV/nuc at 1 AU (top panel) and corresponding modeling results with the PATH code (bottom panel). Notice a rise in particle fluxes around the shock arrival time.

5.1 Transport Equations for Non-relativistic Particles Scattered by Plasma Fluctuations

5.1.1 The Focussed Transport Equation

Electromagnetic fluctuations in a flowing medium such as the solar wind act to scatter particles, in pitch angle, gyrophase, or in energy. Although we do not explicitly restrict our attention to any particular form of electromagnetic waves in this subsection, we will implicitly consider particles scattered in pitch angle by magnetic fluctuations – either Alfvén waves or convected magnetic fluctuations. In this subsection, we derive a general equation for a gyrotropic distribution function that describes non-relativistic particles scattering in a flowing medium. Such a model was developed by Isenberg (1997) based on an approach by Skilling (1971) to describe the propagation of pickup ions in the solar wind. Although particles may eventually scatter towards isotropy in the frame of the medium, we not assume an isotropic distribution in this subsection. Following Isenberg, we begin with the Boltzmann equation for the distribution function $f(\mathbf{x}, \mathbf{v}, t)$ of non-relativistic particles in the inertial frame,1

$$\left(\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \frac{\mathbf{F}}{m} \cdot \nabla_v f = \frac{\delta f}{\delta t} \right)_s + S. \tag{5.1}$$

The force term can be quite general, but we restrict our attention to $\mathbf{F} = q/c(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ i.e., the inertial frame electromagnetic force acting on a particle of charge q, mass m, with c the speed of light. In the Boltzmann equation, S is a particle source term. Of note is that (5.1) has been implicitly separated into mean and fluctuating parts with the fluctuating components being treated as "scattering" terms and relegated to the right-hand-side. The scattering term $\delta f/\delta t)_s$ acts to stochastically scatter particles towards isotropy. In later subsections, we explicitly calculate various forms of the scattering operator. Here, we focus on the left-hand-side of (5.1).

Let us consider a frame of reference that propagates in the inertial "rest" frame at a velocity $\mathbf{U}$. Strictly speaking, this new frame comprises both the background convection velocity and the "average" velocity of the scattering "centers" (Alfvén waves, for example). Certainly in the supersonic solar wind, the convection velocity is much larger than the velocity of the background scattering fluctuations and so the additional velocity of the fluctuations is often neglected. Most importantly, a velocity transformation $\mathbf{U}$ can be identified with the velocity of the background conducting plasma, in which case the motional electric field $\mathbf{E} = -\mathbf{U} \times \mathbf{B}/c$ exactly cancels the electric field and leaves $\mathbf{F} = q\mathbf{v} \times \mathbf{B}/c$. It is important to recognize that the scattering term in this frame conserves energy since all macroscopic electric fields are transformed away. With no electric fields, particles can only scatter in pitch angle. However, energy is not conserved in the "rest" frame and this has important consequences, as we discuss later in considering particle acceleration at shock waves. Let us write

$$\mathbf{v} = \mathbf{c} + \mathbf{U} \Longleftrightarrow \mathbf{c} = \mathbf{v} - \mathbf{U},$$

for which the following transformations hold,

$$\frac{\partial}{\partial t} \to \frac{\partial}{\partial t} + \frac{\partial c_i}{\partial t}\frac{\partial}{\partial c_i} = \frac{\partial}{\partial t} - \frac{\partial U_i}{\partial t}\frac{\partial}{\partial c_i};$$
$$\frac{\partial}{\partial x_j} \to \frac{\partial}{\partial x_j} - \frac{\partial U_i}{\partial x_j}\frac{\partial}{\partial c_i}; \qquad \frac{\partial}{\partial v_i} \to \frac{\partial}{\partial c_i}.$$

On applying these frame transformations to the inertial form of the Boltzmann equation (5.1), we obtain an equation in mixed coordinates for the distribution function $f(\mathbf{x}, \mathbf{c}, t)$,

$$\frac{\partial f}{\partial t} + (U_i + c_i)\frac{\partial f}{\partial x_i} + \left[\frac{q}{m}(\mathbf{c} \times \mathbf{B})_i - \frac{\partial U_i}{\partial t} - (U_j + c_j)\frac{\partial U_i}{\partial x_j}\right]\frac{\partial f}{\partial c_i} = \left.\frac{\delta f}{\delta t}\right)_s . \quad (5.2)$$

The subscripts refer to vector components and the summation convention holds.

Let us now suppose that the particle gyroradius is much smaller than any other spatial scales in the system and similarly that their gyroperiod is smaller than other time scales. Thus, the particle distribution function can be regarded as nearly gyrotropic, and so $f(\mathbf{x}, \mathbf{c}, t)$ is essentially independent of gyrophase i.e., $f(\mathbf{x}, \mathbf{c}, t) \simeq f(\mathbf{x}, c, \mu, t)$, where the particle pitch angle $\mu \equiv \cos\theta = \mathbf{c}\cdot\mathbf{b}/c$ and the direction vector $\mathbf{b} \equiv \mathbf{B}/|B|$ is the unit vector along the large-scale magnetic field. Since we are assuming gyrotropy of the distribution function, we may average (5.2) over gyrophase. By gyrophase averaging, we neglect the action of perpendicular drifts on the distribution function. It is convenient to introduce spherical coordinates ($\theta \equiv$ pitch-angle, $\phi \equiv$ gyrophase, $c \equiv |\mathbf{c}|$),

$$c_x = c\sin\theta\cos\phi; \qquad c_y = c\sin\theta\sin\phi; \qquad c_z = c\cos\theta = c\mu;$$
$$c^2 = c_x^2 + c_y^2 + c_z^2; \qquad \cos\theta = \frac{c_z}{c}; \qquad \mu = \frac{c_i b_i}{c}; \qquad \mathbf{c} = c_x\hat{e}_x + c_y\hat{e}_y + c_z\hat{e}_z,$$

and $\mu = \mu(\mathbf{x})$ and $\phi = \phi(\mathbf{x})$. Consequently, we have the following transformations,

$$\nabla = \nabla + \nabla\mu\frac{\partial}{\partial\mu} + \nabla\phi\frac{\partial}{\partial\phi};$$
$$\frac{\partial}{\partial c_i} = \frac{\partial c}{\partial c_i}\frac{\partial}{\partial c} + \frac{\partial\mu}{\partial c_i}\frac{\partial}{\partial\mu} + \frac{\partial\phi}{\partial c_i}\frac{\partial}{\partial\phi}$$
$$= \frac{c_i}{c}\frac{\partial}{\partial c} + \left(\frac{b_i}{c} - \frac{\mu c_i}{c^2}\right)\frac{\partial}{\partial\mu} + \frac{\partial\phi}{\partial c_i}\frac{\partial}{\partial\phi},$$

which yields, on assuming that $f(\mathbf{x}, c, \mu, \phi, t) = f(\mathbf{x}, c, \mu, t)$,

$$\frac{\partial f}{\partial t} + (U_i + c_i)\left(\frac{\partial f}{\partial x_i} + \frac{\partial \mu}{\partial x_i}\frac{\partial f}{\partial \mu}\right) + \left[\Omega \varepsilon_{ijk} c_j b_k - \frac{\partial U_i}{\partial t} - (U_j + c_j)\frac{\partial U_i}{\partial x_j}\right]$$
$$\times \left(\frac{c_i}{c}\frac{\partial f}{\partial c} + \left(\frac{b_i}{c} - \frac{\mu c_i}{c^2}\right)\frac{\partial f}{\partial \mu}\right) = \left(\frac{\delta f}{\delta t}\right)_s, \tag{5.3}$$

where the gyrofrequency $\Omega = q|\mathbf{B}|/m$ has been introduced and ε is the Levi-Civita tensor. We introduce an averaging operator for ϕ such that $\varepsilon_{ijk} = 1/2\pi \int_0^{2\pi} Q d\phi$ and average (5.3) term-by term. Thus, since

$$\left\langle \frac{\partial f}{\partial t} \right\rangle = \frac{\partial f}{\partial t}; \qquad \left\langle U_i \frac{\partial f}{\partial x_i} \right\rangle = U_i \frac{\partial f}{\partial x_i}; \qquad \left\langle c_i \frac{\partial f}{\partial x_i} \right\rangle = \langle c_i \rangle \frac{\partial f}{\partial x_i},$$

and $\hat{e}_z = \mathbf{b}$, we obtain

$$\langle c_i \rangle \frac{\partial f}{\partial x_i} = c\mu b_i \frac{\partial f}{\partial x_i} \Rightarrow \left\langle (U_i + c_i)\frac{\partial f}{\partial x_i} \right\rangle = (U_i + c\mu b_i)\frac{\partial f}{\partial x_i}.$$

Here we used

$$\langle \mathbf{c} \rangle = c\langle \sin\theta \cos\phi \hat{e}_x + \sin\theta \sin\phi \hat{e}_y + \cos\theta \hat{e}_z \rangle = c\mu\mathbf{b},$$

since $\langle \sin\phi \rangle = \langle \cos\phi \rangle = 0$. Use of

$$\left\langle \frac{\partial \mu}{\partial x_i} \right\rangle = \left\langle \frac{c_j}{c} \right\rangle \frac{\partial b_j}{\partial x_i} = \mu b_j \frac{\partial b_j}{\partial x_i},$$

and $b_j b_j = 1$, or $b_j \partial b_j / \partial x_i = 0$, shows that

$$\left\langle U_i \frac{\partial \mu}{\partial x_i}\frac{\partial f}{\partial \mu} \right\rangle = 0.$$

Now consider

$$\left\langle c_i \frac{\partial \mu}{\partial x_i} \right\rangle = c\left\langle \frac{c_i c_j}{c^2} \right\rangle \frac{\partial b_j}{\partial x_i},$$

and the gyrophase averaged term $\langle c_i c_j / c^2 \rangle$ term-by term. We have

$$\left\langle \frac{c_x^2}{c^2} \right\rangle = \frac{1-\mu^2}{2}\hat{e}_x\hat{e}_x; \qquad \left\langle \frac{c_y^2}{c^2} \right\rangle = \frac{1-\mu^2}{2}\hat{e}_y\hat{e}_y; \qquad \left\langle \frac{c_z^2}{c^2} \right\rangle = \mu^2 \hat{e}_z\hat{e}_z,$$

and the cross terms $\langle c_i c_j / c^2 \rangle = 0$ for all $i \neq j$, $i, j = x, y, z$. Recalling that $\hat{e}_z \hat{e}_z = \mathbf{bb}$ we obtain

$$c \left\langle \frac{c_i c_j}{c^2} \right\rangle \frac{\partial b_j}{\partial x_i} = c \left(\frac{1-\mu^2}{2} \hat{e}_x \hat{e}_x + \frac{1-\mu^2}{2} \hat{e}_y \hat{e}_y + \mu^2 \mathbf{bb} \right) \frac{\partial b_j}{\partial x_i}$$
$$= c \left[\frac{1-\mu^2}{2} (\mathbf{I} - \mathbf{bb}) + \mu^2 \mathbf{bb} \right] \frac{\partial b_j}{\partial x_i}$$
$$= c \left[\frac{1-\mu^2}{2} \left(\delta_{ij} - b_i b_j \right) + \mu^2 \mathbf{bb} \right] \frac{\partial b_j}{\partial x_i}$$
$$= c \frac{1-\mu^2}{2} \delta_{ij} \frac{\partial b_j}{\partial x_i} = c \frac{1-\mu^2}{2} \frac{\partial b_i}{\partial x_i},$$

since $b_j \partial b_j / \partial x_i = 0$. Here we used $\hat{e}_x \hat{e}_x + \hat{e}_y \hat{e}_y + \hat{e}_z \hat{e}_z = \hat{e}_x \hat{e}_x + \hat{e}_y \hat{e}_y + \mathbf{bb} = \mathbf{I}$ or $\hat{e}_x \hat{e}_x + \hat{e}_y \hat{e}_y = \mathbf{I} - \mathbf{bb}$, where $\mathbf{I}$ is the identity matrix. Consequently, we have

$$\left\langle c_i \frac{\partial \mu}{\partial x_i} \frac{\partial f}{\partial \mu} \right\rangle = c \frac{1-\mu^2}{2} \frac{\partial b_i}{\partial x_i} \frac{\partial f}{\partial \mu}.$$

On using the results $\langle \hat{e}_j \rangle = \langle c_j / c \rangle = \mu b_j$ and $\langle \hat{e}_j \hat{e}_i \rangle = (1-\mu^2)/2(\delta_{ij} - b_i b_j) + \mu^2 b_i b_j$ of before, we find

$$\left\langle \left[-\frac{\partial U_i}{\partial t} - U_j \frac{\partial U_i}{\partial x_j} \right] \left(\frac{c_i}{c} \frac{\partial f}{\partial c} + \left(\frac{b_i}{c} - \frac{\mu c_i}{c^2} \right) \frac{\partial f}{\partial \mu} \right) \right\rangle$$
$$= \left(-\frac{\partial U_i}{\partial t} - U_j \frac{\partial U_i}{\partial x_j} \right) \left(\mu b_i \frac{\partial f}{\partial c} + \frac{1-\mu^2}{c} b_i \frac{\partial f}{\partial \mu} \right),$$

and

$$\left\langle c_j \frac{\partial U_i}{\partial x_j} \left(\frac{c_i}{c} \frac{\partial f}{\partial c} + \left(\frac{b_i}{c} - \frac{\mu c_i}{c^2} \right) \frac{\partial f}{\partial \mu} \right) \right\rangle$$
$$= \frac{\partial U_i}{\partial x_j} \left[c \left(\frac{1-\mu^2}{2} \left(\delta_{ij} - b_i b_j \right) + \mu^2 b_i b_j \right) \frac{\partial f}{\partial c} \right.$$
$$\left. + \left[\mu b_i b_j - \mu \left(\frac{1-\mu^2}{2} \left(\delta_{ij} - b_i b_j \right) + \mu^2 b_i b_j \right) \right] \frac{\partial f}{\partial \mu} \right].$$

Finally, the Lorentz force terms yield

$$\Omega \varepsilon_{ijk} \left\langle \frac{c_i c_j}{c^2} \right\rangle b_k c \frac{\partial f}{\partial c} = 0;$$
$$-\Omega \varepsilon_{ijk} \left\langle \frac{c_i c_j}{c^2} \right\rangle b_k \mu \frac{\partial f}{\partial \mu} = 0,$$

because $\langle c_i c_j/c^2 \rangle = 0$ for all $i \neq j$ and $\varepsilon_{ijk} = 0$ if and only if $i \neq j \neq k$. The final term,

$$\Omega \varepsilon_{ijk} \left\langle \frac{c_j}{c} \right\rangle b_i b_k \frac{\partial f}{\partial \mu} = \Omega \varepsilon_{ijk} b_i b_j b_k \mu \frac{\partial f}{\partial \mu} = 0,$$

because $\varepsilon_{ijk} = 0$ if and only if $i \neq j \neq k$ and $\sum_i \sum_j \sum_j \varepsilon_{ijk} = 0$.

On using the above gyrophase-averaged results and collecting terms, we obtain the reduced gyrophase-averaged transport equation

$$\frac{\partial f}{\partial t} + (U_i + c\mu b_i) \frac{\partial f}{\partial x_i} + \left[\frac{1-3\mu^2}{2} b_i b_j \frac{\partial U_i}{\partial x_j} - \frac{1-\mu^2}{2} \nabla \cdot \mathbf{U} \right.$$
$$\left. - \frac{\mu b_i}{c} \left(\frac{\partial U_i}{\partial t} + U_j \frac{\partial U_i}{\partial x_j} \right) \right] c \frac{\partial f}{\partial c} + \frac{1-\mu^2}{2} \left[c \nabla \cdot \mathbf{b} + \mu \nabla \cdot \mathbf{U} \right.$$
$$\left. -3\mu b_i b_j \frac{\partial U_i}{\partial x_j} - \frac{2b_i}{c} \left(\frac{\partial U_i}{\partial t} + U_j \frac{\partial U_i}{\partial x_j} \right) \right] \frac{\partial f}{\partial \mu} = \left\langle \left. \frac{\delta f}{\delta t} \right|_s \right\rangle. \quad (5.4)$$

The transport equation (5.4) is also known as the “focussed transport equation” and this non-relativistic form, derived by Isenberg (1997), differs from the earlier relativistically correct form derived by Skilling (1971) in that it contains the convective derivative of **U** since Skilling assumed that $U \ll c$.

le Roux and Webb (2012) present a particularly nice discussion of the meaning of the terms in the focussed transport equation (5.4). As discussed above, Eq. (5.4) is in the solar wind flow frame, which is noninertial. Since the plasma flow is non-uniform and non-stationary, scattered particles undergo velocity or momentum changes as measured in the flow frame due to pseudoforces associated with the non-uniform non-stationary nature of the flow. Recall from Chap. 2 that the gradient of the flow velocity can be expressed as the sum of the flow divergence, the flow shear, and the flow rotation, i.e.,

$$\frac{\partial U_i}{\partial x_j} = \frac{1}{3} \frac{\partial U_i}{\partial x_i} \delta_{ij} + \frac{1}{2} \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} - \frac{2}{3} \frac{\partial U_i}{\partial x_i} \delta_{ij} \right) + \frac{1}{2} \left(\frac{\partial U_i}{\partial x_j} - \frac{\partial U_j}{\partial x_i} \right)$$
$$= \frac{1}{3} \frac{\partial U_i}{\partial x_i} \delta_{ij} + \sigma_{ij} + \omega_{ij},$$

where σ_{ij} and ω_{ij} denote the shear and rotation tensors of the flow respectively. On expressing the flow gradient terms in the focussed transport equation (5.4) by the general representation above, the method of characteristics shows that

$$\frac{1}{c} \left\langle \frac{\partial c}{\partial t} \right\rangle = -\frac{1}{3} \frac{\partial U_i}{\partial x_i} + \frac{1-3\mu^2}{2} b_i b_j (\sigma_{ij} + \omega_{ij}) - \frac{\mu b_i}{c} \frac{dU_i}{dt}$$
$$= -\frac{1}{3} \frac{\partial U_i}{\partial x_i} + \frac{1-3\mu^2}{2} b_i b_j \sigma_{ij} - \frac{\mu b_i}{c} \frac{dU_i}{dt} \quad (5.5)$$

where dU_i/dt is the convective derivative, and we recognize that the rotation tensor is antisymmetric ($\omega_{ij} = -\omega_{ji}$), so that the sum $b_i b_j \omega_{ij} = 0$. Thus, flow rotation does not contribute to changes in particle speed. Similarly, we find that

$$\left\langle \frac{\partial \mu}{\partial t} \right\rangle = \frac{1-\mu^2}{2} \left[c \frac{\partial b_i}{\partial x_i} - 3\mu b_i b_j \sigma_{ij} - \frac{2b_i}{c} \frac{dU_i}{dt} \right], \tag{5.6}$$

and flow rotation does not contribute to changes in particle pitch-angle either. Expressions (5.5) and (5.6) describe the gyrophase averaged rate of change of the particle velocity c and pitch-angle μ. If particle velocity or momentum is measured in a nonuniform nonstationary plasma (noninertial) flow frame, the magnitude of the particle velocity or momentum will be modified if the flow diverges ($\nabla \cdot \mathbf{U}$), experiences shear (σ_{ij}), or rotation (ω_{ij}), or accelerates (dU_i/dt), while the particle pitch angle varies in response to flow shear, rotation, or acceleration. It is interesting to note (recall the telegrapher equation discussion, Chap. 2) that the shear and rotation tensor terms in Eq. (5.5) are multiplied by the term $-(3\mu^2 - 1)/2$, which is the second-order Legendre polynomial $P_2(\mu)$, whereas the divergence of the flow, $\nabla \cdot \mathbf{U}$ is multiplied by the zeroth-order Legendre polynomial $P_0 = 1$, and the acceleration term dU_i/dt by the first-order Legendre polynomial $P_1(\mu) = \mu$. For distributions f that are close to isotropic, this ordering of the terms associated with the Legendre polynomials gives the order of the importance in terms of energy change with respect to a physical process.

The flow divergence term $\nabla \cdot \mathbf{U}$ in Eq. (5.5) is nothing more than the well known adiabatic momentum change term in the standard cosmic ray transport equation that will be discussed below. Evidently, the divergence of the flow has no effect on the particle pitch angle. Physically, the effect of the divergence of a collisionless flow on energetic particles is consistent with the notion that particles are coupled to the flow through their interaction (scattering) with electromagnetic fields embedded in a highly conductive flow, but when the electromagnetic fields are neglected the divergence of the flow still affects the particle momentum simply because momentum is measured in the frame of a nonuniform plasma flow. As we discuss in more detail below, the rapid (negative) divergence of a flow across a shock wave leads to a convergence of the flow and the compression of electromagnetic fields embedded in the flow. As shown explicitly in the formulation of the focused transport equation, particles respond to the compression of electromagnetic fields embedded in the flow, and experience adiabatic compression. Notice that all of the effects due to a nonuniform nonstationary flow frame vanish if particle momentum is measured in an inertial frame, but if one is interested in what happens to the random component of the particle velocity at a shock, for example, noninertial effects must be taken into account.

Most investigations are currently restricted to the 1D version of the focussed transport equation. If one assumes for example a constant radial flow, such as the solar wind, with $\mathbf{U} = U\hat{\mathbf{r}}$ and a large-scale radial magnetic field pointing away from the Sun, $\mathbf{b} = \hat{\mathbf{r}}$, then Eq. (5.4) simplifies to

$$\frac{\partial f}{\partial t} + (U + \mu c)\frac{\partial f}{\partial r} - \frac{1-\mu^2}{r} Uc \frac{\partial f}{\partial c} + \frac{1-\mu^2}{r}(c + \mu U)\frac{\partial f}{\partial \mu} = \left\langle \frac{\delta f}{\delta t}\bigg|_s \right\rangle. \quad (5.7)$$

Exercises

1. By collecting all the terms associated with the gyrophase-averaging of (5.2), derive the general form of the gyrophase-averaged transport equation (5.4).
2. By assuming a constant radial flow velocity for the solar wind and a radial interplanetary magnetic field, derive the 1D focussed transport equation (5.7).
3. Assume that the one spatial dimensional gyrotropic distribution function can be expressed as

$$f(r, c, \mu) = f_-(r, c)H(-\mu) + f_+(r, c)H(\mu),$$

where $H(x)$ denotes the Heaviside step function and $f_\pm$ refer to anti-sunward (f_+)/sunward (f_-) hemispherical distributions. By substituting $f = f_- H(-\mu) + f_+ H(\mu)$ in the 1D focussed transport equation

$$\frac{\partial f}{\partial t} + (U + \mu c)\frac{\partial f}{\partial r} - \frac{1-\mu^2}{r} Uc \frac{\partial f}{\partial c} + \frac{1-\mu^2}{r}(c + \mu U)\frac{\partial f}{\partial \mu} = \frac{\partial}{\partial \mu}\left(\nu(1-\mu^2)\frac{\partial f}{\partial \mu}\right),$$

and integrating over μ separately from -1 to 0 and then from 0 to 1, show that

$$\frac{\partial f_\pm}{\partial t} + \left(U \pm \frac{c}{2}\right)\frac{\partial f_\pm}{\partial r} - \frac{2U}{r}\frac{c}{3}\frac{\partial f_\pm}{\partial c} + \frac{c}{r}(f_+ - f_-) = \mp \Gamma (f_+ - f_-)$$

where $\Gamma \equiv \nu(\mu = 0)$ gives the rate of scattering across $\mu = 0$. Note that the form of the scattering term is of diffusion in pitch-angle, and this is discussed below. The term ν is the scattering frequency.

5.1.2 *The Diffusive Transport Equation*

The solution of the general gyrophase-averaged transport equation is a formidable task for almost any physically interesting system so considerable effort has been invested in trying to simplify (5.4) by means of several additional assumptions. Let us assume that the scattering operator can be represented by a diffusion operator in pitch-angle,

$$\left\langle \frac{\delta f}{\delta t}\bigg|_s \right\rangle = \frac{\partial}{\partial \mu}\left(\nu\left(1-\mu^2\right)\frac{\partial f}{\partial \mu}\right), \quad (5.8)$$

where ν is a characteristic scattering frequency. The scattering term is discussed further in more general terms in the following subsections.

The dependence of the gyrophase-averaged particle distribution function f on the pitch-angle $\mu = \cos\theta$ with $\mu \in [-1, 1]$ suggests a natural expansion in terms of Legendre polynomials. The orthogonality properties of the complete set of Legendre polynomials allow us to rewrite the focussed transport equation (5.4) as an infinite set of partial differential equations in terms of the polynomial coefficients of the expansion. To ensure tractability, one typically truncates the infinite set at a low order, which is a form of closure. Accordingly, we expand the gyrophase-averaged particle distribution function f as

$$f(\mathbf{x}, t, c, \mu) = \sum_{n=0}^{\infty} \frac{1}{2}(2n+1) P_n(\mu) f_n(\mathbf{x}, t, c), \text{ where } f_n(\mathbf{x}, t, c) = \int_{-1}^{1} f P_n(\mu) d\mu.$$

The orthogonality condition is given by

$$\int_{-1}^{1} P_m(\mu) P_n(\mu) d\mu = \begin{cases} 0 & m \neq n \\ \frac{2}{2n+1} & m = n \end{cases},$$

and some useful recurrence relations that will be used below are

$$\begin{aligned}
&(n+1) P_{n+1}(\mu) = (2n+1)\mu P_n(\mu) - n P_{n-1}(\mu);\\
&(1-\mu^2)\frac{d}{d\mu} P_n(\mu) = n P_{n-1}(\mu) - n\mu P_n(\mu) = \frac{n(n+1)}{2n+1}\left[P_{n-1}(\mu) - P_{n+1}(\mu)\right];\\
&\frac{d}{d\mu} P_{n+1}(\mu) - \mu \frac{d}{d\mu} P_n(\mu) = (n+1) P_n(\mu);\\
&\mu \frac{d}{d\mu} P_n(\mu) - \frac{d}{d\mu} P_{n-1}(\mu) = n P_n(\mu);\\
&\frac{d}{d\mu}\left[P_{n+1}(\mu) - P_{n-1}(\mu)\right] = (2n+1) P_n(\mu).
\end{aligned}$$

We systematically project and expand each of the terms in (5.4) from left to right using the Legendre polynomial $P_m(\mu)$ and the expansion for f.

The first (time-derivative) term becomes

$$\frac{\partial f}{\partial t} : \int_{-1}^{1} P_m \frac{\partial f}{\partial t} d\mu = \sum_{n=0}^{\infty} \frac{1}{2}(2n+1) \int_{-1}^{1} P_m P_n d\mu \frac{\partial f_n}{\partial t} = \frac{\partial f_m}{\partial t},$$

after using the orthogonality relation. Similarly, the second (convective) term becomes

$$U_i \frac{\partial f}{\partial x_i} : U_i \sum_{n=0}^{\infty} \frac{1}{2}(2n+1) \int_{-1}^{1} P_m P_n d\mu \frac{\partial f_n}{\partial x_i} = U_i \frac{\partial f_m}{\partial x_i}.$$

The third term in (5.4) is a little more interesting in that we need to use the first of the recurrence relations. Thus,

$$c\mu b_i \frac{\partial f}{\partial x_i} : cb_i \int_{-1}^{1} \mu P_m \frac{\partial f}{\partial x_i} d\mu = cb_i \int_{-1}^{1} \left[\frac{m+1}{2m+1} P_{m+1} + \frac{m}{2m+1} P_{m-1} \right] \frac{\partial f}{\partial x_i} d\mu.$$

On expanding f, we find

$$\begin{aligned} cb_i \int_{-1}^{1} \mu P_m \frac{\partial f}{\partial x_i} d\mu &= cb_i \int_{-1}^{1} \frac{m+1}{2m+1} P_{m+1} \frac{\partial f}{\partial x_i} d\mu + cb_i \int_{-1}^{1} \frac{m}{2m+1} P_{m-1} \frac{\partial f}{\partial x_i} d\mu \\ &= cb_i \sum_{n=0}^{\infty} \frac{2n+1}{2} \int_{-1}^{1} \frac{m+1}{2m+1} P_{m+1} P_n \frac{\partial f_n}{\partial x_i} d\mu \\ &\quad + cb_i \sum_{n=0}^{\infty} \frac{2n+1}{2} \int_{-1}^{1} \frac{m}{2m+1} P_{m-1} P_n \frac{\partial f_n}{\partial x_i} d\mu. \end{aligned}$$

The first term on the right-hand side contributes only when $n = m + 1$ and the second term only when $n = m - 1$, so yielding

$$\begin{aligned} cb_i \int_{-1}^{1} \mu P_m \frac{\partial f}{\partial x_i} d\mu &= cb_i \frac{2m+3}{2} \frac{m+1}{2m+1} \frac{2}{2m+3} \frac{\partial f_{m+1}}{\partial x_i} \\ &\quad + cb_i \frac{2m-1}{2} \frac{m}{2m+1} \frac{2}{2m-1} \frac{\partial f_{m-1}}{\partial x_i} \\ &= cb_i \frac{m+1}{2m+1} \frac{\partial f_{m+1}}{\partial x_i} + cb_i \frac{m}{2m+1} \frac{\partial f_{m-1}}{\partial x_i}. \end{aligned}$$

The third term can therefore be expressed as

$$c\mu b_i \frac{\partial f}{\partial x_i} : \frac{cb_i}{2m+1} \left[(m+1) \frac{\partial f_{m+1}}{\partial x_i} + m \frac{\partial f_{m-1}}{\partial x_i} \right].$$

The fourth term in the focused transport equation (5.4) is

$$\frac{1-3\mu^2}{2} b_i b_j \frac{\partial U_j}{\partial x_i} c \frac{\partial f}{\partial c} : cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{1}{2} \frac{\partial f}{\partial c} - cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{3\mu^2}{2} \frac{\partial f}{\partial c}.$$

The first term can be rewritten immediately as

$$cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{1}{2} \frac{\partial f}{\partial c} : cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{1}{2} \frac{\partial f_m}{\partial c}.$$

We need to use the first of the recurrence relations to infer

$$\mu^2 P_n = \frac{(n+1)(n+2)}{(2n+1)(2n+3)} P_{n+2} + \left[\frac{(n+1)^2}{2n+3} + \frac{n^2}{2n-1} \right] \frac{P_n}{2n+1} + \frac{n(n-1)}{(2n-1)(2n+1)} P_{n-2}.$$

On using this identity for the second term above, we obtain

$$\begin{aligned} -cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{3\mu^2}{2} \frac{\partial f}{\partial c} &: -cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{3}{2} \int_{-1}^{1} \mu^2 P_m \frac{\partial f}{\partial c} d\mu \\ = -cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{3}{2} \sum_{n=0}^{\infty} \frac{2n+1}{2} &\int_{-1}^{1} \mu^2 P_n P_m \frac{\partial f_n}{\partial c} d\mu \\ = -cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{3}{2} \sum_{n=0}^{\infty} \frac{2n+1}{2} \frac{\partial f_n}{\partial c} &\int_{-1}^{1} \left[\frac{(n+1)(n+2)}{(2n+1)(2n+3)} P_{n+2} P_m \right. \\ &+ \left. \left(\frac{(n+1)^2}{2n+3} + \frac{n^2}{2n-1} \right) \frac{P_n}{2n+1} P_m + \frac{n(n-1)}{(2n-1)(2n+1)} P_{n-2} P_m \right] d\mu. \end{aligned}$$

The first integral contributes only when $n = m - 2$, the second when $n = m$ and the last when $n = m + 2$, so yielding

$$\begin{aligned} -cb_i b_j &\frac{\partial U_j}{\partial x_i} \frac{3}{2} \int_{-1}^{1} \mu^2 P_m \frac{\partial f}{\partial c} d\mu \\ = -cb_i b_j &\frac{\partial U_j}{\partial x_i} \frac{3}{2} \left[\frac{(m-1)m}{(2m+1)(2m-1)} \frac{\partial f_{m-2}}{\partial c} + \left(\frac{(m+1)^2}{2m+3} + \frac{m^2}{2m-1} \right) \right. \\ &\times \left. \frac{1}{2m+1} \frac{\partial f_m}{\partial c} + \frac{(m+2)(m+1)}{(2m+3)(2m+1)} \frac{\partial f_{m+2}}{\partial c} \right]. \end{aligned}$$

On assembling the various terms, we obtain for the fourth term of the focused transport equation,

$$\begin{aligned} \frac{1-3\mu^2}{2} b_i b_j \frac{\partial U_j}{\partial x_i} c \frac{\partial f}{\partial c} : cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{1}{2} \frac{\partial f_m}{\partial c} &- vb_i b_j \frac{\partial U_j}{\partial x_i} \frac{3}{2} \frac{(m-1)m}{(2m+1)(2m-1)} \frac{\partial f_{m-2}}{\partial c} \\ &- cb_i b_j \frac{\partial U_j}{\partial x_i} \frac{3}{2} \left(\frac{(m+1)^2}{2m+3} + \frac{m^2}{2m-1} \right) \frac{1}{2m+1} \frac{\partial f_m}{\partial c} \\ &- vb_i b_j \frac{\partial U_j}{\partial x_i} \frac{3}{2} \frac{(m+2)(m+1)}{(2m+3)(2m+1)} \frac{\partial f_{m+2}}{\partial c}. \end{aligned}$$

The fifth term of (5.4) can be expanded as

$$-c\frac{1-\mu^2}{2}\frac{\partial U_i}{\partial x_i}\frac{\partial f}{\partial c} : \quad -\frac{c}{2}\frac{\partial U_i}{\partial x_i}\frac{\partial f_m}{\partial c} + \frac{c}{2}\frac{\partial U_i}{\partial x_i}\frac{(m-1)m}{(2m+1)(2m-1)}\frac{\partial f_{m-2}}{\partial c}$$
$$+\frac{c}{2}\frac{\partial U_i}{\partial x_i}\left(\frac{(m+1)^2}{2m+3}+\frac{m^2}{2m-1}\right)\frac{1}{2m+1}\frac{\partial f_m}{\partial c}$$
$$+\frac{c}{2}\frac{\partial U_i}{\partial x_i}\frac{(m+2)(m+1)}{(2m+3)(2m+1)}\frac{\partial f_{m+2}}{\partial c}.$$

On using the results of expressing the third term in terms of a Legendre polynomial expansion, we have for the sixth term in (5.4)

$$-b_i\frac{DU_i}{Dt}\mu\frac{\partial f}{\partial c} : \; -\frac{DU_i}{Dt}\frac{b_i}{2m+1}\left[(m+1)\frac{\partial f_{m+1}}{\partial c}+m\frac{\partial f_{m-1}}{\partial c}\right].$$

The computation of the seventh term in the focused transport equation is also straightforward. We can immediately express

$$\frac{c}{2}\frac{\partial b_i}{\partial x_i}(1-\mu^2)\frac{\partial f}{\partial \mu} : \; \frac{c}{2}\frac{\partial b_i}{\partial x_i}\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n\int_{-1}^{1}d\mu\; P_m(1-\mu^2)\frac{\partial P_n}{\partial \mu}.$$

Use of the second recursion relation above yields

$$\frac{c}{2}\frac{\partial b_i}{\partial x_i}(1-\mu^2)\frac{\partial f}{\partial \mu} : \; \frac{c}{2}\frac{\partial b_i}{\partial x_i}\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n\int_{-1}^{1}d\mu\;\frac{n(n+1)}{2n+1}\left[P_mP_{n-1}-P_mP_{n+1}\right],$$

so that, since the first summand contributes only for $n = m + 1$ and the second for $n = m - 1$,

$$\frac{c}{2}\frac{\partial b_i}{\partial x_i}(1-\mu^2)\frac{\partial f}{\partial \mu} : \; \frac{c}{2}\frac{\partial b_i}{\partial x_i}\left[\frac{(m+1)(m+2)}{2m+1}f_{m+1}-\frac{m(m-1)}{2m+1}f_{m-1}\right].$$

Consider now the eighth term in (5.4). We have

$$\frac{1}{2}\frac{\partial U_i}{\partial x_i}(1-\mu^2)\mu\frac{\partial f}{\partial \mu} : \; \frac{1}{2}\frac{\partial U_i}{\partial x_i}\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n\int_{-1}^{1}d\mu\; P_m(1-\mu^2)\mu\frac{\partial P_n}{\partial \mu}.$$

The second of the recursion relations yields

$$
\begin{aligned}
\mu(1-\mu^2)\frac{\partial P_n}{\partial \mu} &= \frac{n(n+1)}{2n+1}\left[\mu P_{n-1} - \mu P_{n+1}\right] \\
&= \frac{n(n+1)}{2n+1}\left[\frac{n-1}{2n-1}P_{n-2} + \frac{n}{2n-1}P_n\right] \\
&\quad -\frac{n(n+1)}{2n+1}\left[\frac{n+2}{2n+3}P_{n+2} + \frac{n+1}{2n+3}P_n\right] \\
&= \frac{n(n+1)}{2n+1}\frac{n-1}{2n-1}P_{n-2} + \left[\frac{n^2(n+1)}{(2n+1)(2n-1)}\right. \\
&\quad \left. - \frac{n(n+1)^2}{(2n+1)(2n+3)}\right]P_n - \frac{n(n+1)}{2n+1}\frac{n+2}{2n+3}P_{n+2},
\end{aligned}
$$

from which we obtain

$$
\begin{aligned}
&\frac{1}{2}\frac{\partial U_i}{\partial x_i}\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n\int_{-1}^{1}d\mu\ P_m(1-\mu^2)\mu\frac{\partial P_n}{\partial \mu} \\
&= \frac{1}{2}\frac{\partial U_i}{\partial x_i}\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n\int_{-1}^{1}d\mu\ P_m\frac{n(n+1)}{2n+1}\frac{n-1}{2n-1}P_{n-2} \\
&\quad +\frac{1}{2}\frac{\partial U_i}{\partial x_i}\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n\int_{-1}^{1}d\mu\ P_m\left[\frac{n^2(n+1)}{(2n+1)(2n-1)} - \frac{n(n+1)^2}{(2n+1)(2n+3)}\right]P_n \\
&\quad -\frac{1}{2}\frac{\partial U_i}{\partial x_i}\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n\int_{-1}^{1}d\mu\ P_m\frac{n(n+1)}{2n+1}\frac{n+2}{2n+3}P_{n+2}.
\end{aligned}
$$

The first term contributes only for $n = m + 2$, the second for $n = m$, and the third for $n = m - 2$, from which we obtain

$$
\begin{aligned}
\frac{1-\mu^2}{2}\mu\frac{\partial U_i}{\partial x_i}\frac{\partial f}{\partial \mu}:\quad & \frac{1}{2}\frac{\partial U_i}{\partial x_i}\frac{(m+1)(m+2)(m+3)}{(2m+1)(2m+3)}f_{m+2} \\
&+\frac{1}{2}\frac{\partial U_i}{\partial x_i}\left[\frac{m^2(m+1)}{(2m+1)(2m-1)} - \frac{m(m+1)^2}{(2m+1)(2m+3)}\right]f_m \\
&-\frac{1}{2}\frac{\partial U_i}{\partial x_i}\frac{m(m-1)(m-2)}{(2m+1)(2m-1)}f_{m-2}.
\end{aligned}
$$

Since

$$
\left[\frac{m^2(m+1)}{(2m+1)(2m-1)} - \frac{m(m+1)^2}{(2m+1)(2m+3)}\right] = \frac{m(m+1)}{(2m-1)(2m+3)},
$$

we can express term eight as

$$\frac{1-\mu^2}{2}\mu\frac{\partial U_i}{\partial x_i}\frac{\partial f}{\partial \mu}: \quad \frac{1}{2}\frac{\partial U_i}{\partial x_i}\frac{(m+1)(m+2)(m+3)}{(2m+1)(2m+3)}f_{m+2}$$
$$+\frac{1}{2}\frac{\partial U_i}{\partial x_i}\frac{m(m+1)}{(2m-1)(2m+3)}f_m$$
$$-\frac{1}{2}\frac{\partial U_i}{\partial x_i}\frac{m(m-1)(m-2)}{(2m+1)(2m-1)}f_{m-2}.$$

We can utilize these results to express term nine in (5.4) as

$$-\frac{1-\mu^2}{2}3\mu b_i b_j\frac{\partial U_j}{\partial x_i}\frac{\partial f}{\partial \mu}: \quad -\frac{3}{2}b_i b_j\frac{\partial U_j}{\partial x_i}\frac{(m+1)(m+2)(m+3)}{(2m+1)(2m+3)}f_{m+2}$$
$$-\frac{3}{2}b_i b_j\frac{\partial U_j}{\partial x_i}\frac{m(m+1)}{(2m-1)(2m+3)}f_m$$
$$+\frac{3}{2}b_i b_j\frac{\partial U_j}{\partial x_i}\frac{m(m-1)(m-2)}{(2m+1)(2m-1)}f_{m-2}.$$

The results from evaluating the seventh term yield

$$-\frac{b_i}{c}\frac{DU_i}{Dt}(1-\mu^2)\frac{\partial f}{\partial \mu}: -\frac{b_i}{c}\frac{DU_i}{Dt}\left[\frac{(m+1)(m+2)}{2m+1}f_{m+1}-\frac{m(m-1)}{2m+1}f_{m-1}\right].$$

Finally, let us consider the specific form of the diffusion term in pitch-angle μ,

$$\frac{\partial}{\partial \mu}\left[\nu\frac{1-\mu^2}{2}\frac{\partial f}{\partial \mu}\right]: \quad \int_{-1}^{1} P_m\frac{\partial}{\partial \mu}\left[\nu\frac{1-\mu^2}{2}\frac{\partial f}{\partial \mu}\right]d\mu$$
$$=\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n\int_{-1}^{1} d\mu\ P_m\frac{\partial}{\partial \mu}\left[\nu\frac{1-\mu^2}{2}\frac{\partial P_n}{\partial \mu}\right].$$

The recursion operator

$$\frac{\partial P_n}{\partial \mu}=P_n'=\frac{n}{1-\mu^2}\left[P_{n-1}-\mu P_n\right],$$

yields

$$\int_{-1}^{1} P_m\frac{\partial}{\partial \mu}\left[\nu\frac{1-\mu^2}{2}\frac{\partial f}{\partial \mu}\right]d\mu=\sum_{n=0}^{\infty}\frac{2n+1}{2}f_n$$
$$\times\int_{-1}^{1} P_m\frac{\partial}{\partial \mu}\left[\nu\frac{1-\mu^2}{2}\frac{n}{1-\mu^2}\left(P_{n-1}-\mu P_n\right)\right]d\mu.$$

Since ν is independent of μ, we find

$$\int_{-1}^{1} P_m \frac{\partial}{\partial \mu}\left[\nu \frac{1-\mu^2}{2}\frac{\partial f}{\partial \mu}\right] d\mu = \nu \sum_{n=0}^{\infty} \frac{2n+1}{2}\frac{n}{2} f_n \int_{-1}^{1} P_m \frac{\partial}{\partial \mu}\left(P_{n-1} - \mu P_n\right) d\mu.$$

On using the following relation,

$$\frac{\partial}{\partial \mu}\left(P_{n-1} - \mu P_n\right) = \frac{\partial}{\partial \mu}\left(P_{n-1} - \frac{n+1}{2n+1}P_{n+1} - \frac{n}{2n+1}P_{n-1}\right)$$
$$= \frac{n+1}{2n+1}\frac{\partial}{\partial \mu}\left(P_{n-1} - P_{n+1}\right),$$

together with the definition

$$\frac{\partial}{\partial \mu}\left(P_{n-1} - P_{n+1}\right) = -(2n+1)P_n,$$

we find

$$\frac{\partial}{\partial \mu}\left(P_{n-1} - \mu P_n\right) = -(n+1)P_n.$$

We therefore have the result that

$$\int_{-1}^{1} P_m \frac{\partial}{\partial \mu}\left[\nu \frac{1-\mu^2}{2}\frac{\partial f}{\partial \mu}\right] d\mu = -\nu \sum_{n=0}^{\infty} \frac{2n+1}{2}\frac{n}{2} f_n (n+1) \int_{-1}^{1} P_m P_n d\mu,$$

and since the integral only contributes for $n = m$, we have

$$\int_{-1}^{1} P_m \frac{\partial}{\partial \mu}\left[\nu \frac{1-\mu^2}{2}\frac{\partial f}{\partial \mu}\right] d\mu = -\nu \frac{m(m+1)}{2} f_m.$$

This completes the evaluation of each of the terms in (5.4).

By gathering the results above together, the complete transformed focused transport equation (5.4) can now be expressed as an infinite set of partial differential equations in the coefficients f_n of the Legendre polynomials,

$$\frac{\partial f_m}{\partial t} + U_i \frac{\partial f_m}{\partial x_i} + \frac{c b_i}{2m+1}\left[(m+1)\frac{\partial f_{m+1}}{\partial x_i} + m \frac{\partial f_{m-1}}{\partial x_i}\right] + c b_i b_j \frac{\partial U_j}{\partial x_i}\frac{1}{2}\frac{\partial f_m}{\partial c}$$
$$- c b_i b_j \frac{\partial U_j}{\partial x_i}\frac{3}{2}\frac{(m-1)m}{(2m+1)(2m-1)}\frac{\partial f_{m-2}}{\partial c} - c b_i b_j \frac{\partial U_j}{\partial x_i}\frac{3}{2}\left(\frac{(m+1)^2}{2m+3} + \frac{m^2}{2m-1}\right)$$
$$\times \frac{1}{2m+1}\frac{\partial f_m}{\partial c} - c b_i b_j \frac{\partial U_j}{\partial x_i}\frac{3}{2}\frac{(m+2)(m+1)}{(2m+3)(2m+1)}\frac{\partial f_{m+2}}{\partial c} - \frac{c}{2}\frac{\partial U_i}{\partial x_i}\frac{\partial f_m}{\partial c}$$
$$+ \frac{c}{2}\frac{\partial U_i}{\partial x_i}\frac{(m-1)m}{(2m+1)(2m-1)}\frac{\partial f_{m-2}}{\partial c} + \frac{c}{2}\frac{\partial U_i}{\partial x_i}\left(\frac{(m+1)^2}{2m+3} + \frac{m^2}{2m-1}\right)\frac{1}{2m+1}\frac{\partial f_m}{\partial c}$$

$$
+\frac{c}{2}\frac{\partial U_i}{\partial x_i}\frac{(m+2)(m+1)}{(2m+3)(2m+1)}\frac{\partial f_{m+2}}{\partial c}-\frac{DU_i}{Dt}\frac{b_i}{2m+1}\left[(m+1)\frac{\partial f_{m+1}}{\partial c}+m\frac{\partial f_{m-1}}{\partial c}\right]
$$
$$
+\frac{c}{2}\frac{\partial b_i}{\partial x_i}\left[\frac{(m+1)(m+2)}{2m+1}f_{m+1}-\frac{m(m-1)}{2m+1}f_{m-1}\right]
$$
$$
+\frac{1}{2}\frac{\partial U_i}{\partial x_i}\frac{(m+1)(m+2)(m+3)}{(2m+1)(2m+3)}f_{m+2}+\frac{1}{2}\frac{\partial U_i}{\partial x_i}\frac{m(m+1)}{(2m-1)(2m+3)}f_m
$$
$$
-\frac{1}{2}\frac{\partial U_i}{\partial x_i}\frac{m(m-1)(m-2)}{(2m+1)(2m-1)}f_{m-2}-\frac{3}{2}b_ib_j\frac{\partial U_j}{\partial x_i}\frac{(m+1)(m+2)(m+3)}{(2m+1)(2m+3)}f_{m+2}
$$
$$
-\frac{3}{2}b_ib_j\frac{\partial U_j}{\partial x_i}\frac{m(m+1)}{(2m-1)(2m+3)}f_m+\frac{3}{2}b_ib_j\frac{\partial U_j}{\partial x_i}\frac{m(m-1)(m-2)}{(2m+1)(2m-1)}f_{m-2}
$$
$$
-\frac{b_i}{c}\frac{DU_i}{Dt}\left[\frac{(m+1)(m+2)}{2m+1}f_{m+1}-\frac{m(m-1)}{2m+1}f_{m-1}\right]=-\nu\frac{m(m+1)}{2}f_m. \tag{5.9}
$$

The infinite set of partial differential equations (5.9) is equivalent to the focused transport equation (5.4) and therefore as challenging to solve. At each order of the expansion, i.e., the pde for a Legendre coefficient of particular order, it is clearly seen that the equation possesses coefficients of a higher order. This is another expression of the closure problem. Closure is typically affected by simply truncating the Legendre polynomial expansion at a finite number of coefficients. This procedure is somewhat arbitrary and one formally needs to establish that the truncation remains sufficiently close to the full solution. This is typically very difficult in practice, and so is rarely done. An example of the subtleties that can arise was discussed in Chap. 2, Sect. 2.8, where an even or an odd truncation of the Legendre polynomial expansion of a simplified Boltzmann equation yielded fundamentally different solutions, with the even truncation capturing the non-propagating characteristic solution and the odd truncation missing that particular mode.

Let us consider the simplest reduction of the set of equations (5.9) by truncating the infinite set of equations at some arbitrary order with the hope that this does not introduce any unphysical character into the reduced model. Typically, truncations are made at the lowest order possible. For the f_1 approximation (i.e. assume $f_n = 0\ \forall\ n \geq 2$), we have, on setting $m = 0$,

$$
\frac{\partial f_0}{\partial t}+U_i\frac{\partial f_0}{\partial x_i}-\frac{c}{3}\frac{\partial U_i}{\partial x_i}\frac{\partial f_0}{\partial c}=-cb_i\frac{\partial f_1}{\partial x_i}+\frac{DU_i}{Dt}b_i\frac{\partial f_1}{\partial c}-c\frac{\partial b_i}{\partial x_i}f_1+2\frac{b_i}{c}\frac{DU_i}{Dt}f_1, \tag{5.10}
$$

and on setting $m = 1$ and neglecting all terms with indices having $i \geq 2$, we find

$$
\frac{\partial f_1}{\partial t}+U_i\frac{\partial f_1}{\partial x_i}+\frac{cb_i}{3}\frac{\partial f_0}{\partial x_i}+\frac{1}{2}cb_ib_j\frac{\partial U_j}{\partial x_i}\frac{\partial f_1}{\partial c}-\frac{9}{10}cb_ib_j\frac{\partial U_j}{\partial x_i}\frac{\partial f_1}{\partial c}-\frac{c}{2}\frac{\partial U_i}{\partial x_i}\frac{\partial f_1}{\partial c}
$$
$$
+\frac{9}{10}c\frac{\partial U_i}{\partial x_i}\frac{\partial f_1}{\partial c}-\frac{DU_i}{Dt}\frac{b_i}{3}\frac{\partial f_0}{\partial c}+\frac{1}{5}\frac{\partial U_i}{\partial x_i}f_1-\frac{3}{5}b_ib_j\frac{\partial U_j}{\partial x_i}f_1=-\nu f_1.
$$

On rearranging the above expression, we obtain

$$\frac{\partial f_1}{\partial t} + U_i \frac{\partial f_1}{\partial x_i} - \frac{2}{5} c b_i b_j \frac{\partial U_j}{\partial x_i} \frac{\partial f_1}{\partial c} + \frac{2}{5} c \frac{\partial U_i}{\partial x_i} \frac{\partial f_1}{\partial c} + \frac{1}{5} \frac{\partial U_i}{\partial x_i} f_1 - \frac{3}{5} b_i b_j \frac{\partial U_j}{\partial x_i} f_1$$
$$= -\nu f_1 - \frac{c b_i}{3} \frac{\partial f_0}{\partial x_i} + \frac{D U_i}{D t} \frac{b_i}{3} \frac{\partial f_0}{\partial c}, \tag{5.11}$$

where the f_0 Legendre coefficients are expressed as source terms in the evaluation of the next order Legendre coefficients f_1. To solve Eq. (5.11) for f_1 in terms of the lower order Legendre coefficient f_0, we make the further assumption that the zeroth-order coefficient f_0 is almost isotropic, implying that $f_1 \ll f_0$. The next assumption that we impose is that $\nu = \tau^{-1}$ is large, i.e., rapid scattering of the charged particles (which is consistent with the assumption that the particle distribution is nearly isotropic), so that the term $\nu f_1 \sim O(f_0)$. Subject to these assumptions, Eq. (5.11) can then be solved, yielding

$$\nu f_1 \simeq -\frac{c b_i}{3} \frac{\partial f_0}{\partial x_i} + \frac{D U_i}{D t} \frac{b_i}{3} \frac{\partial f_0}{\partial c}. \tag{5.12}$$

Suppose first that the background flow possesses no large-scale accelerations or gradients, i.e., $D u_i / D t = 0$, so that f_1 can be expressed as a diffusion term,

$$f_1 = -\frac{c \tau b_i}{3} \frac{\partial f_0}{\partial x_i}. \tag{5.13}$$

For the case that $D U_i / D t = 0$, use of (5.13) in (5.10) yields

$$\frac{\partial f_0}{\partial t} + U_i \frac{\partial f_0}{\partial x_i} - \frac{c}{3} \frac{\partial U_i}{\partial x_i} \frac{\partial f_0}{\partial c} = -c b_i \frac{\partial f_1}{\partial x_i} - c \frac{\partial b_i}{\partial x_i} f_1 = -\frac{\partial}{\partial x_i} (b_i c f_1)$$
$$= \frac{\partial}{\partial x_i} \left(b_i \kappa b_j \frac{\partial f_0}{\partial x_j} \right),$$

where we introduced the diffusion coefficient

$$\kappa = \frac{c^2 \tau}{3}.$$

The diffusion term $b_i \kappa b_j$ is a tensor comprising an isotropic part and an anisotropic part,

$$\mathbf{K} = \begin{pmatrix} \kappa_{11} & 0 & 0 \\ 0 & \kappa_{22} & 0 \\ 0 & 0 & \kappa_{33} \end{pmatrix} + \begin{pmatrix} 0 & \kappa_{12} & \kappa_{13} \\ \kappa_{12} & 0 & \kappa_{23} \\ \kappa_{13} & \kappa_{23} & 0 \end{pmatrix} = \begin{pmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} \\ \kappa_{12} & \kappa_{22} & \kappa_{23} \\ \kappa_{13} & \kappa_{23} & \kappa_{33} \end{pmatrix},$$

where the elements of the tensor are simply $\kappa_{ii} = b_i^2\kappa$, $\kappa_{ij} = b_i b_j \kappa$ for $i \neq j$ and $\kappa_{ij} = \kappa_{ji}$. Use of the diffusion tensor $\mathbf{K}$ allows us to express the convective-diffusive or advective-diffusive transport equation as

$$\frac{\partial f_0}{\partial t} + \mathbf{U} \cdot \nabla f_0 - \frac{c}{3}\nabla \cdot \mathbf{U}\frac{\partial f_0}{\partial c} = \nabla \cdot (\mathbf{K}\nabla f_0) \,. \tag{5.14}$$

Subject to the assumptions imposed in deriving Eq. (5.14), this is the standard form of the transport equation for non-relativistic charged particles experiencing scattering in a turbulent magnetized medium. The physical content of the transport equation (5.14) is interesting when considered term-by-term. The second term shows that the scattered particles that comprise the distribution f_0 essentially co-move with the background flow in which the "scatterers" are embedded. The third term is an energy change term in response to the divergence of the background flow. This is seen by considering

$$\frac{\partial f_0}{\partial t} - \frac{c}{3}\nabla \cdot \mathbf{U}\frac{\partial f_0}{\partial c} = 0 \Leftrightarrow \frac{\partial f_0}{\partial t} - \frac{1}{3}\nabla \cdot \mathbf{U}\frac{\partial f_0}{\partial \xi} = 0,$$

where $\xi = \ln c$. The characteristics for this equation are given by

$$\frac{d\xi}{dt}\left(= \frac{1}{c}\frac{dc}{dt}\right) = -\frac{1}{3}\nabla \cdot \mathbf{U};$$
$$\frac{df_0}{dt} = \text{const.},$$

which yields

$$\xi(t) - \xi(t_0) = -\frac{1}{3}\int_{t_0}^{t} \nabla \cdot \mathbf{U} dt,$$

from some initial time t_0 to a time t. If $\mathbf{U}$ is stationary, then the change in particle velocity is given by

$$\frac{\ln c(t) - \ln c(t_0)}{t - t_0} = -\frac{1}{3}\nabla \cdot \mathbf{U}.$$

According as $\nabla \cdot \mathbf{U}$ is convergent (< 0) or divergent (> 0), particles will gain or lose speed c in the flow. For example, if the particle distribution function upstream of a region of a 1D decelerating flow ($\partial U/\partial x < 0$) is a power law $f \sim c^{-q}$, then the spectrum behind the decelerating flow will be shifted uniformly to the "right" in which each speed $\ln c$ increased by an amount proportional to the velocity gradient. Consequently, the energy of the particle distribution function will increase.

The diffusion term contains much of the physics of the magnetic field structure as well as the scattering properties of the small scale fluctuating field. As a consequence, the term $\mathbf{K}$ contains much more than simply diffusion. The isotropic part of the tensor **K** describes particle diffusion along (parallel) and perpendicular to the magnetic field. The anisotropic terms are generally thought to describe the collective drift of particles due to gradients and curvature in the magnetic field **B** and magnitude $|\mathbf{B}|$. However, in a sense shown below, the particle response to the large-scale magnetic field geometry and gradients is present in all the elements of the tensor **K**. This can be seen by expressing

$$
\begin{aligned}
-cb_i\frac{\partial f_1}{\partial x_i} - c\frac{\partial b_i}{\partial x_i} f_1 &= b_i\frac{\partial}{\partial x_i}\left(\kappa b_j\frac{\partial f_0}{\partial x_j}\right) + \frac{\partial b_i}{\partial x_i}\kappa b_j\frac{\partial f_0}{\partial x_j} \\
&= \kappa b_i b_j\frac{\partial^2 f_0}{\partial x_i \partial x_j} + \kappa b_i\frac{\partial b_j}{\partial x_i}\frac{\partial f_0}{\partial x_j} + \frac{\partial b_i}{\partial x_i}\kappa b_j\frac{\partial f_0}{\partial x_j}. \qquad (5.15)
\end{aligned}
$$

The first term of (5.15) describes the isotropic and the anisotropic diffusive propagation of charged particles. The coefficients of $\partial f_0/\partial x_j$ in the second and third terms of (5.15) are evidently velocity terms that are associated with variations in b_i, i.e., these are drift terms associated either with gradients in **B**, $|\mathbf{B}|$, or large-scale curvature of **B**. Note that

$$
\frac{\partial b_i}{\partial x_i} = \nabla\cdot\mathbf{b} = \nabla\cdot\left(\frac{\mathbf{B}}{|\mathbf{B}|}\right) = -\frac{\mathbf{B}\cdot\nabla|\mathbf{B}|}{|\mathbf{B}|^2},
$$

is non-zero only if $|\mathbf{B}|$ varies spatially. This term is therefore related to the variation in pitch-angle that a single particle experiences as it propagates along a magnetic field that is converging or diverging. Consequently, the term $\nabla\cdot\mathbf{b} = L^{-1}$ defines the so-called focusing length L, and the collective effect of focusing is therefore embedded in the "diffusion" term of the transport equation (5.14). The terms $\partial b_j/\partial x_i$ when $i \neq j$ include the large-scale curvature in **B** since

$$
\frac{\partial b_j}{\partial x_i} = \frac{1}{|\mathbf{B}|}\frac{\partial B_j}{\partial x_i} - \frac{B_j}{|\mathbf{B}|^2}\frac{\partial|\mathbf{B}|}{\partial x_i}.
$$

The terms $\partial b_j/\partial x_i$ also describe gradients in the components of **B**.

If we now include the DU_i/Dt convective derivative that was neglected in the solution of first-order correction f_1, i.e., (5.12), the transport equation for f_0 becomes

$$
\frac{\partial f_0}{\partial t} + U_i\frac{\partial f_0}{\partial x_i} - \frac{\partial U_i}{\partial x_i}\frac{c}{3}\frac{\partial f_0}{\partial c} = \frac{\partial}{\partial x_i}\left(\frac{c^2\tau}{3}b_i b_j\frac{\partial f_0}{\partial x_j}\right) - \frac{\partial}{\partial x_i}\left(\frac{c\tau}{3}b_i b_j\frac{DU_i}{Dt}\frac{\partial f_0}{\partial c}\right).
$$

Use of the definition $\kappa = c^2\tau/3$ and the diffusion tensor **K** allows us to express the transport equation in the presence of large-scale flow gradients and accelerations as

$$
\frac{\partial f_0}{\partial t} + \mathbf{U}\cdot\nabla f_0 - \nabla\cdot\mathbf{u}\frac{c}{3}\frac{\partial f_0}{\partial c} + \nabla\cdot\left(\mathbf{K}\frac{D\mathbf{U}}{Dt}\frac{1}{c}\frac{\partial f_0}{\partial c}\right) = \nabla\cdot(\mathbf{K}\nabla f_0)\,. \qquad (5.16)
$$

The convective transport equation (5.14) and its extensions to relativistic charged particles is one of the most intensively studied equations in space physics and astrophysics as it is the basis for almost all work on energetic charged particle transport, ranging from galactic cosmic rays to solar energetic particles.

5.2 Transport Equation for Relativistic Charged Particles

5.2.1 Derivation of the Focussed Transport Equation

Consider now the extension of the previous two sections to include relativistic charged particles propagating in a non-relativistic background plasma flow with infinite conductivity.[1] It is assumed from the outset that the charged particles experience resonant scattering due to turbulent fluctuations in the background magnetic field. The fluctuations have typically been assumed to be magnetohydrodynamic waves, typically Alfvén waves, which tends to ensure that the scattered particles are trapped by the waves and stream with them. The waves define a frame of reference, the "wave frame," which propagates through the inertial or observer's (rest) frame and this is the frame in which the scattering is executed. In general, the wave frame is non-inertial, since, if we assume that the waves propagate at the local Alfvén speed V_A and they experience convection at the background plasma flow velocity $\mathbf{U}$, the wave frame velocity, $\mathbf{V}_A + \mathbf{U}$ may vary with space and time. This frame as expressed here also assumes that all the waves propagate uniformly in one direction which may not be appropriate. To avoid these complications, we shall assume that the background plasma flow speed sufficiently exceeds the Alfvén speed that we can neglect V_A. This is certainly true in the solar wind where $V_A \simeq 50\,\text{km/s}$ compared to the solar wind radial flow speed of 350–700 km/s.

The collisionless Vlasov equation that is valid for both relativistic and non-relativistic particles may be written as

$$\frac{d}{dt} f(\mathbf{x}, \mathbf{p}, t) \equiv \frac{\partial f}{\partial t} + v_i \frac{\partial f}{\partial x_i} + \frac{dp_i}{dt} \frac{\partial f}{\partial p_i} = 0, \tag{5.17}$$

where $f(\mathbf{x}, \mathbf{p}, t)$ is the distribution function in the rest frame and $d\mathbf{p}/dt$ is the force on the charged relativistic particle. In the wave frame, scattering of the particles does not change the momentum or energy of the particles, so we need to transform (5.17) into the wave frame. The transformations that we need are listed in the footnote.[2]

[1] As noted earlier, the transport equation was derived by Skilling (1971). His treatment is very brief and the development given here is guided by an excellent set of notes developed originally by Dr's G.M. Webb and J.A. le Roux, to whom I am indebted for sharing them with me.

[2] We summarize the various Lorentz transformations that are needed in the derivation of the focussed transport equation. A four-vector has three spatial components and one time component, $(x_0, x_1, x_2, x_3) = (x_0, x^a) = x^\alpha$, where small Roman superscripts denote spatial coordinates of the four-vector and Greek superscripts denote all four components. The length of a four-

We need to derive

$$\frac{d}{dt'} f(\mathbf{x}', \mathbf{p}', t') = 0,$$

where the distribution function and the variables correspond to the wave frame. Nevertheless, we will observe the cosmic rays in the rest or observer's frame, and this will therefore introduce a set of mixed coordinates as was done above. Exploiting the Lorentz invariance of the distribution function, $f(\mathbf{x}, \mathbf{p}, t) = f(\mathbf{x}', \mathbf{p}', t')$, we have

---

vector is $x^\alpha x^\alpha = x_1^2 + x_2^2 + x_3^2 - x_0^2 = x^a x^a - x_0^2$, and is invariant between coordinate systems. The contraction of any two four-vectors is invariant between coordinate systems. The Lorentz transformation matrix (see Jackson 1975, Sect. 11.7) enables one to transform one tensor to another. When the Lorentz matrix operates on a four-vector, it yields

$$x_0' = \gamma \left(x_0 - \beta^a x^a \right);$$
$$x^{a\prime} = x^a + \beta^a \left(\frac{\beta^b x^b}{\beta^2} (\gamma - 1) - \gamma x_0 \right); \gamma = \frac{1}{\sqrt{1 - U^2/c^2}}; \quad \beta^a = U^a/c,$$

where the transformation of the four vector is between reference frames in which the primed variable has velocity U^a relative to the non-primed variable. c denotes the speed of light. The corresponding inverse Lorentz matrix can of course be used. Typical four vectors are time-space $(ct, x^a) = x^\alpha$, and the energy and momentum of a particle $(E/c, p^a) = (mc, p^a) = \gamma m_0 (c, v^a)$, where m_0 is the rest mass of the particle. Since m_0 is constant, $\gamma(c, v^a)$ is a four-vector. Defining the proper time of a particle τ of a particle as $dt/d\tau = \gamma$ allows the four-velocity to be expressed as $dx^\alpha/d\tau$. The various transformations that we need are as follows:

$$t = \gamma \left(t' + \frac{\mathbf{x}' \cdot \mathbf{U}}{c^2} \right); \quad t' = \gamma \left(t - \frac{\mathbf{x} \cdot \mathbf{U}}{c^2} \right);$$
$$\mathbf{p} = \mathbf{p}' + m'\mathbf{U} \left[\frac{\mathbf{v}' \cdot \mathbf{U}}{c^2} (\gamma - 1) + \gamma \right];$$
$$\left(1 - \frac{\mathbf{v} \cdot \mathbf{U}}{c^2} \right) \gamma \mathbf{v}' = \mathbf{v} + \mathbf{U} \left[\frac{\mathbf{v} \cdot \mathbf{U}}{c^2} (\gamma - 1) - \gamma \right];$$
$$\left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2} \right) \gamma \mathbf{v} = \mathbf{v}' + \mathbf{U} \left[\frac{\mathbf{v}' \cdot \mathbf{U}}{c^2} (\gamma - 1) + \gamma \right];$$
$$\mathbf{E}' = \gamma \left(\mathbf{E} + \mathbf{U} \times \mathbf{B} \right) + (1 - \gamma) \frac{\mathbf{E} \cdot \mathbf{U}}{c^2} \mathbf{U};$$
$$\mathbf{B}' = \gamma \left(\mathbf{B} - \frac{1}{c^2} \mathbf{U} \times \mathbf{E} \right) + (1 - \gamma) \frac{\mathbf{B} \cdot \mathbf{U}}{v^2} \mathbf{U}.$$

Note that if $|\mathbf{U}|/c \ll 1$,

$$\gamma = (1 - U^2/c^2)^{-1/2} \simeq 1 + (U^2/2c)^2,$$

so that $\gamma \simeq 1$ is valid to the first order in U/c.

$$\frac{d}{dt} f(\mathbf{x}, \mathbf{p}, t) \frac{dt}{dt'} = 0.$$

The Lorentz transformation for time between the observer and wave frames yields to first-order in $|\mathbf{U}|/c$

$$t \simeq t' + \frac{\mathbf{x}' \cdot \mathbf{U}}{c^2}.$$

Consequently, we have

$$\frac{dt}{dt'} \simeq \left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right) \Longrightarrow \left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right) \frac{df}{dt}(\mathbf{x}, \mathbf{p}, t) = 0.$$

We now need to introduce a transformation so that the particle momentum is measured in the wave frame. This requires that the various partial derivatives in the Vlasov equation are transformed from the observer's frame to the wave frame i.e., $(\mathbf{x}, \mathbf{p}, t) \mapsto (\mathbf{x}, \mathbf{p}', t)$. This requires the use of the inverse Lorentz transformation for particle momentum (Footnote), which to first order in $|\mathbf{U}|/c$ yields $\gamma \simeq 1$ and

$$\mathbf{p}' = \mathbf{p} - m'\mathbf{U},$$

where $m' = \gamma' m_0$ and $\gamma' = 1/\sqrt{1 - v'^2/c^2}$ for the relativistic particle in the observer's frame. Considering the time derivative yields

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t} + \frac{\partial p_i'}{\partial t} \frac{\partial}{\partial p_i'} = \frac{\partial}{\partial t} + \frac{\partial}{\partial t}(p_i - m'U_i) \frac{\partial}{\partial p_i'} = \frac{\partial}{\partial t} - m' \frac{\partial U_i}{\partial t} \frac{\partial}{\partial p_i'}.$$

The spatial derivative transforms as

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial x_i} + \frac{\partial p_j'}{\partial x_i} \frac{\partial}{\partial p_j'} = \frac{\partial}{\partial x_i} - m' \frac{\partial U_j}{\partial x_i} \frac{\partial}{\partial p_j'}.$$

Finally, instead of the inverse transform, we use $\mathbf{p}' = \mathbf{p} - m\mathbf{U}$ to obtain

$$\frac{\partial}{\partial p_i} = \frac{\partial p_j'}{\partial p_i} \frac{\partial}{\partial p_j'} = \frac{\partial}{\partial p_i} \left(p_j - mU_j\right) \frac{\partial}{\partial p_j'} = \delta_{ij} \frac{\partial}{\partial p_j'} - U_j \frac{\partial m}{\partial p_i} \frac{\partial}{\partial p_j'}.$$

Introducing the basis vector for spherical coordinates allows us to express

$$\frac{\partial m}{\partial p_i} = \hat{e}_{pi} \frac{\partial m}{\partial p},$$

and since

$$m = \gamma m_0 = m_0 \left(1 - \frac{v^2}{c^2}\right)^{-1/2} = m_0 \left(1 - \frac{p^2}{m^2 c^2}\right)^{-1/2} \Rightarrow m = \left(1 + \frac{p^2}{m_0^2 c^2}\right)^{1/2},$$

we have

$$\frac{dm}{dp} = \frac{p}{mc^2}.$$

We then obtain

$$\frac{\partial}{\partial p_i} = \delta_{ij}\frac{\partial}{\partial p'_j} - \frac{v_i U_j}{c^2}\frac{\partial}{\partial p'_j}.$$

On retaining only terms of $O(U/c)$, we obtain

$$\left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right)\frac{\partial f}{\partial t} = \left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right)\left(\frac{\partial f}{\partial t} - m\frac{\partial U_i}{\partial t}\frac{\partial f}{\partial p'_i}\right)$$
$$\simeq \left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right)\frac{\partial f}{\partial t} - m\frac{\partial U_i}{\partial t}\frac{\partial f}{\partial p'_i}.$$

Consider now the convective term, $\left(1 + \mathbf{v}' \cdot \mathbf{U}/c^2\right) v_i \partial f/\partial x_i$. To the first order in U/c, using the Lorentz transformation for the velocity, and $\gamma \simeq 1$ gives

$$\mathbf{v} = \frac{\mathbf{v}' + \mathbf{U}}{1 + \mathbf{v}' \cdot \mathbf{U}/c^2}.$$

This then yields

$$\left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right)\frac{v'_i + U_i}{1 + \mathbf{v}' \cdot \mathbf{U}/c^2}\left(\frac{\partial f}{\partial x_i} - m'\frac{\partial U_j}{\partial x_i}\frac{\partial f}{\partial p'_j}\right) = \left(v'_i + U_i\right)\frac{\partial f}{\partial x_i}$$
$$- m'\left(v'_i + U_i\right)\frac{\partial U_j}{\partial x_i}\frac{\partial f}{\partial'_j}.$$

Consider now the momentum change term $\left(1 + \mathbf{v}' \cdot \mathbf{U}/c^2\right)(dp_i/dt)\partial f/\partial p_i$. We assume that the momentum change is due to electromagnetic fields only. Thus, we have the Lorentz force

$$\frac{dp_i}{dt} = q\left(E_i + \varepsilon_{ijk} v_j B_k\right),$$

where q is the particle charge, $\mathbf{B}$ the external magnetic field, $\mathbf{E}$ the electric field, and ε_{ijk} is the Levi-Civita tensor. The first order Lorentz transformation for $\mathbf{E}$ is simply

$$\mathbf{E}' \simeq \mathbf{E} + \mathbf{U} \times \mathbf{B} \Longleftrightarrow \mathbf{E} \simeq \mathbf{E}' - \mathbf{U} \times \mathbf{B},$$

which yields

$$\frac{dp_i}{dt} = q\left(E'_i + \varepsilon_{ijk}(v_j - U_j)B_k\right).$$

To address the transformation of the velocity, the Lorentz transformation yields $\gamma \simeq 1$ and

$$\left(1 - \frac{\mathbf{v}\cdot\mathbf{U}}{c^2}\right)\mathbf{v}' = \mathbf{v} - \mathbf{U},$$

at $O(U/v)$, from which we find

$$\frac{dp_i}{dt} = q\left[E_i' + \left(1 - \frac{\mathbf{v}\cdot\mathbf{U}}{c^2}\right)\varepsilon_{ijk}v_j' B_k\right],$$

so that

$$\left(1 + \frac{\mathbf{v}'\cdot\mathbf{U}}{c^2}\right)\frac{dp_i}{dt}\frac{\partial f}{\partial p_i} = q\left(1 + \frac{\mathbf{v}'\cdot\mathbf{U}}{c^2}\right)E_i'\frac{\partial f}{\partial p_i} + q\left(1 + \frac{\mathbf{v}'\cdot\mathbf{U}}{c^2}\right)\left(1 - \frac{\mathbf{v}\cdot\mathbf{U}}{c^2}\right)\varepsilon_{ijk}v_j' B_k\frac{\partial f}{\partial p_i}.$$

The Lorentz transformation for time and its inverse yield

$$\frac{dt'}{dt} = \gamma\left(1 - \frac{\mathbf{v}\cdot\mathbf{U}}{c^2}\right); \quad \frac{dt}{dt'} = \gamma\left(1 + \frac{\mathbf{v}'\cdot\mathbf{U}}{c^2}\right),$$

from which we obtain

$$\gamma^2\left(1 - \frac{\mathbf{v}\cdot\mathbf{U}}{c^2}\right)\left(1 + \frac{\mathbf{v}'\cdot\mathbf{U}}{c^2}\right) = 1,$$

or

$$\left(1 - \frac{\mathbf{v}\cdot\mathbf{U}}{c^2}\right)\left(1 + \frac{\mathbf{v}'\cdot\mathbf{U}}{c^2}\right) = 1$$

in the limit $U/c \ll 1$. We may therefore derive

$$\left(1 + \frac{\mathbf{v}'\cdot\mathbf{U}}{c^2}\right)\frac{dp_i}{dt}\frac{\partial f}{\partial p_i} = q\left[\left(1 + \frac{\mathbf{v}'\cdot\mathbf{U}}{c^2}\right)E_i' + \varepsilon_{ijk}v_j' B_k\right]\left(\delta_{ij} - \frac{v_i U_j}{c^2}\right)\frac{\partial f}{\partial p_j'}.$$

Now consider the Lorentz transformation of the magnetic field. To first order, we have

$$\mathbf{B}' = \mathbf{B} - \frac{1}{c^2}\mathbf{U}\times\mathbf{E},$$

but since $\mathbf{E} = -\mathbf{U}\times\mathbf{B}$, $\mathbf{B}' = \mathbf{B} + \mathbf{U}\times(\mathbf{U}\times\mathbf{B})/c^2$, this implies that

$$\mathbf{B}' = \mathbf{B},$$

at this order. This, together with

$$\frac{dp'_i}{dt} = q\left(E'_i + \varepsilon_{ijk} v'_j B'_k\right),$$

allows us to write

$$\left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right) \frac{dp_i}{dt} \frac{\partial f}{\partial p_i} = q \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2} E'_i \left(\delta_{ij} - \frac{v_i U_j}{c^2}\right) \frac{\partial f}{\partial p'_j} + \frac{dp'_i}{dt} \left(\delta_{ij} - \frac{v_i U_j}{c^2}\right) \frac{\partial f}{\partial p'_j}.$$

Consider the term $q\varepsilon_{ijk} v'_j B'_k (v_i U_j / c^2) \partial f / \partial p'_j$. The first order velocity transformation yields

$$\mathbf{v} = \frac{\mathbf{v}' + \mathbf{U}}{1 + \mathbf{v}' \cdot \mathbf{U}/c^2},$$

so that

$$\begin{aligned} q\varepsilon_{ijk} v'_j B'_k \frac{v_i U_j}{c^2} \frac{\partial f}{\partial p'_j} &= \frac{q}{1 + \mathbf{v}' \cdot \mathbf{U}/c^2} \varepsilon_{ijk} v'_j B'_k \frac{v'_i + U_i}{c^2} U_j \frac{\partial f}{\partial p'_j} \\ &\simeq \frac{q}{1 + \mathbf{v}' \cdot \mathbf{U}/c^2} \frac{1}{c^2} \varepsilon_{ijk} v'_j B'_k v'_i U_j \frac{\partial f}{\partial p'_j} \\ &= \frac{q}{1 + \mathbf{v}' \cdot \mathbf{U}/c^2} \frac{1}{c^2} \left(\mathbf{v}' \times \mathbf{B}'\right) \cdot \mathbf{v}' U_j \frac{\partial f}{\partial p'_j} = 0, \end{aligned}$$

after neglecting the $U_i U_j / c^2$ term in the second line. The term $q(\mathbf{v}' \cdot \mathbf{U}/c^2) E'_i v_i (U_j / c^2) \partial f / \partial p'_j$ is $O\left((U/c)^2\right)$ and so is neglected. We therefore obtain

$$\left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right) \frac{dp_i}{dt} \frac{\partial f}{\partial p_i} \simeq \left[q \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2} E'_i + \frac{dp'_i}{dt}\right] \frac{\partial f}{\partial p'_i},$$

where

$$\frac{dp'_i}{dt} = q\left(E'_i + \varepsilon_{ijk} v'_j B'_k\right).$$

On combining the results above, we obtain the Vlasov equation in mixed coordinates,

$$\begin{aligned} &\left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right) \frac{\partial f}{\partial t} + \left(\mathbf{v}' + \mathbf{U}\right) \cdot \nabla f + q \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2} E'_i \frac{\partial f}{\partial p'_i} \\ &- \left[m' \left(\frac{\partial U_i}{\partial t} + U_j \frac{\partial U_i}{\partial x_j}\right) + p'_j \frac{\partial U_i}{\partial x_j} - \frac{dp'_i}{dt}\right] \frac{\partial f}{\partial p'_i} = 0, \end{aligned} \tag{5.18}$$

with $f(\mathbf{x}, \mathbf{p}, t)$. However, since the coordinates $(\mathbf{x}, \mathbf{p}', t)$ are in the mixed coordinate system, we need to introduce the transformation $f(\mathbf{x}, \mathbf{p}, t) \mapsto f'(\mathbf{x}, \mathbf{p}', t)$. Recall that

$$f(\mathbf{x}, \mathbf{p}, t) d^3\mathbf{x} d^3\mathbf{p} = f'(\mathbf{x}, \mathbf{p}', t) d^3\mathbf{x}' d^3\mathbf{p}',$$

and that $d^3\mathbf{x} = \gamma d^3\mathbf{x}'$. Consider the transformation of the volume element in momentum space. On using $\mathbf{p} = \mathbf{p}' + m'\mathbf{U}$, $dm'/dp = p/(m'c^2)$, and $dp/dp_i = p_i/p$, we have

$$\begin{aligned} d^3\mathbf{p} &= dp_x dp_y dp_z \\ &= \left(dp'_x + U_x \frac{p'}{m'c^2} dp'\right)\left(dp'_y + U_y \frac{p'}{m'c^2} dp'\right)\left(dp'_z + U_z \frac{p'}{m'c^2} dp'\right) \\ &= dp'_x dp'_y dp'_z \left(1 + \frac{p'_x U_x}{m'c^2}\right)\left(1 + \frac{p'_y U_y}{m'c^2}\right)\left(1 + \frac{p'_z U_z}{m'c^2}\right) \\ &= dp'_x dp'_y dp'_z \left(1 + \frac{\mathbf{p}' \cdot \mathbf{U}}{m'c^2}\right) + O\left(\frac{U^2}{c^2}\right) \\ &\simeq dp'_x dp'_y dp'_z \left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right). \end{aligned}$$

Thus, we have the transformation

$$f(\mathbf{x}, \mathbf{p}, t) = \frac{f'(\mathbf{x}, \mathbf{p}', t)}{\gamma\,(1 + \mathbf{v}' \cdot \mathbf{U}/c^2)},$$

which to first order in U/c, $\gamma \simeq 1$ yields

$$f(\mathbf{x}, \mathbf{p}, t) = \frac{f'(\mathbf{x}, \mathbf{p}', t)}{(1 + \mathbf{v}' \cdot \mathbf{U}/c^2)} \equiv f''(\mathbf{x}, \mathbf{p}', t).$$

On setting $f(\mathbf{x}, \mathbf{p}, t) = f''(\mathbf{x}, \mathbf{p}', t)$ in (5.18), we have the final form of the transformed equation,

$$\begin{aligned} &\left(1 + \frac{\mathbf{v}' \cdot \mathbf{U}}{c^2}\right)\frac{\partial f''}{\partial t} + \left(\mathbf{v}' + \mathbf{U}\right) \cdot \nabla f'' + q\frac{\mathbf{v}' \cdot \mathbf{U}}{c^2} E'_i \frac{\partial f''}{\partial p'_i} \\ &-\left[m'\left(\frac{\partial U_i}{\partial t} + U_j \frac{\partial U_i}{\partial x_j}\right) + p'_j \frac{\partial U_i}{\partial x_j}\right]\frac{\partial f''}{\partial p'_i} + \frac{\partial}{\partial p'_i}\left(F'_i f''\right) = 0. \end{aligned} \tag{5.19}$$

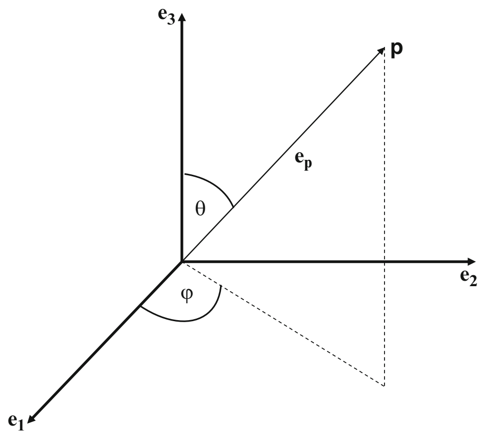


Fig. 5.1 The coordinates for a particle gyrating about a mean magnetic field **B** oriented along the z-axis. The particle momentum is given by the vector **p**, the pitch-angle by θ, and the gyrophase by ϕ. The directional vector $\mathbf{b} \equiv \mathbf{B}/|\mathbf{B}| = \mathbf{e}_3$

In deriving (5.19), we used

$$\begin{aligned}
\frac{\partial}{\partial p'_i}\left(\frac{dp'_i}{dt}\right) &= \frac{\partial}{\partial p'_i}\left(q\varepsilon_{ijk}v'_j B'_k\right) = q\varepsilon_{ijk}\frac{\partial v'_j}{\partial p'_i}B'_k \\
&= q\varepsilon_{ijk}B'_k e'_{pi}\frac{\partial}{\partial p'}\left(e'_{pj}v'\right) \\
&= q\varepsilon_{ijk}e'_{pi}e'_{pj}\frac{\partial v'}{\partial p'}B'_k \\
&= qe'_{pi}(\varepsilon_{ijk}e'_{pj}B'_k)\frac{1}{m'\gamma'^2} \\
&= \frac{q}{m'\gamma'^2}\mathbf{e}'_p\cdot\left(\mathbf{e}'_p\times\mathbf{B}'\right) = 0.
\end{aligned}$$

Just as we did in the derivation of the focussed transport equation for non-relativistic particles, we shall assume that the particle distribution function is nearly gyrotropic, making $f(\mathbf{x}, \mathbf{v}, t) \simeq f(\mathbf{x}, v, \mu, t)$ where the particle pitch angle is $\mu \equiv \cos\theta$ as before. For the sake of notational convenience, we henceforth drop the "prime" on the variables and distribution function. The averaging procedure proceeds in much the same way as before. For completeness, we provide some of the details in the derivation although using a slightly more general notation. The local geometry of a charged particle gyrating about the mean magnetic field **B** is illustrated in Fig. 5.1. The coordinates (x_1, x_2, x_3) refer to a magnetic field system and $\mathbf{e}_3 = \mathbf{b} \equiv \mathbf{B}/|\mathbf{B}|$. Since the magnetic field is not assumed to be uniform, the unit vectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 = \mathbf{b})$ are functions of **x**. As before, $\mu = \cos\theta = \mathbf{e}_p \cdot \mathbf{b}$.

Recall that the momentum can be expressed in spherical coordinates as

$$\frac{\partial f}{\partial \mathbf{p}} = \mathbf{e}_p\frac{\partial f}{\partial p} + \mathbf{e}_\theta\frac{1}{p}\frac{\partial f}{\partial \theta} + \mathbf{e}_\phi\frac{1}{p\sin\theta}\frac{\partial f}{\partial \phi}$$

where

$$\mathbf{p} = p\left(\sin\theta\cos\phi\mathbf{e}_1 + \sin\theta\sin\phi\mathbf{e}_2 + \cos\theta\mathbf{b}\right);$$
$$\mathbf{e}_p = \frac{\partial\mathbf{p}}{\partial p} = \sin\theta\cos\phi\mathbf{e}_1 + \sin\theta\sin\phi\mathbf{e}_2 + \cos\theta\mathbf{b};$$
$$\mathbf{e}_\theta = \frac{1}{p}\frac{\partial\mathbf{p}}{\partial\theta} = \cos\theta\cos\phi\mathbf{e}_1 + \cos\theta\sin\phi\mathbf{e}_2 - \sin\theta\mathbf{b};$$
$$\mathbf{e}_\phi = \frac{1}{p\sin\theta}\frac{\partial\mathbf{p}}{\partial\phi} = -\sin\phi\mathbf{e}_1 + \cos\phi\mathbf{e}_2.$$

As before, we require the following integrals,

$$\left\langle\mathbf{e}_p\right\rangle = \frac{1}{2\pi}\int_0^{2\pi}\mathbf{e}_p d\phi = \cos\theta\mathbf{b};$$
$$\left\langle\mathbf{e}_p\mathbf{e}_p\right\rangle = \frac{1}{2\pi}\int_0^{2\pi}\mathbf{e}_p\mathbf{e}_p d\phi = \frac{1}{2}\sin^2\theta\left[\mathbf{I}-\mathbf{bb}\right] + \cos^2\theta\mathbf{bb},$$

after using $\langle\cos^2\phi\rangle = \langle\sin^2\phi\rangle = 1/2$.

Consider the time derivative

$$\begin{aligned}\frac{\partial f}{\partial t} &= \frac{\partial f}{\partial t} + \frac{\partial\mu}{\partial t}\frac{\partial f}{\partial\mu}\\ &= \frac{\partial f}{\partial t} + \frac{\partial}{\partial t}\left(\mathbf{e}_p\cdot\mathbf{b}\right)\frac{\partial f}{\partial\mu} = \frac{\partial f}{\partial t}\end{aligned}$$

after using $b_i\partial b_i/\partial t = 0$ as before. Evidently, $\langle\partial f/\partial t\rangle = \partial f/\partial t$. Now,

$$\begin{aligned}\left\langle\frac{\mathbf{p}\cdot\mathbf{U}}{mc^2}\right\rangle\frac{\partial f}{\partial t} &= \frac{1}{2\pi}\int_0^{2\pi}\frac{p_iU_i}{mc^2}\left[\frac{\partial f}{\partial t} + e_{pj}\frac{\partial b_j}{\partial t}\frac{\partial f}{\partial\mu}\right]d\phi\\ &= \frac{U_i}{mc^2}\frac{1}{2\pi}\int_0^{2\pi}p_i\frac{\partial f}{\partial t}d\phi + \frac{U_i}{mc^2}\frac{\partial b_j}{\partial t}\frac{1}{2\pi}\int_0^{2\pi}p_ie_{pj}\frac{\partial f}{\partial\mu}d\phi\\ &= \frac{pU_i}{mc^2}\frac{\partial f}{\partial t}\frac{1}{2\pi}\int_0^{2\pi}e_{pi}d\phi + \frac{pU_i}{mc^2}\frac{\partial b_j}{\partial t}\frac{\partial f}{\partial\mu}\frac{1}{2\pi}\int_0^{2\pi}e_{pi}e_{pj}d\phi\\ &= \frac{p\mu}{mc^2}U_ib_i\frac{\partial f}{\partial t} + \frac{pU_i}{mc^2}\frac{\partial b_j}{\partial t}\frac{\partial f}{\partial\mu}\left[\frac{1}{2}\sin^2\theta(\delta_{ij}-b_ib_j) + \cos^2\theta b_ib_j\right]\\ &= \frac{v\mu}{c}\left(\frac{\mathbf{U}}{c}\cdot\mathbf{b}\right)\frac{\partial f}{\partial t} + \frac{1}{2}\frac{v(1-\mu^2)}{c}\left(\frac{\mathbf{U}}{c}\cdot\frac{\partial\mathbf{b}}{\partial t}\right)\frac{\partial f}{\partial\mu}.\end{aligned}$$

On considering the convective term,

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial x_i} + e_{pj}\frac{\partial b_j}{\partial x_i}\frac{\partial f}{\partial \mu},$$

we derive the gyrophase averaged expression

$$\left\langle \left(U_i + \frac{p_i}{m}\right)\frac{\partial f}{\partial x_i}\right\rangle = U_i\frac{\partial f}{\partial x_i} + \frac{p}{m}\frac{\partial f}{\partial x_i}\langle e_{pi}\rangle + U_i\frac{\partial b_j}{\partial x_i}\frac{\partial f}{\partial \mu}\langle e_{pj}\rangle + \frac{p}{m}\frac{\partial b_j}{\partial x_i}\frac{\partial f}{\partial \mu}\langle e_{pi}e_{pj}\rangle$$
$$= (U_i + v\mu b_i)\frac{\partial f}{\partial x_i} + \frac{1}{2}v(1-\mu^2)\frac{\partial b_i}{\partial x_i}\frac{\partial f}{\partial \mu}.$$

On expressing

$$\begin{aligned}\frac{\partial f}{\partial p_i} &= \frac{\partial p}{\partial p_i}\frac{\partial f}{\partial p} + \frac{\partial \mu}{\partial p_i}\frac{\partial f}{\partial \mu}\\ &= e_{pi}\frac{\partial f}{\partial p} + \frac{\partial}{\partial p_i}(e_{pj}b_j)\frac{\partial f}{\partial \mu}\\ &= e_{pi}\frac{\partial f}{\partial p} + b_j\frac{\partial}{\partial p_i}\left(\frac{p_j}{p}\right)\frac{\partial f}{\partial \mu}\\ &= e_{pi}\frac{\partial f}{\partial p} + \frac{b_j}{p}\left(\delta_{ij} - \frac{p_j}{p}\frac{\partial p}{\partial p_i}\right)\frac{\partial f}{\partial \mu}\\ &= e_{pi}\frac{\partial f}{\partial p} + \frac{b_j}{p}(\delta_{ij} - e_{pi}e_{pj})\frac{\partial f}{\partial \mu},\end{aligned}$$

we may consider

$$\left\langle -m\frac{dU_i}{dt}\frac{\partial f}{\partial p_i}\right\rangle = -m\frac{dU_i}{dt}\frac{\partial f}{\partial p}\langle e_{pi}\rangle - m\frac{dU_i}{dt}\frac{b_j}{p}\delta_{ij}\frac{\partial f}{\partial \mu} + m\frac{dU_i}{dt}\frac{b_j}{p}\frac{\partial f}{\partial \mu}\langle e_{pi}e_{pj}\rangle$$
$$= -m\mu\frac{dU_i}{dt}b_i\frac{\partial f}{\partial p} - \frac{m}{p}(1-\mu^2)\frac{dU_i}{dt}b_i\frac{\partial f}{\partial \mu}.$$

We can similarly evaluate

$$\left\langle -p_k\frac{\partial U_i}{\partial x_k}\frac{\partial f}{\partial p_i}\right\rangle = -p\left[\frac{1}{2}(1-\mu^2)\frac{\partial U_i}{\partial x_i} + \frac{1}{2}(3\mu^2-1)b_ib_j\frac{\partial U_i}{\partial x_j}\right]\frac{\partial f}{\partial p}$$
$$+ \frac{1}{2}(1-\mu^2)\left[\frac{\partial U_i}{\partial x_i} - 3b_ib_j\frac{\partial U_i}{\partial x_j}\right]\mu\frac{\partial f}{\partial \mu}.$$

Finally,

$$\frac{q}{m}(\mathbf{p}\times\mathbf{B})\cdot\frac{\partial f}{\partial \mathbf{p}} = \frac{q}{m}(\mathbf{p}\times\mathbf{B})\cdot\left[\mathbf{e}_p\frac{\partial f}{\partial p} + \left(\frac{\mathbf{b}}{p} - \mu\mathbf{e}_p\right)\frac{\partial f}{\partial \mu}\right] = 0,$$

since $(\mathbf{p}\times\mathbf{B})\cdot\mathbf{b} = 0$ and $(\mathbf{p}\times\mathbf{B})\cdot\mathbf{e}_p = 0$.

On combining the results above, we obtain the focussed transport equation or, equivalently, the Boltzmann equation for a gyrotropic particle distribution,

$$\left(1+\frac{v\mu}{c}\frac{\mathbf{U}\cdot\mathbf{b}}{c}\right)\frac{\partial f}{\partial t}+(\mathbf{U}+v\mu\mathbf{b})\cdot\nabla f+\frac{1-\mu^2}{2}\left[\frac{v}{c}\left(\frac{\mathbf{U}}{c}\cdot\frac{\partial\mathbf{b}}{\partial t}\right)+v\nabla\cdot\mathbf{b}\right.$$
$$\left.-\frac{2m}{p}\left(\frac{d\mathbf{U}}{dt}\cdot\mathbf{b}\right)+\mu\left(\nabla\cdot\mathbf{U}-3\mathbf{bb}:\nabla\mathbf{U}\right)\right]\frac{\partial f}{\partial\mu}-\left[\frac{\mu m}{p}\left(\frac{d\mathbf{U}}{dt}\cdot\mathbf{b}\right)\right.$$
$$\left.+\frac{1}{2}(1-\mu^2)\nabla\cdot\mathbf{U}+\frac{1}{2}(3\mu^2-1)\mathbf{bb}:\nabla\mathbf{U}\right]p\frac{\partial f}{\partial p}=\left(\frac{\delta f}{\delta t}\right)_s. \tag{5.20}$$

The righthand term is the scattering term, due charged particles scattering in pitch-angle due to the stochastically fluctuating magnetic field. Certainly for parallel propagation, the scattering fluctuations are typically assumed to be Alfvén waves. The scattering of charged particles conserves particle energy in the wave frame. In the transformation from the observer's frame (the rest frame) to the wave frame, the macroscopic electric fields are transformed away by the background velocity **U** because the plasma is infinitely conductive. Electric fields associated with the waves disappear in a frame moving with the waves. In the absence of electric fields, charged particles can only experience scattering in pitch angle. Energy is not, however, conserved in the observer's frame.

On assuming that $d\mathbf{U}/dt = 0$ and neglecting terms $O(U/c)$, we recover the usual form of the focussed transport equation,

$$\frac{\partial f}{\partial t}+(\mathbf{U}+v\mu\mathbf{b})\cdot\nabla f+\left[\frac{1-3\mu^2}{2}(\mathbf{bb}:\nabla\mathbf{U})-\frac{1}{2}(1-\mu^2)\nabla\cdot\mathbf{U}\right]p\frac{\partial f}{\partial p}$$
$$+\frac{1-\mu^2}{2}\left[v\nabla\cdot\mathbf{b}+\mu\nabla\cdot\mathbf{U}-3\mu\mathbf{bb}:\nabla\mathbf{U}\right]\frac{\partial f}{\partial\mu}=\left(\frac{\delta f}{\delta t}\right)_s. \tag{5.21}$$

The focussed transport equation (5.21) can be reduced to the convective-diffusive equation if the distribution function $f(\mathbf{x},p,\mu,t)\simeq f(\mathbf{x},p,t)$ i.e., if the scattering experienced by the particle is sufficiently strong that the distribution is nearly isotropic. The analysis of Sect. **2** carries over directly with "c" being replaced by "p", and the general convective-diffusive transport equation is given by

$$\frac{\partial f}{\partial t}+\mathbf{U}\cdot\nabla f-\frac{p}{3}\nabla\cdot\mathbf{U}\frac{\partial f}{\partial p}=\nabla\left(\mathbf{K}\nabla f\right). \tag{5.22}$$

This is the standard form of the transport equation for relativistic charged particles experiencing scattering in a non-relativistic turbulent plasma.

Exercises

1. Derive the following averaging relations:

$$\langle \mathbf{e}_p \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{e}_p d\phi = \cos\theta \mathbf{b};$$

$$\langle \mathbf{e}_p \mathbf{e}_p \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{e}_p \mathbf{e}_p d\phi = \frac{1}{2} \sin^2\theta \,[\mathbf{I} - \mathbf{bb}] + \cos^2\theta \mathbf{bb}.$$

2. Complete the derivation of

$$\left\langle \left(U_i + \frac{p_i}{m} \right) \frac{\partial f}{\partial x_i} \right\rangle = (\mathbf{U} + v\mu \mathbf{b}) \cdot \nabla f + \frac{1}{2} v(1-\mu^2) \nabla \cdot \mathbf{b} \frac{\partial f}{\partial \mu}.$$

3. Show that

$$\frac{\partial}{\partial p_i}(e_{pj} b_j) = \frac{b_j}{p}(\delta_{ij} - e_{pi} e_{pj}).$$

4. Complete the derivation of

$$\left\langle -p_k \frac{\partial U_i}{\partial x_k} \frac{\partial f}{\partial p_i} \right\rangle = -p \left[\frac{1}{2}(1-\mu^2) \frac{\partial U_i}{\partial x_i} + \frac{1}{2}(3\mu^2 - 1) b_i b_j \frac{\partial U_i}{\partial x_j} \right] \frac{\partial f}{\partial p}$$
$$+ \frac{1}{2}(1-\mu^2) \left[\frac{\partial U_i}{\partial x_i} - 3 b_i b_j \frac{\partial U_i}{\partial x_j} \right] \mu \frac{\partial f}{\partial \mu}.$$

5.3 The Magnetic Correlation Tensor

As will be discussed in detail below, the magnetic correlation tensor plays a central role in determining the transport properties of particles experiencing pitch-angle scattering by turbulent magnetic field fluctuations. A very detailed discussion of different forms of the magnetic correlation tensor has been presented by Shalchi (2009).[3] The general form of the two-point, two-time magnetic correlation tensor has the form

$$R_{ij}(\mathbf{r}, t, \mathbf{r}_\prime, t_0) = \langle \delta B_i(\mathbf{r}, t), \delta B_j(\mathbf{r}', t_0) \rangle,$$

where r' denotes a different spatial location and $\langle \cdot \rangle$ an ensemble average. It is convenient to consider the correlation tensor using a Fourier representation

[3]See also Tautz and Lerche (2011).

$$\delta B_i(\mathbf{r},t) = \int \delta B_i(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{r}}d^3k,$$

from which we find

$$R_{ij}(\mathbf{r},t,\mathbf{r}',t_0) = \int d^3k \int d^3k' \left\langle \delta B_i(\mathbf{k},t)\delta B_j(\mathbf{k}',t_0)\right\rangle e^{i\mathbf{k}\cdot\mathbf{r}+i\mathbf{k}'\cdot\mathbf{r}'}. \tag{5.23}$$

As is typically assumed, we suppose that the magnetic turbulence is homogeneous, so that the correlation function depends only on the separation $|\mathbf{r}-\mathbf{r}'|$ between two points. Then we can express $\left\langle \delta B_i(\mathbf{k},t), \delta B_j(\mathbf{k}',t_0)\right\rangle$ as $\left\langle \delta B_i(\mathbf{k},t), \delta B_j(\mathbf{k}',t_0)\right\rangle \delta(\mathbf{k}+\mathbf{k}')$, which allows us to integrate (5.23) as

$$R_{ij} = \int d^3k \left\langle \delta B_i(\mathbf{k},t), \delta B_j(-\mathbf{k},t_0)\right\rangle e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}.$$

From the definition of the Fourier transform, $\delta B_j(-\mathbf{k}) = \delta B_j^*(\mathbf{k})$ where * denotes the complex conjugate. This allows us to introduce the usual definition of the correlation tensor,

$$P_{ij}(\mathbf{k},t,t_0) = \left\langle \delta B_i(\mathbf{k},t)\delta B_j^*(\mathbf{k},t_0)\right\rangle,$$

and the correlation tensor $P_{ij}(\mathbf{k},t,t_0)$ is expressed in wave number space. The correlation tensor (5.23) then reduces to

$$R_{ij}(\mathbf{r},t,\mathbf{r}',t_0) = \int d^3k P_{ij}(\mathbf{k},t,t_0)e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}.$$

On setting $t_0 = 0$ and $\mathbf{r}' = 0$, we have

$$P_{ij}(\mathbf{k},t) = \langle \delta B_i(\mathbf{k},t)\delta B_j^*(\mathbf{k},0)\rangle,$$

with

$$R_{ij}(\mathbf{r},t) = \int d^3k P_{ij}(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{5.24}$$

Although we restrict ourselves to stationary turbulence, we note that the inclusion of temporal effects in the correlation tensor is typically accomplished by assuming that the correlation tensor has a separable form in the spatial and temporal components,

$$P_{ij}(\mathbf{k},t) = P_{ij}(\mathbf{k},0)\Gamma(\mathbf{k},t),$$

where $\Gamma(\mathbf{k},t)$ is a dynamical correlation function and $P_{ij}(\mathbf{k},0) \equiv P_{ij}(\mathbf{k})$ is the magnetostatic correlation tensor.

For completeness, we first consider isotropic turbulence. The general form of an isotropic rank-2 tensor is[4]

$$P_{ij}(\mathbf{k}) = A(k)\delta_{ij} + B(k)k_i k_j + C(k)\sum_k \varepsilon_{ijk}k_k.$$

Recall that ε_{ijk} is the Levi-Civita or unit alternating tensor and has values $\varepsilon_{ijk} = 0$ if any of i, j, and k are repeated, $\varepsilon_{ijk} = +1$ or -1 when i, j, and k are all different and in cyclic or acyclic order respectively.

Since $\nabla \cdot \delta\mathbf{B} = 0$,

$$\sum_i k_i \delta B_i(\mathbf{k}) = 0,$$

which yields

$$\sum_{i,j} \langle k_i \delta B_i k_j \delta B_j^* \rangle = \sum_{i,j} k_i k_j P_{ij} = 0.$$

If we substitute the general form P_{ij} of an isotropic rank-2 tensor, it therefore follows immediately that for magnetic turbulence

$$\begin{aligned} 0 &= A(k)\sum_{i,j} k_i k_j \delta_{ij} + B(k)\sum_{i,j} k_i^2 k_j^2 + C(k)\sum_{i,j,k} \varepsilon_{ijk}k_i k_j k_k \\ &= A(k)k^2 + B(k)k^4, \end{aligned}$$

and hence that

$$B(k) = -\frac{A(k)}{k^2}.$$

The general form of the magnetic isotropic tensor is therefore

$$P_{ij}(\mathbf{k}) = A(k)\left(\delta_{ij} - \frac{k_i k_j}{k^2}\right) + C(k)\sum_k \varepsilon_{ijk}k_k.$$

Since

$$P_{ij}(\mathbf{k}) = \langle \delta B_i \delta B_j^* \rangle = \langle \delta B_i^* \delta B_j \rangle^* = \langle \delta B_j \delta B_i^* \rangle^* = P_{ji}^*(\mathbf{k}),$$

[4] Batchelor (1953)

we have

$$\begin{aligned}
P^*_{ji} &= A^*(k)\left(\delta_{ji} - \frac{k_j k_i}{k^2}\right) + C^*(k)\sum_k \varepsilon_{jik} k_k \\
&= A^*(k)\left(\delta_{ij} - \frac{k_i k_j}{k^2}\right) - C^*(k)\sum_k \varepsilon_{ijk} k_k \\
&= P_{ij}(\mathbf{k}) = A(k)\left(\delta_{ij} - \frac{k_i k_j}{k^2}\right) + C(k)\sum_k \varepsilon_{ijk} k_k,
\end{aligned}$$

after using $\varepsilon_{jik} = -\varepsilon_{ijk}$. We therefore have $A(k) = A^*(k)$, i.e., $A(k)$ is real, and $C(k) = -C^*(k)$, implying that $C(k)$ is imaginary. Quite generally, we can express

$$C(k) = iA(k)\frac{\sigma(k)}{k},$$

to obtain

$$P_{ij}(\mathbf{k}) = A(k)\left[\delta_{ij} - \frac{k_i k_j}{k^2} + i\sigma(k)\sum_k \varepsilon_{ijk}\frac{k_k}{k}\right], \tag{5.25}$$

where $A(k)$ and $\sigma(k)$ are real, and $\sigma(k)$ is known as the *magnetic helicity*. Appropriate models for $A(k)$ and $\sigma(k)$ must be given.

Let us reconsider now the correlation tensor in the presence of magnetic turbulence that is axisymmetric with respect to a preferred direction; typically the z-axis along which the uniform mean magnetic field is assumed to be oriented. In this case, it can be shown (not done here, see Matthaeus and Smith (1981)) that the isotropic form of the correlation tensor also holds for axisymmetric turbulence,

$$P_{ij}(\mathbf{k}) = A(k_\parallel, k_\perp)\left[\delta_{ij} - \frac{k_i k_j}{k^2} + i\sigma(k_\parallel, k_\perp)\sum_k \varepsilon_{ijk}\frac{k_k}{k}\right].$$

In most applications to cosmic ray or energetic particle transport, the magnetic helicity term is neglected, as is the parallel component of the turbulent magnetic field δB_z. In this case, the correlation tensor reduces to

$$P_{ij}(\mathbf{k}) = A(k_\parallel, k_\perp)\left[\delta_{ij} - \frac{k_i k_j}{k^2}\right], \tag{5.26}$$

where $i, j = x, y$ and $P_{iz} = 0 = P_{zj}$.

To complete the correlation tensor for use in a transport equation describing particle scattering in a turbulent magnetic field, we need to specify both the geometry of the magnetic turbulence and the spectrum of the magnetic field fluctuations. This will allow us to model the function $A(k_{\|}, k_{\perp})$. Three possible geometries, besides the isotropic case discussed already, are possible in the interplanetary (and possibly interstellar) environment. The first is the *slab model*, which is a one-dimensional model in that the turbulent magnetic field depends only on the z-coordinate

$$\delta B_i^{slab}(\mathbf{r}) = \delta B_i^{slab}(z),$$

allowing us to express the function

$$A^{slab}(k_{\|}, k_{\perp}) = g^{slab}(k_{\|}) \frac{\delta(k_{\perp})}{k_{\perp}}.$$

For the slab model, the wave vectors are parallel to the mean magnetic field, i.e., $\mathbf{k} \parallel \mathbf{B}_0$.

Alternatively, we can consider a 2D or *perpendicular turbulence model* in which the turbulent field is a function of the perpendicular coordinates (x, y) only, i.e.,

$$\delta B_i^{2D}(\mathbf{r}) = \delta B_i^{2D}(x, y),$$

so that

$$A^{2D}(k_{\|}, k_{\perp}) = g^{2D}(k_{\perp}) \frac{\delta(k_{\|})}{k_{\perp}}.$$

In this case, the wave vectors are orthogonal to the mean magnetic field, $\mathbf{k} \perp \mathbf{B}_0$, and therefore lie in a 2D plane perpendicular to the mean field.

Finally, one can construct a *two-component model* that corresponds to a superposition of the slab and 2D models. This model is quasi-3D and

$$\delta B_i^{comp}(\mathbf{r}) = \delta B_i^{2D}(x, y) + \delta B_i^{slab}(z).$$

Because we have

$$\langle \delta B_i^{slab}(z) \delta B_i^{*,2D}(x, y) \rangle = 0,$$

the correlation tensor has the form

$$P_{ij}^{comp}(\mathbf{k}) = P_{ij}^{slab}(\mathbf{k}) + P_{ij}^{2D}(\mathbf{k}).$$

In addition to the underlying geometry of the assumed interplanetary or interstellar turbulence, we need to specify the wave number dependence of $A(k_{\|}, k_{\perp})$, i.e., the wave number spectrum. For the slab model, this requires that we prescribe $g^{slab}(k_{\|})$ and similarly $g^{2D}(k_{\perp})$ for the 2D model. A typical spectrum observed in the solar wind has three distinct regions: (i) an *energy containing range* at small wave numbers (i.e., large scales), and is typically of the form k^{-1}. The energy

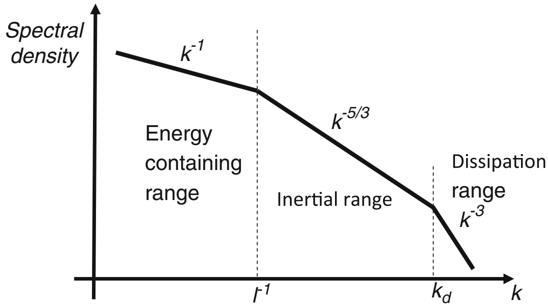


Fig. 5.2 Schematic of the wave number spectrum observed typically in the solar wind, illustrating the energy containing range, the inertial range, and the dissipation range. The bendover scale ℓ^{-1} and the dissipation scale k_d are identified

range is defined by a *bendover* or *turnover* scale such that $g^{slab/2D}(k_{slab/2D} \leq \ell_{slab/2D}^{-1})$ = energy range of the spectrum, depending on whether the turbulence is of the slab or 2D kind. (ii) At larger wave number scales, energy in turbulent fluctuations is transferred locally from larger to smaller scales in a self-similar manner. This part of the spectrum is called the *inertial range* and typically has the form $k^{-5/3}$, which is the Kolmogorov form of the spectrum.[5] For the inertial range, we introduce a dissipation wave number $k_{d,slab/2D}$ and defined the spectrum by $g^{slab/2D}(\ell_{slab/2D}^{-1} \leq k_{slab/2D} \leq k_{d,slab/2D})$ = inertial range of the spectrum. (iii) Finally, for large wave numbers or small scales, the turbulence loses energy through dissipation, and so this part of the spectrum is called the *dissipation range*, and is much steeper than the rest of the spectrum, typically k^{-3}. The dissipation range may be defined as $g^{slab/2D}(k_{d,slab/2D} \leq k_{slab/2D})$ = dissipation range of the spectrum (see Fig. 5.2 for a schematic illustration and Fig. 5.3 for several examples observed in the solar wind).

In most studies of energetic particle transport, the dissipation range plays very little role and is therefore neglected typically. The energy and inertial ranges are however critical in determining particle transport properties and a useful analytic form of the wave number spectrum for magnetic (and other) fluctuations is

$$g^i(k_i) \sim \left(1 + k_i^2 \ell_i^2\right)^{-\nu}, \quad i \equiv \text{slab or 2D}.$$

This form of the spectrum contains both the energy range (modeled as a constant) and an inertial range with slope $k^{-2\nu}$ defined by a bendover scale ℓ_i.

An important quantity used to characterize turbulence and closely related to the bendover scale is the correlation length, defined by the following integral,

$$\ell_{c,ij} \left\langle \delta B_j^2 \right\rangle = \int_0^\infty R_{jj}(\mathbf{r}, 0) dr_i.$$

The correlation length represents the characteristic length scale for the spatial decorrelation of turbulence. Hence, $\ell_{c,ij} \delta B_j^2$ is simply the area under the correlation function R_{ii}. Clearly, the correlation length depends intimately on the nature of the turbulence and wave number spectrum through the correlation function.

[5] Kolmogorov (1941).

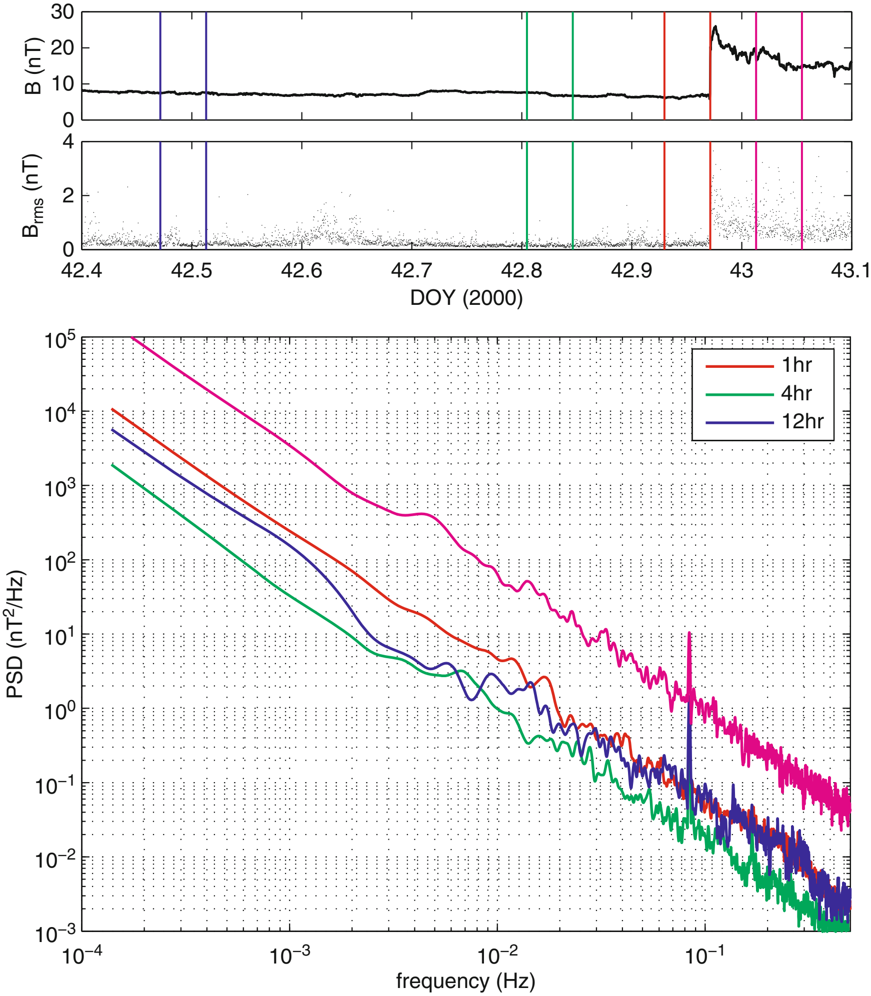


Fig. 5.3 Example of spectra upstream and downstream of a perpendicular interplanetary shock wave (Zank et al. 2006)

Consider now the correlation function for slab turbulence, assuming that $\Gamma(\mathbf{r}, t) = 1$, i.e., magnetostatic turbulence. Turbulent magnetic fluctuations vary only along the direction of the mean magnetic field z, so

$$R_{ij}^{slab} = \langle \delta B_i(z) \delta B_j^*(0) \rangle,$$

assuming $z(0) = 0$ because of homogeneous turbulence. On using the form of the axisymmetric magnetic correlation tensor, and the results from the geometric form of $A(k_\parallel, k_\perp)$, we find

$$P_{ij}(\mathbf{k}) = g^{slab}(k_\parallel)\frac{\delta(k_\perp)}{k_\perp}\left(\delta_{ij} - \frac{k_i k_j}{k^2}\right)$$
$$= g^{slab}(k_\parallel)\frac{\delta(k_\perp)}{k_\perp}\delta_{ij} \quad \text{if } i, j = x, y,$$

and $P_{iz} = 0 = P_{zj}$. If we assume the general form of the turbulence spectrum above, we can express g^{slab} as

$$g^{slab}(k_\parallel) = \frac{C(\nu)}{2\pi}\ell_{slab}\left\langle\delta B_{slab}^2\right\rangle\left(1 + k_\parallel^2\ell_{slab}^2\right)^{-\nu}, \tag{5.27}$$

where the normalization constant has to be determined. Thus, using cylindrical coordinates $k_x = k_\perp\cos\theta$, $k_y = k_\perp\sin\theta$, $k_z = k_\parallel$ to express the wave vector, we find

$$\left\langle\delta B_{slab}^2\right\rangle = \left\langle\delta B_x^2\right\rangle + \left\langle\delta B_y^2\right\rangle = R_{xx}(0) + R_{yy}(0) = \int d^3k\left[P_{xx}(\mathbf{k}) + P_{yy}(\mathbf{k})\right]$$
$$= 2\int_0^{2\pi}\int_0^{\infty}\int_{-\infty}^{\infty} g^{slab}(k_\parallel)\frac{\delta(k_\perp)}{k_\perp}k_\perp d\theta dk_\perp dk_\parallel$$
$$= 8\pi\int_0^{\infty} g^{slab}(k_\parallel)dk_\parallel.$$

On using (5.27), we find

$$C^{-1}(\nu) = 4\ell_{slab}\int_0^{\infty}\left(1 + k_\parallel^2\ell_{slab}^2\right)^{-\nu} dk_\parallel,$$
$$= 2\int_0^{\infty} t^{-1/2}(1-t)^{-\nu}dt,$$

after using the change of variables $t = k_\parallel^2\ell_{slab}^2$. This integral is the beta function (related to the gamma function $\Gamma(x)$) defined by $B(x, y) \equiv \int_0^\infty t^{x-1}/(1+t)^{x+y}dt$, $x > 0$, $y > 0$, and $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$. Thus, setting $x = \frac{1}{2}$, $y = \nu - \frac{1}{2}$ yields

$$C(\nu) = \frac{1}{2\sqrt{\pi}}\frac{\Gamma(\nu)}{\Gamma(\nu - 1/2)},$$

since $\Gamma(1/2) = \sqrt{\pi}$.

The slab correlation function can now be calculated using (5.24)

$$R_{xx}^{slab}(z) = \langle\delta B_x(z)\delta B_x^*(0)\rangle = \int d^3k P_{xx}^{slab}\cos(k_\parallel z)$$
$$= 4\pi\int_0^{\infty} g^{slab}(k_\parallel)\cos(k_\parallel z)dk_\parallel$$
$$= 2C(\nu)\left\langle\delta B_{slab}^2\right\rangle\int_0^{\infty}(1 + x^2)^{-\nu}\cos(ax)dx,$$

where $x \equiv k_\parallel \ell_{slab}$ and $a \equiv z/\ell_{slab}$. The last integral is of a standard tabulated form

$$\int_0^\infty (1+x^2)^{-\nu}\cos(ax)dx = \frac{\sqrt{\pi}}{\Gamma(\nu)}\left(\frac{2}{a}\right)^{1/2-\nu} K_{\nu-1/2}(a),$$

where $K_\beta(z)$ is the modified Bessel function of imaginary argument. The perpendicular correlation function $R_\perp = R_{xx} + R_{yy}$ can therefore be expressed as

$$R_\perp^{slab} = \frac{2\left\langle \delta B_{slab}^2\right\rangle}{\Gamma(\nu-1/2)}\left(\frac{2\ell_{slab}}{z}\right)^{1/2-\nu} K_{\nu-1/2}\left(\frac{z}{\ell_{slab}}\right). \tag{5.28}$$

Shalchi provides two useful asymptotic forms[6] for the slab correlation function in the limits $z \ll \ell_{slab}$ and $z \gg \ell_{slab}$, respectively

$$\begin{aligned}
&K_{\nu-1/2}(z \ll \ell_{slab}) \simeq \frac{1}{2}\Gamma(\nu-1/2)\left(\frac{2\ell_{slab}}{z}\right)^{\nu-1/2},\\
&\Longrightarrow R_\perp^{slab}(z \ll \ell_{slab}) = \left\langle \delta B_{slab}^2\right\rangle \quad \text{if } \nu > 1/2;\\
&K_{\nu-1/2}(z \gg \ell_{slab}) \simeq \sqrt{\frac{\pi \ell_{slab}}{2z}}e^{-z/\ell_{slab}},\\
&\Longrightarrow R_\perp^{slab}(z \ll \ell_{slab}) = \frac{\sqrt{\pi}}{\Gamma(\nu-1/2)}\left\langle \delta B_{slab}^2\right\rangle\left(\frac{2\ell_{slab}}{z}\right)^{1-\nu} e^{-z/\ell_{slab}}.
\end{aligned}$$

The bendover scale ℓ_{slab} is the characteristic length scale for the spatial decorrelation of the turbulence for the exponentially decaying correlation function in the limit $z \gg \ell_{slab}$.

The slab correlation length can also be computed, and this illustrates the relationship between $\ell_{c,slab}$ and the bendover scale length ℓ_{slab}. Recall from the definition of $\ell_{c,slab}$

$$\begin{aligned}
\ell_{c,slab}\left\langle \delta B_{slab}^2\right\rangle &= \int_0^\infty R_\perp^{slab}(z)dz\\
&= 2\pi\int_{-\infty}^\infty dk_\parallel g^{slab}(k_\parallel)\int_{-\infty}^\infty dz e^{ik_\parallel z}\\
&= (2\pi)^2\int_{-\infty}^\infty dk_\parallel g^{slab}(k_\parallel)\delta(k_\parallel)\\
&= (2\pi)^2 g^{slab}(0) = 2\pi C(\nu)\ell_{slab}\left\langle \delta B_{slab}^2\right\rangle,
\end{aligned}$$

[6]Useful limits of these and many other related functions are tabulated in Abramowitz and Stegun (1974). For this case, A. Shalchi used the formulae (9.6.9) and (9.7.2).

since $\int dz e^{ik_\parallel z} = 2\pi\delta(k_\parallel)$. Thus the slab correlation length and the bendover scale are related via

$$\ell_{c,slab} = 2\pi C(\nu)\ell_{slab},$$

which if we assume a Kolmogorov power law for the inertial range, $\nu = 5/6$, we have $C(5/6) = 0.1188$ and hence $\ell_{c,slab} \simeq 0.75\ell_{slab}$.

The 2D magnetostatic correlation function is a little more laborious to compute. Since $\delta B_i(\mathbf{r}) = \delta B_i(x, y)$, the 2D correlation tensor is given by

$$R_{ij}^{2D}(x, y) = \langle \delta B_i(x, y)\delta B_j(0, 0)\rangle,$$

or

$$R_{xx}(x, y) = \int d^3k P_{xx}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}} = \int d^3k P_{xx}(\mathbf{k})e^{ik_x x + ik_y y},$$

and we have

$$P_{ij}^{2D}(\mathbf{k}) = g^{2D}(k_\perp)\frac{\delta(k_\parallel)}{k_\perp}\left(\delta_{ij} - \frac{k_i k_j}{k^2}\right) \quad \text{if } i, j = x, y\ ,$$
$$\text{or} \quad = 0 \quad \text{if } i \text{ or } j = z.$$

For the wave spectrum, we assume the same normalized form as for the slab case except that we introduce the 2D counterparts ℓ_{2D} and $\left\langle \delta B_{2D}^2 \right\rangle$,

$$g^{2D}(k_\perp) = \frac{C(\nu)}{2\pi}\ell_{2D}\left\langle \delta B_{2D}^2 \right\rangle \left(1 + k_\perp^2 \ell_{2D}^2\right)^{-\nu}.$$

On introducing cylindrical coordinates for the wave vector and position

$$k_x = k_\perp \cos\Psi, \quad k_y = k_\perp \sin\Psi;$$
$$x = r\cos\Phi, \quad y = r\sin\Phi,$$

we find

$$\begin{aligned} R_{xx}^{2D}(x, y) &= \int_0^{2\pi}\int_0^\infty\int_{-\infty}^\infty g^{2D}(k_\perp)\frac{\delta(k_\parallel)}{k_\perp}\left(1 - \frac{k_\perp^2}{k^2}\cos^2\Psi\right)e^{i\mathbf{k}\cdot\mathbf{r}}k_\perp d\Psi dk_\perp dk_\parallel \\ &= \int_0^\infty dk_\perp g^{2D}(k_\perp)\int_0^{2\pi} d\Psi \sin^2\Psi \exp\left[ik_\perp r(\cos\Phi\cos\Psi + \sin\Phi\sin\Psi)\right] \\ &= \int_0^\infty dk_\perp g^{2D}(k_\perp)\int_0^{2\pi} d\Psi \sin^2\Psi \exp\left[ik_\perp r\cos(\Phi - \Psi)\right]. \end{aligned}$$

A standard simplification of these integrals makes use of a series expansion in terms of Bessel functions,

$$e^{ix\sin\alpha} = \sum_{n=-\infty}^{\infty} J_n(x)e^{in\alpha}, \quad e^{ix\cos\alpha} = \sum_{n=-\infty}^{\infty} J_n(x)e^{in(\alpha+\pi/2)},$$

which allows us to obtain

$$R_{xx}^{2D}(x, y) = \int_0^\infty dk_\perp g^{2D}(k_\perp) \sum_{n=-\infty}^{\infty} J_n(k_\perp r) \int_0^{2\pi} d\Psi \sin^2\Psi e^{-in\Psi} e^{in(\Phi+\pi/2)}.$$

The corresponding expression for R_{yy} is given by

$$R_{yy}^{2D}(x, y) = \int_0^\infty dk_\perp g^{2D}(k_\perp) \sum_{n=-\infty}^{\infty} J_n(k_\perp r) \int_0^{2\pi} d\Psi \cos^2\Psi e^{-in\Psi} e^{in(\Phi+\pi/2)},$$

meaning that

$$R_\perp^{2D}(x, y) = \int_0^\infty dk_\perp g^{2D}(k_\perp) \sum_{n=-\infty}^{\infty} J_n(k_\perp r) \int_0^{2\pi} d\Psi e^{-in\Psi} e^{in(\Phi+\pi/2)}.$$

Since

$$\int_0^{2\pi} d\Psi e^{\pm in\Psi} = 2\pi\delta_{n0},$$

the 2D perpendicular correlation function reduces to

$$R_\perp^{2D}(x, y) = 2\pi \int_0^\infty dk_\perp g^{2D}(k_\perp) J_0(k_\perp r),$$

which can be further expressed as (using as before $x \equiv k_\perp \ell_{2D}$ and $a \equiv r/\ell_{2D}$)

$$R_\perp^{2D}(r) = 4C(\nu)\left\langle \delta B_{2D}^2 \right\rangle \int_0^\infty (1 + x^2)^{-\nu} J_0(ax)dx. \tag{5.29}$$

As before, it is instructive to consider the limits $a = 0$ and $a \to \infty$. The former limit yields ($J_0(0) = 1$)

$$\int_0^\infty (1 + x^2)^{-\nu} dx = (4C(\nu))^{-1} \Rightarrow R_\perp^{2D}(r = 0) = \left\langle \delta B_{2D}^2 \right\rangle.$$

The latter limit yields ($\int_0^\infty J_0(y)dy = 1$)

$$\begin{aligned}\int_0^\infty (1+x^2)^{-\nu} J_0(ax)dx &= \frac{1}{a}\int_0^\infty \left(1+\frac{y^2}{a^2}\right)^{-\nu} J_0(y)dy \\ &\simeq \frac{1}{a}\int_0^\infty J_0(y)dy = \frac{1}{a}; \\ &\Rightarrow R_\perp^{2D}(r \gg \ell_{2D}) = 4C(\nu)\left\langle \delta B_{2D}^2\right\rangle \frac{\ell_{2D}}{r}.\end{aligned}$$

Note that the spatial decorrelation length for the turbulence is determined by the 2D bendover scale ℓ_{2D}. Notice too that although the same forms of the wave number spectrum were used for both the slab and 2D cases, the correlation functions are nonetheless different, with the 2D correlation function decaying more slowly with increasing distance compared to the slab case (which falls off exponentially).

As before, we can relate the 2D correlation length $\ell_{c,2D}$ to the bendover scale ℓ_{2D}. In this case, we need to introduce a minimum wave number, $x_{min} \equiv \ell_{2D}/L_{2D}$, to avoid a divergent integral,

$$\begin{aligned}\ell_{c,2D} &= \frac{1}{\left\langle \delta B_{2D}^2\right\rangle}\int_0^\infty R_\perp(r)dr \\ &= 4C(\nu)\int_{x_{min}}^\infty dx(1+x^2)^{-\nu}\int_0^\infty dr J_0\left(\frac{xr}{\ell_{2D}}\right) \\ &= 4C(\nu)\ell_{2D}\int_{x_{min}}^\infty \frac{dx}{x}(1+x^2)^{-\nu} \\ &\simeq 4C(\nu)\ell_{2D}\left(\int_{x_{min}}^1 \frac{dx}{x} + \int_1^\infty x^{-2\nu-1}dx\right) \\ &\simeq 4C(\nu)\ell_{2D}\left(\frac{1}{2\nu} + \ln\frac{L_{2D}}{\ell_{2D}}\right).\end{aligned}$$

The wave spectrum used here is normalized correctly only if $L_{2D} \gg \ell_{2D}$, and in the limit of an infinitely large box, $L_{2D} \to \infty$, the correlation length is infinite.

5.4 Quasi-linear Transport Theory of Charged Particle Transport: Derivation of the Scattering Tensor

We have so far prescribed a very simple diffusion in pitch-angle expression to describe the scattering of particles by in situ magnetic fluctuations. In this and the next section, we derive expressions that describe the scattering of energetic particles in low-frequency magnetic turbulence.

Since we consider particles that can have high energies, we begin with the momentum form of the Vlasov or collisionless Boltzmann equation

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla f + q \left(\mathbf{E} + \frac{\mathbf{p} \times \mathbf{B}}{m} \right) \cdot \frac{\partial f}{\partial \mathbf{p}} = 0, \tag{5.30}$$

for particles of mass m and charge q. Following le Roux et al. (2004)[7] we use a quasi-linear approach to derive a Fokker-Planck kinetic transport equation for the diffusion of charged particles experiencing scattering in pitch-angle and momentum space due to the presence of Alfvénic/slab and quasi-2D turbulence in the solar wind. Quasi-linear theory proceeds essentially by assuming that charged particle gyro-orbits are only weakly perturbed by electromagnetic fluctuations. Typically, there are three ways to proceed. One can proceed from the formalism discussed in the derivation of the Fokker-Planck equation from the Chapman-Kolmogorov equation, assuming a Markovian process, and evaluate the diffusion coefficients directly. A second approach, which we follow here, is to directly expand Eq. (5.30) to determine the diffusion coefficients. A third approach is to work directly from the Newton-Lorentz equations for particle motion in a fluctuating electromagnetic field and directly compute momentum and spatial diffusion coefficients from the Taylor-Green-Kubo (TGK) forms,[8]

$$D_{\mu\mu}(\mu) \equiv \int_0^\infty dt \langle \dot{\mu}(t) \dot{\mu}(t) \rangle;$$

$$D_{ij}(\mu) \equiv \int_0^\infty dt \langle v_i(t) v_j(t) \rangle,$$

where μ is the cosine of the particle pitch angle and $\mathbf{v}$ is the particle velocity.

Several assumptions are made explicitly to ensure the validity of the quasi-linear approximation. The first is that the electromagnetic fluctuations are of small amplitude. This ensures that particles follow approximately undisturbed helical orbits on a particle correlation time τ_c^p, which is the characteristic time for a particle to gyrate on an undisturbed trajectory before being disturbed by incoherent or random fluctuations. This obviously means that the particle correlation time is much less than the characteristic time scale for particle pitch-angle scattering τ_μ i.e., $\tau_c^p \ll \tau_\mu$. The time scale over which particle orbits are significantly distorted by pitch-angle scattering is therefore much longer than the particle correlation time scale on which a coherent helical orbit is maintained.

In Eq. (5.30), we may expand the electromagnetic fields, $\mathbf{E}$ and $\mathbf{B}$, the flow velocity $\mathbf{u}$, and the distribution f into mean and fluctuating parts using a *mean field decomposition*, i.e., a field or scalar Q is may be decomposed as $Q = Q_0 + \delta Q$

[7]See also le Roux and Webb (2007).

[8]See Shalchi (2009) for a general discussion of this approach.

such that the ensemble average $\langle Q \rangle = Q_0$ and $\langle \delta Q \rangle = 0$. It does not necessarily follow that $\delta Q \ll Q_0$, although in quasi-linear theory, this assumption is made to eliminate second-order and higher correlations. Hence,

$$\mathbf{E} = \mathbf{E}_0 + \delta\mathbf{E}, \quad \langle \delta\mathbf{E} \rangle = 0;$$
$$\mathbf{B} = \mathbf{B}_0 + \delta\mathbf{B}, \quad \langle \delta\mathbf{B} \rangle = 0;$$
$$\mathbf{u} = \mathbf{u}_0 + \delta\mathbf{u}, \quad \langle \delta\mathbf{u} \rangle = 0;$$
$$f = f_0 + \delta f, \quad \langle \delta f \rangle = 0.$$

The fields are assumed to vary smoothly on the large scale L, and randomly varying fluctuations occur on the smaller correlation length scale $\ell_c \ll L$. The power spectrum of fluctuations ranges from scales on the order of the correlation length to smaller than the particle gyroradius r_g. In the analysis here, we assume an infinitely extended wave number power spectrum for simplicity, rather than include the details of the dissipation range part of the spectrum. The total electric field, in the MHD approximation, is

$$\mathbf{E} = -\mathbf{u} \times \mathbf{B},$$

where $\mathbf{u}$ and $\mathbf{B}$ are measured in the observer's frame. Applying the small amplitude assumption to the mean field decomposition of the electric field $\mathbf{E}$ yields

$$\mathbf{E}_0 = -\mathbf{u}_0 \times \mathbf{B}_0, \quad \text{and} \quad \delta\mathbf{E} = -\mathbf{u}_0 \times \delta\mathbf{B} - \delta\mathbf{u} \times \mathbf{B}_0,$$

after neglecting quadratically small terms ($\delta\mathbf{u} \ll \mathbf{u}_0$, and $\delta\mathbf{B} \ll \mathbf{B}_0$). We will neglect the induced turbulent electric field $\delta\mathbf{E}$ (although see le Roux et al. for the case where this term is retained). We will make the assumption that the particle distribution is co-moving with the background plasma frame, so that the mean motional electric field term is zero, $\mathbf{E}_0 = 0$.

The mean field decomposition above is substituted into the collisionless Boltzmann equation (5.30). The ensemble averaged form of this equation is then subtracted from the full, unaveraged transport equation (5.30), yielding a transport equation for the rapidly fluctuating variable δf. This equation contains the differences of second-order terms and their corresponding ensemble averages. Since we assume from the outset that $\delta f \ll f_0$, $\delta\mathbf{B} \ll \mathbf{B}_0$, the quadratic terms are small and can be neglected (Exercise). The linearized equation for the correction δf is

$$\frac{\partial}{\partial t}\delta f + \frac{\mathbf{p}}{m} \cdot \nabla \delta f + (\mathbf{p} \times \varOmega) \cdot \frac{\partial \delta f}{\partial \mathbf{p}} = -q\frac{\mathbf{p} \times \delta\mathbf{B}}{m} \cdot \frac{\partial f_0}{\partial \mathbf{p}}, \tag{5.31}$$

where $\varOmega = q\mathbf{B}_0/m$ is the particle gyrofrequency. The corresponding mean-field equation for the distribution function f_0 is given by

$$\frac{\partial f_0}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla f_0 + (\mathbf{p} \times \varOmega) \cdot \frac{\partial f_0}{\partial \mathbf{p}} = -q\left\langle \frac{\mathbf{p} \times \delta\mathbf{B}}{m} \cdot \frac{\partial \delta f}{\partial \mathbf{p}} \right\rangle, \tag{5.32}$$

where the right-hand nonlinear term describes the perturbing effect of the fluctuating magnetic field on the scattered particle distribution. As we illustrate below, this term introduces a diffusion coefficient in pitch-angle space. The closure of (5.32) can be affected by solving the quasi-linear equation (5.31) for δf, and then evaluating the ensemble-averaged term in (5.32).

Consider a homogeneous, infinitely extended plasma system with Cartesian coordinates (x, y, z) with the z-axis aligned with the mean magnetic field $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$. Since the turbulence is comprised of slab turbulence with wave vectors along the mean magnetic field and 2D turbulence with fluctuations and wave vectors transverse to $\mathbf{B}_0$, we have

$$\delta\mathbf{B}(x, y, z) = \delta B_x(x, y)\hat{\mathbf{x}} + \delta B_y(x, y)\hat{\mathbf{y}} + \delta B_x(z)\hat{\mathbf{x}} + \delta B_y(z)\hat{\mathbf{y}},$$

where the 2D component $\delta B_{x/y}(x, y)$ describes the magnetic field fluctuations that convect with the background flow. The second set of terms $\delta B_{x/y}(z)$ comprises the slab or Alfvénic component. For notational convenience, we express magnetic field variations as $\delta B_{x/y}$ and this includes both the slab and 2D components.

The Cartesian form of the momentum coordinates $\mathbf{p} = (p_x, p_y, p_z)$ in the mean-field aligned co-moving coordinate system (p_z is along the mean-field direction) can be expressed in terms of spherical coordinates, so that $\mathbf{p} = p(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, where p is the particle momentum magnitude, θ the particle pitch-angle, and ϕ the particle phase angle. Consider the right-hand side of (5.31),

$$\begin{aligned}(\mathbf{p}\times\delta\mathbf{B})\cdot\nabla_p f_0 &= p(\delta B_z\sin\theta\sin\phi - \delta B_y\cos\theta, \delta B_x\cos\theta - \delta B_z\sin\theta\cos\phi,\\ &\qquad \delta B_y\sin\theta\cos\phi - \delta B_x\sin\theta\sin\phi)\cdot\nabla_p f_0\\ &= (\mathbf{p}\times\delta\mathbf{B})_x\frac{\partial f_0}{\partial p_x} + (\mathbf{p}\times\delta\mathbf{B})_y\frac{\partial f_0}{\partial p_y} + (\mathbf{p}\times\delta\mathbf{B})_z\frac{\partial f_0}{\partial p_z}.\end{aligned}$$

On using the results,

$$\begin{aligned}\frac{\partial}{\partial p_x} &= \sin\theta\cos\phi\frac{\partial}{\partial p} + \cos\theta\cos\phi\frac{1}{p}\frac{\partial}{\partial\theta} - \frac{\sin\phi}{p\sin\theta}\frac{\partial}{\partial\phi};\\ \frac{\partial}{\partial p_y} &= \sin\theta\sin\phi\frac{\partial}{\partial p} + \cos\theta\sin\phi\frac{1}{p}\frac{\partial}{\partial\theta} + \frac{\cos\phi}{p\sin\theta}\frac{\partial}{\partial\phi};\\ \frac{\partial}{\partial p_z} &= \cos\theta\frac{\partial}{\partial p} - \sin\theta\frac{1}{p}\frac{\partial}{\partial\theta},\end{aligned} \tag{5.33}$$

we find that

$$-\frac{q}{m}\left(\mathbf{p}\times\delta\mathbf{B}\cdot\nabla_p f_0\right) = -\frac{|\Omega|}{B}\left(\delta B_x\sin\phi - \delta B_y\cos\phi\right)\frac{\partial f_0}{\partial\theta},$$

and the coefficient of $\partial f_0/\partial p$ is identically zero. In deriving this result, we have invoked the further assumption that the ensemble averaged distribution function is gyrotropic i.e., is independent of the particle phase angle ϕ. Thus, effects such as diffusion perpendicular to the mean magnetic field and gradient and curvature drifts are neglected in this description of particle transport. This is equivalent to assuming that the particle gyroradius $r_g \ll \ell_c$, the correlation length of the turbulent fluctuations. Equivalently, this requires that the particle gyroperiod $\tau_g = \Omega^{-1} \ll \tau_c^p$.

The evolution equation for δf is a first-order quasi-linear equation and therefore can be solved using the method of characteristics. Accordingly, we have the following set of seven ordinary differential equations to solve,

$$\frac{d}{dt}\delta f = -\frac{|\Omega|}{B}\left(\delta B_x \sin\phi - \delta B_y \cos\phi\right)\frac{\partial f_0}{\partial \theta}; \tag{5.34}$$

$$\frac{d\mathbf{r}}{dt} = \frac{\mathbf{p}}{m}; \tag{5.35}$$

$$\frac{d\mathbf{p}}{dt} = \mathbf{p} \times \Omega. \tag{5.36}$$

For particles located initially at $\mathbf{r}_0 = \mathbf{r}(t_0) = (x_0, y_0, z_0)$ with momentum p_0 and phase angle ϕ_0, we can solve the above odes to obtain

$$\begin{aligned}
&\phi(t') = \phi_0 - \Omega(t' - t_0); \quad x(t') = x_0 - r_g\left(\sin\phi(t') - \sin\phi_0\right);\\
&y(t') = y_0 + r_g\left(\cos\phi(t') - \cos\phi_0\right); \quad z(t') = z_0 + v\cos\theta(t' - t_0);\\
&\delta f(\mathbf{r}, \mathbf{p}, t) = \int_{t_0}^{t}\left(-\frac{|\Omega|}{B}\left(\delta B_x \sin\phi' - \delta B_y \cos\phi'\right)\frac{\partial f_0'}{\partial\theta}\right)dt' + \delta f(\mathbf{r}_0, \mathbf{p}_0, t_0),
\end{aligned} \tag{5.37}$$

where $r_g = v\sin\theta/\Omega$ is the particle gyroradius, and $\phi' \equiv \phi(t')$, $\delta B_i(\mathbf{r}(t'), t')$, and $f_0 = f_0(\mathbf{r}(t'), \mathbf{p}(t'), t')$. The particles evidently follow undisturbed helical orbits along $\mathbf{B}_0$ since p and θ are unchanged during the interaction period, this being less than the characteristic time scale for particles to interact with small-amplitude turbulence, viz. τ_c^p. Consequently, τ_c^p must be restricted so that $t' - t_0$ remains sufficiently small that $\delta f \ll f_0$.

The above expressions can be rewritten in terms of the time difference $\Delta t \equiv t - t'$, where t denotes the observation time and t' is the time during which the particle executes a helical trajectory. Hence, $\Delta t \in [t - t_0, 0]$ so this substitution implies that we follow the particle trajectory backward in time. Rewriting the solution for δf yields

$$\delta f(\mathbf{r}, \mathbf{p}, t) = \int_0^{t-t_0}\left(-\frac{\Omega}{B}\left(\delta B_x \sin\phi - \delta B_y \cos\phi\right)\frac{\partial f_0}{\partial\theta}\right)d(\Delta t) + \delta f(\mathbf{r}_0, \mathbf{p}_0, t_0),$$

where $\phi = \phi(t-\Delta t)$, $\delta B_i(\mathbf{r}(t-\Delta t), t-\Delta t)$, and $f_0 = f_0(\mathbf{r}(t-\Delta t), \mathbf{p}(t-\Delta t), t-\Delta t)$. The expressions for the undisturbed particle orbits are now independent of the initial values, and are given by

$$\begin{aligned}\phi(t-\Delta t) &= \phi(t)+\Omega(\Delta t); \quad x(t-\Delta t)=x(t)-r_g\left(\sin\phi(t-\Delta t)-\sin\phi(t)\right);\\ y(t-\Delta t) &= y(t)+r_g\left(\cos\phi(t-\Delta t)-\cos\phi(t)\right); \quad z(t-\Delta t)\\ &= z(t)+-v\cos\theta(\Delta t).\end{aligned}$$

Note that $t-t_0 \gg \tau_c^p$ and thus $|\mathbf{r}-\mathbf{r}_0| \gg \ell_c$. If $\lambda_\parallel$ denotes the parallel mean free path for the spatial diffusion of particles, then the assumption of small amplitude turbulence implies that $\ell_c \ll \lambda_\parallel$. The overall ordering of scales is therefore $r_g \ll \ell_c \ll \lambda_\parallel \ll L$.

Having obtained the solution δf, we can evaluate the ensemble-averaged collision term on the right-hand-side of (5.32). Introducing

$$\Psi_1 \equiv -\cos\phi\delta B_y + \sin\phi\delta B_x; \qquad \Psi_2 \equiv \sin\phi\delta B_y + \cos\phi\delta B_x,$$

we have

$$\begin{aligned}\frac{q}{m}\left\langle \mathbf{p}\times\delta\mathbf{B}\cdot\frac{\partial}{\partial\mathbf{p}}\delta f\right\rangle &= \frac{|\Omega|}{B}\left\langle(-\cos\phi\delta B_y+\sin\phi\delta B_x)\frac{\partial}{\partial\theta}\delta f\right\rangle\\ &\quad+\frac{|\Omega|}{B}\left\langle\frac{\cos\theta}{\sin\theta}(\sin\phi\delta B_y+\cos\phi\delta B_x)\frac{\partial}{\partial\phi}\delta f\right\rangle\\ &= \frac{|\Omega|}{B}\frac{\partial}{\partial\theta}\left\langle\delta f\Psi_1\right\rangle+\frac{|\Omega|}{B}\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\phi}\left\langle\delta f\Psi_2\right\rangle\\ &\quad-\frac{|\Omega|}{B}\frac{\cos\theta}{\sin\theta}\left\langle\delta f\frac{\partial\Psi_2}{\partial\phi}\right\rangle\\ &= \frac{|\Omega|}{B}\left(\frac{\partial}{\partial\theta}\left\langle\delta f\Psi_1\right\rangle+\frac{\cos\theta}{\sin\theta}\left\langle\delta f\Psi_1\right\rangle\right)\\ &\quad+\frac{|\Omega|}{B}\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\phi}\left\langle\delta f\Psi_2\right\rangle\\ &= \frac{|\Omega|}{B}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\left\langle\delta f\Psi_1\right\rangle\right)+\frac{|\Omega|}{B}\frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\phi}\left\langle\delta f\Psi_2\right\rangle,\end{aligned}$$

after using $\partial\Psi_2/\partial\phi = -\Psi_1$. Since f_0 is independent of gyrophase, we neglect the last term. Thus, in spherical coordinates, we have the relation

$$\begin{aligned}&-q\left\langle\frac{\mathbf{p}\times\delta\mathbf{B}}{m}\cdot\frac{\partial\delta f}{\partial\mathbf{p}}\right\rangle\\ &= -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\frac{|\Omega|}{B}\sin\theta\left\langle\left(\delta B_x(\mathbf{r},t)\sin\phi(t)-\delta B_y(\mathbf{r},t)\cos\phi(t)\right)\delta f\right\rangle\right).\end{aligned}$$

On substituting for δf and using the relations,

$$\begin{aligned}\cos\phi(t) &= \cos\left(\phi(t-\Delta t) - \Omega(\Delta t)\right)\\ &= \cos\phi(t-\Delta t)\cos(\Omega\Delta t) + \sin\phi(t-\Delta t)\sin(\Omega\Delta t);\\ \sin\phi(t) &= \sin\left(\phi(t-\Delta t) - \Omega(\Delta t)\right)\\ &= \sin\phi(t-\Delta t)\cos(\Omega\Delta t) - \cos\phi(t-\Delta t)\sin(\Omega\Delta t).\end{aligned}$$

we obtain a diffusion equation in particle pitch angle,

$$-q\left\langle \frac{\mathbf{p}\times\delta\mathbf{B}}{m}\cdot\frac{\partial\delta f}{\partial\mathbf{p}}\right\rangle = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(D_{\theta\theta}\frac{\partial f_0}{\partial\theta}\right), \tag{5.38}$$

where the diffusion coefficient $D_{\theta\theta}$ is given by

$$\begin{aligned}D_{\theta\theta}(\mathbf{r},t) = \left(\frac{\Omega}{B}\right)^2 \sin\theta \int_0^\infty &\left(R_{yy}c^2 - (R_{xy} - R_{yx})cs + R_{xx}s^2\right)\\ &\times\cos(\Omega\Delta t)d(\Delta t)\\ &+\left(\frac{\Omega}{B_0}\right)^2\sin\theta\int_0^\infty\left(R_{yy}cs + R_{xy}c^2 - R_{yx}s^2 - R_{xx}cs\right)\sin(\Omega\Delta t)d(\Delta t).\end{aligned}$$

Here, $c \equiv \cos\phi(t-\Delta t) = \cos\left(\phi(t)+\Omega\Delta t\right)$ and $s = \sin\phi(t-\Delta t) = \sin\left(\phi(t)+\Omega\Delta t\right)$, and R_{ij} is the two-point, two-time correlation function for the magnetic fluctuations along the unperturbed particle orbit, i.e.,

$$R_{ij}\left(\Delta\mathbf{r}(\Delta t),\Delta t\right) \equiv \langle\delta B_i(0)\delta B_j(\Delta\mathbf{r}(\Delta t),\Delta t)\rangle.$$

We then have

$$R_{ij}\left(\mathbf{r},\mathbf{r}(t-\Delta t),t,t-\Delta t\right) = \left\langle\delta B_i(\mathbf{r},t),\delta B_j(\mathbf{r}(t-\Delta t),t-\Delta t)\right\rangle,$$

where the components of $\mathbf{r}(t-\Delta t)$ are determined above.

Observe that in deriving the diffusion form of the particle transport equation, we moved the pitch-angle derivative of the distribution function f_0 from under the integral in the expression for δf. There is an important implication embedded in the time scales associated with the ordering of particle scattering and diffusion, $\tau_c^p \ll \tau_\mu$. This ordering implies that $R_{ij} \to 0$ on a much shorter time scale than the time scale on which the particle orbit deviates from an undisturbed trajectory, implying that the integrand contributes only over the time τ_c^p rather than τ_μ to the time integration. Since the gyrotropic-independent distribution function f_0 varies on a time scale comparable to the pitch-angle diffusion time τ_μ, derivatives of f_0 can be taken out from under the integral. The second implication is that we can then extend the integral describing pitch-angle diffusion to ∞ ($t_0 \to -\infty$ in the expression for δf).

The turbulence responsible for scattering the particles has been assumed to be axisymmetric about the mean magnetic field $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$. The axisymmetry condition for the correlation matrix $\mathbf{R}(\delta\mathbf{r})$ under an arbitrary rotation ϕ' about $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$ is expressed by

$$\mathbf{R}(\delta\mathbf{r}) = \mathbf{O}\mathbf{R}(\mathbf{O}^T\delta\mathbf{r})\mathbf{O}^T,$$

where both the left- and right-hand sides are independent of ϕ', and $\mathbf{O}$ is the rotation matrix

$$\mathbf{O} = \begin{pmatrix} \cos\phi' & \sin\phi' & 0 \\ -\sin\phi' & \cos\phi' & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and $\mathbf{O}^T$ is the transpose. Hence, the elements of the left and right matrices

$$\mathbf{R}(\delta\mathbf{r}) = \begin{pmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{pmatrix} = \mathbf{O}\mathbf{R}(\mathbf{O}^T\delta\mathbf{r})\mathbf{O}^T$$

$$= \begin{pmatrix} R_{xx}c^2+R_{xy}sc+R_{yx}sc+R_{yy}s^2 & -R_{xx}sc+R_{xy}c^2-R_{yx}s^2+R_{yy}sc & R_{xz}c+R_{yz}s \\ -R_{xx}sc-R_{xy}s^2+R_{yx}c^2+R_{yy}sc & R_{xx}s^2-R_{xy}sc+R_{yx}sc+R_{yy}c^2 & -R_{xz}s+R_{yz}c \\ R_{zx}c+R_{zy}s & -R_{zx}s+R_{zy}c & R_{zz} \end{pmatrix}, \tag{5.39}$$

are independent of ϕ'. Inspection of the axisymmetric matrix conditions show that the integrands of the diffusion coefficient $D_{\theta\theta}$ are therefore independent of ϕ'. Consequently, using $\phi' = \phi(t-\Delta t)$, we have

$$\begin{aligned} \delta x &\equiv x(t-\Delta t) - x(t) \\ &= r_g\left[\sin\phi(t) - \sin\phi(t-\Delta t)\right] \\ &= r_g\left[\sin\phi(t-\Delta t)(\cos(\Omega\Delta t)-1) - \cos\phi(t-\Delta t)\sin(\Omega\Delta t)\right] \\ &= -r_g\sin(\Omega\Delta t), \end{aligned}$$

etc. if we set $\phi(t-\Delta t) = 0$. This therefore yields

$$\delta\mathbf{r} = \left[-r_g\sin(\Omega\Delta t), r_g(1-\cos(\Omega\Delta t)), -v\cos\theta(\Delta t)\right],$$

from which we find

$$\begin{aligned} \left(\mathbf{O}^T\delta\mathbf{r}\right)_x &= -r_g\left[\cos\phi(t-\Delta t)\sin\Omega\Delta t + \sin\phi(t-\Delta t)(1-\cos\Omega\Delta t)\right]; \\ \left(\mathbf{O}^T\delta\mathbf{r}\right)_y &= r_g\left[-\sin\phi(t-\Delta t)\sin\Omega\Delta t + \cos\phi(t-\Delta t)(1-\cos\Omega\Delta t)\right]; \\ \left(\mathbf{O}^T\delta\mathbf{r}\right)_y &= -v\cos\theta\Delta t, \end{aligned}$$

which corresponds to the unperturbed helical trajectories derived by substituting the trigonometric expansions for $\cos\phi(t)$ etc. as done above. Thus, for axisymmetric turbulence, the R_{ij} terms in the pitch-angle diffusion coefficient are independent of $\phi(t-\Delta t)$, so we may without loss of generality set $\phi(t-\Delta t)=\pi/2$, significantly simplifying the expression for the diffusion coefficient,

$$D_{\theta\theta}=\sin\theta\left(\frac{\Omega}{B_0}\right)^2\int_0^\infty \big(\cos(\Omega\Delta t)R_{xx}-\sin(\Omega\Delta t)R_{yx}\big]\,d(\Delta t). \tag{5.40}$$

The integral (5.40), divided by B_0^2, is essentially the particle decorrelation time. In addition, setting $\phi(t-\Delta t)=\pi/2$ allows the arguments of the two-point, two-time correlation functions to be expressed as

$$x(t-\Delta t)=x(t)+r_g[\cos(\Omega\Delta t)-1];\quad y(t-\Delta t)=y(t)-r_g\sin(\Omega\Delta t);$$
$$z(t-\Delta t)=z(t)-v\cos(\theta\Delta t).$$

By introducing a mean magnetic field $\mathbf{B}_0=B_0\hat{\mathbf{z}}$ into Eq. (5.32), and using $\mu=\cos\theta$, the cosine of the particle pitch-angle, we obtain the simplest 1D form of the collisionless transport equation as

$$\frac{\partial f_0}{\partial t}+\mu v\frac{\partial f_0}{\partial z}=\frac{\partial}{\partial\mu}\left(D_{\mu\mu}\frac{\partial f_0}{\partial\mu}\right), \tag{5.41}$$

where the Fokker-Planck diffusion coefficient in pitch-angle space is given by

$$D_{\mu\mu}=(1-\mu^2)\left(\frac{\Omega}{B_0}\right)^2\int_0^\infty \big(\cos(\Omega\Delta t)R_{xx}-\sin(\Omega\Delta t)R_{yx}\big)\,d(\Delta t). \tag{5.42}$$

For slab turbulence, the pitch-angle scattering diffusion coefficient can be further simplified since $R_{yx}=0$, yielding the standard expression

$$D_{\mu\mu}=(1-\mu^2)\left(\frac{\Omega}{B}\right)^2\int_0^\infty R_{xx}^{slab}\cos(\Omega\Delta t)d(\Delta t). \tag{5.43}$$

Using the results of the previous section, we may evaluate $D_{\mu\mu}$ for slab turbulence. Recall that

$$\begin{aligned} R_{xx}&=\int d^3kP_{xx}^{slab}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}\\ &=\int_0^{2\pi}\int_0^\infty\int_{-\infty}^\infty g^{slab}(k_\parallel)\frac{\delta(k_\perp)}{k_\perp}e^{i\mathbf{k}\cdot\mathbf{r}}k_\perp d\theta dk_\perp dk_\parallel\\ &=4\pi\int_0^\infty g^{slab}(k_\parallel)e^{ik_\parallel z}dk_\parallel\\ &=4\pi\int_0^\infty g^{slab}(k_\parallel)e^{ik_\parallel\mu v}dk_\parallel, \end{aligned}$$

where we used $z = -v\cos\theta t$. On replacing Δt by t in (5.43), we have

$$\begin{aligned} D_{\mu\mu} &= 4\pi(1-\mu^2)\left(\frac{\Omega}{B}\right)^2 \int_0^\infty \int_0^\infty g^{slab}(k_\parallel) e^{-i(k_\parallel \mu v - \Omega)t} dk_\parallel dt \\ &= 4\pi^2(1-\mu^2)\left(\frac{\Omega}{B}\right)^2 \int_0^\infty g^{slab}(k_\parallel)\delta(k_\parallel \mu v - \Omega) dk_\parallel \\ &= 4\pi^2(1-\mu^2)\left(\frac{\Omega}{B}\right)^2 g^{slab}\left(k_\parallel = \frac{\Omega}{\mu v}\right). \end{aligned}$$

Thus, for slab turbulence, energetic charged particles diffuse in pitch angle due to their scattering with waves that satisfy the resonance condition $\mu v k_\parallel = \Omega$.

Exercises

1. Rewrite the Vlasov equation (5.30) using a mean field expansion for the electromagnetic variables, assuming that the particle distribution function is co-moving with the plasma (thus ensuring that $\mathbf{E}_0 = 0$), and neglecting the fluctuating electric field term. Hence derive (5.31) and (5.32).
2. Derive the relations (5.33) and hence show that

$$-\frac{q}{m}\left(\mathbf{p} \times \delta\mathbf{B} \cdot \nabla_p f_0\right) = -\frac{\Omega}{B}\left(\delta B_x \sin\phi - \delta B_y \cos\phi\right)\frac{\partial f_0}{\partial \theta}.$$

5.5 Diffusion Perpendicular to the Mean Magnetic Field: The Nonlinear Guiding Center Theory

To determine the transport of energetic particles perpendicular to a mean magnetic field is not possible within a gyrophase averaged formulation of the Fokker-Planck equation. Instead, we can compute directly the perpendicular spatial diffusion coefficient $\kappa_\perp$ from the Fokker-Planck coefficients. Recall that the mean square displacement is given by

$$\left\langle (\Delta x)^2 \right\rangle = \left\langle (x(t) - x(0))^2 \right\rangle,$$

for an averaging operator $\langle \ldots \rangle$. Several forms of diffusion can be described if we suppose that the following temporal scaling holds for the spatial variance

$$\left\langle (\Delta x)^2 \right\rangle \sim t^\sigma.$$

The following regimes are typically identified:

1. $0 < \sigma < 1$: subdiffusion;
2. $\sigma = 1$: regular Markovian diffusion;

3. $1 < \sigma < 2$: superdiffusion, and
4. $\sigma = 2$: free streaming or ballistic particle motion.

There have been suggestions that energetic particles can be subdiffusive, and at early times in an impulsive solar energetic particle event, particles are typically free streaming. Over long time scales, however, particle motion is more typically diffusive.

The diffusion coefficient is defined as

$$\kappa_{xx} = \lim_{t\to\infty} \frac{\left\langle (\Delta x)^2 \right\rangle}{2t},$$

where we assume that x is normal to the mean magnetic field. To estimate the spatial variance, we appeal to the Taylor-Green-Kubo (TGK) formalism. In general, the variance is given by

$$\begin{aligned}
\left\langle (\Delta x)^2 \right\rangle (t) &= \left\langle \left(\int_0^t v_x(\tau) d\tau \right)^2 \right\rangle \\
&= \int_0^t d\tau \int_0^t d\xi \langle v_x(\tau) v_x(\xi) \rangle \\
&= \int_0^t d\tau \int_0^\tau d\xi \langle v_x(\tau) v_x(\xi) \rangle + \int_0^t d\tau \int_\tau^t d\xi \langle v_x(\tau) v_x(\xi) \rangle .
\end{aligned}$$

On assuming temporal homogeneity, i.e., that the velocity correlation depends only on the time difference, then we choose

$$\langle v_x(\tau) v_x(\xi) \rangle = \langle v_x(\tau - \xi) v_x(0) \rangle$$

for the first integral, and

$$\langle v_x(\tau) v_x(\xi) \rangle = \langle v_x(\xi - \tau) v_x(0) \rangle$$

for the second, to obtain

$$\begin{aligned}
\left\langle (\Delta x)^2 \right\rangle (t) &= \int_0^t d\tau \int_0^\tau d\xi \langle v_x(\tau - \xi) v_x(0) \rangle + \int_0^t d\tau \int_\tau^t d\xi \langle v_x(\xi - \tau) v_x(0) \rangle \\
&= \int_0^t d\tau \int_0^\tau d\xi \langle v_x(\xi) v_x(0) \rangle + \int_0^t d\tau \int_0^{t-\tau} d\xi \langle v_x(\xi) v_x(0) \rangle ,
\end{aligned}$$

after using the transformations $\tau - \xi \to \xi$ and $\xi - \tau \to \xi$ in the respective integrals. These integrals can be simplified using partial integration and applying Leibnitz' rule to obtain

$$\begin{aligned}
\left\langle(\Delta x)^2\right\rangle(t) = {} & \tau\int_0^\tau d\xi\langle v_x(\xi)v_0(0)\rangle\Big|_0^t - \int_0^t d\tau\tau\langle v_x(\tau)v_x(0)\rangle \\
& + \tau\int_0^{t-\tau} d\xi\langle v_x(\xi)v_x(0)\rangle\Big|_0^t + \int_0^t d\tau\tau\langle v_x(t-\tau)v_x(0)\rangle \\
= {} & t\int_0^t d\xi\langle v_x(\xi)v_x(0)\rangle - \int_0^t d\tau\tau\langle v_x(\tau)v_x(0)\rangle \\
& + \int_0^t d\tau\tau\langle v_x(t-\tau)v_x(0)\rangle \\
= {} & \int_0^t d\tau(t-\tau)\langle v_x(\tau)v_x(0)\rangle + \int_0^t d\tau\tau\langle v_x(t-\tau)v_x(0)\rangle \\
= {} & 2\int_0^t d\tau(t-\tau)\langle v_x(\tau)v_x(0)\rangle.
\end{aligned}$$

The running diffusion coefficient $d_{xx}(t)$ is defined as

$$\begin{aligned}
d_{xx}(t) &= \frac{1}{2}\frac{d}{dt}\left\langle(\Delta x)^2\right\rangle(t) \\
&= \frac{1}{2}\frac{d}{dt}2\int_0^t (t-\tau)\langle v_x(\tau)v_x(0)\rangle \\
&= \int_0^t d\tau\langle v_x(\tau)v_x(0)\rangle.
\end{aligned}$$

The limit $d_{xx}(t \to \infty)$ defines diffusive particle transport, therefore

$$\kappa_{xx} = \int_0^\infty d\tau\langle v_x(\tau)v_x(0)\rangle,$$

which is the Kubo formula for the diffusion coefficient.

A detailed discussion of guiding center motion of energetic charged particles can be found in many plasma text books and so is not repeated here. Instead, if we assume that background magnetic field is varying slowly, that for any of the slab, 2D, or composite turbulence models discussed above, the guiding center velocity (assuming $\mathbf{B} = B_0\hat{\mathbf{z}} + \delta\mathbf{B}$) is given by

$$v_x^g(t) \simeq v_z(t)\frac{\delta B_x}{B_0}; \qquad v_y^g \simeq v_z(t)\frac{\delta B_y}{B_0}.$$

Note that the assumption of slab, 2D, or composite turbulence models implies that $\delta B_z = 0$. Particle motion is thus a superposition of the particle gyromotion and the stochastic motion of the particle's guiding center, which follows the random motion of magnetic field lines. The gyromotion can be neglected when computing a diffusion coefficient.

The first systematic derivation of the perpendicular diffusion coefficient was proposed by Matthaeus et al. (2003) and is called the nonlinear guiding center (NLGC) theory. Improvements and extensions to the original model have been made[9] but the original development is very instructive in its simplicity. To ensure agreement with numerical simulations of particles experiencing scattering in low frequency turbulence, we introduce a parameter a (typically taken to be $1/3$) that allows for slight deviations from purely guiding center motion, and take

$$v_x^g = a v_z \frac{\delta B_x}{B_0}.$$

This is reasonable since the magnetic field can occasionally experience variation on scales that are not necessarily slowly varying. The TGK expression for the perpendicular diffusion coefficient is

$$\begin{aligned}\kappa_{xx} &= \int_0^\infty dt\, \langle v_x^g(t) v_x^g(0)\rangle \\ &= \frac{a^2}{B_0^2}\int_0^\infty dt\, \langle v_z(t)\delta B_x(t) v_z(0)\delta B_x^*(0)\rangle .\end{aligned}$$

The fourth-order correlation introduces a closure problem. This is frequently resolved by the assumption that the fourth-order correlation can be replaced by the product of second-order correlations (motivated by the example of Gaussian statistics), which yields

$$\kappa_{xx} = \frac{a^2}{B_0^2}\int_0^\infty dt\, \langle v_z(t) v_z(0)\rangle \langle \delta B_x(t)\delta B_x^*(0)\rangle .$$

Since the particle velocity along the field is mediated by pitch-angle scattering, we may suppose that particle distribution becomes approximately isotropic on diffusion time scales and that there is a decorrelation time scale associated with the parallel velocity. The decorrelation time will be related to the parallel mean free path, so we can use an exponential model to describe the two-point velocity correlation function,

$$\langle v_z(t) v_z(0)\rangle = \frac{v^2}{3} e^{-vt/\lambda_\parallel}.$$

[9]Well summarized by Shalchi (2009)

The magnetic correlation function $R_{xx}(t) = \langle \delta B_x(t)\delta B_x^*(0)\rangle$ can be expressed as a Fourier transform

$$\delta B_x(\mathbf{x},t) = \int d^3k \delta B_x(\mathbf{k},t) e^{i\mathbf{k}\cdot\mathbf{x}} \Rightarrow R_{xx}(t) = \int d^3k \left\langle \delta B_x(t)\delta B_x^*(0) e^{i\mathbf{k}\cdot\Delta\mathbf{x}}\right\rangle,$$

under the assumption of homogeneous turbulence.

At this point, it is still unclear how to further decompose the ensemble averaged integrand in the magnetic correlation function. Corrsin (1959) suggested that at long diffusion times, the probability distribution of particle displacements and the probability distribution of the Eulerian velocity field would become statistically independent of each other – this is Corrsin's independence hypothesis. At large values of the diffusion time, the independence hypothesis asserts that the joint average in R_{xx} can be expressed as the product of two separate averages, i.e.,

$$\left\langle \delta B_x(t)\delta B_x^*(0) e^{i\mathbf{k}\cdot\Delta\mathbf{x}}\right\rangle = \langle \delta B_x(t)\delta B_x^*(0)\rangle \left\langle e^{i\mathbf{k}\cdot\Delta\mathbf{x}}\right\rangle.$$

Applying Corrsin's independence hypothesis then yields

$$R_{xx}(t) = \int d^3k P_{xx}(\mathbf{k},t) \left\langle e^{i\mathbf{k}\cdot\Delta\mathbf{x}}\right\rangle,$$

requiring only that we estimate the characteristic function $\left\langle e^{i\mathbf{k}\cdot\Delta\mathbf{x}}\right\rangle$. The simplest approximation is to assume a Gaussian distribution of the particles, so that

$$\left\langle e^{i\mathbf{k}\cdot\Delta\mathbf{x}}\right\rangle = \exp\left[-\frac{1}{2}\left\langle(\Delta x)^2\right\rangle k_x^2 - \frac{1}{2}\left\langle(\Delta y)^2\right\rangle k_y^2 - \frac{1}{2}\left\langle(\Delta z)^2\right\rangle k_z^2\right].$$

Since we are considering time scales that correspond to large values of the diffusion time, we can approximate the parallel and perpendicular transport as diffusion, so that $\left\langle(\Delta x)^2\right\rangle = 2t\kappa_{xx}$ for example, yielding

$$\left\langle e^{i\mathbf{k}\cdot\Delta\mathbf{x}}\right\rangle = \exp\left[-\kappa_{xx}k_x^2 t - \kappa_{yy}k_y^2 t - \kappa_{zz}k_z^2 t\right].$$

Subject to these six assumptions, we obtain a nonlinear integral equation for the perpendicular diffusion coefficient

$$\kappa_{xx} = \frac{a^2}{B_0^2}\int d^3k \int_0^\infty dt P_{xx}(\mathbf{k},t)\exp\left[-vt/\lambda_\parallel - \kappa_{xx}k_x^2 t - \kappa_{yy}k_y^2 t - \kappa_{zz}k_z^2 t\right].$$

On expressing the correlation tensor $P_{xx}(\mathbf{k},t)$ as the product of a stationary tensor $P_{xx}(\mathbf{k})$ and a dynamical correlation tensor $\Gamma(\mathbf{k},t)$, i.e., $P_{xx}(\mathbf{k},t) = \Gamma(\mathbf{k},t)P_{xx}(\mathbf{k})$, and assuming the exponential form,

$$\Gamma(\mathbf{k},t) = e^{-\gamma(\mathbf{k})t},$$

allows the time integral to be solved

$$\kappa_{xx} = \frac{a^2 v^2}{3B_0^2} \int d^3k \frac{P_{xx}}{v/\lambda_\parallel + \kappa_{xx}k_x^2 + \kappa_{yy}k_y^2 + \kappa_{zz}k_z^2 + \gamma(\mathbf{k})}. \tag{5.44}$$

The nonlinear integral equation (5.44) is the central result of the NLGC theory, describing the diffusion of energetic particles perpendicular to the mean magnetic field where $\delta B_z = 0$. The particle transport results from a combination of pitch-angle scattering along the magnetic field while the underlying magnetic field is experiencing random diffusive motion. The superposition of parallel transport and random magnetic field transport of the particle guiding center leads to a nonlinear diffusion of particle normal to the large-scale magnetic field. As indicated, more sophisticated treatments of the NLGC theory have been developed since. The nonlinear integral equation (5.44) can be solved approximately and analytically for the slab, 2D, and composite turbulence models in the magnetostatic limit.[10]

5.6 Hydrodynamic Description of Energetic Particles

In deriving the cosmic ray transport equation, we have assumed that the underlying energetic particle distribution function is isotropic to zeroth order. We further assumed that the energetic particle number density and momentum is sufficiently small that the background flow in which the "scattering centers" (Alfvén waves or MHD turbulence) are convected is not altered by the energetic particle population, nor is the convection electric field. Energetic particles therefore behave essentially as massless particles that may possess a significant internal energy, which will be expressed through an isotropic or scalar pressure, say P_c, and energy density E_c, and an energy flux $\mathbf{F}_c$.[11] In this case, the general system of MHD equations will be modified by the inclusion of the cosmic rays, through

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{u}) = 0; \tag{5.45}$$

$$\frac{\partial \mathbf{G}}{\partial t} + \nabla \cdot \boldsymbol{\Pi} = 0; \tag{5.46}$$

$$\frac{\partial W}{\partial t} + \nabla \cdot \mathbf{S} = 0, \tag{5.47}$$

[10] Zank et al. (2004) and Shalchi et al. (2004).

[11] Webb (1983) and Zank (1988).

5.7 Application 1: Diffusive Shock Acceleration

It is quite straightforward to see that a particle gains energy by interacting once with a shock, most easily seen for a superluminal shock perpendicular to the magnetic field. In this case, we can suppose that a particle conserves its first adiabatic moment,

$$\frac{p_{\perp,1}^2}{B_1} = \frac{p_{\perp,2}^2}{B_2},$$

where the subscripts 1,2 denote upstream and downstream of the shock. At a perpendicular shock, the jump in magnetic field B_2/B_1 is equal to the shock compression ratio, showing that the perpendicular momentum of an energetic particle can be increased by a factor of 2 or less. This is not a particularly large energy gain, and the effect is of course annulled by the expansion of the downstream medium to the original density. Since the process is purely kinematic and reversible, the energetic particle spectrum is essentially the preacceleration spectrum shifted in energy. The situation is quite different when diffusive effects are included since the number of times that a particle interacts with a shock then becomes a random variable and some particles, by interacting many times with the shock, achieve very high energies. The stochastic character of particles interacting with the shock diffusively corresponds to an increase in entropy for the energetic particle distribution (as it does for the thermal background plasma), with the result that the accelerated particle spectrum is relatively independent of the details of the preacceleration spectrum. We discuss the macroscopic approach to the diffusive acceleration of energetic particles at a shock based on the transport equation that we have derived above. This approach was pioneered by Krymsky (1977), Axford et al. (1977), and Blandford and Ostriker (1978), and is well reviewed by Drury (1983) and Forman and Webb (1985).

The shock is taken to be an infinite plane separating a uniform upstream and downstream state, and we choose a frame in which the shock front is stationary. We shall suppose that all quantities depend only on the x spatial coordinate (a 1D problem) and that the flow velocity is steady, given by

$$u(x) = \begin{cases} u_1 & x < 0 \\ u_2 & x > 0 \end{cases},$$

where u_1 and u_2 are the upstream and downstream constant velocities. To determine the boundary conditions that the energetic particle distribution must satisfy at the shock, we require first that the particle number density must be conserved across the shock i.e., particles are neither created nor lost at the shock, so that

$$[f] = f|_{0-}^{0+} = 0, \tag{5.55}$$

where $x = 0-$ and $x = 0+$ denote locations infinitesimally close to the shock on the upstream and downstream side respectively. The second condition (the transport equation governing particle transport is second-order) that we require is that the normal component of the particle current is continuous if there is no source at the surface, and changes by an amount equal to the particle injection rate at the surface. To determine the current, observe that the transport equation

$$\frac{\partial f}{\partial t} + \mathbf{u} \cdot \nabla f - \frac{p}{3} \nabla \cdot \mathbf{u} \frac{\partial f}{\partial p} = \nabla \cdot (\kappa \cdot \nabla f),$$

can be expressed as

$$\frac{\partial f}{\partial t} + \nabla \cdot \left[-\kappa \cdot \nabla f - \frac{p}{3} \frac{\partial f}{\partial p} \mathbf{u} \right] + \frac{p}{3} \mathbf{u} \cdot \nabla \frac{\partial f}{\partial p} + \mathbf{u} \cdot \nabla f = 0$$

$$\Rightarrow \frac{\partial f}{\partial t} + \nabla \cdot \mathbf{S} + \frac{1}{p^2} \frac{\partial}{\partial p} \left[\frac{p^3}{3} \mathbf{u} \cdot \nabla f \right] = 0, \tag{5.56}$$

where

$$\mathbf{S} = -\kappa \cdot \nabla f - \frac{p}{3} \frac{\partial f}{\partial p} \mathbf{u}$$

is the energetic particle streaming in space and $\mathbf{J}_p = (p/3)\mathbf{u} \cdot \nabla f$ is the streaming in momentum space. Equation (5.56) expresses the transport equation in fully conservative form in phase space, averaged over ϕ and with the distribution function close to isotropy. Because cosmic rays are highly mobile ($v \gg u$), the omnidirectional density f cannot change abruptly, hence the normal component of the net streaming $\mathbf{S}$ must be the same on both sides of any surface of discontinuity. On assuming a steady state and integrating across a sharp discontinuity, we obtain the second boundary condition that energetic particles must satisfy across a shock,

$$[\mathbf{S}] = \mathbf{S} \cdot \mathbf{n}|_{0-}^{0+} = \frac{Q(p)}{4\pi p^2} \Leftrightarrow - \left[\kappa \cdot \nabla f + \frac{p}{3} \frac{\partial f}{\partial p} \mathbf{u} \right] \cdot \mathbf{n} \Bigg|_{0-}^{0+} = \frac{Q(p)}{4\pi p^2}. \tag{5.57}$$

Here, $\mathbf{n}$ is the shock normal, and $Q(p)$ is the particle injection rate at the shock. This form of the boundary conditions includes the effects of shock drift acceleration. Note that the transport equation and the derived boundary conditions are appropriate to relativistic particles i.e., only in the limit that the velocity W (where W is the speed of the scattering frame or the observer's frame relative to the frame in which the electric field vanishes) is much less than the particle velocity v ($W \ll v$), as well as particle drift (through the antisymmetric part of the spatial diffusion tensor κ). That the boundary conditions apply in the limit that $W/v \ll 1$ implies that the boundary conditions (5.55) and (5.57) are valid only for particles of speed $v \gg u_1 \sec \theta_{Bn}$, where θ_{Bn} is the angle between the upstream magnetic field and the

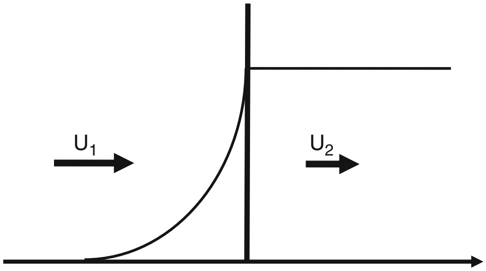


Fig. 5.4 General form of the solution (5.58) illustrating the spatial exponential growth of the distribution function upstream of the shock and the constant ambient value far upstream

shock normal. Furthermore, the transport equation was derived in the limit of near isotropy in the scattering frame, meaning that the particle distribution upstream and downstream of the shock must remain close to isotropy. These conclusions can be weakened slightly for the non-relativisitic form of the transport equation derived above, but isotropy remains a critical assumption. This latter condition is not always met at shocks where energetic particle distributions are often observed to be highly anisotropic.

Consider the 1D transport equation with a constant upstream and downstream velocity and solve the transport equation on either side of the shock, imposing continuity of $f(x, p)$ as $x \to \pm\infty$. The transport equation becomes

$$u_i \frac{df_i}{dx} - \frac{d}{dx}\left(\kappa(x, p)\frac{df_i}{dx}\right) = 0,$$

where $i = 1, 2$ (upstream, downstream) and $\kappa(x, p)$ is the diffusion coefficient parallel to the shock normal. The general solution is

$$f_i(x, p) = A_i(p) + B_i(p) \exp \int_0^x \frac{u}{\kappa(s, p)} ds;$$
$$f_i(x, p) = f(\pm\infty, p) + [f(0, p) - f(\pm\infty, p)]\frac{e^{b(x)} - e^{b(\pm\infty)}}{1 - e^{b(\pm\infty)}},$$

where $b(x) \equiv \int_0^x (u/\kappa)\, dx$. If $b(\pm\infty)$ are unbounded, the spatial dependence is then given by

$$f(x, p) = f(-\infty, p) + [f(0, p) - f(-\infty, p)] \exp \int_0^x \frac{u}{\kappa(s, p)} ds \qquad x < 0;$$
$$= f(0, p) \qquad x > 0. \tag{5.58}$$

The general solution (5.58) is illustrated in Fig. 5.4. The general solution $f(x, p)$ has a possible constant background of upstream particles $f(-\infty, p)$ plus an accelerated population that increases toward the shock on a diffusive scale length $\kappa(x, p)/u_1$ but remains constant downstream.

The momentum spectrum of the energetic particle population is determined by the streaming boundary condition (5.57) at the shock,

$$-u_2 \frac{p}{3} \frac{df(0,p)}{dp} + u_1 \frac{p}{3} \frac{df(0,p)}{dp} + u_1 [f(0,p) - f(-\infty,p)] = \frac{Q(p)}{4\pi p^2},$$

where we have used the result $u_1 [f(0,p) - f(-\infty,p)] = \kappa \partial f/\partial x$ and have allowed for the injection of $Q(p)$ particles at the shock per unit momentum per cm^2 s at the shock. This then yields the ordinary differential equation in momentum

$$p\frac{df}{dp}(0,p) + \frac{3u_1}{u_1 - u_2} f(0,p) = \frac{3}{u_1 - u_2}\left[u_1 f(-\infty,p) + \frac{Q(p)}{4\pi p^2}\right],$$

illustrating that the source of the energetic particles is the background particle population $f(-\infty, p)$ convected through the shock and locally injected particles. Which particle population is more important depends on the relative flux and the characteristic energies. On solving the equation for the particle spectrum, we obtain the central result of diffusive shock acceleration theory,

$$f(0,p) = \frac{3}{u_1 - u_2} p^{-q} \int_{p_{inj}}^{p} (p')^q \left[u_1 f(-\infty,p') + \frac{Q(p')}{4\pi p'^2}\right] \frac{dp'}{p'}, \tag{5.59}$$

where $q = 3r/(r-1)$ and $r = u_1/u_2$ is the shock compression ratio. Here, p_{inj} is the injection momentum. The upper limit on particle momentum is particularly important if time-dependent particle acceleration is considered, such as at interplanetary shock waves where the shock propagation time and evolution need to be considered carefully since this places constraints on the time available for a particle to experience acceleration.[13] Time dependent diffusive shock acceleration is discussed below. The spectrum of particles at energies well above the source energy is therefore a power law $\propto p^{-q}$. The characteristic compression ratio for a strong shock is $r = 4$ for a gas with adiabatic index $\gamma_g = 5/3$, implying that $q = 4$, which is very close to the index of 4.3 inferred for the source of galactic cosmic rays. For weak shocks, the power law is steeper, indicating fewer high energy particles.

A very important point to note is that the spectral slope of the accelerated particle spectrum is independent of the details of the scattering process i.e., the diffusion coefficient, depending only the kinematics of the flow. The reason a power law results is because the momentum gained by the particle on each shock interaction is proportional to the momentum it already has and to the probability of its escaping from the acceleration region. This is very nicely discussed by Bell (1978) from a microscopic perspective.

In (5.59), the accelerated particle spectrum p^{-q} is formed from the spectrum of sources at lower momenta $p' < p$. If no source of particles is present for momenta above some p_a, then $f(0,p) \propto p^{-q}$ for all $p > p_a$. If the spectrum of the source

[13] Zank et al. (2000).

is steeper than p^{-q}, then at large p, the accelerated spectrum will still approach the p^{-q} power law, but if the source is flatter (harder) than p^{-q}, the reaccelerated spectrum at high energies will have the same slope as the source i.e., the new spectrum will not reflect the characteristics of the last acceleration. In general then, a shock with $q = 3r/(r-1)$ will produce a power law spectrum with p^{-q} if the source spectrum is mono-energetic or has a spectral slope steeper than q, but if the source is harder such that $q' < q$, the spectrum tends to $p^{-q'}$ at large energies (Exercise).

The basic time scale associated with diffusive shock acceleration is of the order of κ/u^2. The importance of the acceleration time scale has to do with the maximum energy to which a particle can be accelerated by a shock wave. Observationally, galactic cosmic rays possess a source spectrum that is a power law $\sim p^{-4.3}$ over many decades up until about 10^{14} eV/nucleon, at which point the spectrum begins to steepen (the *knee*). The maximum energy to which a galactic cosmic ray can be accelerated is related presumably to either the time available to accelerate the particle (the lifetime of shock wave responsible for particle acceleration) or to the size of the acceleration region (both of which are possibly related). Similarly, energetic particles accelerated in *solar energetic particle (SEP)* events have a maximum energy. To estimate the maximum energy, whether at a supernova drive shock wave or at an interplanetary shock requires that we know the particle acceleration time scale, and that this then be related to, for example, the characteristic time scale associated with the shock wave.[14] To make the estimate for the time scale of diffusive shock acceleration more precise, we consider a steady planar shock at which a steady mono-energetic source of particles at the shock is turned on at $t = 0$.[15] We then seek time dependent solutions of the cosmic ray transport equation across a discontinuous shock with $f(t=0,x,p) = 0$ and source $Q\delta(p-p_0)$ at the shock, located at $x = 0$. On introducing the Laplace transform

$$g(s,x,p) = \int_0^\infty e^{-st} f(t,x,p)dt,$$

the transport equation upstream ($i = 1$) and downstream ($i = 2$) of the shock becomes

$$sg + u_i \frac{dg}{dx} = \kappa_i \frac{d^2 g}{dx^2},$$

assuming for simplicity that κ is independent of x. The solutions that satisfy the boundary condition

$$g \to 0 \quad \text{as} \quad x \to \pm\infty \quad \text{are} \quad g \propto \exp(\beta_i x),$$

[14] Zank et al. (2000).

[15] Axford (1981).

where

$$\beta_i = \frac{u_i}{2\kappa_i}\left[1 - (-1)^i\left(1 + \frac{4\kappa_i s}{u_i^2}\right)^{1/2}\right].$$

The boundary conditions at the shock are given by

$$[f] = 0; \qquad \left[\kappa\frac{\partial f}{\partial x} + \frac{p}{3}u\frac{\partial f}{\partial p}\right] = -n\delta(p - p_0),$$

where the square brackets denote as usual a jump in the enclosed quantity. On writing $g_0(s, p) = g(s, 0, p)$ for the Laplace transform of the spectrum at the shock, we find that

$$\kappa_1\beta_1 g_0 - \kappa_2\beta_2 g_0 + \frac{u_1 - u_2}{3}p\frac{dg_0}{dp} = \frac{1}{s}n\delta(p - p_0).$$

On letting $A_i = \sqrt{1 + 4\kappa_i s/u_i^2} - 1$, we can rewrite this as

$$\frac{1}{2}\left(u_1 A_1 + u_2 A_2\right) g_0 + u_1 g_0 + \frac{u_1 - u_2}{3}p\frac{dg_0}{dp} = \frac{1}{s}n\delta(p - p_0),$$

which has the solution

$$g_0(s, p) = \frac{3n}{s(u_1 - u_2)}\left(\frac{p}{p_0}\right)^{-q}\exp\left[-\int_{p_0}^{p}\frac{3}{2}\frac{u_1 A_1 + u_2 A_2}{u_1 - u_2}\frac{dp'}{p'}\right].$$

By formally inverting the transform, the time-dependent spectrum of accelerated particles at the shock is given by

$$f_0(t, 0, p) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} g_0(s, p)e^{ts}ds.$$

To obtain the asymptotic behavior at large times, we consider the contribution of the simple pole at $s = 0$, which gives the steady spectrum,

$$f_0(\infty, 0, p) = f_0(\infty, p) = \frac{3n}{u_1 - u_2}\left(\frac{p}{p_0}\right)^{-q}, \quad p \geq p_0, \quad q = \frac{3r}{r - 1},$$

in agreement with the steady-state result. Obviously,

$$f_0(t, p_0) = \frac{3n}{u_1 - u_2} = f_0(\infty, p_0).$$

At a general time $t > 0$ and momentum $p > p_0$, we can express the spectrum formally as

$$f_0(t, p) = f_0(\infty, p) \int_0^t \phi(t')dt',$$

where

$$\phi(t) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \exp\left[ts - h(s)\right] ds,$$

and

$$h(s) = \frac{3}{2} \int_{p_0}^{p} \frac{u_1 A_1 + u_2 A_2}{u_1 - u_2} \frac{dp}{p}.$$

The function $\phi(t, p_0, p)$ is the probability distribution function for the time taken to accelerate a particle from momentum p_0 to p. In fact, since

$$\int_0^\infty \phi(t) \exp(-ts)dt = \exp[-h(s)],$$

and $h(0) = 0$, we have that

$$\int_0^\infty \phi(t)dt = 1,$$

indicating that the distribution is properly normalized. Hence, $\exp[-h(s)]$ can be thought of as the moment generating function for $\phi(t)$. Recall that to obtain the mean we can differentiate $h(s)$ with respect to s and then set $s = 0$ to obtain an expression for the mean acceleration time

$$\begin{aligned} \langle t \rangle &= \int_0^\infty t\phi(t)dt = \frac{\partial}{\partial s} h(0) \\ &= \frac{3}{u_1 - u_2} \int_{p_0}^{p_1} \left(\frac{\kappa_1}{u_1} + \frac{\kappa_2}{u_2} \right) \frac{dp}{p}. \end{aligned} \tag{5.60}$$

Thus, the important conclusion is that the time scale for the acceleration of particles of momentum p at a shock not mediated by cosmic rays is simply

$$\tau_{acc}(p) = \frac{3}{u_1 - u_2} \left(\frac{\kappa_1}{u_1} + \frac{\kappa_2}{u_2} \right). \tag{5.61}$$

Exercises

1. Suppose that an upstream energetic particle distribution proportional to p^{-a} is convected into a shock with compression ratio r from upstream. In the absence of particle injection at the shock itself, calculate the reaccelerated downstream energetic particle spectrum, and explain what happens if $a < q = 3r/(r-1)$ or $a > q$.
2. Suppose that a shock of compression ratio r accelerates n cm^{-3} particles injected as a monoenergetic source $\delta(p - p_0)$ at the shock, so producing a downstream energetic particle spectrum $\propto p^{-q}$. Now suppose the shock propagates out of the system and the compressed gas relaxes back to the ambient state. Let another shock propagate into the system and suppose that this shock reaccelerates the decompressed accelerated power law spectrum that was accelerated earlier. Assume no additional injection of particles into the diffusive shock acceleration process. Compute the energetic particle distribution reaccelerated at the second shock. Again, suppose that the second shock disappears out of the system and the energetic particle decompresses again. Derive the energetic particle spectrum if a third shock reaccelerates the previously accelerated spectrum of particles. What can you infer about the effect of multiple accelerations and decompressions of a spectrum of energetic particles by multiple shock waves?

5.8 Application 2: The Modulation of Cosmic Rays by the Solar Wind

The fundamental concepts underlying the modulation of galactic cosmic rays by the solar wind can be developed on the basis of a simplified form of the cosmic ray transport equation. The solar wind flows supersonically and nearly radially outward from the sun and carries the heliospheric magnetic field. The large-scale magnetic field follows the Parker spiral. On smaller scales, as discussed, the solar wind convects magnetic irregularities – magnetic turbulence – that are responsible for scattering galactic cosmic rays. The charged particles gyrate about the mean magnetic field but experience pitch-angle scattering due to the magnetic turbulence, meaning that the cosmic ray transport equation is a suitable description of particle transport for galactic cosmic rays attempting to enter the heliosphere. That cosmic rays experience scattering in the outwardly flowing solar wind means that they experience considerable difficulty in reaching the inner heliosphere. Consequently, the intensity of cosmic rays in the inner heliosphere will be much lower than in the outer heliosphere.

To ensure a tractable description, consider the cosmic ray transport equation in the absence of a large-scale magnetic field and adopt a 1D spherically symmetric geometry. For a constant radial solar wind speed u, the steady-state spherically symmetric 1D cosmic ray transport equation becomes first-order, we find that the scattering of PUIs in a turbulent magnetofluid introduces a term analogous to that of heat conduction. The second-order correct set of equations describing PUIs in a multi-fluid context leads to the introduction of the viscous terms that define the PUI stress tensor. We systematically derive the system of multi-fluid equations that describe a background Maxwellian proton and electron plasma plus a non-Maxwellian PUI population. Because we assume Maxwellian distributions for the background protons and electrons, the background plasma contributes no heat flux or stress tensor terms. For completeness, we derive a "single-fluid" description analogous to the equations of magnetohydrodynamics (MHD) that describes a PUI mediated plasma. The "single-fluid" model possesses collisionless heat flux and viscous stress terms, unlike the MHD equations. In Section 3, we derive the dispersion relation for linear waves in a PUI mediated plasma and discuss the general properties of waves in a multi-fluid PUI mediated plasma. In Section 4, we present numerical solutions to the dispersion relation for the supersonic solar wind, the IHS plasma, and the plasma in the VLISM. The multi-fluid waves are also compared to the more familiar two-fluid plasma modes. In Section 5, we present an analysis and numerical solutions of the linearized single-fluid model, presenting results for the VLISM only. In the final section, we discuss and summarize our results.

2. DERIVATION OF THE MULTI-FLUID MODEL

2.1. *First-order Correct Multi-fluid Model: Heat Conduction*

In deriving a multi-fluid model that includes PUIs self-consistently, we shall assume that the distribution function for the background protons and electrons are each Maxwellian, which ensures the absence of any heat flux or stress tensor terms for the background plasma. The exact form of the continuity, momentum, and energy equations governing the thermal electrons and protons are therefore given by

$$\frac{\partial n_e}{\partial t} + \nabla \cdot (n_e \mathbf{u}_e) = 0; \tag{4}$$

$$m_e n_e \left(\frac{\partial \mathbf{u}_e}{\partial t} + \mathbf{u}_e \cdot \nabla \mathbf{u}_e \right) = -\nabla P_e - e n_e (\mathbf{E} + \mathbf{u}_e \times \mathbf{B}); \tag{5}$$

$$\frac{\partial P_e}{\partial t} + \mathbf{u}_e \cdot \nabla P_e + \gamma_e P_e \nabla \cdot \mathbf{u}_e = 0, \tag{6}$$

for the electrons, and

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{u}_s) = 0; \tag{7}$$

$$m_p n_s \left(\frac{\partial \mathbf{u}_s}{\partial t} + \mathbf{u}_s \cdot \nabla \mathbf{u}_s \right) = -\nabla P_s + e n_s (\mathbf{E} + \mathbf{u}_s \times \mathbf{B}); \tag{8}$$

$$\frac{\partial P_s}{\partial t} + \mathbf{u}_s \cdot \nabla P_s + \gamma_s P_s \nabla \cdot \mathbf{u}_s = 0, \tag{9}$$

for the protons. Here $n_{e/s}$, $\mathbf{u}_{e/s}$, and $P_{e/s}$ are the usual macroscopic fluid variables for the electron/proton number density, velocity, and pressure respectively, $\gamma_{e/s}$ the electron/proton adiabatic index, $\mathbf{E}$ the electric field, $\mathbf{B}$ the magnetic field, and e the charge of an electron.

Pickup ions initially form an unstable distribution that excites Alfvénic fluctuations. The self-generated fluctuations and in situ turbulence serve to scatter PUIs in pitch angle. The Alfvén waves and magnetic field fluctuations both propagate and convect with the bulk velocity of the system $\mathbf{U} = \mathbf{U}(\mathbf{u}_e, \mathbf{u}_s, \mathbf{u}_p, n_e, n_s, n_p, m_e, m_p)$, where n_p and $\mathbf{u}_p$ refer to PUI variables. The PUIs are governed by the Fokker–Planck transport equation with a (for now unspecified) collisional term $\delta f/\delta t|_c$,

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \frac{e}{m_p} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_v f = \left. \frac{\delta f}{\delta t} \right|_c, \tag{10}$$

for average electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$. We assume that the velocity $\mathbf{v}$ of PUIs is always non-relativistic. The transport Equation (10) has to be transformed into a frame that ensures there is no change in PUI momentum and energy due to scattering. For the present, assume that the cross-helicity σ is nonzero and let

$$\mathbf{v} = \mathbf{c} + \mathbf{U} + \sigma \mathbf{V}_A \quad \Longleftrightarrow \quad \mathbf{c} = \mathbf{v} - \mathbf{U} - \sigma \mathbf{V}_A, \tag{11}$$

where $\mathbf{V}_A$ is the Alfvén velocity and $\mathbf{c}$ is the random velocity. The transport equation is therefore

$$\begin{aligned} &\frac{\partial f}{\partial t} + (U_i + \sigma V_{Ai} + c_i) \frac{\partial f}{\partial x_i} + \left[\frac{e}{m_p} (\mathbf{E} + \mathbf{U} \times \mathbf{B})_i \right. \\ &\quad + \frac{e}{m_p} (\mathbf{c} \times \mathbf{B})_i - \frac{\partial U_i}{\partial t} - (U_j + \sigma V_{Aj} + c_j) \frac{\partial U_i}{\partial x_j} \\ &\quad \left. - \sigma \left(\frac{\partial V_{Ai}}{\partial t} - \sigma (U_j + \sigma V_{Aj} + c_j) \frac{\partial V_{Ai}}{\partial x_j} \right) \right] \frac{\partial f}{\partial c_i} = \left. \frac{\delta f}{\delta t} \right|_c . \end{aligned} \tag{12}$$

The velocity $\mathbf{U}$ is still unspecified, so we choose $\mathbf{U}$ such that $\mathbf{E}' \equiv \mathbf{E} + \mathbf{U} \times \mathbf{B} = 0$. This assumption corresponds to choosing

$$\mathbf{U}_\perp = \mathbf{U} - \mathbf{U}_\parallel = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \equiv \mathbf{U}, \tag{13}$$

since we choose $\mathbf{U}_\parallel = 0$ ($\mathbf{U}_\parallel$ is parallel to $\mathbf{B}$ and therefore arbitrary). This corresponds to expressing (10) in the guiding center frame. The transformation to the velocity $\mathbf{U}$ then yields

$$\begin{aligned} &\frac{\partial f}{\partial t} + (U_i + c_i) \frac{\partial f}{\partial x_i} + \left[\frac{e}{m_p} (\mathbf{c} \times \mathbf{B})_i - \frac{\partial U_i}{\partial t} \right. \\ &\quad \left. - (U_j + c_j) \frac{\partial U_i}{\partial x_j} \right] \frac{\partial f}{\partial c_i} = \left. \frac{\delta f}{\delta t} \right|_c , \end{aligned} \tag{14}$$

after setting the cross-helicity $\sigma = 0$. By taking moments of (14), we can derive the evolution equations for the macroscopic PUI variables, such as the number density $n_p = \int f d^3c$, velocity $n_p u_{p_i} = \int c_i f d^3c$, and so on. Although unspecified for now, we shall assume that moments of the collisional term $\delta f/\delta t|_c$ are zero. This can be checked against the particular scattering model that we use below. The zeroth moment of (14) yields the continuity equation for PUIs,

$$\frac{\partial n_p}{\partial t} + \frac{\partial}{\partial x_i} (n_p (U_i + u_{p_i})) = 0, \tag{15}$$

where $\mathbf{u}_p$ is the PUI bulk velocity in the guiding center frame. For the first moment, we multiply (14) by c_j and integrate over velocity space. This yields, after a little algebra,

$$\frac{\partial}{\partial t}(n_p(U_j + u_{pj})) + \nabla \cdot [n_p \mathbf{U}(U_j + u_{pj}) + n_p \mathbf{u}_p U_j] + \frac{\partial}{\partial x_i}\int c_i c_j f d^3c = \frac{e}{m_p} n_p \varepsilon_{jkl} u_{pk} B_l, \quad (16)$$

where ε_{ijk} is the Levi–Civeta tensor.

To close Equation (16), we need to evaluate the PUI distribution function f, which requires that we solve (14). In solving (14), we assume (1) that the PUI distribution is gyrotropic, and (2) that scattering of PUIs is sufficiently rapid to ensure that the PUI distribution is nearly isotropic. We can therefore average (14) over gyrophase, obtaining the so-called focused transport equation for non-relativistic particles (Isenberg 1997). Details of the derivation can be found in Chapter 5 of Zank (2013), and the explicit expression is given in Appendix A. To solve the gyrophase-averaged transport equation requires that we specify the scattering or collisional operator. We make the simplest possible choice, which is the isotropic pitch-angle diffusion operator,

$$\frac{\partial}{\partial \mu}\left(\nu_s(1-\mu^2)\frac{\partial f}{\partial \mu}\right), \quad (17)$$

where $\mu = \cos\theta$ is the cosine of the particle pitch-angle θ and $\nu_s = \tau_s^{-1}$ is the scattering frequency. The form of the scattering operator (17) allows us to solve the focused transport equation (A1) using a Legendre polynomial expansion of the distribution function f. This is summarized in Appendix A and details can be found in Chapter 5 of Zank (2013). The first-order correct solution to the gyrophase-averaged form of Equation (14), i.e., (A1), is

$$f \simeq f_0 + \mu f_1; \quad (18)$$

$$f_0 = f_0(\mathbf{x}, c, t); \quad (19)$$

$$f_1 = -\frac{c\tau_s}{3} b_i \frac{\partial f_0}{\partial x_i} + \frac{DU_i}{Dt}\frac{\tau_s}{3} b_i \frac{\partial f_0}{\partial c}, \quad (20)$$

where $c = |\mathbf{c}|$ is the particle random speed, $\mathbf{b} \equiv \mathbf{B}/B$ is a directional unit vector defined by the magnetic field, and $D/Dt \equiv \partial/\partial t + U_i \partial/\partial x_i$ is the convective derivative. Both f_0 and f_1 are functions of position, time, and particle random speed c, i.e., independent of pitch-angle μ (and of course gyrophase ϕ). Of particular importance is the retention of the large-scale velocity $\mathbf{U}$ acceleration and shear terms. These terms are often neglected in the derivation of the transport equation describing f_0 (for relativistic particles, the transport equation is the familiar cosmic ray transport equation). Thus, the second term in (20) is typically neglected, although it is known as the relativistic heat inertia term in the relativistic transport theory of cosmic rays (Webb 1985, 1987, 1989). As will be seen below, retaining these terms is absolutely essential to derive the correct multi-fluid formulation for PUIs. By introducing

$$\int c_i c_j f d^3c = \int (c_i - u_{pi})(c_j - u_{pj}) f d^3c + n_p u_{pi} u_{pj} \equiv \int c'_i c'_j f d^3c + n_p u_{pi} u_{pj} \simeq \int c'_i c'_j (f_0 + \mu f_1) d^3c + n_p u_{pi} u_{pj},$$

we can show that

$$\frac{\partial}{\partial x_i}\int c'_i c'_j f_0 d^3c = \frac{1}{m_p}\frac{\partial}{\partial x_i}(\delta_{ij} P_p), \quad \text{and} \quad \int c'_i c'_j \mu f_1 d^3c = 0, \quad (21)$$

where

$$P_p \equiv m_p \frac{4\pi}{3}\int c'^2 f_0 c'^2 dc. \quad (22)$$

Consequently, the PUI stress tensor is identically zero at first-order and there exists only an isotropic pressure tensor $\delta_{ij} P_p$. We show in the following section that retaining the second-order terms in the Legendre polynomial expansion of the gyrophase-averaged equation (A2) does in fact yield a non-zero collisionless stress tensor. The PUI momentum equation to first-order can therefore be expressed as

$$\frac{\partial}{\partial t}(n_p(U_j + u_{pj})) + \frac{\partial}{\partial x_i}\left[n_p(\mathbf{U} + \mathbf{u}_p)(U_j + u_{pj}) + \frac{1}{m_p}\delta_{ij} P_p\right] = \frac{e}{m_p} n_p \varepsilon_{jkl} u_{pk} B_l. \quad (23)$$

To derive the transport equation for P_p, we multiply (14) by $(1/2)c^2$ and integrate over d^3c. We then use (18)–(20) to evaluate the various integrals. Introducing $\mathbf{c}' \equiv \mathbf{c} - \mathbf{u}_p$ as before, we find

$$\int \frac{1}{2} c^2 f_0 d^3c = \frac{3}{2}\frac{1}{m_p} P_p + \frac{1}{2} n_p u_p^2,$$

for example. Similarly, we find that the heat flux $\mathbf{q}(\mathbf{x}, t)$ can be expressed as

$$q_i(\mathbf{x}, t) \equiv \int \frac{1}{2} c'^2 c'_i f d^3c' = \frac{1}{2}\int c^2 c_i f d^3c - \frac{5}{2}\frac{1}{m_p} u_{pi} P_p - \frac{1}{2} n_p u_p^2 u_{pi}. \quad (24)$$

It then follows that

$$\int \frac{1}{2} c'^2 c'_i f_0 d^3c = \pi \int c'^3 \mu b_i f_0 c'^2 dc' = 0,$$

and

$$\int \frac{1}{2} c'^2 c'_i \mu f_1 d^3c' = -\frac{2\pi}{3}\int c'^2 \kappa_{ij} \frac{\partial f_0}{\partial x_j} c'^2 dc' = -\frac{1}{2} K_{ij} \frac{\partial P_p}{\partial x_j} = q_i(\mathbf{x}, t). \quad (25)$$

In (25), we introduced the spatial diffusion coefficient

$$\kappa_{ij} \equiv b_i \frac{c^2 \tau_s}{3} b_j, \quad (26)$$

together with PUI speed-averaged form K_{ij}. The collisionless heat flux for PUIs is therefore described in terms of the PUI pressure gradient and consequently the averaged spatial diffusion introduces a PUI diffusion time and length scale into the multi-fluid system. The diffusion coefficient, i.e., the coefficient for the PUI heat flux, is proportional to the particle scattering time τ_s, and therefore a function of the background turbulent intensity. A separate calculation, possibly based on quasi-linear theory for the parallel diffusion coefficient or the nonlinear guiding center theory for the perpendicular diffusion coefficient, is necessary to obtain reasonable estimates of the scattering time (Matthaeus et al. 2003; Zank et al. 2004).

The remaining terms are straightforwardly evaluated. We find

$$\frac{e}{m_p}\varepsilon_{ijk}B_k\int\frac{1}{2}c^2c_j\frac{\partial f}{\partial c_i}d^3c=-\frac{e}{m_p}\varepsilon_{ijk}n_pu_{p_i}B_ku_{p_j};$$
$$-\frac{DU_i}{Dt}\int\frac{1}{2}c^2\frac{\partial f}{\partial c_i}d^3c=n_pu_{p_i}\frac{DU_i}{Dt};$$
$$\frac{\partial U_i}{\partial x_j}\int\frac{1}{2}c^2c_j\frac{\partial f}{\partial c_i}d^3c=\frac{5}{2}P_p\frac{\partial U_i}{\partial x_i}+\frac{1}{2}n_pu_p^2\frac{\partial U_i}{\partial x_i}+n_pu_{p_i}u_{p_j}\frac{\partial U_i}{\partial x_j}.$$

On combining these results, we obtain, after some algebra, the transport equation for the PUI pressure

$$\frac{\partial P_p}{\partial t}+(\mathbf{u}_p+\mathbf{U})\cdot\nabla P_p+\frac{5}{3}P_p\nabla\cdot(\mathbf{u}_p+\mathbf{U})=\frac{1}{3}\nabla\cdot(\mathbf{K}\cdot\nabla P_p),\tag{27}$$

illustrating that the PUI heat flux yields a spatial diffusion term in the PUI equation of state. The PUI system of equations is properly closed and correct to the first-order. The second-order correct PUI equations, which includes the PUI stress tensor, is given in the following subsection. For completeness, the PUI total energy equation has the form

$$\frac{\partial}{\partial t}\left(\frac{3}{2}P_p+\frac{1}{2}n_p(u_p+U)^2\right)+\frac{\partial}{\partial x_i}\left[\frac{1}{2}n_p(u_p+U)^2(u_{p_i}+U_i)+\frac{5}{2}P_p(u_{p_i}+U_i)-\frac{1}{2}K_{ij}\frac{\partial P_p}{\partial x_j}\right]=\frac{e}{m_p}\varepsilon_{ijk}n_pu_{p_j}B_k(u_{p_i}+U_i).\tag{28}$$

The full system of PUI equations is given by (15), (23), and (27) or (28). It is not particularly illuminating to work in the guiding center frame, and we may simplify (15), (23), and (27), (28), by letting

$$\mathbf{U}_p=\mathbf{u}_p+\mathbf{U}.$$

The right-hand side (RHS) of Equations (23) and (28) is proportional to $\mathbf{u}_p\times\mathbf{B}$, which becomes

$$\left(\mathbf{U}_p-\mathbf{U}\right)\times\mathbf{B}=\mathbf{E}+\mathbf{U}_p\times\mathbf{B},$$

since $\mathbf{E}$ was perpendicular to $\mathbf{B}$ by construction initially. Hence the PUI fluid equations can be written in the more familiar form

$$\frac{\partial n_p}{\partial t}+\nabla\cdot(n_p\mathbf{U}_p)=0;\tag{29}$$

$$\frac{\partial}{\partial t}(n_p\mathbf{U}_p)+\nabla\cdot[n_p\mathbf{U}_p\mathbf{U}_p+\mathbf{I}P_p]=\frac{e}{m_p}n_p(\mathbf{E}+\mathbf{U}_p\times\mathbf{B});\tag{30}$$

$$\frac{\partial}{\partial t}\left(\frac{3}{2}P_p+\frac{1}{2}n_pU_p^2\right)+\nabla\cdot\left[\frac{1}{2}n_pU_p^2\mathbf{U}_p+\frac{5}{2}P_p\mathbf{U}_p-\frac{1}{2}\mathbf{K}\cdot\nabla P_p\right]=\frac{e}{m_p}n_p\mathbf{U}_p\cdot(\mathbf{E}+\mathbf{U}_p\times\mathbf{B}),\tag{31}$$

which is the form we use below. Similarly, we have

$$\frac{\partial P_p}{\partial t}+\mathbf{U}_p\cdot\nabla P_p+\frac{5}{3}P_p\nabla\cdot\mathbf{U}_p=\frac{1}{3}(\nabla\cdot\mathbf{K}\cdot\nabla P_p).\tag{32}$$

The full thermal electron-thermal proton–PUI multi-fluid system is therefore given by Equations (4)–(9) and (29)–(31) or (32), together with Maxwell's equations

$$\frac{\partial\mathbf{B}}{\partial t}=-\nabla\times\mathbf{E};\tag{33}$$

$$\nabla\times\mathbf{B}=\mu_0\mathbf{J};\tag{34}$$

$$\nabla\cdot\mathbf{B}=0;\tag{35}$$

$$\mathbf{J}=e(n_s\mathbf{u}_s+n_p\mathbf{U}_p-n_e\mathbf{u}_e),\tag{36}$$

where $\mathbf{J}$ is the current and μ_0 the permeability of free space.

2.2. Second-order Correct Multi-fluid Model: the Stress Tensor

As shown above, the zeroth- and first-order solutions for the pressure tensor yields an isotropic scalar pressure $P_{ij}=P_p\delta_{ij}$ only. Consider now the second-order Legendre polynomial expansion of f,

$$f\simeq f_0+\mu f_1+\frac{1}{2}(3\mu^2-1)f_2.\tag{37}$$

As before, we need to evaluate

$$\int c_ic_jf\,d^3c=\int c_i'c_j'\left(f_0+\mu f_1+\frac{1}{2}(3\mu^2-1)f_2\right)\times d^3c'+n_pu_{p_i}u_{p_j},$$

from which we find

$$\int c_i'c_j'f_0d^3c'=\frac{1}{m_p}P_p\delta_{ij};$$
$$\int c_i'c_j'\mu f_1d^3c'=0.$$

Although not discussed explicitly above, since the PUI pressure is defined in the frame of the bulk PUI velocity $\mathbf{u}_p$, the distribution function over which the integral is taken needs to evaluated in this frame. Since the expression (A7) for f_2 is a function of the guiding center velocity $\mathbf{U}$, we need to transform to the frame $\mathbf{U}'=\mathbf{U}+\mathbf{u}_p$. On using the solution (A7) for f_2, we obtain

$$\int c_x'^2\frac{1}{2}(3\mu^2-1)f_2d^3c'=\int c_y'^2\frac{1}{2}(3\mu^2-1)f_2d^3c'=\frac{\eta}{15}\left(b_ib_j\frac{\partial U_j'}{\partial x_i}-\frac{1}{3}\frac{\partial U_i'}{\partial x_i}\right);\tag{38}$$

$$\int c_z'^2\frac{1}{2}(3\mu^2-1)f_2d^3c'=-\frac{2\eta}{15}\left(b_ib_j\frac{\partial U_j'}{\partial x_i}-\frac{1}{3}\frac{\partial U_i'}{\partial x_i}\right);\tag{39}$$

$$\int c_i' c_j' \frac{1}{2}(3\mu^2 - 1) f_2 d^3c' = 0, \quad (i \neq j), \tag{40}$$

where the coefficient of viscosity η is defined as

$$\eta \equiv \frac{4\pi}{15} \int \frac{\partial}{\partial c'}(c'^4 c \tau_s) f_0 dc' \tag{41}$$

$$\simeq \frac{4\pi}{3} \int c'^2 \tau_s f_0 c'^2 dc' \tag{42}$$

$$\simeq \frac{P_p \tau_s}{m_p}. \tag{43}$$

Equation (41) is the formal definition of the coefficient of viscosity for the PUI gas. If we assume (probably reasonably) that $|\mathbf{c}| \gg |\mathbf{u}_p|$, then we obtain (42), which may be regarded as a PUI pressure moment weighted by the PUI scattering time. Finally, if we assume that τ_s is independent of c, we then obtain the "classical" form (43) of the viscosity coefficient. The pressure tensor may therefore be expressed as

$$(P_{ij}) = P_p(\delta_{ij}) + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \frac{\eta}{15}\left(b_k b_\ell \frac{\partial U_k'}{\partial x_\ell} - \frac{1}{3}\frac{\partial U_m'}{\partial x_m}\right). \tag{44}$$

The pressure tensor may be written in a more revealing form if we introduce a "viscosity matrix,"

$$(M_{k\ell}) \equiv (\eta_{k\ell}) = \left(\frac{\eta}{15} b_k b_\ell\right) \simeq \left(\frac{1}{15}\frac{P_p \tau_s b_k b_\ell}{m_p}\right), \tag{45}$$

and note that $\eta_{ij} = \eta_{ji}$ and $\eta/15 = \eta_{11} + \eta_{22} + \eta_{33} = \eta_{ij}\delta_{ij}$ (since $b^2 = 1$). Then

$$\frac{\eta}{15}\left(b_k b_\ell \frac{\partial U_k'}{\partial x_\ell} - \frac{1}{3}\frac{\partial U_m'}{\partial x_m}\right) = \frac{\eta_{k\ell}}{2}\left(\frac{\partial U_k'}{\partial x_\ell} + \frac{\partial U_\ell'}{\partial x_k}\right) - \frac{1}{3}\eta_{k\ell}\delta_{k\ell}\frac{\partial U_m'}{\partial x_m} = \frac{\eta_{k\ell}}{2}\left(\frac{\partial U_k'}{\partial x_\ell} + \frac{\partial U_\ell'}{\partial x_k} - \frac{2}{3}\delta_{k\ell}\frac{\partial U_m'}{\partial x_m}\right), \tag{46}$$

which yields the pressure tensor as the sum of an isotropic scalar pressure P_p and the stress tensor, i.e.,

$$(P_{ij}) = P_p(\delta_{ij}) + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \frac{\eta_{k\ell}}{2} \times \left(\frac{\partial U_k'}{\partial x_\ell} + \frac{\partial U_\ell'}{\partial x_k} - \frac{2}{3}\delta_{k\ell}\frac{\partial U_m'}{\partial x_m}\right) \equiv P_p \mathbf{I} + \Pi_p. \tag{47}$$

The stress tensor is a generalization of the "classical" form in that several coefficients of viscosity are present, and of course the derivation here is for a collisionless charged gas of PUIs experiencing only pitch-angle scattering by turbulent magnetic fluctuations. Use of the pressure tensor (47) yields a "Navier–Stokes"-like modification of the PUI momentum equation,

$$\frac{\partial}{\partial t}(n_p \mathbf{U}_p) + \nabla \cdot \left[n_p \mathbf{U}_p \mathbf{U}_p + \frac{1}{m_p}\mathbf{I} P_p\right] = \frac{e}{m_p} n_p(\mathbf{E} + \mathbf{U}_p \times \mathbf{B}) - \frac{1}{m_p}\nabla \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \frac{\eta_{k\ell}}{2}\left(\frac{\partial U_{p k}}{\partial x_\ell} + \frac{\partial U_{p \ell}}{\partial x_k} - \frac{2}{3}\delta_{k\ell}\frac{\partial U_{p m}}{\partial x_m}\right) = \frac{e}{m_p} n_p(\mathbf{E} + \mathbf{U}_p \times \mathbf{B}) - \frac{1}{m_p}\nabla \cdot \Pi_p, \tag{48}$$

where we used $\mathbf{U}_p = \mathbf{u}_p + \mathbf{U} \equiv \mathbf{U}'$ as before. The full momentum equation with the second-order stress tensor correction is included for completeness but in the linearized wave analysis below, we use only the first-order correct equations, i.e., only the heat conduction term is included.

2.3. Reduced "Single-fluid" Model

For some problems, such as the investigation of turbulence in the outer heliosphere, IHS, or VLISM, the full multi-fluid model is far too complicated to solve. By making the key assumption that $\mathbf{U}_p \simeq \mathbf{u}_s$, we can reduce the multi-fluid system above to an MHD-like set of model equations. The assumption that $\mathbf{U}_p \simeq \mathbf{u}_s$ is quite reasonable since (1) the bulk flow velocity of a plasma is always dominated by the protons, and (2) the pick-up process itself forces newly created PUIs to essentially co-move with the background plasma flow. Accordingly, we let $\mathbf{U}_p \simeq \mathbf{u}_s = \mathbf{U}_i$ be the overall proton (i.e., thermal background protons and PUIs) velocity. The thermal proton and PUI continuity and momentum equations are therefore trivially combined as

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{U}_i) = 0; \tag{49}$$

$$m_p n_i\left(\frac{\partial \mathbf{U}_i}{\partial t} + \mathbf{U}_i \cdot \nabla \mathbf{U}_i\right) = -\nabla(P_s + P_p) + e n_i(\mathbf{E} + \mathbf{U}_i \times \mathbf{B}) - \nabla \cdot \Pi_p, \tag{50}$$

where $n_i = n_s + n_p$. Since the PUIs are not thermally equilibrated with the background plasma ($T_s \neq T_p$), we need to deal separately with the P_s and P_p equations. These become

$$\frac{\partial P_s}{\partial t} + \mathbf{U}_i \cdot \nabla P_s + \gamma_s P_s \nabla \cdot \mathbf{U}_i = 0; \tag{51}$$

$$\frac{\partial P_p}{\partial t} + \mathbf{U}_i \cdot \nabla P_p + \gamma_p P_p \nabla \cdot \mathbf{U}_i = \frac{1}{3}\nabla \cdot (\mathbf{K} \cdot \nabla P_p). \tag{52}$$

By combining the proton Equations (49)–(52) with the electron Equations (4)–(6), we can obtain an MHD-like system of equations. On defining new macroscopic variables,

$$\begin{aligned} \rho &\equiv m_e n_e + m_p n_i; \\ q &\equiv -e(n_e - n_i); \\ \rho \mathbf{U} &\equiv m_e n_e \mathbf{u}_e + m_p n_i \mathbf{U}_i; \\ \mathbf{J} &\equiv -e(n_e \mathbf{u}_e - n_i \mathbf{U}_i), \end{aligned} \tag{53}$$

we can express

$$\begin{aligned} n_e &= \frac{\rho - (m_p/e)q}{m_p(1-\xi)} \simeq \rho/m_p; \\ n_i &= \frac{\rho + \xi(m_p/e)q}{m_p(1+\xi)} \simeq \rho/m_p; \\ \mathbf{u}_e &= \frac{\rho\mathbf{U} - (m_p/e)\mathbf{J}}{\rho - (m_p/e)q} \simeq \mathbf{U} - \frac{m_p}{e}\frac{\mathbf{J}}{\rho}; \\ \mathbf{u}_i &= \frac{\rho\mathbf{U} + \xi(m_p/e)\mathbf{J}}{\rho + \xi(m_p/e)q} \simeq \mathbf{U}, \end{aligned} \tag{54}$$

where the smallness of the mass ratio $\xi \equiv m_e/m_p \ll 1$ has been exploited. Use of the approximations (54) allows us to combine

The full thermal electron–thermal proton–PUI multi-fluid system is therefore given by Eqs. (2)–(4) and (31)–(33) or (30), together with Maxwell's equations,

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}; \tag{34}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}; \tag{35}$$

$$\nabla \cdot \mathbf{B} = 0; \tag{36}$$

$$\mathbf{J} = e\left(n_s \mathbf{u}_s + n_p \mathbf{U}_p - n_e \mathbf{u}_e\right), \tag{37}$$

where $\mathbf{J}$ is the current and μ_0 the permeability of free space. The diffusion tensor is assumed to be of a simple diagonal form (i.e., we do not include the off-diagonal terms associated with drift and curvature–see the discussion in Zank (2014) and we specify

$$\mathbf{K} = \begin{pmatrix} \kappa_\perp & 0 & 0 \\ 0 & \kappa_\perp & 0 \\ 0 & 0 & \kappa_\parallel \end{pmatrix}; \quad \kappa_\perp = \eta \frac{1}{3\Omega_p} C_0^2, \quad \kappa_\parallel = \frac{1}{3\Omega_p} C_0^2. \tag{38}$$

We parametrize the perpendicular component of the heat conduction tensor by a term $\eta < 1$. In estimating the diffusion coefficients (38) from (29), we choose a characteristic PUI speed for the region of interest and assume that the scattering time can be approximated by a time scale greater than the corresponding gyroperiod.

Single-fluid-like model

For many problems, the complete multi-component model derived above is far too complicated to solve. The multi-fluid system (2)–(4) and (31)–(33) or (30), together with Maxwell's equations can be considerably reduced in complexity by making the key assumption that $\mathbf{U}_p \simeq \mathbf{u}_s$. The assumption that $\mathbf{U}_p \simeq \mathbf{u}_s$ is quite reasonable since (i) the bulk flow velocity of the plasma is dominated by the background protons since the PUI component scatters off fluctuations moving with the background plasma speed and (ii) the large-scale motional electric field forces newly created PUIs to essentially co-move with the background plasma flow perpendicular to the mean magnetic field. Accordingly, we let $\mathbf{U}_p \simeq \mathbf{u}_s = \mathbf{U}_i$ be the bulk proton (i.e., thermal background protons and PUIs) velocity. The thermal proton and PUI continuity and momentum equations are therefore trivially combined as

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{U}_i) = 0; \tag{39}$$

$$m_p n_i \left(\frac{\partial \mathbf{U}_i}{\partial t} + \mathbf{U}_i \cdot \nabla \mathbf{U}_i \right) = -\nabla(P_s + P_p) + e n_i (\mathbf{E} + \mathbf{U}_i \times \mathbf{B}) - \nabla \cdot \Pi_p, \tag{40}$$

where $n_i = n_s + n_p$. Since the PUIs are not thermally equilibrated with the background plasma ($T_s \neq T_p$), we need to deal separately with the P_s and P_p equations. These become

$$\frac{\partial P_s}{\partial t} + \mathbf{U}_i \cdot \nabla P_s + \gamma_s P_s \nabla \cdot \mathbf{U}_i = 0; \tag{41}$$

$$\frac{\partial P_p}{\partial t} + U_i \frac{\partial P_p}{\partial x_i} + \frac{5}{3} P_p \frac{\partial U_i}{\partial x_i} = \frac{1}{3} \frac{\partial}{\partial x_i} \left(K_{ij} \frac{\partial P_p}{\partial x_j} \right) - \frac{2}{3} \Pi_{ij} \frac{\partial U_j}{\partial x_i}. \tag{42}$$

We can combine the proton Eqs. (39)–(42) with the electron Eqs. (2)–(4) to obtain an MHD-like system of equations. On defining the macroscopic variables,

$$\rho \equiv m_e n_e + m_p n_i; \quad q \equiv -e(n_e - n_i); \quad \rho \mathbf{U} \equiv m_e n_e \mathbf{u}_e + m_p n_i \mathbf{U}_i; \; \mathbf{J} \equiv -e(n_e \mathbf{u}_e - n_i \mathbf{U}_i), \tag{43}$$

we can express

$$n_e = \frac{\rho - (m_p/e) q}{m_p (1 - \xi)} \simeq \rho / m_p; \quad n_i = \frac{\rho + \xi (m_p/e) q}{m_p (1 + \xi)} \simeq \rho / m_p;$$
$$\mathbf{u}_e = \frac{\rho \mathbf{U} - (m_p/e) \mathbf{J}}{\rho - (m_p/e) q} \simeq \mathbf{U} - \frac{m_p}{e} \frac{\mathbf{J}}{\rho}; \quad \mathbf{u}_i = \frac{\rho \mathbf{U} + \xi (m_p/e) \mathbf{J}}{\rho + \xi (m_p/e) q} \simeq \mathbf{U}, \tag{44}$$

where the smallness of the mass ratio $\xi \equiv m_e/m_p \ll 1$ has been exploited. Use of the approximations (44) allows us to combine the continuity and momentum equations in the usual way and to rewrite the thermal electron and proton pressure in terms of the single-fluid macroscopic variables. Thus,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0; \tag{45}$$

$$\rho \left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} \right) = -\nabla (P_e + P_s + P_p) + \mathbf{J} \times \mathbf{B} - \nabla \cdot \Pi; \tag{46}$$

$$\frac{\partial P_s}{\partial t} + \mathbf{U} \cdot \nabla P_s + \gamma_s P_s \nabla \cdot \mathbf{U} = 0; \tag{47}$$

$$\frac{\partial P_e}{\partial t} + \mathbf{U} \cdot \nabla P_e + \gamma_e P_e \nabla \cdot \mathbf{U} = \frac{m_p}{e\rho} \mathbf{J} \cdot \nabla P_e + \frac{\gamma_e m_p}{e} P_e \nabla \cdot \left(\frac{\mathbf{J}}{\rho} \right), \tag{48}$$

where

$$\Pi_{k\ell} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \frac{\eta_{k\ell}}{2} \left(\frac{\partial U_k}{\partial x_\ell} + \frac{\partial U_\ell}{\partial x_k} - \frac{2}{3} \delta_{k\ell} \frac{\partial U_m}{\partial x_m} \right).$$

Since we may assume that the current density is much less than the momentum flux, i.e., $|\mathbf{J}| \ll |\rho \mathbf{U}|$, we can simplify (48) further by neglecting the RHS. By assuming that $\gamma_e = \gamma_s = \gamma$, we can combine the thermal proton and electron equations in a single thermal plasma pressure equation with $P \equiv P_e + P_s$,

$$\frac{\partial P}{\partial t} + \mathbf{U} \cdot \nabla P + \gamma P \nabla \cdot \mathbf{U} = 0. \tag{49}$$

Note that at this point, no assumptions about either the thermal electron or proton pressures (or temperatures) have been made.

Finally, we need an equation for the electric field $\mathbf{E}$. To do so, we multiply the respective momentum equations by the electron or proton charge, sum, and use the approximations (44) to obtain

$$\xi \left(\frac{m_p}{e}\right)^2 \frac{1}{\rho} \left[\frac{\partial \mathbf{J}}{\partial t} + \nabla \cdot (\mathbf{JU} + \mathbf{UJ})\right] = \frac{m_p}{e\rho} (\nabla P_e - \mathbf{J} \times \mathbf{B} - \xi \nabla (P_s + P_p) - \xi \nabla \cdot \Pi) + \mathbf{E} + \mathbf{U} \times \mathbf{B}.$$

The generalized Ohm's law is therefore

$$\mathbf{E} = -\mathbf{U} \times \mathbf{B} - \frac{m_p}{e\rho} (\nabla P_e - \mathbf{J} \times \mathbf{B} - \xi \nabla P_p), \tag{50}$$

where we have retained the PUI pressure since in principle it can be a high-temperature component of the plasma system and ξP_p may be comparable to the P_e term. For typical cases of interest, however, the P_p term can be neglected in Ohm's law (50). Neglect of the electron pressure and Hall current term then yields the usual form of Ohm's law.

The reduced single-fluid model equations may therefore be summarized as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0; \tag{51}$$

$$\rho \left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U}\right) = -\nabla (P + P_p) + \mathbf{J} \times \mathbf{B} - \nabla \cdot \Pi; \tag{52}$$

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho U^2 + \frac{3}{2}(P + P_p) + \frac{1}{2\mu_0}B^2\right) + \nabla \cdot \left[\frac{1}{2}\rho U^2 \mathbf{U} + \frac{5}{2}(P + P_p)\mathbf{U} + \frac{1}{\mu_0}B^2\mathbf{U} - \frac{1}{\mu_0}\mathbf{U} \cdot \mathbf{BB} + \Pi \cdot \mathbf{U}_p - \frac{1}{2}\mathbf{K} \cdot \nabla P_p\right] = 0; \tag{53}$$

$$\frac{\partial P}{\partial t} + \mathbf{U} \cdot \nabla P + \gamma P \nabla \cdot \mathbf{U} = 0; \tag{54}$$

$$\mathbf{E} = -\mathbf{U} \times \mathbf{B}; \quad \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}; \quad \mu_0 \mathbf{J} = \nabla \times \mathbf{B}; \quad \nabla \cdot \mathbf{B} = 0. \tag{55}$$

The single-fluid description (51)–(55) differs from the standard MHD model in that a separate description for the PUI pressure is required. Instead of the conservation of energy Eq. (53), one could use the PUI pressure Eq. (42) for continuous flows. PUIs introduce both a collisionless heat conduction and viscosity into the system.

The model Eqs. (51)–(55), despite being appropriate to non-relativistic PUIs, are identical to the so-called two-fluid MHD system of equations used to describe cosmic ray-mediated plasmas (Webb 1983). However, the derivation of the two models is substantially different in that the cosmic ray number density is explicitly neglected in the two-fluid cosmic ray model and a Chapman–Enskog derivation is not used in deriving the cosmic ray hydrodynamic equations. Nonetheless, the sets of equations that emerge are the same indicating that the cosmic ray two-fluid equations do in fact include the cosmic ray number density explicitly.

The single-fluid-like model may be extended to include, e.g., anomalous cosmic rays (ACRs) as well as PUIs. In this case, the ACRs are relativistic particles. The same analysis carries over, and one has an obvious extension of the model Eqs. (51)–(55) with the inclusion of the ACR pressure. Thus, the extension of (51)–(55) is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0; \tag{56}$$

$$\rho \left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U}\right) = -\nabla (P + P_p + P_A) + \mathbf{J} \times \mathbf{B} - \nabla \cdot \Pi_p - \nabla \cdot \Pi_A; \tag{57}$$

$$\frac{\partial P}{\partial t} + \mathbf{U} \cdot \nabla P + \gamma P \nabla \cdot \mathbf{U} = 0; \tag{58}$$

$$\frac{\partial P_p}{\partial t} + \mathbf{U} \cdot \nabla P_p + \gamma_p P_p \nabla \cdot \mathbf{U} = \frac{1}{3} \nabla \cdot (\mathbf{K_p} \cdot \nabla P_p) - (\gamma_p - 1)\Pi_p : (\nabla \mathbf{U}); \tag{59}$$

$$\frac{\partial P_A}{\partial t} + \mathbf{U} \cdot \nabla P_A + \gamma_A P_A \nabla \cdot \mathbf{U} = \frac{1}{3} \nabla \cdot (\mathbf{K_A} \cdot \nabla P_A) - (\gamma_A - 1)\Pi_A : (\nabla \mathbf{U}); \tag{60}$$

$$\mathbf{E} = -\mathbf{U} \times \mathbf{B}; \quad \frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}; \quad \mu_0 \mathbf{J} = \nabla \times \mathbf{B}; \quad \nabla \cdot \mathbf{B} = 0, \tag{61}$$

where we have introduced the ACR pressure P_A, the corresponding stress tensor Π_A, the ACR diffusion tensor $\mathbf{K}_A$ and adiabatic index γ_A ($4/3 \le \gamma_A \le 5/3$). The coupled system (56)–(61) is the simplest continuum model to describe a non-equilibrated plasma comprising a thermal proton–electron plasma with suprathermal particles (e.g., PUIs or even solar energetic particles) and relativistic energy (anomalous) cosmic rays. The system includes both the collisionless heat flux and viscosity associated with the suprathermal and relativistic particle distributions.

On reverting to Eqs. (51)–(55), we can recover the standard form of the MHD equations if we set the heat conduction spatial diffusion tensor $\mathbf{K} = 0$ and the coefficient of viscosity $(\eta_{kl}) = 0$, which corresponds to assuming $\tau_s \to 0$. If the total thermodynamic pressure $P_{\text{total}} = P + P_p$ is introduced, then we recover the standard MHD equations (dropping the subscript "total"), i.e.,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0; \qquad (62)$$

$$\rho \frac{\partial \mathbf{U}}{\partial t} + \rho \mathbf{U} \cdot \nabla \mathbf{U} + (\gamma - 1)\nabla e + (\nabla \times \mathbf{B}) \times \mathbf{B} = 0; \qquad (63)$$

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho U^2 + e + \frac{B^2}{2\mu_0}\right) + \nabla \cdot \left[\left(\frac{1}{2}\rho U^2 + \gamma e\right)\mathbf{U} + \frac{1}{\mu_0}\mathbf{B} \times (\mathbf{U} \times \mathbf{B})\right] = 0; \qquad (64)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B}); \quad \nabla \cdot \mathbf{B} = 0, \qquad (65)$$

with an equation of state $e = \alpha n k_B T/(\gamma - 1)$. The choice of $\alpha = 2$ (or greater if incorporating the contribution of cosmic rays, etc.) corresponds to a plasma population comprising protons and electrons.

In setting $\mathbf{K} = 0$ and $(\eta_{kl}) = 0$, we have implicitly assumed that PUIs are completely coupled to the thermal plasma. With $\mathbf{K} \neq 0$, heat conduction reduces the effective coupling of energetic particles to the thermal plasma, and their contribution to the total pressure is not as large. This will have important consequences for numerical models of, e.g., the large-scale heliosphere since they incorporate PUIs into the MHD equations, without distinguishing PUIs from thermal plasma and therefore neglect heat conduction. Consequently, the total pressure is over-estimated.

Conclusions

Observations by Voyager 1 and 2 and the IBEX spacecraft indicate that plasma in the outer heliosphere (the super- and subsonic solar wind) and the VLISM possesses characteristics of a multi-component plasma, being essentially a non-equilibrated distribution of background thermal protons and electrons and PUIs of various origins. Limitations of space prevent discussion of all the observational threads that lead to this conclusion, and we list and discuss above only a few. In the supersonic solar wind region of the outer heliosphere, the anomalous heating of the solar wind (Williams et al. 1995) has been interpreted in terms of the dissipation of PUI-driven turbulence that leads to the heating of the solar wind plasma (Zank et al. 1996; 2012; Matthaeus et al. 1996, 1999; Smith et al. 2001; Adhikari et al. 2015a). In the inner heliosheath and the VLISM, the observed plasma characteristics of the HTS (Zank et al. 1996; Richardson 2008; Richardson et al. 2008) and the ENA observations made by IBEX (Zank et al. 2010; Desai et al. 2012, 2014; Zirnstein et al. 2014) have been similarly interpreted in terms of a multi-component plasma distribution comprising various PUI populations. Estimates of the collisional frequency between thermal plasma components and PUIs in the outer supersonic solar wind (> ~10 AU), IHS, and VLISM show that equilibration cannot be achieved in these regions. Illustrated in Fig. 4 is a schematic of the solar wind–LISM interaction region with colors indicating regions that have to be described in terms of a multi-component plasma. The three colors for the different regions indicate that each region has a distinct multi-component plasma description reflecting the different origins of the PUI population for each. In the

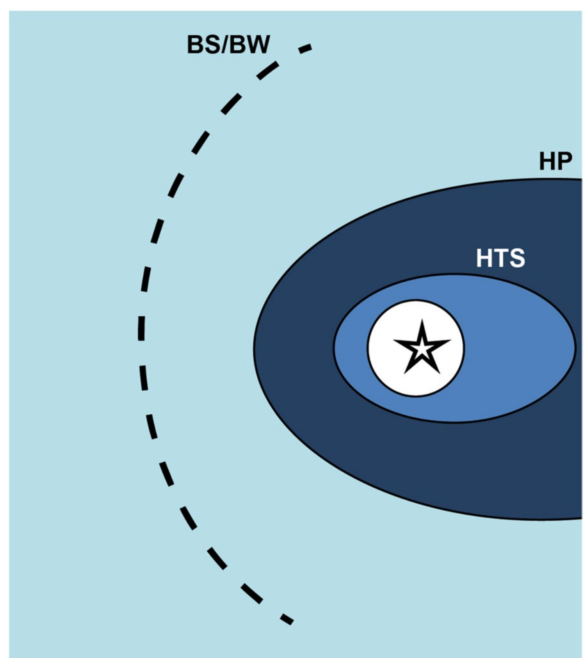


Fig. 4 Schematic of the solar wind–LISM interaction showing the boundaries. The *colored regions* require a non-equilibrated multi-component plasma description. The *different colors* indicate that the non-equilibrated PUI component(s) originates from different physical processes. The region in *white* surrounding the Sun corresponds to the ionization cavity where PUIs are not present in sufficient numbers to effectively mediate the plasma. See text and Table 1 for details