PARTICLE ACCELERATION AT ASTROPHYSICAL SHOCKS: A THEORY OF COSMIC RAY ORIGIN

Roger BLANDFORD

Theoretical Astrophysics, 130-33 Caltech, Pasadena, CA 91125, U.S.A.

and

David EICHLER

Astronomy Program, University of Maryland, College Park, MD 20742, U.S.A.

and

Department of Physics, Ben Gurion University, Be'er Sheva, Israel

NORTH-HOLLAND—AMSTERDAM

PHYSICS REPORTS (Review Section of Physics Letters) 154, No. 1 (1987) 1–75. North-Holland, Amsterdam

PARTICLE ACCELERATION AT ASTROPHYSICAL SHOCKS: A THEORY OF COSMIC RAY ORIGIN

Roger BLANDFORD

Theoretical Astrophysics, 130-33 Caltech, Pasadena, CA 91125, U.S.A.

and

David EICHLER

Astronomy Program, University of Maryland, College Park, MD 20742, U.S.A.

and

Department of Physics, Ben Gurion University, Be'er Sheva, Israel

Received February 1987

Contents:

Single orders for this issue

PHYSICS REPORTS (Review Section of Physics Letters) 154, No. 1 (1987) 1–75.

Copies of this issue may be obtained at the price given below. All orders should be sent directly to the Publisher. Orders must be accompanied by check.

Single issue price Dfl. 54.00, postage included.

Abstract:

The theory of first order Fermi acceleration at collisionless astrophysical shock fronts is reviewed. Observations suggest that shock waves in different astrophysical environments accelerate cosmic rays efficiently. In the first order process, high energy particles diffuse through Alfvén waves that scatter them and couple them to the background plasma. These particles gain energy, on the average, every time they cross the shock front and bounce off approaching scattering centers. Calculations demonstrate that the distribution function transmitted by a plane shock is roughly a power law in momentum with slope similar to that inferred in galactic cosmic ray sources. The generation of the scattering Alfvén waves by the streaming cosmic rays is described and it is argued that the wave amplitude is probably non-linear within sufficiently strong astrophysical shocks. Hydromagnetic scattering can operate on the thermal particles as well, possibly establishing the shock structure. This suggests a model of strong collisionless shocks in which high energy particles are inevitably produced very efficiently. Observable consequences of this model, together with its limitations and some alternatives, are described. Cosmic ray origin and astrophysical shocks can no longer be considered separately.

1. Introduction

Ever since the discovery of cosmic rays by Elster, Geitel, Wilson and Hess, physicists and astronomers have speculated upon their origin [1]. Although the low energy particles ($\lesssim$100 MeV/n) incident on the earth's atmosphere are mostly produced within the solar system (their counterparts in interstellar space being unable to penetrate the solar wind), the higher energy particles must be produced beyond the solar system. For many years it has been widely [2, 3], though not universally [4], believed that their *energy* is derived from supernovae. These are the explosions of dying stars which release about 10^{51} erg of energy into our galaxy's interstellar medium every ~30 years. Supernovae are naturally associated with particle acceleration because they dominate the heat input for the interstellar gas.

However, cosmic ray physicists have found it more difficult to agree upon the *mechanism* whereby a large fraction of the energy of a supernova explosion is channeled into comparatively few high energy particles. In the past decade, one particular scheme – first order Fermi acceleration – has become increasingly attractive as an explanation as a consequence of both observational and theoretical developments. In this view, high energy particles are a natural by-product of the passage of a high Mach number, collisionless shock wave. They are accelerated systematically by scattering off the converging flow on either side of the shock front and, in the case of a supernova remnant, they can absorb the kinetic energy of the surrounding blast wave with quite high efficiency.

There have been parallel developments in space physics. The interplanetary medium is filled with a supersonic wind flowing out from the sun. Strong shocks are formed when this wind runs into the planetary magnetospheres and when fast moving streams collide with slower streams. Energetic particle acceleration is a feature of these shocks. In the last few years spacecraft have amassed a wealth of data on different types of shock structure, and the first order Fermi process has been observed *in situ*.

Going beyond our galaxy, we find that efficient particle acceleration is associated with many of the most active extragalactic sources (quasars and radio galaxies). As these also appear to involve supersonic flows and strong shocks we believe that first order Fermi acceleration is also important there.

In this review we attempt to summarize current research on particle acceleration at shock fronts. We do not discuss competitive acceleration schemes; neither do we discuss acceleration sites such as planetary magnetotails, the aurorae, and radio pulsar magnetospheres where strong shock waves are absent or irrelevant. (For good reviews see, for example, the books by Melrose [5, 6] and the conference proceedings [7, 8, 9].) Our endeavor is to discuss shock wave acceleration as a physical process rather than from an astrophysical or space physical vantage point. Nevertheless, we do provide

an abbreviated summary of relevant astronomical and solar system observations in section 2. In section 3 we introduce the theory of particle acceleration by describing the general Fermi mechanism which is the basis of most astrophysical particle acceleration schemes. In section 4, we provide the relevant background on the fluid dynamical structure of shock waves and demonstrate how the Fermi process can operate on test particles at a discontinuous shock front. We believe that this is an inadequate model of a real shock wave although the test particle theory does predict an asymptotic high energy spectrum in accord with the observations reviewed in section 2. In section 5 we discuss the scatterers responsible for coupling the accelerated particles to the background plasma. These scatterers are generally described as Alfvén waves – transverse hydromagnetic modes allowed by the tension in the background magnetic field. Their growth is probably driven by the anisotropy in the cosmic ray distribution function and their decay by non-linear processes. In section 6, we discuss the non-linear theory of particle acceleration at shock fronts. Here the effects of the cosmic rays and the waves on the fluid are included self-consistently at the hydrodynamic level. We argue that cosmic rays form an essential component of the structure of astrophysical shocks and explain how such a small fraction of the incident particles can absorb such a disproportionate fraction of the energy. In section 7, we summarize this theoretical progress, describe its relatively successful confrontation with the observations and suggest some future directions for research.

In writing this review we have drawn freely upon earlier accounts [10, 11, 12] and in particular the excellent articles by Axford [13, 14, 15], Drury [16], and Kennel, Edmiston and Hada [17]. We will refer to these articles for issues discussed thoroughly in them and emphasize recent developments.

2. Observational background

High energy particles seem to be associated with most astrophysical objects and particle acceleration seems to be most efficient when strong shock waves are also present. These shock waves vary in length scale over some 15 orders of magnitude, which suggests some basic physical process of acceleration that is rather insensitive to the astronomical circumstances under which it operates. Observations of these diverse acceleration sites plus detailed spacecraft measurements of low energy cosmic rays in the solar system, pose a great challenge to a general theory of particle acceleration.

In this section we briefly review the relevant observations of cosmic-ray-producing systems involving shock fronts roughly in the order of their distance from Earth. More extensive discussions are referenced below.

2.1. Earth's bow shock

The solar wind is a somewhat erratic, supersonic radial outflow of ionized, magnetized plasma emanating from the sun's corona. At the radius of the Earth's orbit, typical values for the density ρ, temperature T_e, speed V and field strength B of the wind are $\rho \sim 10^{-23}\,\mathrm{g\,cm^{-3}}$, $T_e \sim 10^5\,\mathrm{K}$, $V \sim 400\,\mathrm{km\,s^{-1}}$, $B \sim 3 \times 10^{-5}\,\mathrm{G}$, respectively. The Earth's magnetic field is approximately dipolar and so the magnetic pressure will decrease with distance r from the Earth $\propto r^{-6}$. The solar wind accelerates through a bow shock, when the magnetic pressure becomes comparable with the momentum flux in the solar wind $\sim \rho V^2$ at $r \sim 10$ Earth radii. This has long been identified as a site of suprathermal particle generation [18, 19, 20]. We are fortunate to have a laboratory at hand where we can study particle

acceleration at a shock front under conditions of density, temperature, velocity and field strength that are very similar to those present in interstellar space. However, there is one important difference. The size of the Earth's bow shock is $\sim 10^5$ km which is only 30 times the Larmor radius of a ~10 keV proton and it is easy for suprathermal particles to escape transversely before they can be accelerated to relativistic energy unlike the case with supernova remnants which are $\sim 10^{10}$ times larger. Bow shocks associated with the magnetospheres of other planets, especially Jupiter and Saturn, have also yielded valuable information.

Our understanding of particle and wave intensities in the foreshock (i.e., ahead of the bow shock) region has improved dramatically in recent years as a consequence of data obtained by the three ISEE spacecraft [21]. A schematic illustration of the typical structure of the bow shock is shown in fig. 1. On the "dusk" side, when the solar wind magnetic field is nearly tangential to the shock front, the shock is termed quasi-perpendicular and is well localized and thin (of order an ion Larmor radius ~10 000 km). Roughly one per cent of the incident solar wind particles are reflected upstream in a beam with a speed of $\sim 1000\,\mathrm{km\,s^{-1}}$, ~2–3 times their incident speed. These "reflected" particles gyrate about the magnetic field lines and are convected by the solar wind towards the "dawn" side, where the angle between the magnetic field and the shock normal henceforth θ_{Bn} is less than ~45°. The shock is then termed "quasi-parallel" and is obscured by an extended region of magnetic turbulence [22]. A fairly isotropic or "diffuse" component of particles with energies ranging up to 100 keV appears on field lines

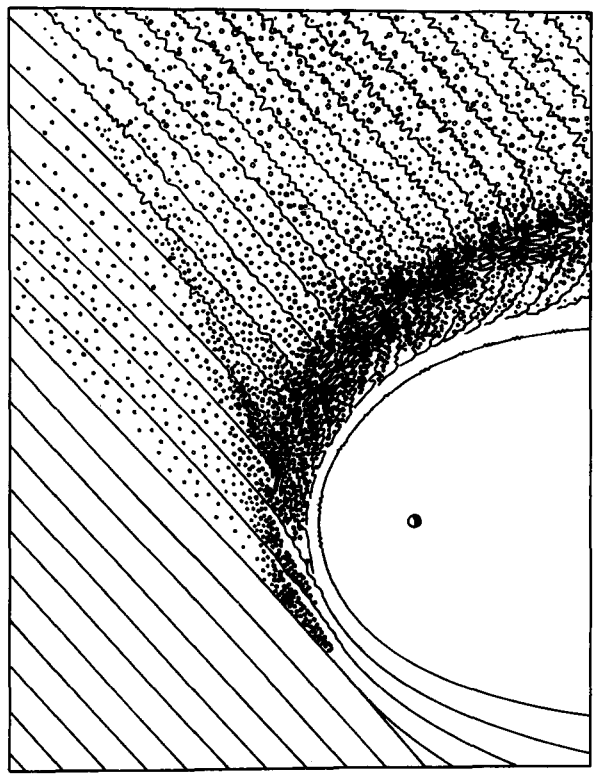

Fig. 1. The earth's bow shock is displayed with a typical magnetic field configuration of the solar wind. The earth rotates counterclockwise. The electrons (hollow circles) are found farther upstream than the ions (filled circles). This is attributed to the fact that the electrons have higher velocities, and that each population is created by the shock and then penetrates upstream (adapted from [261]).

connected to the quasi-parallel region. Their energy spectrum appears above 30 keV to be exponential in energy per charge with a typical e-folding value ~15 keV per unit charge [23, 24]. The variations in the e-folding energy per charge from one observation period to the next are much larger than those between different ion species at any given time, suggesting that energy per charge is the most relevant variable characterizing ion injection (cf. section 4.5 below). The ion density declines away from the shock front but it is still measurable at ~200 Earth radii where the ions are essentially free streaming away from Earth. The charge states of the particles in the diffuse component indicate temperatures typical of the solar corona. This provides further evidence that those energetic particles come from the solar wind.

While the solar wind magnetic field typically points in the direction shown in fig. 1, it may occasionally be parallel to the solar wind velocity, i.e., radially from the sun. A given field line then remains connected to the bow shock for a long time, and the convection of particles away from the shock is inhibited. Also, the entire head of the bow shock has a quasi-parallel geometry. Under such circumstances, particle acceleration by the bow shock is observed to be much more effective, and particle fluxes are observed to rise [25]. The high fluxes measured when the solar wind field is radial are often called "radial" events. The rise time for the particle flux at 30 keV/electron charge is about 15 minutes, and is very roughly proportional to energy per charge [26, 27]. The fact that the measured flux rises first at lower energies suggests that the spacecraft is basically *at* the acceleration site. If the change in magnetic field direction merely connected the spacecraft to a remote source the faster particles should arrive first. The radial events create an opportunity to test time dependent theories quantitatively (cf. section 4.4).

Wave modes are also observed associated with these diffuse ions. Alfvén waves propagate away from the shock roughly parallel to the field at the Alfvén speed ~30 km s^{-1} relative to the solar wind. However, they are convected into the shock at the much faster solar wind velocity ~400 km s^{-1} [27]. Thus, the waves far upstream of the bow shock must have been made there by the fast particles. Their amplitudes can become non-linear, i.e. $\delta B \sim B$, and their wavelengths are in the same range as the Larmor radii of the diffuse particles with which they interact resonantly. Langmuir (i.e., longitudinal electrostatic waves), ion-acoustic (longitudinal sound waves driven by the electron pressure, which are strongly damped unless the ions are much cooler than the electrons) and whistler (electromagnetic waves propagating below the electron plasma and gyrofrequencies) modes are detected in this region. Energetic electrons are also accelerated and observed as beams ahead of the ion foreshock on account of their greater speeds. These electrons are responsible for driving the Langmuir and whistler waves. The ion acoustic waves by contrast appear to be associated with the suprathermal protons.

As mentioned in the Introduction, energetic particles are produced in the earth's magnetotail by reconnection processes that do not involve strong shocks. When the solar wind field connects an upstream spacecraft to the magnetotail, particles produced in the tail can be observed upstream. Anagostopoulos, Sarris and Krimigis [28] have measured proton fluxes in the energy range 50–960 keV using the IMP spacecraft, and they argue that many of the typical upstream events that have been ascribed to Fermi acceleration at the bow shock are in fact particles accelerated in the magnetotail. Magnetotail particles undeniably escape upstream and this source is likely to be important at high energy.

2.2. *Interplanetary shock waves*

Suprathermal particles (primarily p, α, C, N, O, Fe) with energy in the range 0.5–10 MeV/n have been detected in interplanetary space with fluxes many orders of magnitude above the Galactic cosmic

ray background. In many cases, the intensities increase with radial distance from the sun and the "events" are frequently associated with the passage of shock waves. These high energy particles are therefore generally believed to be accelerated within the shock fronts. These are two distinct types of traveling interplanetary shock waves.

Corotating interaction regions are usually characterized by a pair of forward and reverse shock waves. This pattern recurs with the solar rotation period (fig. 2). These regions are formed when a fast solar wind stream plows into slower moving gas. The shocks recorded by Pioneer spacecraft [29, 30] in the inner solar system were fairly weak and quasi-perpendicular. Intense hydromagnetic fluctuations are also found in association with the accelerated particles. The particle spectra are generally exponential in rigidity (momentum per charge) rather than power laws. Shocks further out in the heliosphere have been observed by the Voyager spacecraft during a period of more intense solar activity. Some are strong, quasi-parallel, and appear to be extremely efficient in accelerating particles [31].

Energetic Storm Particle (ESP) shock waves are driven by energetic magnetic storms on the sun. The most powerful of such events can have large Mach numbers and be quasi-parallel [32]. Recent observations are summarized by Lee [33]. A relatively strong shock with a fast mode Mach number $M_f = 2.6$ and $\theta_{Bn} = 40°$ was analyzed by Kennel et al. [34, 35]. The fast quasi-parallel shocks appear to

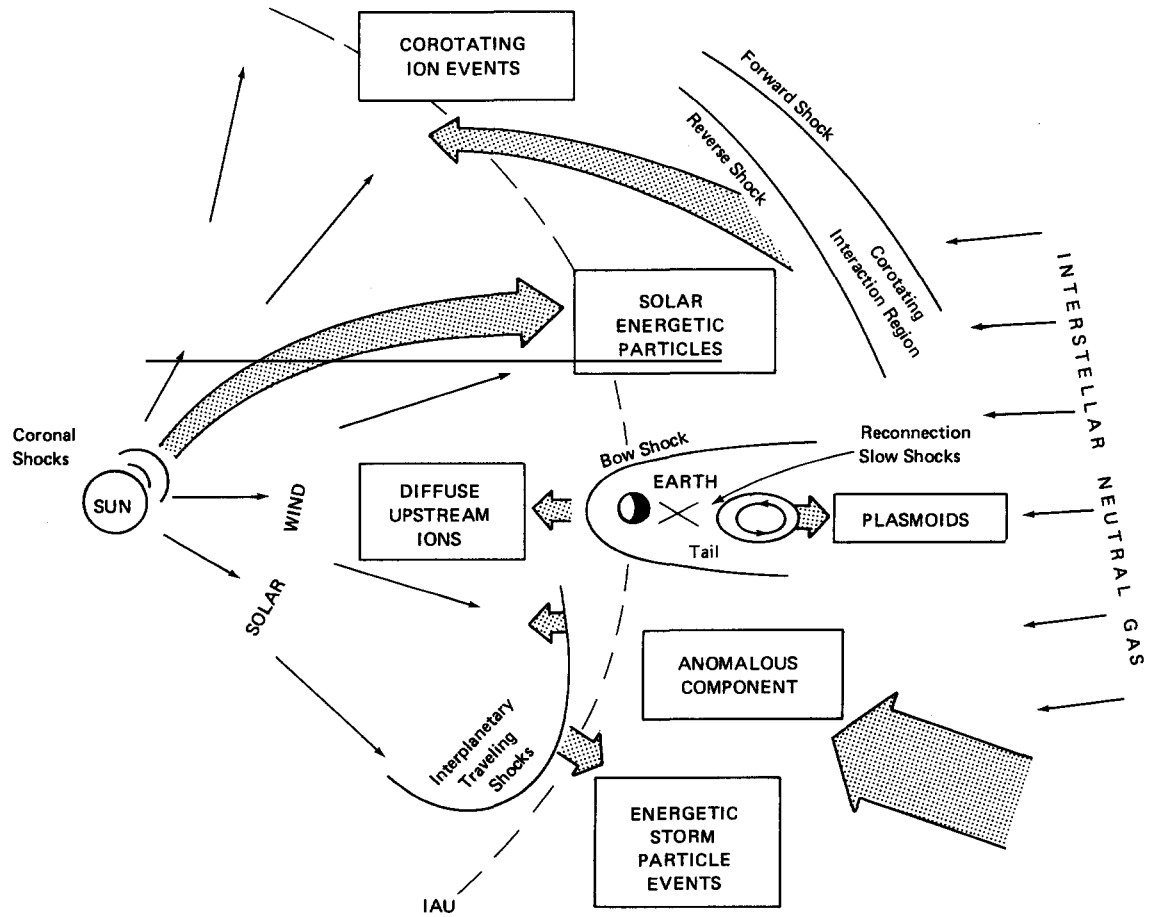

Fig. 2. Various particle acceleration sites within the heliosphere (adapted from [262]).

be highly efficient at accelerating suprathermal ions just as in the bow shock, but without the complication of large shock curvature. Detailed, quantitative measurements of particle and wave spectra from this type of shock exhibit the following features.

(i) There is a definite magnetic jump over a distance ~3 000 km, comparable with an ion Larmor radius. This defines a subshock in the gas.

(ii) There is a density jump in the thermal electrons over a longer scale ~30 000 km.

(iii) Suprathermal ions have scale heights ahead of the shock which increase from $\sim 10^5$ km at 30 keV to $\sim 10^7$ km at 1 MeV. The ion distribution function at the shock is roughly a power law in kinetic energy T with $dN/dT \propto T^{-\gamma}$, $2 \lesssim \gamma \lesssim 2.5$, steepening at the highest energy. Roughly 1% of the incident total momentum flux is converted into energetic ions by the shock.

(iv) Intense Langmuir (~50 kHz) and whistler mode (~20 Hz) turbulence, driven by the upstream electrons is observed.

(v) Intense ion acoustic (~1 kHz) and Alfvén wave turbulence driven by the upstream ions is also observed. The Alfvén waves which are responsible for scattering the energetic particles can have a non-linear amplitude $\delta B/B \sim 0.2$.

(vi) The pre-shock gas is decelerated (in the shock frame) by the pressure gradient exerted by the energetic ions.

The efficiency with which ESP shocks accelerate particles appears to increase with the strength of the shock. Shocks produced by the most powerful solar flares appear to accelerate particles with an efficiency exceeding 30 percent [36, 37, 38].

2.3. Solar flares

Solar flares are explosions above the sun's photosphere caused by a sudden release of magnetic energy (up to 10^{33} erg). They generate shock waves in the sun's corona as well as in the solar wind. These shock waves are after associated with type II radio bursts, which allow one to trace the time history of the shock as it moves into regions of lower density because they emit at the local plasma frequency (or twice that value). Solar flares also produce energetic particles, many of which are probably accelerated in the shocks [39, 40]. There is also evidence that other acceleration mechanisms are at work during flares, but they will not be discussed here.

An advantage offered by observations of energetic solar flare particles is that the charge states and composition can be measured for individual events. These charge states are consistent with the hypothesis that the particles are mostly produced in the sun's corona at a temperature of 10^6 K and undergo little stripping during and after the acceleration process although they are fractionated in a charge to mass ratio-dependent manner. The acceleration and subsequent escape of the particles must be quite rapid. The composition of particles produced in large flares is quite similar to that of Galactic cosmic rays [41].

2.4. Stellar wind termination shocks

The solar wind must be decelerated by the interstellar medium when its momentum flux declines to the value of the interstellar pressure ($\sim 10^{-12}$ dyne cm^{-2}). This is believed to occur at roughly 50 astronomical units ($\approx 7 \times 10^9$ km) from the sun [42]. As the solar wind is highly supersonic up to this point, it must form a shock there. Such a shock would be much larger than other shocks in the solar system and stronger than most of them, although small by interstellar standards. It has been proposed

[43] that this shock is responsible for the anomalous component of cosmic rays – a population that is more energetic than the cosmic rays typically produced in interplanetary space but less energetic than galactic cosmic rays. The energy per charge extends up to 100 MeV which corresponds to the total pole to equator potential difference across the termination shock (or equivalently across the sun).

Far more powerful winds are associated with stellar winds from hot massive stars (speeds $\sim 3000\,\mathrm{km\,s^{-1}}$ and discharges $\sim 10^{-6} M_\odot\,\mathrm{yr}^{-1}$). These are believed to terminate at strong shocks where cosmic rays may be accelerated. The power of the galaxy in these winds is estimated to be comparable with the galactic cosmic ray power (section 2.5) and so the acceleration efficiency must be very high if they are to contribute to the galactic flux [44, 45].

2.5. Galactic cosmic rays and supernovae

Galactic cosmic rays, observed directly at Earth, have kinetic energies T (measured henceforth in GeV/n) ranging from ~ 0.1 to $\sim 10^{11}$ [45]. They comprise protons (the dominant component), alpha particles, and heavier nuclei ($\sim 1\%$) together with electrons ($\sim 2\%$), positrons and antiprotons. The proton intensity $I(T)$ ($\mathrm{cm^{-2}\,s^{-1}\,sterad^{-1}\,(GeV/n)^{-1}}$) appears to be isotropic (to a few parts in 10^4 at $T \sim 100$ GeV), time steady over geological timescales $\gtrsim 10^6$ yr and can be described as a power law in kinetic energy,

$$I(T) \propto T^{-2.7}, \qquad 3 \lesssim T \lesssim 10^5 \tag{2.1}$$

(see fig. 3). At energies below ~ 3 GeV the observed spectrum is strongly influenced by solar modulation, and the measured intensity falls below an extrapolation of this power law. At energies above $\sim 10^5$ GeV, the spectrum appears to flatten and then steepen to a power law with exponent ~ -3.1 possibly flattening again so that the exponent increases to ~ -2 above $\sim 10^9$ GeV [47, 48]. Cosmic rays in the energy range 10^5–10^9 GeV may be predominantly Fe [49, 50].

The cosmic ray energy density in interstellar space is dominated by mildly relativistic protons and is therefore uncertain but probably lies in the range $(1\text{–}2) \times 10^{-12}\,\mathrm{erg\,cm^{-3}}$. This is slightly larger than the energy density in the microwave background and in starlight but comparable with the estimated thermal energy density [51]. The interstellar magnetic field has a mean energy density variously estimated to lie in the range 10^{-14}–$3 \times 10^{-13}\,\mathrm{erg\,cm^{-3}}$ and may act as a pressure valve which allows cosmic rays to escape intermittently from the galactic disk when they have built up a pressure comparable with the magnetic stress [52]. Alternatively, cosmic rays may be convected steadily away from the galactic plane in a wind driven by supernova explosions in the disk [15, 53].

Other primary species are observed to have power law spectra with similar slopes to the proton spectrum. However, the abundances at a given kinetic energy per nucleon differ from standard solar system abundances in the sense that H, He and CNO are suppressed by factors of order 10, 20 and 5, respectively with respect to Fe [54]. This trend is often included in an inverse correlation of abundance with first ionization potential [55, 56]. Similar to that inferred for solar cosmic rays [41] isotopic anomalies are apparent in a few elements, most notably Ne, Mg and Si [57] but otherwise are not as prominent as might be expected if the accelerated particles had just been synthesized in a supernova explosion. r-process element enrichment, reported by some earlier authors, no longer seems to be present.

Secondary species, particularly Li, Be and B are produced by spallation reactions in which primary cosmic rays (especially C, N, O) collide with interstellar hydrogen atoms. The observed elemental

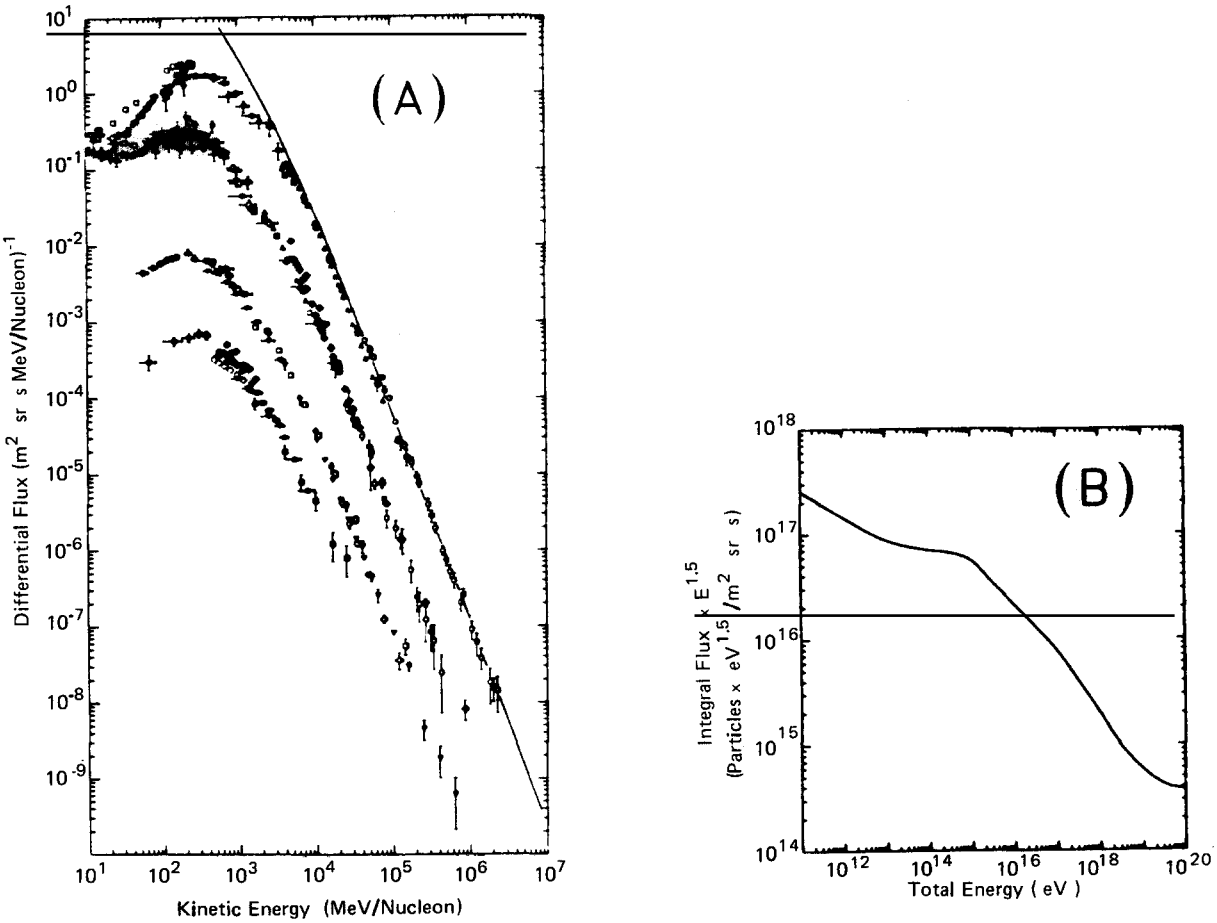

Fig. 3. (a) Different energy spectra of cosmic rays species measured at 1 AU based on satellite and balloon data. In order of decreasing flux, the data describe hydrogen, helium, carbon and iron. The continuous line shows the spectrum extrapolated to interstellar space after removing the effects of solar modulation (adapted from [46]). (b) Integral energy spectrum at high energy, multiplied by $E^{1.5}$ (adapted from [46]).

abundances and their energy dependences are consistent with the primary cosmic rays having traversed an exponential distribution of column densities within the galaxy of mean escape column density λ given by [59, 60]

$$\lambda \sim 8\,\mathrm{g\,cm^{-2}}, \qquad T \lesssim 5$$
$$\sim 20\,\mathrm{g\,cm^{-2}} T^{-0.5}, \qquad 5 \lesssim T \lesssim 100\,. \tag{2.2}$$

Two conclusions can be drawn immediately. Firstly, as the column density through the galactic disk is $\lambda_g \sim 2\,\mathrm{mg\,cm^{-2}}$, we can compute the steady flux of cosmic rays leaving the galactic disk to be $\sim(\lambda_g/\lambda)u_{CR}c \sim 10^{-5}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$. Averaging over the galaxy gives a galactic cosmic ray power $\sim 3 \times 10^{40}\,\mathrm{erg\,s^{-1}}$ good to a factor ~ 3. If, as is generally believed, the supernova rate in the galaxy is $S_{SN} \sim 3 \times 10^4\,\mathrm{Myr^{-1}}$ [61], then each supernova explosion must channel $\sim 3 \times 10^{49}$ erg or ~ 3 per cent of its total energy release into cosmic ray acceleration. (Supernovae provide the dominant power input to

the galaxy, with the possible exception of shocks in spiral arms [62]. So any alternative source of cosmic rays such as stellar wind termination shocks [63, 64] must be even more efficient than this.)

Secondly, as the escape time decreases with increasing energy, the injected cosmic ray spectrum, $S(T)$, at high energy ($T \gtrsim 5\,\mathrm{GeV/n}$) must be flatter than that observed, specifically, $S(T) \propto I(T)\lambda^{-1} \tilde{\propto} T^{-2.2}$. Note that if the slope of the injection spectrum $S(T)$ is close to -2, the contribution of particles at $E \gg 1\,\mathrm{GeV}$ to the total cosmic ray power ($\propto \int \mathrm{d}T\, T\, S(T)$) becomes significant. A further inference that can be drawn is that cosmic rays have to be accelerated rapidly out of the thermal pool. There cannot be a suprathermal pool of incipient cosmic rays already existing in the interstellar medium, say at 1 or 2 MeV, because the ionization loss grammage decreases rapidly with decreasing subrelativistic kinetic energy, $\lambda_{\mathrm{ion}} \propto T$, and so if such a pool did exist the power necessary to replenish it would greatly exceed that invoked to accelerate the cosmic rays [65]. Several other constraints along these lines, derived from the effects of such ionization losses on molecular cloud chemistry and the spectra of spallation products, tighten the argument, ruling out pre-injection at higher energy as well. An independent argument against pre-injection due originally to Fermi is that ionization losses, which vary substantially among the elements, would greatly distort cosmic ray composition (unless the injection energy lies between 1 and 2 MeV [66]).

An independent estimate of the galactic cosmic ray power comes from the gamma ray observations which indicate total galactic gamma ray luminosity in the energy range 100 MeV–1 GeV of $L_\gamma \sim 8 \times 10^{38}\,\mathrm{erg\,s^{-1}}$ [67]. These γ-rays are mostly created by the decay of neutral pions produced in inelastic scattering by interstellar nuclei. Estimating the inelastic cross section for p–p scattering by $\sigma \sim 30\,\mathrm{mb}$ and noting that γ-rays account for roughly 1/3 of the energy release, we again arrive at an estimated cosmic ray power $\sim(3m_{\mathrm{p}}/\lambda\sigma)L_\gamma \sim 2 \times 10^{40}\,\mathrm{erg\,s^{-1}}$, consistent with our earlier estimate.

There are two commonly used chronometers for galactic cosmic rays. Firstly, the isotope ^{10}Be has a β-decay half-life of $\sim$3 Myr. The low observed abundance implies that the age of local cosmic rays is $\sim$20 Myr [68]. Secondly, the electron spectrum appears to steepen above $\sim$100 GeV [69]. This is probably caused by energy loss through the emission of synchroton radiation in the Galactic magnetic field and the inverse Compton scattering of interstellar starlight. (Synchrotron radio emission is a good tracer of relativistic electron density within our galaxy). The lifetime of a $\sim$100 GeV electron within typical Galactic field $\sim$1–2 μG is also $\sim$30 Myr. If this is the cosmic ray lifetime, then cosmic rays occupy a volume of mean density $\sim 2 \times 10^{-25}\,\mathrm{g\,cm^{-3}}$, approximately one tenth of the mean value in the galactic disk. Cosmic rays are therefore believed to be confined to a disk of thickness roughly ten times larger than that of the hydrogen disk. These estimates take no account of spatial and temporal inhomogeneity, which are both thought to be important features of the interstellar medium and global models of galactic cosmic rays must allow for them. (See ref. [15] for further review of existing attempts to do this.)

In fact, the energetics allow us to be more specific about when the acceleration must occur. Cosmic rays are believed to be strongly scattered by interstellar gas (as indicated by their observed isotropy). If they are accelerated following a supernova, then their energy should decline inversely with the cube root of their number density as long as they remain relativistic. This in turn implies that the cosmic rays cannot be accelerated too soon after the supernova explosion. (The fairly "normal" cosmic ray abundances are consistent with this conclusion.) If accelerated, relativistic cosmic rays lose energy as the thermal gas decompresses back to the ambient interstellar pressure then $T \propto R^{-3/5}$. Cosmic rays accelerated within 10^2 years of the explosion will, according to this prescription, lose 96 per cent of their energy. It therefore seems unlikely that most of the cosmic ray power is generated at early times. However the very highest energy particles may have to be accelerated in young remnants [151].

A typical supernova explosion ejects $M \sim 1\text{–}10 M_\odot$ of the stellar envelope with speeds V_0 between $3\,000\,\mathrm{km\,s^{-1}}$ and $10\,000\,\mathrm{km\,s^{-1}}$. When the ejecta has expanded out to a radius $\sim (M/\rho)^{1/3} \sim 3\,\mathrm{pc}$ it will have swept up its own mass of interstellar matter and will start to decelerate. This takes several hundred years to occur and many of the prominent observed remnants are believed to have reached this stage in their evolution.

Older remnants are thought to expand according to the Sedov–Taylor similarity law [70, 71]

$$R \sim (M/\rho)^{1/5} (V_0 t)^{2/5} \tag{2.3}$$

where R is the radius of the blast wave and t is the age. Once interstellar gas has been passed by the blast wave, it is quickly accelerated to a radial velocity $\sim 3/4$ of the shock velocity. The interior of the remnant is roughly isobaric with pressure $p \sim 3MV_0^2/8\pi R^3$. As the blast wave is decelerating, the stellar ejecta must likewise be decelerated by passing through a "reverse" shock which may also accelerate cosmic rays. The duration of this phase depends upon the ambient density. If, as is now usually believed, most of the volume of the interstellar medium comprises a low density ($\lesssim 10^{-27}\,\mathrm{g\,cm^{-3}}$) gas [50, 72, 73] then radiative losses are relatively unimportant and the remnant expands until the internal pressure approaches the ambient external pressure $\sim 10^{-12}\,\mathrm{dyne\,cm^{-2}}$. In this case, the interstellar medium must be filled with overlapping $\gtrsim 10^6\,\mathrm{yr}$ old, $R \lesssim 200\,\mathrm{pc}$ spheres. (A supernova remnant expanding into a uniform interstellar medium cools more rapidly and it will not expand beyond a radius $\sim 20\,\mathrm{pc}$.) If cosmic ray acceleration were still effective at radii $R \gtrsim (3L_{\mathrm{CR}}/4\pi u_{\mathrm{CR}} S_{\mathrm{SN}})^{1/3} \sim 50\,\mathrm{pc}$, then individual cosmic rays would be accelerated several times before they leave the galaxy. This would have the further consequence that the spectrum of the secondary particles would resemble that of the primary particles instead of being steeper as is observed. Perhaps cosmic ray acceleration is only efficient in strong shocks associated with younger remnants, cf. section 6 below. So we must conclude that there are both upper and lower bounds on the size of a supernova remnant R_{acc} when the *bulk* of cosmic ray acceleration must occur, we estimate very roughly i.e., $1\,\mathrm{pc} \lesssim R_{\mathrm{acc}} \lesssim 50\,\mathrm{pc}$ [74, 75] (but see [76, 77]). (Note, though, that supernova blast waves cannot accelerate cosmic rays to energies $T \gtrsim 10^6$ because there Larmor radii $\sim 10^{-6} T$ pci would be too large for them to be confined to the remnant, cf. section 4.4 below.) In order to be more quantitative, we must first understand the physics of the mechanism.

There is further circumstantial evidence that particle acceleration does occur in supernova remnants from the fact that we observe these remnants most easily by the radio synchrotron emission produced by freshly accelerated relativistic electrons. If we exclude filled (plerionic) supernova remnants, then most remnants have spectral indices $\alpha \sim 0.6$ from which we infer underlying relativistic electron spectra $I(E) \tilde{\propto} E^{-2.2}$ [78, 79].

Most supernova remnants are also X-ray sources. The X-ray emission is believed to be produced by thermal bremsstrahlung of the hot post-shock gas and the detailed X-ray spectra can be rendered consistent with the post-shock temperature expected from the shock jump conditions [80]. The very existence of this emission and its sharp boundary provides strong evidence for the presence of collisionless shocks in the interstellar medium [71].

The connection between supernova shocks and cosmic rays (at least cosmic ray electrons) is strikingly illustrated by the superposition of the X-ray and radio profiles of supernova remnants which sometimes coincide even at the scale of the detailed irregularities of the blast wave (fig. 4).

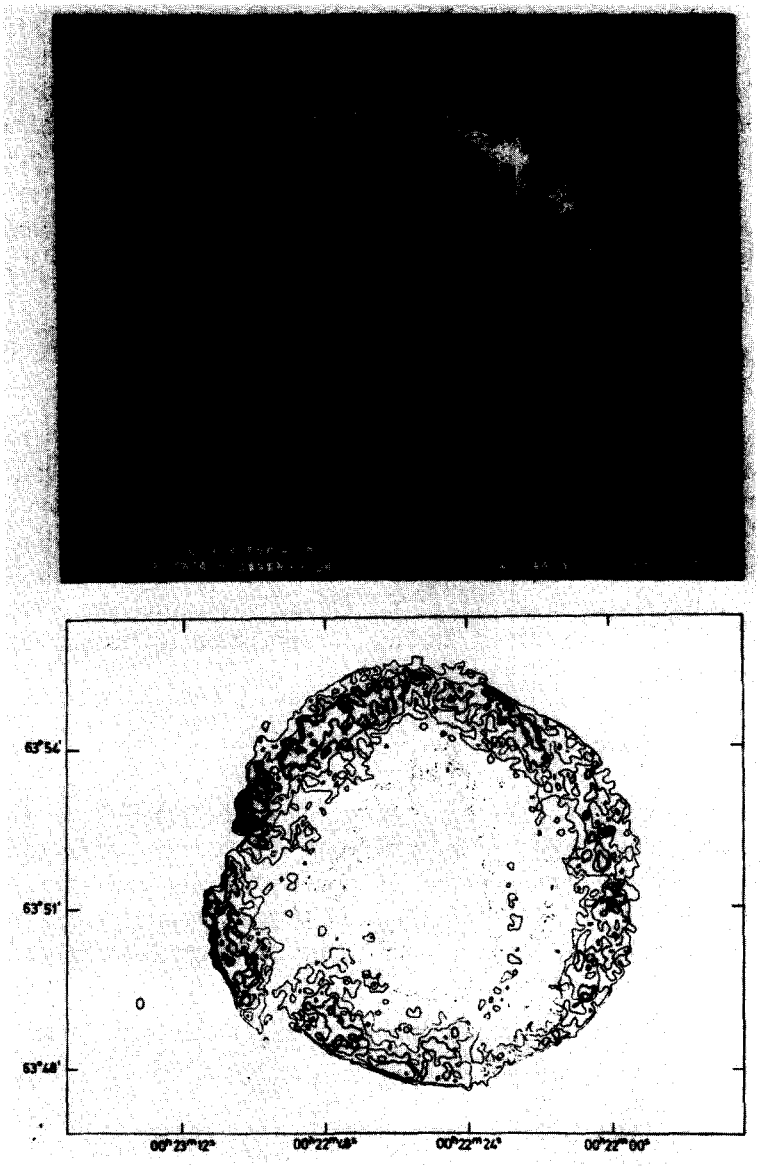

Fig. 4. The Tycho supernova remnant is displayed in X-ray (upper) and radio (lower) maps. The radio emission is believed to be synchrotron emission from cosmic ray electrons. The X-rays are emitted by shock-heated electrons. That the two profiles are so similar in detail, e.g. the irregularity in the lower left hand sector, suggests that both populations are energized by the same event (presumably the shock wave) on a much shorter timescale than that of the expansion. Provided courtesy of Paul Gorenstein.

2.6. *Binary X-ray sources*

Ultrahigh energy radiation has been reported from several binary X-ray sources. The UHE neutral quanta are probably photons produced by cosmic ray interactions within the binary system (though there have also been reports of underground muons from Cyg X-3, which are inexplicable by

conventional physics). The signals have been reported at energies ranging from 10^{12} to 10^{16} eV. The reality of these signals is to varying degrees underscored by the periodicity of the signals which, in all cases, matches some periodicity in the systems as determined by X-ray measurements. In three of the sources, the signals have the spin periodicity of the neutron star. On the other hand, the neutron stars, in most cases, are rotating too slowly to accelerate the primary particles via the pulsar mechanism. The implied power in UHE particles is comparable to the X-ray luminosity. The X-rays are likely to come from shocks in the accreting material near the neutron star surface and shock acceleration could take place as well. However, many puzzles remain, and better observations are eagerly awaited. For a recent review see ref. [145].

2.7. *Galactic wind termination shock*

It is widely, though not universally, believed that the power deposited in the interstellar medium by supernovae drives a galactic wind with mass loss rate $\approx 10 M_\odot\,\mathrm{yr}^{-1}$ and speed $\sim 500\,\mathrm{km\,s}^{-1}$ which terminates at a distance ~ 100 kpc from the galaxy when its momentum flux balances the pressure of the intergalactic medium in the local group or meets winds from other galaxies in the local group [47, 74]. This may be the acceleration site for high energy ($T \gtrsim 10^6$) cosmic rays. By analogy with the solar wind termination shock, the maximum energy per charge would be of order the potential drop across the galactic disk, some 10^{17} V. It is easier to account for the highest energy cosmic rays in this way if they comprise iron, because their charge is larger [15, 81].

2.8. *Extragalactic radio sources and active galactic nuclei*

It has been known for over thirty years that a subset of external galaxies are powerful radio sources [82]. The radio emission comes from two general regions – a compact source of size in the range 10^{18}–10^{20} cm in the nucleus [83] and an extended component that ranges in size from 10^{22}–10^{25} cm [84]. Either component may dominate but usually both are present at some level. The extended components generally have a characteristic double structure – two radio-emitting lobes situated on either side of the

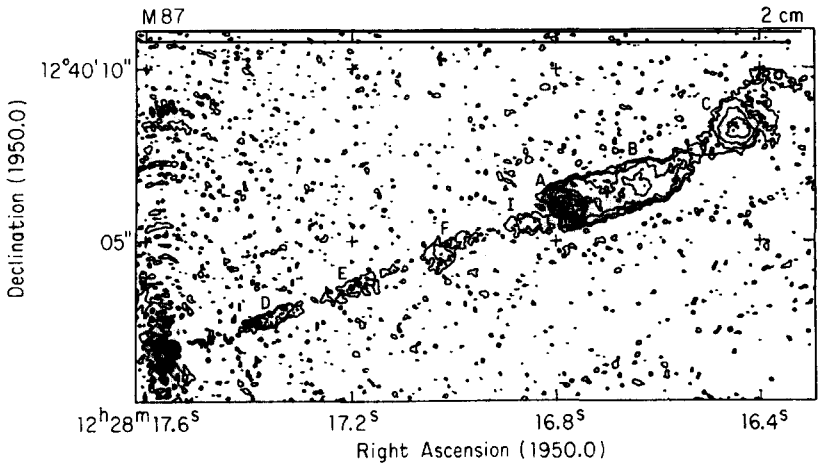

Fig. 5. Possible shock wave formed in the extragalactic jet associated with the galaxy M87. The radio contours, from [263] show a sharp feature elongated roughly perpendicular to the presumed jet outflow from bottom left to top right. The intense synchrotron emission is generally attributed to particle acceleration and field amplification in the vicinity of this shock wave.

optical galaxy and presumably moving away from it into the surrounding medium. More recently it has been discovered that the extended components are fueled with mass, momentum and energy through thin jets that emanate from the nucleus [85, 86].

The radio emission through which these sources are observed is thought to be synchrotron emission from electrons with energy in the ~1–100 GeV range which must be accelerated locally. In the more powerful sources, the jets are believed to terminate at "radio hot spots" by means of a strong shock, and it seems that shock acceleration is occurring here. Shocks may also be produced in the jets themselves through the non-linear development of surface instabilities, encounters with dense clouds in the ambient medium, or variations in the jet speed. In particular, if the external pressure changes more rapidly than the internal pressure of the jet can adjust, then the jet may develop Mach disk shock structures similar to those observed in under-expanded rocket exhausts. It is tempting to interpret the quasiperiodic "knots" of emission seen in some jets (e.g., M87) [87] in these terms. The compact sources are also believed to involve expansion flows (fig. 5). We can obtain crude estimates of the relativistic electron and magnetic energy densities within the source region which form the basis of quantitative dynamical models also involving shock waves [88, 89].

In fact in many of the compact and extended radio sources, the plasma appears to be moving at relativistic speeds necessitating the development of a theory of relativistic shock acceleration. Furthermore, polarization maps can be used in some cases to argue that positrons rather than protons maintain charge neutrality within the plasma.

Active galactic nuclei generally exhibit intense power-law continua extending from the infrared to the X-ray. These are believed to originate from regions $\lesssim 10^{18}$ cm across and our understanding of the physical conditions in these sources is even poorer than is the case for the extended radio components. However, the more rapidly variable examples are thought to be synchrotron sources. It is then generally argued that the radiative cooling times of the electrons are much shorter than the source transit times, again requiring local acceleration and suggesting that shock waves may be responsible [90, 91].

2.9. Intergalactic shock waves

An intergalactic medium is required to confine extragalactic radio sources, to emit the extended thermal bremsstrahlung X-rays from rich clusters of galaxies [92] and (arguably) to confine quasar absorption line clouds [93]. It is also expected to exist simply because galaxy formation must be a somewhat inefficient process. Estimates of the density and temperature within this medium range from $\sim 10^{-32}$–10^{-30} g cm^{-3} and from 3×10^5–3×10^8 K, respectively.

There are several possible sources of strong shock waves in the intergalactic medium. Galaxies, especially those in rich clusters, can move supersonically and will be preceded by bow shocks. (Additional shocks may be found at the termination of galactic winds, cf. section 2.7.) The large double radio sources are also believed to be expanding supersonically and indeed some of the observed relativistic electrons may be accelerated in an intergalactic medium shock wave. The intergalactic medium must have been cooled by the expansion of the universe and so if it is now hot, it must have been reheated at some stage in the past. Shock waves seem to be the only efficient way to do this. These shock waves may be derived from the rapid release of energy by supernovae during the first $\sim 10^8$ yr of the galaxies' lifetime and may be intimately connected with the process of galaxy formation itself [94]. Alternatively, they may result from the non-linear development of velocity perturbations in the expanding medium.

Shock waves in radio sources and intergalactic media can extend over tremendous distances exceeding 10^{25} cm. They may therefore be important for accelerating the highest energy cosmic rays ($10^9 \lesssim T \lesssim 10^{11}$) whose Larmor radii ($\sim 3 \times 10^{21} B^{-1} T$ cm) are far too large for them to be accelerated within the interstellar medium or at galactic wind termination shocks [95], provided that the intergalactic magnetic field strength is $\gtrsim 10^{-8}$ G.

3. The diffusion approximation

In this section we introduce general concepts in particle acceleration theory. We start with the Vlasov equation and eventually derive the general particle transport equation that describes particle acceleration by magnetohydrodynamical turbulence in a moving fluid background. Some relevant physical processes are discussed along the way.

3.1. Particle distribution function

In the majority of astrophysical particle acceleration schemes, cosmic rays acquire their energies through a succession of small increments and can therefore be considered as diffusing and/or being convected through momentum space. There will also be a spatial diffusion associated with these changes in momentum and so it proves convenient to describe the particles by the six-dimensional distribution function of the position $\boldsymbol{x}$, relativistic momentum $\boldsymbol{p}$ and time t, $f(\boldsymbol{x}, \boldsymbol{p}; t)$ where $f\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{p}$ is the number of particles lying in $\mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{p}$. f is Lorentz invariant and satisfies the relativistic Vlasov equation [96, 97]

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f}{\partial \boldsymbol{x}} + \frac{\partial}{\partial \boldsymbol{p}} \cdot (\boldsymbol{F} f) = 0 \tag{3.1}$$

where $\boldsymbol{v}$ is the velocity and $\boldsymbol{F} = \mathrm{d}\boldsymbol{p}/\mathrm{d}t$ is the force acting on the particle. When $\boldsymbol{F}$ is the electromagnetic force, the third term in eq. (3.1) is usually rewritten as $\boldsymbol{F} \cdot (\partial/\partial \boldsymbol{p}) f$. Below, we discuss the various physical processes that contribute to this six-dimensional transport. In particular, we derive a collision operator $(\partial f/\partial t)_{\mathrm{c}}$ that approximately expresses the electromagnetic forcing term $\boldsymbol{F} \cdot (\partial f/\partial \boldsymbol{p})$ and changes eq. (3.1) to a Boltzmann equation. This is then further transformed into a transfer equation.

3.2. Fermi acceleration

We begin with the theory of Fermi [98] published in 1949. Fermi made the first serious attempt at explaining the power law nature of the cosmic ray spectrum using a line of reasoning of quite general applicability. He noted that if cosmic rays are injected steadily into some acceleration region and gain energy at a rate that is proportional to their energy while at the same time their escape from the acceleration region is a Poisson process with energy-independent escape probability, then the stationary particle distribution will be power law. Specifically, if relativistic particles gain energy at a rate $\mathrm{d}E/\mathrm{d}t = E/\tau_{\mathrm{acc}}$ and the probability per unit time of escape is τ_{esc}^{-1} then in a steady state, the number of particles per unit energy satisfies

$$\mathrm{d}N/\mathrm{d}E \propto E^{-(1+\tau_{\mathrm{acc}}/\tau_{\mathrm{esc}})}\,. \tag{3.2}$$

To use a simple analogy, suppose gamblers continually join a game of chance in which they can either increase their winnings by a small fraction f with probability $(1-p)$ or lose everything with probability $p(\ll 1)$, then the number of gamblers that win more than some amount w before inevitably losing is proportional to $w^{-p/f}$ (provided that they don't break the bank, which happens if $f > p$).

In the particular realization of this idea that Fermi first proposed, the cosmic rays gained energy by colliding with interstellar clouds. If the clouds move with speed u, then the relative energy increase in one collision for a relativistic particle is $\sim\pm(u/c)$. However, energy-increasing collisions are favored over energy-decreasing collisions by a fraction also $\sim(u/c)$ and so $\tau_{\rm acc} \sim (c/u)^2$ collision times. (The numerical coefficient depends upon the kinematical assumptions made about the collision and can be straightforwardly calculated [99].)

There are three problems with the Fermi mechanism. Firstly, the acceleration rate depends upon the *square* of the scatterer's velocity and for this reason this mechanism (and others like it) are relatively inefficient and known as "second-order" processes. Secondly, in order to account for the observed spectrum, Fermi required that $\tau_{\rm acc} \sim 1.7\tau_{\rm esc}$. There is no good reason why this should be so and there is every expectation that the ratio $\tau_{\rm acc}/\tau_{\rm esc}$ should vary throughout the galaxy and indeed with energy. The third problem is that the acceleration can only overcome ionization losses for protons if the kinetic energy T exceeds $\sim$100 MeV (as $(\mathrm{d}T/\mathrm{d}t)_{\rm ion} \propto T^{-1/2}$). In fact, Fermi realized this and proposed an ingenious injection mechanism for creating new cosmic ray protons. He suggested that a cosmic ray proton would generate on average one new proton above the injection threshold by knock-on collision before escaping the galaxy. Unfortunately, as Fermi also realized, this mechanism cannot operate for heavy elements which cannot be boosted without fragmentation, and as we now know, are in fact overabundant in cosmic rays. In more modern variants on this type of Fermi acceleration, the clouds have been replaced by hydromagnetic waves [100, 5] though the difficulties remain [101].

3.3. *Momentum space diffusion*

As the fractional energy changes in a single collision are small, we can treat Fermi acceleration using the Fokker–Planck formalism (e.g., see refs. [96, 102]). The probability that an electron changes its momentum by $\Delta\boldsymbol{p}$ in time interval Δt is presumed to depend on the current values of $\boldsymbol{p}$ and $\boldsymbol{x}$ and not on the complete past history of the particle. In other words, we are dealing with a Markov process. The distribution function at time $t + \Delta t$ is then given by

$$f(\boldsymbol{p}, x + v\,\Delta t, t + \Delta t) = \int \mathrm{d}\,\Delta\boldsymbol{p}\,\psi(\boldsymbol{p} - \Delta\boldsymbol{p}, \Delta\boldsymbol{p})\, f(\boldsymbol{p} - \Delta\boldsymbol{p}, \boldsymbol{x}, t) \tag{3.3}$$

where $\psi(\boldsymbol{p}, \Delta\boldsymbol{p})\,\mathrm{d}\,\Delta\boldsymbol{p}$ is the element of probability for changing the momentum $\boldsymbol{p}$ to $\boldsymbol{p} + \Delta\boldsymbol{p}$ in time Δt. We expand both the integrand and the left hand side in a Taylor series and use the fact that $\int \mathrm{d}\,\Delta\boldsymbol{p}\,\psi(\boldsymbol{p}, \Delta\boldsymbol{p}) = 1$ to obtain the Fokker–Planck equation

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\cdot\frac{\partial f}{\partial \boldsymbol{x}} = \frac{\partial}{\partial \boldsymbol{p}}\cdot\left\{-\left\langle\frac{\Delta\boldsymbol{p}}{\Delta t}\right\rangle f + \frac{1}{2}\frac{\partial}{\partial \boldsymbol{p}}\cdot\left\langle\frac{\overline{\Delta\boldsymbol{p}}\,\overline{\Delta\boldsymbol{p}}}{\Delta t}\right\rangle f\right\} \tag{3.4}$$

where the Fokker–Planck coefficients are

$$\left\langle \frac{\Delta \boldsymbol{p}}{\Delta t} \right\rangle = \frac{1}{\Delta t} \int \mathrm{d}\,\Delta \boldsymbol{p}\, \psi(\boldsymbol{p}, \Delta \boldsymbol{p})\, \Delta \boldsymbol{p}$$

and (3.5)

$$\left\langle \frac{\Delta \boldsymbol{p}\, \Delta \boldsymbol{p}}{\Delta t} \right\rangle = \frac{1}{\Delta t} \int \mathrm{d}\,\Delta \boldsymbol{p}\, \psi(\boldsymbol{p}, \Delta \boldsymbol{p})\, \Delta \boldsymbol{p}\, \Delta \boldsymbol{p}\,.$$

This equation describes a biased random walk of the particles in momentum space.

In Fermi processes, particles both gain and lose energy by scattering off moving scatterers. Now there is a simplification to the Fokker–Planck equation when the recoil of the scatterer can be ignored (as is usually true). In this case, the principle of detailed balance assures us that $\psi(\boldsymbol{p}, -\Delta \boldsymbol{p}) = \psi(\boldsymbol{p} - \Delta \boldsymbol{p}, \Delta \boldsymbol{p})$. We again make a Taylor expansion and integrate over $\mathrm{d}\,\Delta \boldsymbol{p}$ to obtain

$$\frac{\partial}{\partial \boldsymbol{p}} \cdot \left\{ \left\langle \frac{\Delta \boldsymbol{p}}{\Delta t} \right\rangle - \frac{1}{2} \frac{\partial}{\partial \boldsymbol{p}} \cdot \left\langle \frac{\Delta \boldsymbol{p}\, \Delta \boldsymbol{p}}{\Delta t} \right\rangle \right\} = 0\,. \tag{3.6}$$

If the Fokker–Planck coefficients vanish for small values of $\boldsymbol{p}$ as is also generally true,

$$\left\langle \frac{\Delta \boldsymbol{p}}{\Delta t} \right\rangle = \frac{1}{2} \frac{\partial}{\partial \boldsymbol{p}} \cdot \left\langle \frac{\Delta \boldsymbol{p}\, \Delta \boldsymbol{p}}{\Delta t} \right\rangle. \tag{3.7}$$

The Fokker–Planck equation then simplifies to

$$\frac{\partial f}{\partial t} + (\boldsymbol{v} \cdot \nabla) f = \frac{\partial}{\partial \boldsymbol{p}} \cdot D_{pp} \cdot \frac{\partial f}{\partial \boldsymbol{p}} \tag{3.8}$$

where the momentum space diffusion coefficient is

$$\boldsymbol{D}_{pp} = \tfrac{1}{2} \langle \Delta \boldsymbol{p}\, \Delta \boldsymbol{p} / \Delta t \rangle\,. \tag{3.9}$$

As we might expect, Fick's law of diffusion applies in momentum space if the scattering process is time-reversible. Terms describing loss and escape can be added to the right hand side of eq. (3.8). Further discussion of eq. (3.8) and its properties can be found in Melrose [5, 6].

In the simplest case of isotropic scattering by an isotropic distribution of scatterers moving with speed $V \ll c$, there will be a change in the momentum of the particle $\Delta p = -(\boldsymbol{p} \cdot \boldsymbol{V})/v$ caused by transforming into the frame of the scatterer and an independent change caused by transforming back again to the original frame. This will happen every L/v where L is the collision mean free path. Hence $\boldsymbol{D}_{pp} = D_{pp}$ times the unit tensor where

$$D_{pp} = \tfrac{1}{2} \times 2 \times \frac{\overline{(\boldsymbol{p} \cdot \boldsymbol{V})^2}}{v^2} \frac{v}{L} = \frac{1}{3} \frac{p^2 \langle V^2 \rangle}{vL} \tag{3.10}$$

and where the overbar denotes an angle average. Note that the rate of gain of energy for relativistic particles satisfies $\mathrm{d}E/\mathrm{d}t \propto \mathrm{d}D_{pp}/\mathrm{d}E \propto E$ as in a traditional Fermi process. However this scaling does not hold at non-relativistic energy.

3.4. Pitch angle scattering and spatial transport

The Fermi process requires that in a steady state particles leave the acceleration region and so we must also discuss spatial transport. The simplest formalism (which we have already invoked) is the escape time formalism which consists of adding a term $-f/\tau_{esc}$ to the right hand side of eq. (3.8). This ignores spatial gradients. The acceleration region is regarded as a "leaky box" at whose walls the particles have a small but finite transmission probability. (This appears to be a surprisingly good approximation in the case of galactic cosmic rays.) We can then use this approach to study Fermi acceleration in a leaky box model of the galaxy. We seek a steady state solution for relativistic particles using the diffusion coefficient (3.10). If τ_{esc} is energy-independent, then there is a power law solution,

$$\mathrm{d}N/\mathrm{d}E \propto p^2 f \propto E^{-n}$$

where

$$n = \frac{3}{2}\left(1 + \frac{4cL}{3\langle V^2\rangle \tau_{esc}}\right)^{1/2} - \tfrac{1}{2}\,. \tag{3.11}$$

Note how the (necessary) inclusion of diffusion in momentum space changes the result from the original Fermi theory except when $\tau_{acc} \ll \tau_{esc}$ [99]. It is an elementary exercise to solve the time-dependent equation (3.3) for arbitrary initial conditions.

Spatial gradients are included in a more sophisticated approach, which is certainly necessitated by observations of particles in the interplanetary medium. We assume that particles are continuously scattered by magnetic inhomogeneities or Alfvén waves which are incompressible hydromagnetic transverse modes driven by the tension $B^2/4\pi$ in the field. They therefore travel with the Alfvén speed

$$v_A = \boldsymbol{B}/\sqrt{4\pi\rho}\,. \tag{3.12}$$

Cosmic rays can also be scattered by fast magnetosonic wave modes. However, in the plasmas that most concern us the gas pressure exceeds the magnetic pressure and the magnetosonic modes will be Landau damped. Alfvén waves have frequencies below the ion Larmor frequency $\Omega = eB/m_p c$. A suprathermal particle traveling much faster than the Alfvén speed will therefore see the wave as an essentially magnetostatic disturbance. If its Larmor radius, $r_L = pc/ZeB$ is comparable with the wavelength of the disturbance, then there is a possible resonant interaction which results in a change in the particle's pitch angle $\theta = \cos^{-1}\{(\boldsymbol{p}\cdot\boldsymbol{B})/pB\} \equiv \cos^{-1}\mu$. Successive interactions with randomly phased waves cause the particle's pitch angle to random walk and after $\sim(B/\delta B)^2$ wavelengths, the particle's velocity will have been reversed so that it can be regarded as having been scattered.

More formally, we can compute the pitch angle scattering rate for Alfvén waves traveling in one direction along the ambient magnetic field [103, 104, 105]. (This is probably not too bad an approximation as Alfvén waves propagating across the field are expected to have a lower growth rate than those propagating along the field.) If we transform to the frame of the wave, then the electric field vanishes and we can consider the motion of a particle interacting with a single circularly polarized wave of magnetic amplitude $\boldsymbol{B}_1$ perpendicular to $\boldsymbol{B}$. The particle energy is strictly conserved and so the pitch angle changes according to

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \frac{Ze}{pc}\,[\boldsymbol{v}_\perp \times \boldsymbol{B}_1]\,. \tag{3.13}$$

To lowest order, we evaluate the right hand side of eq. (3.13) along the unperturbed trajectory

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \frac{Zev(1-\mu^2)^{1/2}B_1}{pc}\cos[(kv\mu - \Omega)t + \psi] \tag{3.14}$$

where k is the wave k-vector and ψ is the phase of the wave, assumed to be randomly distributed. Averaging over ψ, we obtain

$$\langle \Delta\mu^2 \rangle = \frac{Z^2e^2v^2(1-\mu^2)B_1^2}{2p^2c^2}\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \cos[(kv\mu - \Omega)(t'-t'')]\,. \tag{3.15}$$

This expression confirms that only resonant waves with $k = \Omega/v\mu$ interact strongly with the particles. (When $\boldsymbol{k}$ is not parallel to $\boldsymbol{B}$, higher cyclotron resonances are also important.) In the limit $t\to\infty$, eq. (3.15) can be converted into an expression for the pitch angle diffusion coefficient,

$$\nu \equiv \left\langle \frac{\Delta\theta^2}{\Delta t} \right\rangle = \frac{\pi Z^2e^2vB_1^2}{p^2c^2\mu}\,\delta(k - \Omega/v\mu)\,. \tag{3.16}$$

The total energy density in an Alfvén wave, which is divided equally between particle motion and field energy, is $B_1^2/4\pi$. (Only one of the two modes of circular polarization will interact with particles with given sign of μ and Z.) We then integrate eq. (3.16) over all waves to obtain

$$\nu = \frac{\pi}{4}\left(\frac{k\mathscr{E}_k}{B^2/8\pi}\right)\Omega \tag{3.17}$$

where $\mathscr{E}_k\,\mathrm{d}k$ is the wave energy density in interval $\mathrm{d}k$ evaluated at $k = \Omega/v\mu$.

The evolution of the particle distribution function in the presence of the waves satisfies the Fokker–Planck equation (3.8) which specialises to

$$\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial}{\partial\mu}\left((1-\mu^2)\,\nu\,\frac{\partial f}{\partial\mu}\right) \tag{3.18}$$

when f does not depend on gyrational phase and changes of particle energy can be ignored. Note that as the scattering is elastic in the wave frame, it is only necessary to calculate the first order perturbations to the particles' motion in order to derive all the Fokker–Planck coefficients (cf. section 3.3).

In the above derivation of (3.18), spatial uniformity, $\partial f/\partial \boldsymbol{x} = 0$ was assumed. This identifies the electromagnetic force term that appeared in the Vlasov equation (3.1), given the stated approximations, as the effective collision operator in a Boltzmann equation. This can be generalized to accommodate spatial gradients associated with particle streaming to give

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\cdot\frac{\partial f}{\partial \boldsymbol{x}} = \left(\frac{\partial f}{\partial t}\right)_{\mathrm{c}} = \frac{\partial}{\partial\mu'}\,(1-\mu'^2)\,\nu\,\frac{\partial f}{\partial\mu'}\,. \tag{3.19}$$

Here the primes in the collision operator denote the fact that the pitch angle is to be measured in the frame of the waves, whereas v refers to the velocity in the laboratory frame. (We ignore a time dilation factor, second order in the wave speed.) The collision operator used here represents "small angle scattering", i.e., diffusion in pitch angle. (Large angle scatterings can be included by adding the BGK scattering operator $-\nu^*[f-\langle f\rangle_\mu]$ [106], where ν^* is an effective scattering frequency.)

Several features of eq. (3.19), which can also be derived directly from the Vlasov equation [5, 6], deserve comment. Firstly we emphasize that we have assumed small wave amplitudes, random phases and that $\boldsymbol{k}$ is parallel to $\boldsymbol{B}$, any of which may be violated in a real cosmic ray shock. Secondly, the resonant k-vector increases as the pitch angle approaches 90°. When $\mu < v_A/v$, there are no resonant Alfvén waves and formally there should be no scattering between forward and backward hemispheres in velocity space. At first sight this seems to rule out multiple shock crossings. (Actually, when the sound speed v_s exceeds the Alfvén speed, v_A, Alfvén waves resonant with particles with $\mu < v_s/v$ will suffer ion cyclotron damping by thermal ions and there is an even larger range of pitch angle within which scattering is absent [107, 108].) The problem of scattering when the pitch angle is close to 90° has been addressed by several authors [109, 110, 111, 112]. It is now generally believed that, unless $v_s \gg v_A$, non-linear effects, interpretable as particle mirroring and resonance broadening allow passage through $\mu = 0$. This is indeed consistent with interplanetary observations at high energies; although at lower energy ($E \lesssim 10$ MeV), the inferred mean free path exceeds that predicted by quasi-linear theory (cf. also section 5.2). Scattering by waves propagating above the ion Larmor frequency (e.g., whistlers) may also be important.

When the particle mean free paths are sufficiently short, it is more appropriate to treat the spatial transport in the diffusion approximation. As we show formally in the next section, a term $-\partial/\partial\boldsymbol{x}(\boldsymbol{D}(\boldsymbol{x}, p)\,\partial f/\partial\boldsymbol{x})$ must be added to the right hand side of eq. (3.8). The diffusion tensor $\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{p})$ is anisotropic because of the presence of a particular magnetic field direction. If we identify ν as a collision frequency then we can identify the spatial diffusion coefficient along the mean direction of the field by $D_\parallel \simeq v^2/3\nu$ which is inversely proportional to the level of resonant wave turbulence (cf. section 3.5). The perpendicular component of the diffusion coefficient is likewise estimated by

$$D_\perp \simeq v^2\nu/3\Omega^2 \sim D_\parallel(\delta B/B)^4 \tag{3.20}$$

(refs. [5, 6]). There are also off-diagonal components, $D_{xy} = -D_{yx} \simeq v^2/3\Omega$ [113] which contribute to the transport as gradient B drift [114]. Spatial transport perpendicular to the ambient field can also be enhanced non-resonantly by field line wandering [115, 116]. In this process, two field lines that start off close to each other will separate at a rate fixed by their separation and the level of field fluctuations. Particles on these field lines with Larmor radii much smaller than the separation, s, will also separate to the extent that they move systematically along the field lines in a particular direction. In the limit that $\mathscr{E}_k$ diminishes with k (but not faster than k^{-2}), the dominant contribution is by waves with $k \sim s^{-1}$ and a simple random walk argument gives

$$\mathrm{d}s_\perp/\mathrm{d}z \sim (8\pi k\mathscr{E}_k/B^2)_{k=s^{-1}}\,. \tag{3.21}$$

3.5. Convection-diffusion equation

We have so far considered particle transport in a stationary medium. In order to discuss cosmic ray acceleration at a shock front, we must generalize the above results to allow for motion of the scattering

medium. This has been done by several authors. A sufficiently complete and rigorous derivation is due to Skilling [117] (cf. also ref. [118]).

We require a transport equation for the distribution function $f(\boldsymbol{x}, \boldsymbol{p}, t)$ measured in a suitable inertial frame. Let us anticipate our application to shock fronts by assuming *ab initio* that

$$\frac{\partial}{\partial t} = \mathrm{O}\left(u \frac{\partial}{\partial x}\right) = \mathrm{O}\left(\frac{v^2}{\nu} \frac{\partial^2}{\partial x^2}\right) \tag{3.22}$$

and $u \ll v$, where $\boldsymbol{u}(\boldsymbol{x}, t)$ is the velocity of the frame in which the scattering is presumed to be elastic. In particular if the scattering is due to Alfvén waves, then $\boldsymbol{u}$ is the velocity of the wave frame in which the electric field vanishes. If the Alfvén speed $v_{\mathrm{A}} \ll u$ then $\boldsymbol{u}$ is effectively also the speed of the background medium.

The procedure which we follow to derive the particle transport equation is to transform the Vlasov equation (3.1) from the inertial frame to the non-inertial (primed) frame moving with the scattering waves expressing the transformed distribution function in terms of the original $\boldsymbol{x}, t$ coordinates in the inertial frame and the transformed momentum $\boldsymbol{p}'$, measured in the wave frame, i.e.,

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{p}, t) &= f'(\boldsymbol{x}, \boldsymbol{p}' = \boldsymbol{p} - E'\boldsymbol{u}, t) \\ &= f'(\boldsymbol{x}, \boldsymbol{p}' = \boldsymbol{p}, t) - E'\boldsymbol{u} \frac{\partial}{\partial \boldsymbol{p}'} f'(\boldsymbol{x}, \boldsymbol{p}' = \boldsymbol{p}, t) \end{aligned} \tag{3.23}$$

to first order. The transformed Vlasov equation is then

$$\frac{\partial f'}{\partial t} + (\boldsymbol{v}' + \boldsymbol{u}) \cdot \nabla f' - (\boldsymbol{p}' \cdot \nabla)\boldsymbol{u} \cdot \frac{\partial f'}{\partial \boldsymbol{p}'} = -\frac{\partial}{\partial \boldsymbol{p}'} \cdot \left\{ f' \frac{\mathrm{d}\boldsymbol{p}'}{\mathrm{d}t'} \right\} \tag{3.24}$$

again working to Galilean order, $u \ll c$.

The Larmor radius is, by assumption, much smaller than the scale length and so the distribution function should be effectively independent of the gyrational phase ϕ' measured in the wave frame. We can then average over this coordinate. We introduce a unit vector field $\boldsymbol{n}(\boldsymbol{x}, t)$ parallel to the local magnetic field and denote the local pitch angle of a particle at $\boldsymbol{x}$ (i.e., the angle between $\boldsymbol{p}'$ and $\boldsymbol{n}$) as $\theta' = \cos^{-1} \mu'$. In this case f' is a function of x, v', μ', t. Now particles moving along a specific direction in the inertial frame will have variable pitch angles measured with respect to the local field. If we introduce coordinates x, y, z with z measured along $\boldsymbol{n}$, then

$$(\partial \mu'/\partial \boldsymbol{x})_v = \sin\theta' \cos\phi' \, \partial n_x/\partial x \,, \qquad (\partial \mu'/\partial y)_v = \sin\theta' \sin\phi' \, \partial n_y/\partial y \,. \tag{3.25}$$

Hence,

$$\langle (\boldsymbol{v}' \cdot \nabla) f' \rangle_{\phi'} = \mu_0' v' (\boldsymbol{n} \cdot \nabla) f + \left\langle v_x' \frac{\partial f'}{\partial \mu'} \frac{\partial \mu'}{\partial x} + v_y' \frac{\partial f'}{\partial \mu'} \frac{\partial \mu'}{\partial y} \right\rangle_{\phi} \tag{3.26}$$

where $\mu' = \cos\theta'$. Inserting the derivatives from eq. (3.25) and averaging, gives

$$\langle (\boldsymbol{v}' \cdot \nabla) f' \rangle_{\phi'} = \mu' v' (\boldsymbol{n} \cdot \nabla) f' + \tfrac{1}{2}(1 - \mu'^2)\, v' \, (\nabla \cdot \boldsymbol{n}) \, \partial f'/\partial \mu' \,. \tag{3.27}$$

The first term on the left hand side is the contribution associated with gradients in the distribution function along the field. The second term is present when the magnetic field direction changes. Particles on neighboring field lines must be compared at slightly different pitch angles (measured with respect to the local field direction) when evaluating the convective derivative in the Vlasov equation.

Likewise,

$$\langle (\boldsymbol{p}'\cdot\nabla)\boldsymbol{u}\cdot\partial f'/\partial \boldsymbol{p}'\rangle_{\phi'} = \langle p'\{\sin\theta'\cos\phi'\,\partial/\partial x + \sin\theta'\sin\phi'\,\partial/\partial y + \cos\theta'\,\partial/\partial z\} \times\{u_x\sin\theta'\cos\phi' + u_y\sin\theta'\sin\phi' + u_z\cos\theta'\}\,\partial f'/\partial p'\rangle_{\phi'} \tag{3.28}$$

where we use the fact that f' is isotropic to lowest order. Performing these averages, gives

$$\left\langle (p'\cdot\nabla)\boldsymbol{u}\cdot\frac{\partial f'}{\partial \boldsymbol{p}'}\right\rangle_{\phi'} = \left\{\frac{(1-\mu'^2)}{2}(\nabla\cdot\boldsymbol{u}) + \frac{(3\mu'^2-1)}{2}(\boldsymbol{n}\cdot\nabla)(\boldsymbol{n}\cdot\boldsymbol{u})\right\} p'\frac{\partial f'}{\partial p'}\,. \tag{3.29}$$

The dominant contributor to the right hand side of eq. (3.24) involves the Lorentz force $e(\boldsymbol{v}'\times\boldsymbol{B})$ acting on the particle gyrating in the magnetostatic field. However this vanishes after performing the average over gyrational phase. We are then left with the pitch angle scattering caused by the Alfvén waves which we express using the Fokker–Planck form, eq. (3.18). Note that there are no derivatives in momentum p' as the scattering is presumed to be elastic in the wave frame.

Collecting terms, the pitch angle-averaged transport equation becomes

$$\frac{\partial f'}{\partial t} + (\mu' v'\boldsymbol{n} + \boldsymbol{u})\cdot\nabla f' - \left\{\frac{(1-\mu'^2)}{2}(\nabla\cdot\boldsymbol{u}) + \frac{(3\mu'^2-1)}{2}(\boldsymbol{n}\cdot\nabla)(\boldsymbol{n}\cdot\boldsymbol{u})\right\} p'\frac{\partial f'}{\partial p'} + \frac{(1-\mu'^2)}{2}v'(\nabla\cdot\boldsymbol{n})\frac{\partial f'}{\partial \mu'} = \frac{\partial}{\partial\mu'}\left\{\frac{(1-\mu'^2)}{2}\nu'\frac{\partial f'}{\partial\mu'}\right\}. \tag{3.30}$$

The distribution function in a scattering medium will be approximately isotropic. We therefore expand f' in powers of the ratio of the mean free path to the scale length using (3.22),

$$f'(\boldsymbol{x}, p', \mu', t) = f'_0 + f'_1 + f'_2$$

where

$$f'_n = \mathrm{O}\left(\frac{v}{\nu}\frac{\partial}{\partial x}\right)^n f'_0 = \mathrm{O}\left(\frac{u}{v}\right)^n f'_0\,. \tag{3.31}$$

We substitute this expression in eq. (3.25). The lowest order equation is

$$\frac{\partial}{\partial\mu'}\left\{\frac{(1-\mu'^2)}{2}\nu'\frac{\partial f'_0}{\partial\mu'}\right\} = 0\,. \tag{3.32}$$

We then require that the distribution function be non-singular at $\mu' = \pm 1$ and so confirm that f'_0 is isotropic. To first order,

$$\mu' v' (\boldsymbol{n}\cdot\nabla) f_0 = \frac{\partial}{\partial \mu'} \left\{ \frac{(1-\mu'^2)}{2}\, \nu' \frac{\partial f_1'}{\partial \mu'} \right\}. \tag{3.33}$$

This equation can be integrated to obtain

$$\frac{\partial f_1'}{\partial \mu'} = -\frac{v'}{\nu'} (\boldsymbol{n}\cdot\nabla) f_0'. \tag{3.34}$$

Finally, we take the second order equation,

$$\frac{\partial f_0'}{\partial t} + (\boldsymbol{u}\cdot\nabla) f_0' + \mu' v' (\boldsymbol{n}\cdot\nabla) f_1' - \left\{ \frac{(1-\mu'^2)}{2} \nabla\cdot\boldsymbol{u} + \frac{(3\mu'^2-1)}{2} (\boldsymbol{n}\cdot\nabla)(\boldsymbol{n}\cdot\boldsymbol{u}) \right\} p' \frac{\partial f_0'}{\partial p'}$$
$$+ \frac{(1-\mu'^2)}{2} v' (\nabla\cdot\boldsymbol{n}) \frac{\partial f_1'}{\partial \mu'} = \frac{\partial}{\partial \mu'} \left\{ \frac{(1-\mu'^2)}{2}\, \nu' \frac{\partial f_2'}{\partial \mu'} \right\} \tag{3.35}$$

and average over μ'. We substitute eq. (3.34) to obtain

$$\partial f_0'/\partial t + (\boldsymbol{u}\cdot\nabla) f_0' - \nabla\cdot\{\boldsymbol{n} D_{\parallel} (\boldsymbol{n}\cdot\nabla) f_0'\} = \tfrac{1}{3} (\nabla\cdot\boldsymbol{u})\, p'\, \partial f_0'/\partial p' \tag{3.36}$$

where

$$D_{\parallel}(\boldsymbol{x}, \boldsymbol{p}', t) = \langle \tfrac{1}{2}(1-\mu'^2) v'^2/\nu' \rangle_{\mu'} \tag{3.37}$$

is the diffusion coefficient with the average being taken over μ' and we have used,

$$v' \langle \mu' f_1' \rangle_{\mu'} = \frac{v'}{2} \left\langle (1-\mu'^2) \frac{\partial f_1'}{\partial \mu'} \right\rangle_{\mu'} = -D_{\parallel} (\boldsymbol{n}\cdot\nabla) f_0'. \tag{3.38}$$

Note that it is not necessary to assume, as in some earlier treatments, that $f_1' \propto \mu'$.

Equation (3.36) is the transport equation for the isotropic part of the distribution function $f_0'(\boldsymbol{x}, p, t)$ and to the order to which we are working, it does not matter whether we measure it in the wave frame or in the inertial frame. We will henceforth suppress the prime and the zero subscript. The left hand side contains the expected convection and diffusion terms. It is also not necessary to assume that the particles are scattered through small angles and their angular distribution evolves according to the Fokker–Planck equation, as is true for our derivation. Essentially the same equation has been derived for photons interacting with electrons through large angle Thomson scattering [267]. The right hand side takes account of the energy-loss (gain) suffered by particles in an expanding (converging) flow.

We can rewrite eq. (3.36) as a conservation law

$$\frac{\partial f}{\partial t} + \nabla\cdot\{\boldsymbol{u} f - \boldsymbol{n} D_{\parallel} (\boldsymbol{n}\cdot\nabla) f\} + \frac{1}{p^2} \frac{\partial}{\partial p} p^2 \left\{ -\frac{p}{3} (\nabla\cdot\boldsymbol{u}) f \right\} = 0 \tag{3.39}$$

from which we isolate the speed of the particles in phase space.

We shall also need the mean flux of particles of a given momentum measured in the inertial frame. This can be evaluated directly from eqs. (3.23) and (3.38)

$$F(\boldsymbol{x}, p, t) = v \langle \mu\, f(\boldsymbol{x}, p, t)\rangle_\mu = v'\langle \mu'\, f'(\boldsymbol{x}, \boldsymbol{p}' = \boldsymbol{p} - E'\boldsymbol{u}, t)\rangle = -\boldsymbol{n}D_\parallel (\boldsymbol{n}\cdot\nabla) f - \tfrac{1}{3}\boldsymbol{u}p\,\partial f/\partial p\,. \tag{3.40}$$

The second term corrects for the so-called Compton–Getting effect, i.e., the differential Doppler shifting of the particle energy on transforming from the wave to the inertial frame. (Particles of speed v in the inertial frame moving in the direction of $\boldsymbol{u}$, have a smaller energy in the wave frame than those moving in the opposite direction, so in calculating the particle flux in the inertial frame, we must compare particles of slightly different energies in the wave frame.)

Equation (3.36) gives the transport equation in the form which we shall mostly need. However, it can be generalized to take account of the possibility that particles be scattered by Alfvén waves propagating in both directions along the field with corresponding scattering rates ν_+ and ν_- [117]. This means that there is no single frame in which the μ average can be performed. The final kinetic equation is given by

$$\frac{\partial f}{\partial t} + \left(\frac{1}{3p^2}\frac{\partial}{\partial p}(p^3\boldsymbol{w})\right)\cdot\nabla f - \nabla\cdot\{\boldsymbol{n}D_\parallel(\boldsymbol{n}\cdot\nabla) f\} = \tfrac{1}{3}(\nabla\cdot\boldsymbol{w})\, p\frac{\partial f}{\partial p} + \frac{1}{p^2}\frac{\partial}{\partial p}p^2 D_{pp}\frac{\partial f}{\partial p}$$

where

$$\begin{aligned}\boldsymbol{w} &= \boldsymbol{u} + \left\langle \tfrac{3}{2}(1-\mu^2)\frac{\nu_+ - \nu_-}{\nu_+ + \nu_-}\right\rangle_\mu v_A \boldsymbol{n}\\ D_{pp} &= 2\varepsilon^2 v_A^2 \left\langle (1-\mu^2)\frac{\nu_+ - \nu_-}{\nu_+ + \nu_-}\right\rangle_\mu .\end{aligned} \tag{3.41}$$

$\boldsymbol{w}$ is clearly a compromise velocity between those of the two wave frames and it coincides with the velocity of one wave frame if the oppositely directed Alfvén waves are absent. The second term on the right hand side describes Fermi acceleration by Alfvén waves and may be recovered by elementary arguments as in section 3.3. (We do not require that $v_A \ll u$ however.) Furthermore, if we keep some extra time-dependent terms in the derivation we can verify that eq. (3.36) is still valid when $\partial/\partial t = \mathrm{O}(v\,\partial/\partial x)$ when there are only a few scatterings per scale height. However it must be remembered that this approach can be correct only when $u \ll v$ as we must expand in powers of (u/v). Note that our approach neglects spatial diffusion perpendicular to the magnetic field by ignoring the variation of the distribution function with gyrational phase (cf. section 3.4).

3.6. The fluid limit

We can derive fluid equations from eq. (3.39). This is particularly simple when the motion is along the field. If we just integrate over momentum space, we obtain a conservation law for the particles. If we first multiply by the particle energy and then integrate over momentum space, we obtain

$$\partial W_c/\partial t + \partial F_c/\partial x = u\,\mathrm{d}P_c/\mathrm{d}x \tag{3.42}$$

where $F_c = u(W_c + P_c) + Q_c$ is the cosmic ray energy flux and where W_c, the energy density in the cosmic rays, is given by:

$$W_c = \int 4\pi mc^2 (1 + p^2/m^2c^2)^{1/2} fp^2\,\mathrm{d}p\,. \tag{3.43}$$

P_c, the cosmic ray pressure, is given by:

$$P_c = \frac{4\pi}{3} \int \frac{f p^4 \, dp}{m[1 + p^2/m^2c^2]^{1/2}} \tag{3.44}$$

and Q_c, the "heat flux", is given by:

$$Q_c = -\int 4\pi D_{\parallel} mc^2 \, [1 + p^2/m^2c^2]^{1/2} \, \frac{\partial f}{\partial x} \, p^2 \, dp \, . \tag{3.45}$$

In this subsection, x is measured along the direction of motion.

The pressure term in (3.42) is obtained by integration by parts assuming that $p^4 f$ vanishes as $p \to \infty$. Note that (3.42) is just the energy equation obeyed by any fluid when viscosity and inertia are neglected. By using eq. (3.39) together with the equations of mass and energy momentum conservation to study shock structure, one is essentially keeping information about the particle spectrum that is discarded by the fluid equations.

When higher order terms are included, terms resembling viscosity appear, which are conceptually useful in describing how thermal particles might be energized within a viscous subshock. Krymsky [119] has attempted such a non-relativistic derivation of viscosity from kinematic considerations, starting with eq. (3.19). Assuming $v \, du/\nu u \, dx \ll 1$, he obtains eq. (3.36) with an additional term on the r.h.s.

$$\frac{u^2}{3\nu} \left(\frac{du}{dx}\right)^2 \frac{1}{v^2} \frac{\partial}{\partial v} \left(v^2 \frac{\partial f}{\partial v}\right). \tag{3.46}$$

This resembles a second order acceleration term with an acceleration rate proportional to $(du/dx)^2$. This can be interpreted as being analogous to Ohmic dissipation with the inertial force $m \, du/dx$ that results from a particle wandering from one fluid element to another one with a different velocity. However, the shock thickness is expected to be of order one mean free path v/ν, and the thermal velocity is expected to be of order u, so the expansion parameters used to obtain a fluid picture may not be small near a shock. We must resort to a kinetic picture for quantitative calculations.

4. Test particle approximation

In this section we discuss the solution of (3.36) for a discontinuous velocity profile $u(x)$. This describes the acceleration process as seen by a population of test particles and is suitable for deriving the asymptotic particle spectrum at high energy.

4.1. Rankine–Hugoniot relations

Before we investigate the motion of cosmic rays at a shock front, we must first describe the fluid structure of the front. This is described in great detail in standard texts [70] and we will only summarize the properties that we need. A shock front is a surface of discontinuity across which there is a steady flow of mass, momentum and energy. Of course, no shock wave is truly steady or discontinuous and all that we require is that the distance over which the flow variables vary, the shock thickness, be much

smaller than the corresponding scales ahead of and behind the shock and that the overall flow pattern does not change substantially in the time it takes a fluid element to cross the shock. The fluid structure is determined by the requirement that the fluxes of mass, momentum and energy be continuous across the shock front. As material is actually crossing the shock front this tells us that the velocity parallel to the front be continuous. (This is untrue of hydromagnetic shocks however.) Only the velocity perpendicular to the front is relevant. (Equivalently, we can perform a Galilean transformation into the frame in which the flow is normal to the shock.)

Let the normal component of the fluid velocity (in the frame in which the shock is stationary) behind (ahead of) the shock be u_- (u_+) and let the density, internal energy per unit mass, pressure and enthalpy per unit mass be ρ, e, p and h respectively. Continuity of the fluxes of mass, momentum and energy then implies that

$$\begin{aligned} &\rho_- u_- = \rho_+ u_+ \\ &p_- + \rho_- u_-^2 = p_+ + \rho_+ u_+^2 \\ &h_- + \tfrac{1}{2}u_-^2 = h_+ + \tfrac{1}{2}u_+^2 . \end{aligned} \tag{4.1}$$

We eliminate the velocities to obtain the equation for the "shock adiabat"

$$h_+ - h_- = (p_+ - p_-)(\rho_- + \rho_+)/2\rho_-\rho_+ . \tag{4.2a}$$

Shock waves bring about a discontinuous increase in the entropy of the fluid which for normal substances implies that shocks are compressive and decelerate the flow velocity from being supersonic with respect to its internal sound speed to being subsonic. The decrease in the bulk kinetic energy is compensated by the increase in the enthalpy and the decrease in the bulk momentum flux, ρu^2 is taken up in the pressure.

If the fluid is a perfect gas with specific heat ratio γ, then $h = \gamma p/(\gamma - 1)\rho$ and the jump conditions take on a simpler form,

$$\frac{1}{r} = \frac{\gamma - 1}{\gamma + 1} + \frac{2}{(\gamma + 1)M^2} \tag{4.3a}$$

$$\frac{p_+}{p_-} = \frac{2\gamma M^2}{\gamma + 1} - \frac{\gamma - 1}{\gamma + 1} \tag{4.4}$$

where $r = u_-/u_+$ is the compression ratio and $M = (\rho_- u_-^2/\gamma p)^{1/2}$ is the pre-shock Mach number. Note that the compression ratio and pressure ratio tend to unity as $M \to 1$.

When energy is lost by radiation, heat flux to upstream infinity, or cosmic rays accelerated to such energies that they cease to interact with the rest of the flow, then eq. (4.2a) must be generalized to

$$h_- + \tfrac{1}{2}u_-^2 = h_+ + \tfrac{1}{2}u_+^2 + Q/\rho_- u_- \tag{4.2b}$$

where Q is the rate of energy loss per unit area. The compression ratio is then given by

$$\frac{1}{r} = \frac{1}{\gamma+1}\{\gamma + M^{-2} - [(1-M^{-2})^2 + 2(\gamma^2-1)Q/\rho_- u_-^3]^{1/2}\}\,. \tag{4.3b}$$

In many astrophysical shocks, magnetic fields and cosmic rays contribute substantially to the transport of momentum and energy and therefore alter these jump conditions. We ignore this in this section confining our attention to the test particle approximation. However, we shall need to consider the passive behavior of magnetic fields at a shock front. Astrophysical plasmas are generally believed to have such a high electrical conductivity that the electric field frame of the moving plasma can be ignored. In a frame in which the plasma moves with speed $\boldsymbol{u}$, this implies

$$\boldsymbol{E} = -\frac{\boldsymbol{u}}{c} \times \boldsymbol{B}\,. \tag{4.5}$$

The magnetic jump conditions then follow from the continuity of the parallel (referred to the shock normal) component of $\boldsymbol{B}$ and the perpendicular component of $\boldsymbol{E}$ at the shock front,

$$B_{-\parallel} = B_{+\parallel}\,, \qquad u_- B_{-\perp} = u_+ B_{+\perp}\,. \tag{4.6}$$

4.2. Scatter-free interaction

We now describe the behavior of a charged particle at a shock front. We first consider the electrodynamical motion of a charged particle of sufficiently large energy that its gyroradius exceeds the thickness of the shock front. If we ignore the possibility of scattering, then the motion across a plane shock front is most easily computed by transforming into a frame in which the fluid velocity is parallel to the magnetic field. (This transformation is possible only if $B_\parallel > B_\perp u/c$ in the frame in which the fluid moves in a direction perpendicular to the shock front.) The electric field vanishes in this frame and so the particle cannot change its energy although the pitch angle of its gyrational motion will change on crossing the shock. Transforming back to the original frame gives a relative change in the particle energy of only $O(u/c)$, regardless of how many times a plane shock is crossed. This is clearly not the basis of a general efficient acceleration mechanism.

There are three circumstances when the above analysis breaks down. Firstly when the shock velocity is nearly perpendicular to the field, an energetic particle must cross the shock several times changing its energy by a small fraction every gyration. The motion is approximately adiabatic and the invariant $p^2 \sin^2\theta/B$ (where θ is the pitch angle) will be approximately conserved. For an explicit demonstration in the case of a perpendicular shock see ref. [121]. For an analysis of integrations of the equation of motion that demonstrate that this is a surprisingly good approximation on average for nearly perpendicular shocks (see refs. [122, 123, 124]). Secondly when the shock front is curved, the electric field cannot be transformed away by a single global transformation and successive crossings of the shock front can lead to further particle acceleration. This is important at planetary bow shocks and interplanetary shock waves where the gyroradii of energetic particles can be a substantial fraction of the transverse size of the shock. Finally when the shock speed approaches the speed of light, we see that the relative gain in energy is O(1) for a single shock passage and the efficiency of particle acceleration can be quite high even in the absence of scattering [125]. For further discussion of the scatter-free case see [15] and references therein.

4.3. Steady-state solution with scattering

The acceleration process automatically becomes more efficient if the particles are scattered on both sides of the shock front. This must occur on the average if the shock is propagating within a region of space where the particles are confined, as we know to be true of the galaxy. In this case a typical particle will cross the shock $O(v/u)$ times increasing its energy by a fraction $O(u/v)$ on each crossing giving an average fractional increase in the energy O(1). Furthermore, this acceleration is a Fermi process (cf. section 3.3) and a power-law distribution function will automatically be generated at high energy. The slope of this distribution function is determined directly by the kinematics and not by some chance coincidence between the escape and the acceleration timescales.

Particle scattering in astrophysical shocks is probably effected by hydromagnetic disturbances that move through the fluid (cf. sections 3.4, 3.5). For the moment however we regard the scattering centers as being convected by the fluid and ignore the influence of the cosmic rays upon the flow (cf. section 6 below). If the scattering mean free path is small compared with both the longitudinal and the lateral scales of the flow then we can treat the spatial transport in the diffusion approximation. Similarly, if the particle speeds are large compared with the flow velocity then the fractional change in particle energy at the shock front is small and the momentum space distribution can be described by a distribution function as in section 3.

We use a spatial coordinate x in the frame in which a plane shock front is stationary and locate the shock at $x=0$ with the fluid moving in the direction of increasing x. We also introduce a spatial diffusion coefficient normal to the shock front, D. If diffusion is only effective parallel to a weak magnetic field that makes an angle η with the shock normal, then the effective diffusion coefficient is reduced by a factor $\cos^2\eta$ (unless as we discussed above $\theta \gtrsim \pi/2 - u/v$). We assume that both the flow field and the particle distribution function are stationary. The transport equation (3.36) for high energy particles then specialises to

$$u\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} D \frac{\partial f}{\partial x}$$

where

$$\begin{aligned} u &= u_- ; \qquad x<0 \\ u &= u_+ ; \qquad x>0 . \end{aligned} \tag{4.7}$$

This equation is straightforwardly integrated ahead of the shock where the convection opposes the diffusion to give

$$f(x,p) = f_-(p) + [f_0(p) - f_-(p)] \exp\left\{\int_0^x \frac{u_-\,\mathrm{d}x}{D(x,p)}\right\}; \qquad x<0 \tag{4.8}$$

with $f_-(p) = f(-\infty, p)$ and $f_0(p) = f(0, p)$. It is not possible to balance diffusion against convection behind the shock if $f(\infty, p)$ is to remain finite and the only possible solution in a steady state has f spatially constant,

$$f(x, p) = f_0(p); \qquad x > 0. \tag{4.9}$$

We must now join these two solutions at the shock front. There are two junction conditions. Firstly, the flux of particles crossing the shock either way is $O(fv)$ which is $O(v/u)$ times the net flux of particles in a steady state. So to $O(u/v)$, which is the order to which we have derived that transport equation, f must be isotropic and continuous [126]. Secondly, the flux of particles across the shock front at a given momentum must also be continuous. Using eq. (3.40) in the frame of the shock, we obtain

$$\left[-D\frac{\partial f}{\partial x} - \frac{u}{3}p\frac{\partial f}{\partial p}\right]_{0_-}^{0_+} = 0. \tag{4.10}$$

Equivalently, we can derive this condition by noting that the transport equation (3.36) can be rewritten in the form

$$\frac{\partial}{\partial x}\left\{-D\frac{\partial f}{\partial x} - u\frac{p}{3}\frac{\partial f}{\partial p}\right\} = -u\frac{\partial}{\partial x}\frac{\partial f p^3}{\partial p^3}. \tag{4.11}$$

The right hand side of this equation is finite, through the continuity of f and so the quantity in braces, which is the flux at a given momentum must be continuous at the shock front. (The junction condition (4.10) is actually only correct if the shock is one dimensional. If, for some reason, there is a gradient along the shock front (y-direction), then an additional gradient drift-term, $-(pcvB_2/3eB^2)(\partial f/\partial y)$ must be included [127]. This will be relevant when the diffusion length $\sim D/u$ becomes comparable with the size of the shock.)

We can now join the two solutions (4.8) and (4.9) at the shock using relation (4.10) to obtain a differential equation for the transmitted distribution function $f_+(p)$,

$$\frac{\mathrm{d}f_+}{\mathrm{d}\ln p} = \frac{3r}{r-1}(f_- - f_+) \tag{4.12}$$

of which the solution is

$$f_+(p) = qp^{-q}\int_0^p \mathrm{d}p'\, f_-(p')\, p'^{(q-1)} \tag{4.13}$$

where

$$q = 3r/(r-1). \tag{4.14}$$

We have demonstrated that particles of incident energy p' are Fermi-accelerated by the converging flow at the shock front to transmit a power law distribution function in relativistic momentum of logarithmic slope q [15, 128, 129, 130, 131, 132]. This result is independent of the functional form of $D(x, p)$ (save that it be positive). The power law is dictated by the kinematic structure of the shock front.

Now we emphasize that this result only holds as long as the test particle approximation is valid. As we argue in section 6 below, the acceleration of the majority of the cosmic rays must be calculated in a shock structure that depends non-linearly upon the cosmic rays themselves. The transmitted spectrum is

then dependent upon the shock parameters in a somewhat more elaborate fashion. However, the test particle approximation may well be appropriate at moderate to high particle energy and, as the test particle calculation provides a paradigm for the more general problem, we now present some alternative derivations that clarify its physical content.

Firstly, following Bell, consider the interaction of test particles with a shock front from the kinetic theory viewpoint [16, 130, 131, 137]. There is a steady flux of particles incident upon the shock from the upstream side. For particles of speed v the ratio of the net transmitted flux to this incident flux is given by

$$P = \frac{n_0 u_+}{n_0 \frac{1}{2}\int_0^1 \mathrm{d}\mu\, v\mu} = \frac{4u_-}{rv} \tag{4.15}$$

where n_0 is the total particle density at the shock front and μ now measures the angle between the velocity and the shock normal at the shock front. P can be regarded as the probability that a particle having once crossed the shock will not return. Now, when a particle crosses the shock, there will be a change of momentum associated with the change from the frame of the upstream scatterers to that of the downstream scatterers

$$\Delta p = \frac{p\mu(u_- - u_+)}{v} = \frac{p\mu u}{v} - (1 - r^{-1})\,. \tag{4.16}$$

The *flux-averaged* gain in momentum per crossing, can be computed from

$$\langle \Delta \ln p \rangle = \frac{u_-(1 - r^{-1})\int_0^1 \mathrm{d}\mu\, \mu^2}{v\int_0^1 \mathrm{d}\mu\, \mu} = \frac{2u_-(1 - r^{-1})}{3v}\,. \tag{4.17}$$

The total increase in $\ln p$ is related to the number of times the shock has been crossed. After n crossings the mean increase will be $n\langle \Delta \ln p \rangle$ and the dispersion about this value smaller by a factor $O(n^{-1/2})$. So if there are N particles remaining out of some original sample of particles of mass m injected at the shock front with a given momentum $p_0 \gg mu_-$ then we have that

$$-\frac{1}{N}\frac{\mathrm{d}N}{\mathrm{d}\ln p} = \frac{P}{2\langle \Delta \ln p \rangle} \tag{4.18}$$

where the factor 2 takes into account the fact that particles gain momentum from both forward and reverse shock crossings. Substituting for P and $\langle \Delta \ln p \rangle$ and integrating we compute the transmitted distribution function

$$f_+(p) \propto \frac{1}{p^2}\frac{\mathrm{d}N}{\mathrm{d}p} \propto p^{-3r/(r-1)} \tag{4.19}$$

in agreement with eq. (4.14).

A second derivation of the formula for the critical slope (eq. (4.14)) generalises an argument given by Michel [137]. All particles incident upon the shock must gain energy by varying amounts that depend upon the number of times they have crossed the shock front. The distribution function of the transmitted particles when the incident particles are more energetic must be a featureless power-law in momentum because (as long as $v \gg u$) there is no momentum scale in the problem except for the

momentum of the incident particle. (It might be thought that mc provides a scale and that the argument is only valid at subrelativistic particle energies. However, as is clear from the following argument, if we describe the particle by its relativistic momentum, the rest mass does not enter directly into the consideration and the power-law distribution function extends up to relativistic momenta. Of course, when we consider a realistic shock model and include the back reaction of the accelerated particles on the background fluid, the particle rest mass does become a crucial scale in the problem, cf. section 6.)

Now consider a volume V of fluid containing many particles with momentum p_0 incident upon the shock. Let those particles that have crossed the shock m times increase their momenta from p_0 to $p(m)$ again ignoring the fluctuations about the mean momentum gain in eq. (4.18). If they constitute a small fraction ε_m of all the incident particles they can be regarded as filling a volume $\varepsilon_m V$ ahead of the shock. Now, instead of letting these particles cross the shock, we could just as well have accelerated them from momentum p_0 to momentum p by compressing them adiabatically, as the shock is, by assumption, moving slowly compared with the particle speeds. They would then occupy a volume $(p_0/p)^3 \varepsilon_m V$ behind the shock. If we require that the total downstream volume occupied by all the particles be V/r where r is the compression ratio, we can take the continuum limit

$$\varepsilon \to 4\pi p^2 \,\mathrm{d}p\, f_+(p) \Big/ \int_{p_0}^{\infty} 4\pi p^2 \,\mathrm{d}p\, f_+(p) \tag{4.20}$$

to obtain the equation

$$\frac{\int_{p_0}^{\infty} 4\pi p^2 \,\mathrm{d}p\, V\, f_+(p)\,(p_0/p)^3}{\int_0^{\infty} 4\pi p^2 \,\mathrm{d}p\, f_+(p)} = \frac{V}{r} \tag{4.21}$$

where $f_+(p)$ is the transmitted distribution function. If we substitute a power law for $f_+(p)$ (i.e., $f_+(p) \propto p^{-q}$; $p \geqq p_0$), then we recover eq. (4.14).

What is remarkable about this argument is that it tells us that we could obtain the same distribution function as that transmitted by a shock, on the average, using reversible compressions. We could imagine a machine which compresses the incoming cosmic ray fluid in a periodic manner (by a factor $(t/P)^{1-r}$, where P is the period and t is the time, for $0 \leqq t < P$ and then repeating for successive periods). Provided that the period is sufficiently long that the accelerated particles cannot mix within an individual cycle, running the machine backwards will recover the original monoenergetic distribution at the cost of no net work. However the distribution function transmitted by a shock is not just a power law in momentum when averaged over a long time, but is a steady power law. In order to simulate this with our machine, we would have to reduce the period P so that the particles could mix effectively downstream. It is at this point that the entropy is introduced into the flow. The mixing is an irreversible process and corresponds to changing from the fine-grained to the coarse-grained distribution function (cf. [138, 139]). The entropy is introduced when the particles in the different partial volumes are mixed. (For related arguments to this, see [138, 139, 140].)

A third alternative derivation demonstrates that the acceleration mechanism does not depend upon the nature of the scattering ahead of the shock [78]. Consider a spherical stellar cloud that suddenly finds itself in a high pressure environment. A strong, spherical shock front will propagate radially inward with speed u relative to the enclosed stationary medium. The gas behind the shock must be strongly scattering although the gas ahead of the shock need not scatter at all. Now consider a particle

with momentum p trapped inside the contracting shock. It will bounce several times off the post-shock gas (moving with speed $u(1-1/r)$), each time increasing its momentum. Eventually, it will be transmitted downstream and will suffer no further acceleration. Let the volume enclosed by the shock when this happens be V. Now if the particle were to bounce off the shock, then the momentum with which it is transmitted would satisfy the familiar relation $p \propto V^{-1/3}$. However as it is scattered by the post-shock gas this relation must be amended to

$$p \propto V^{-(1-1/r)/3} . \tag{4.22}$$

Now let there be N particles with momentum in the interval Δp within the contracting shock. Particles must be transmitted at just the right rate to keep the phase space density of the remaining particles constant (as the shock moves slowly compared with the particles). So,

$$N \propto p^2 \Delta p V \propto p^3 V . \tag{4.23}$$

The distribution function of the transmitted particles then satisfies

$$\frac{1}{p^2}\frac{\mathrm{d}N}{\mathrm{d}p} \propto V \propto p^{-3r/(r-1)} \tag{4.24}$$

recovering our original result. This argument does not depend upon the assumption that the shock is spherical.

These alternative arguments also serve to demonstrate the generality of the acceleration mechanism for test particles. The crucial assumptions that we have used are that the fluid velocity changes relatively abruptly, that the scattering be elastic and random and that the distribution function be nearly isotropic and 3 dimensional. (In 1D and 2D, $q = r/(r-1)$ and $2r/(r-1)$ respectively.)

Now the total space density of particles behind the shock is increased by the compression ratio and so if we want to make a fair comparison between the pre- and post-shock distribution functions, we should expand the post-shock particle distribution function adiabatically back to the pre-shock density. The particles will lose energy and their resulting distribution function will be given by

$$f_{\mathrm{d}}(p) = f_+(r^{1/3}p) \tag{4.25}$$

$$= q\left(\frac{q}{q-3}\right)^{-q/3} p^{-q} \int_0^{(q/(q-3))^{1/3}p} \mathrm{d}p'\, f_-(p')\, p'^{q-1} . \tag{4.26}$$

(Other forms of decompression are possible [142], but provided that there is no energy dependence in the degree of cooling, the power-law spectrum should be preserved.) If eq. (4.26) applies and $r<4$, then the mean final momentum of the particles is increased by a factor $3r^{-1/3}/(4-r)$ which always exceeds unity for a compressive shock. If $r \geqq 4$ the momentum increase is fixed by the maximum value at which the scattering is efficient.

If we consider test particle acceleration at a shock of Mach number M, then combining eqs. (4.3) and (4.14) we see that the slope of the transmitted distribution function is given by

$$q = \frac{3(\gamma+1)M^2}{2(M^2-1)} = \frac{4M^2}{M^2-1} \tag{4.27}$$

where the second equality holds for $\gamma = \frac{5}{3}$, the usual assumption in astrophysical shocks.

We can also compute the pressure increase for the test particle distribution. If the incident particles have relativistic energy, then their pressure is increased by a factor $q/(q-4) = M^2$ for $\gamma = \frac{5}{3}$. For mildly relativistic particles, the pressure jump is in fact larger. This compares with the pressure increase of $(5M^2-1)/4$ for the thermal particles. This simple result already indicates that shock acceleration can be very efficient and that the test particle theory is probably inadequate.

The weak shock limit is also of interest. It is a well known result of shock theory [70] that the entropy jump at a weak shock is third order in the velocity jump, Δu. So on decompression back to the original density the mean energy per particle in the background medium is unchanged to $O(\Delta u^2)$. However, as long as the scattering remains elastic in the fluid frame, the mean energy of a cosmic ray increases to order Δu^2, in fact by fractional amounts $2\,\Delta u^2/9v_s^2$ and $5\,\Delta u^2/9v_s^2$, where v_s is the sound speed, for non- and ultra-relativistic particles respectively [132, 143]. This is another indication of the efficiency of the Fermi mechanism.

A rather more difficult case to analyze arises when the particle speed becomes comparable with the shock speed [144, 145]. The difficulty is that it is no longer correct to assume that the momentum distribution function is nearly isotropic in the local center of momentum frame of the fluid. It appears that if the downstream scattering is sufficiently efficient to maintain isotropy in the frame of the post-shock fluid then the acceleration becomes less efficient. This in turn implies that relativistic shocks will not be able to transmit a power law distribution function at high particle energy.

4.4. Time dependence

So far, we have restricted our attention to steady solutions. Astrophysical strong shocks can form suddenly (e.g. after a solar flare or when field lines suddenly become connected to a planetary bow shock) and so it is also important to consider time-dependent solutions to the transport equation. These can sometimes be derived using a Laplace transform in time. A variety of exact solutions have been given, e.g., [146] making simplifying though not entirely realistic assumptions concerning the diffusion coefficient and the injection. They confirm the scaling apparent from the equation that if (as expected) the mean free path ahead of the shock increases with energy and exceeds that behind the shock, the time to accelerate a particle up to momentum p is given approximately by $D(p)/u_-^2$. In fact as the diffusion coefficient will generally decrease away from the shock front the acceleration time will be somewhat longer than estimated using the diffusion coefficient at the shock front [15, 147, 148, 149, 150]. These calculations confirm that there is a genuine difficulty in accelerating cosmic rays as far as the "break" in the galactic spectrum at $\sim 10^{15}$ eV using supernova remnants. (This difficulty is alleviated though not entirely removed if these cosmic rays comprise iron nuclei.)

A different time-dependent problem of relevance to the acceleration of galactic cosmic rays is to determine the integrated cosmic ray spectrum accelerated by an expanding supernova blast wave of declining strength. Taking into account the finite acceleration time and the expansion losses, Bogdan and Völk [151, 152] find that the average spectrum is a slowly varying power law with slope $4.1 \lesssim q \lesssim 4.3$ as expected. (In order to solve this problem self-consistently, it is necessary to include the influence of the cosmic rays themselves on the dynamics of the remnant [132, 153].)

For shocks in the solar wind, the acceleration is also expected to be limited by the age of the shock

[154]. This is consistent with the observations that the intensities typically increase with distance from the sun and that, at least in the inner heliosphere, quasi-perpendicular shocks provide harder spectra than quasi-parallel ones. The latter trend can be understood inasmuch as the particles have much smaller diffusive mobility away from quasi-perpendicular shocks; hence their acceleration time is shorter than is the case for quasi-parallel shocks.

4.5. Escape

Another simplifying assumption that we have made is that the shock is a planar discontinuity in a uniform infinite flow. As we mentioned in section 2.1, this is a poor approximation at planetary bow shocks and at sufficiently large cosmic ray energies in the interstellar medium. We must therefore modify our transport equation to take account of non-diffusive escape away from the shock front. Specifically, particles will escape without further acceleration when the diffusion scale length $\sim D/u$ at a given energy ahead of the shock becomes comparable with the radius of curvature of the shock front.

In one simple model [155] (cf. also [156] for related ideas) of escape applicable to quasi-parallel shocks, it is assumed that non-relativistic particles random walk one gyroradius across field lines every collision mean free path. This implies that their field-parallel and field-perpendicular diffusion coefficients satisfy an approximate relation

$$D_\perp D_\parallel \simeq v^4/18\Omega^2 = \tfrac{2}{9}(Ec/ZeB)^2 \tag{4.28}$$

independent of the value of the mean free path (cf. section 3.4). Here, E is the particle energy, and Ze is the charge.

Now consider shock acceleration subject to the boundary condition at a square-shaped shock of side $2L$. Let us assume that the particles escape freely when $|z| > L$ or $|y| > L$. In this simple geometry, there is a $\cos(\pi z/2L)\cos(\pi y/2L)$ dependence in the solutions to f upstream and downstream of the shock. In the limit of high energy, where diffusion dominates convection, the differential equation for f is simply

$$-2D_\perp\left(\frac{\pi}{2L}\right)^2 f + D_\parallel \frac{\partial^2}{\partial x^2} f \approx 0\,, \tag{4.29}$$

whence

$$f \propto \mathrm{e}^{k_\parallel x}\,;$$

with

$$k_\parallel = \pm(2D_\perp/D_\parallel)^{1/2}(\pi/2L) \tag{4.30}$$

where the sign of k is chosen on either side of the shock to keep f finite as $x \to \pm\infty$.

The jump condition at the shock (eq. (4.10))

$$\tfrac{1}{3}(u_- - u_+)\, p\, \partial f/\partial p = D_\parallel[\partial f/\partial x|_{x=+\varepsilon} - \partial f/\partial x|_{x=-\varepsilon}] \tag{4.31}$$

then implies that at the shock,

$$\partial \ln f/\partial \ln p = -3\pi(2D_{\parallel}D_{\perp})^{1/2}/L(u_{-}-u_{+})\,. \tag{4.32}$$

Using relation (4.28), we see that the spectrum is an exponential in energy per charge $f \propto \exp[-(E/Z)/(E/Z)_0]$ where the e-folding energy per charge is given by

$$(E/Z)_0 = eBL(u_{-}-u_{+})/\pi c\,. \tag{4.33}$$

A more detailed analysis [155, 157] indicates that this form of the spectrum is valid even when convective transport is of comparable importance to diffusion. Reasonable parameters for L and B at the earth's bow shock give a value for $(E/Z)_0$ of ~20 kV, in good agreement with observations of radial events (cf. section 2.1).

The geometry of particle escape at the bow shock is surely more complicated because of the free streaming boundary sufficiently far upstream of the shock. This boundary is probably shaped like a nose-cone that protrudes ahead of the shock and connects to its flanks. The point is that its scale is set by the bow shock, and perpendicular diffusion limits the maximum energy of accelerated protons to a few times $(E/Ze)_0$.

For the earth's bow shock $(E/Ze)_0$ is roughly 20 kV; for a supernova remnant $(E/Ze)_0 \sim 10^5$ GV; for a strong extragalactic radio source $(E/Ze)_0 \sim 10^{11}$ GV. In fact these are upper bounds. When the field is not parallel to the shock normal, particles are convected laterally away from the shock (in the shock frame) with the field lines. Field line meandering may enhance this effect (cf. section 2.4). In fact when a particle is able to escape transversely in an acceleration time, its diffusion length ahead of the shock will be larger than the lateral diffusion length by a factor of order the ratio of the mean free path to the gyroradius. The magnetic field geometry may then make it even less likely for the particle to return to the shock front. The general argument can be used to explain why protons are not accelerated to relativistic energies at planetary bow shocks and why the highest energy cosmic rays cannot have been accelerated by supernova remnants.

An analytical solution incorporating momentum-dependent diffusion and escape that illustrates some of these ideas has been given in [158].

4.6. *Adiabatic losses*

When the shock propagates within a diverging flow, sufficiently energetic particles can be adiabatically decelerated concomitantly with the shock acceleration. Such is the case in corotating interaction regions in the solar wind. At finite distances from the shock, assuming spherical symmetry or 1D diffusion along a flux tube, the transport equation takes the form

$$-u\frac{\partial f}{\partial R} + \frac{1}{3R^2}\frac{\partial}{\partial R}(R^2u)\,p\frac{\partial f}{\partial p} + \frac{1}{R^2}\frac{\partial}{\partial R}\left(R^2D\frac{\partial f}{\partial R}\right) = 0\,. \tag{4.34}$$

For simplicity, assume that convection can be ignored (i.e., $|u\,\partial f/\partial R| \ll |(1/R^2)(R^2D\,\partial f/\partial R)|$), the particles are non-relativistic, u is constant upstream of the shock, and $D = K\upsilon R$. The jump condition at the shock is

$$-\frac{p}{3}\frac{\partial f}{\partial p}-\frac{D}{u}\frac{\partial f}{\partial R}=-\frac{p}{3r}\frac{\partial f}{\partial p}. \tag{4.35}$$

Equations (4.34) and (4.35) have a separable solution which remains finite at $R=0$. The asymptotic form upstream of the shock in the limit of large p is

$$f\propto R^{2/(r-1)}\exp-\{6Krv/u(r-1)^2\} \tag{4.36}$$

upstream of the shock. This, and more complex solutions, were given by Fisk and Lee [159]. The results, principally the exponential spectrum in velocity, are in good agreement with observations of energetic particles at corotating shocks in the solar wind (cf. section 2.2), given appropriate values for D, u and r. Similar solutions have been derived for models of stellar wind terminal shocks [16, 63].

In *convergent* flows, e.g., accretion onto a neutron star, there is adiabatic energy *gain*. Solutions to the analogous equation to (4.34) have been given [161].

4.7. Other losses

Escape is the simplest loss mechanism and it can be incorporated into the test particle theory by adding an energy-dependent term $-f/\tau_e(p)$ to the right hand side of the transport equation (3.10). Simple analytical models confirm the assertions of the previous subsection. However, escape is not the only loss process to which cosmic rays are susceptible. Non-relativistic particles are subject to ionization loss and relativistic electrons are subject to radiative loss through the bremsstrahlung and synchrotron processes and inverse Compton scattering. Catastrophic losses, e.g. nuclear collisions are incorporated in the same way as escape. Gradual loss processes are handled by adding a resistive Fokker–Planck term (cf. section 3.3). Again several analytic solutions have been given, usually requiring that the spatial diffusion coefficient be energy-independent. They confirm that acceleration is effective as long as the energy loss time exceeds the acceleration timescale at that energy $\sim D(E)/u^2$ [15].

Synchrotron loss is important in limiting the energies of shock-accelerated relativistic electrons in extragalactic radio sources. If we assume a steady state and a strong shock with compression ratio $r=4$, then the transport equation ahead of the shock becomes

$$u\frac{\partial f}{\partial x}-\frac{\partial}{\partial x}D\frac{\partial f}{\partial x}=\frac{1}{p^2}\frac{\partial}{\partial p}\left(\frac{p^4}{\tau_s m_e c}f\right) \tag{4.37}$$

where $\tau_s=6\pi m_e c/B^2\sigma_T$ is the synchrotron cooling time for a mildly relativistic electron. Shock acceleration can account for the X-ray emission that has been seen from a few extragalactic jets. Substituting $B\sim 10^{-5}$ G, $u\sim 0.1c$, $D\sim 10r_g c$ (cf. section 6), we find that it should be possible to accelerate electrons up to energies $\sim 3\times 10^5$ GeV so that they can emit hard X-ray synchrotron radiation. Another example of acceleration proceeding in the presence of a competitive loss process arises when an interstellar shock propagates through a neutral hydrogen cloud and ionization loss has to be included.

4.8. Grains and photons

Fermi acceleration at shock fronts may not be confined to electrons and ions. In fact it has been suggested that interstellar grains may also be accelerated efficiently and that this may provide an

efficient mechanism for injecting cosmic rays [75]. A small interstellar grain has a mass of $\sim 10^{-14}$ g a speed of $\sim 10\,\mathrm{km\,s^{-1}}$ and carries ~20 extra electrons. It therefore has the rigidity of a 10^5 GeV cosmic ray and can be scattered by resonant Alfvén waves in just the same way as an individual charged particle as long as it is traveling faster than the magnetosonic speed. As grains are accelerated, they will probably be broken up by their high speed motion through the ambient gas and as their charges are relatively insensitive to their sizes their rigidities will decrease making acceleration even more efficient. This provides a possible scheme for accounting for the observed over-abundance of galactic cosmic rays in the refractory elements. Unfortunately this mechanism is irrelevant to solar cosmic rays which exhibit the same pattern of abundances.

It is also possible to accelerate photons at a shock front. In this case, the scattering mechanism is Thomson scattering [161, 162, 163]. Photons can be accelerated more efficiently by the converging bulk flow on either side of a radiation-dominated shock than by the random motions of the thermal electrons when the shock speed significantly exceeds the electron thermal velocity. It is then possible to transmit steep power law radiation spectra though as the shocks themselves are necessary optically thick, these spectra do not necessarily correspond directly to what would be observed. Such shocks could conceivably be of importance in the early universe just prior to recombination and also in active galactic nuclei.

5. Wave spectrum

In this section, we consider the generation and interaction of the waves believed to be responsible for the scattering of the cosmic rays. As we will show in the following section, the efficiency of cosmic ray production depends upon the spatial diffusion coefficient, which in turn depends upon the scattering rate. The scattering rate is also important at high energy as it determines the maximum energy to which a particle can be accelerated (cf. section 4.6). There have been two approaches to determining the scattering rate. In the first section, we consider scattering in the weak turbulence limit, i.e., when the amplitudes of the scattering Alfvén waves are small and the quasi-linear coefficient of section 3.4 can be used. Strong turbulence is discussed in the following section.

5.1. Quasi-linear calculations

If we use eq. (3.11), and substitute into eq. (3.32) for the diffusion coefficient we obtain

$$D_{\parallel} = \frac{4v^2}{3\pi I(p)\,\Omega}$$

where

$$I(p) = [8\pi k \varepsilon_k / B^2]_{k=eB/pc} \tag{5.1}$$

is the dimensionless intensity of the resonant waves. The Alfvén waves may have a "standard" turbulence spectrum, unaffected by the presence of the cosmic rays. Two commonly proposed forms are the "Kolmogorov" ($\varepsilon_k \propto k^{-5/3}$) and the "Kraichnan" ($\varepsilon_k \propto k^{-3/2}$) spectra, which have been argued to exist in the ambient interstellar and interplanetary media [11]. We see that $D_{\parallel}$ increases with increasing particle momentum as $D_{\parallel} \propto vp^{1/3}$ and $D_{\parallel} \propto vp^{1/2}$ respectively.

However, shock acceleration can still operate even in the absence of externally produced Alfvén waves as these waves may be excited by the cosmic rays themselves [11, 105, 164]. In fact, when the streaming speed of the cosmic rays through the ambient medium exceeds the Alfvén speed, there is a quasi-linear instability in which the amplitude of resonant Alfvén waves grows exponentially with time. This condition is automatically satisfied in the shock acceleration mechanism. The waves are emitted by the gyrating particles as coherent cyclotron radiation in a medium with a high dielectric constant. We now compute the growth rate.

Consider a right-handed circularly polarized Alfvén wave propagating parallel to a uniform field $B_{0\hat{z}}$ with $\boldsymbol{B}_1 = B_1(\hat{\boldsymbol{x}} + \mathrm{i}\hat{\boldsymbol{y}}) \propto \exp\{\mathrm{i}(kz - \omega t)\}$. It is simplest to compute the current due to the resonant particles in the wave (primed) frame in which $\omega'_{\mathrm{R}} \equiv \mathrm{Re}\,\omega' = 0$. In this frame, the magnetic field is stationary and the electric field vanishes, except for a small perturbation associated with the wave growth, which will be assumed to be slow compared to other time scales in the problem. The linearized Vlasov equation (3.1) for the perturbation f'_1 to the particle distribution function f'_0 is

$$\frac{\partial f'_1}{\partial t} + \boldsymbol{v}' \cdot \frac{\partial f'_1}{\partial \boldsymbol{x}} + Ze(\boldsymbol{v}' \times \boldsymbol{B}_0)\frac{\partial f'_1}{\partial \boldsymbol{p}'} = -Ze(\boldsymbol{v}' \times \boldsymbol{B}'_1)\frac{\partial f'_0}{\partial \boldsymbol{p}'}\,. \tag{5.2}$$

(Note that we are expanding the distribution function in powers of the perturbing wave amplitude not in powers of (u/v) as in section 3. f_0 is independent of the gyrational phase ϕ of the velocity vectors measured from the $\hat{x}$ direction and so the only contribution to the right hand side is $-\Omega(1-\mu'^2)^{1/2}(\partial f'_0/\partial\mu')(-\mathrm{i}\,\mathrm{e}^{\mathrm{i}\phi})(B'_1/B_0)$. As there is no electric field in the wave frame (to the order of our approximation), there is no change in the particle energies in this frame; the particles are simply scattered in pitch angle. We solve (5.2) for $f'_1 = B'_1 = 0$ at $t = -\infty$. Equivalently, we assume that the frequency ω' has a small positive imaginary part $\mathrm{i}\omega'_{\mathrm{I}}$, and obtain

$$f'_1 = \frac{\Omega(1-\mu'^2)^{1/2}}{\mu' k v' - \Omega - \mathrm{i}\omega'_{\mathrm{I}}}\frac{\partial f'_0}{\partial \mu'}\,\mathrm{e}^{\mathrm{i}\phi}\,\frac{B'_1}{B_0}\,. \tag{5.3}$$

The growth factor $\exp(\omega_{\mathrm{I}} t)$ has been absorbed in B_1. We next compute the current due to f_1 perpendicular to both $\boldsymbol{B}_0$ and $\boldsymbol{B}_1$. We call this $j'_\perp$,

$$j'_\perp = -Ze\int \mathrm{d}^3p'\, f'_1 v'\,(1-\mu'^2)^{1/2}\sin\phi \tag{5.4}$$

$$= -\pi\mathrm{i}Ze\int \frac{\mathrm{d}\mu'\,(1-\mu'^2)\,\mathrm{d}p'\,p'^2 v'}{(\mu' k v' - \Omega - \mathrm{i}\omega'_{\mathrm{I}})}\,\frac{\partial f'_0}{\partial \mu'}\,\Omega\,\frac{B'_1}{B_0}\,. \tag{5.5}$$

When the growth rate is small, we can invoke the Plemelj formula [96] to isolate the resonant contribution to the current in phase with the wave magnetic field,

$$j'_{\mathrm{res}} = -\pi^2 Ze\int \mathrm{d}\mu'\,(1-\mu'^2)\,\mathrm{d}p'\,p'^2 v'\,\frac{\partial f'_0}{\partial \mu'}\,\Omega\,\delta(\mu' k v' - \Omega)\,\frac{B'_1}{B_0}\,. \tag{5.6}$$

Although there is no energy transfer between the waves and the particles in the wave frame, there is a momentum transfer or equivalently a force density, $j'_{\mathrm{res}} B'_1/c$. Next transform to the fluid frame in which the wave frequency has a real part $\omega_{\mathrm{R}} = k v_{\mathrm{ph}}$, and $B_1 = B'_1$ to the order to which we are working. (v_{ph}, the wave phase velocity, is generally, though not necessarily, equal to the Alfvén speed v_{A}.) In this frame the force is unchanged but the particles do work on the waves at a rate per unit volume

$v_{\rm ph} j'_{\rm res} B_1/c$. The total wave energy density (magnetic plus kinetic) is $B_1^2/4\pi$. From this we identify the wave growth rate $\sigma \equiv \omega_{\rm I}$ as

$$\sigma = \frac{1}{2}\left(\frac{v_{\rm ph} j'_{\rm res} B_1}{c}\right)\Big/\left(\frac{B_1^2}{4\pi}\right) \tag{5.7}$$

$$= 2\pi^3 \frac{Zev_{\rm ph}}{B_0 c} \int {\rm d}\mu' (1-\mu'^2)\, {\rm d}p'\, p'^2\, v' \frac{\partial f'_0}{\partial \mu'} \Omega\delta(\mu' k v' - \Omega)\,. \tag{5.8}$$

The factor 1/2 in eq. (5.7) is included because we are calculating the amplitude growth rate. The integral in eq. (5.8) must be carried out in the wave frame. The growth rate is positive if, in the wave frame, the particles stream in the opposite direction to the motion of the background fluid. Equivalently the waves grow if, in the fluid frame, the mean streaming velocity of the particles exceeds the wave velocity. In this case the growth rate is approximately

$$\sigma \simeq \Omega[v_{\rm drift}/v_{\rm ph} - 1] n_{\rm cr}/n_{\rm i} \tag{5.9}$$

where $n_{\rm cr}(\approx 4\pi f p^3)$ is an estimate of the total number of resonant cosmic rays and $n_{\rm i}$ is the background ion density.

If we now ignore the distinction between the wave frame and the fluid frame ($v_{\rm drift} \gg v_{\rm A}$), we can use eq. (3.34) to express the growth rate in terms of the spatial gradient (now measured for convenience in the fluid frame)

$$\sigma = +2\pi^3 \frac{Zev_{\rm ph}}{B_0 c} \int {\rm d}\mu\, (1-\mu^2)\, {\rm d}p\, p^2 \frac{v^2}{\nu(p,\mu)} \Omega \frac{\partial f}{\partial z} \delta(\mu k v - \Omega) \tag{5.10}$$

$$\simeq + \frac{2\pi^{5/2} Ze}{\rho^{1/2} c}\left[p^3 D_{\parallel} \frac{\partial f}{\partial z}\right]_{p=eB/kc}, \tag{5.11}$$

where we have used eq. (3.37) to approximate the integral.

Next, we apply this result to obtain the self-consistent intensity of Alfvén waves in the test particle approximation [130, 131]. For the remainder of this section (only), the shock normal is taken to be the $\hat{z}$ direction. The wave spectral intensity will be stationary and obey the equation

$$u_- \, {\rm d}I/{\rm d}z = 2\sigma I\,, \tag{5.12}$$

where $I \to 0$ as $z \to -\infty$.

If we now substitute the diffusion coefficient of eq. (5.1) in expression (5.11) for σ we find that we can solve directly for the diffusion coefficient

$$D(z,p) = \frac{\rho^{1/2} c}{4\pi^{5/2} Zep^3 (f - f_-)}\,. \tag{5.13}$$

We then solve for the distribution function in the precursor using eq. (4.7)

$$f = \frac{f_0 z_0 - f_- z}{z_0 - z}\,; \qquad z \leqq 0$$

where

$$z_0 = \frac{\rho^{1/2} c}{4\pi^{5/2} Zep^3 u_- (f_0 - f_-)} \tag{5.14}$$

is a scale length determined by the transmitted distribution function. The corresponding dimensionless wave intensity is

$$I = \frac{4}{3\pi} \frac{v^2}{\Omega u_- (z_0 - z)}; \qquad z \leqq 0. \tag{5.15}$$

We see that as long as the distribution function is steeper than $f \propto p^{-3}$ (as will surely be the case at high enough energy), then the diffusion coefficient, $D = u_-(z_0 - z)$, will increase with momentum for $|z| \lesssim z_0(p)$, but will then increase $\propto |z|$ for all momenta for which $|z| \gtrsim z_0(p)$. (In practice, the diffusion coefficient will be bounded above because of the presence of pre-existing Alfvén waves ahead of this shock.)

We have ignored damping processes in this analysis. Four important damping mechanisms have been considered.

a) If there is a significant fraction of neutral atoms and ion-neutral damping in the upstream plasma, then there will be a (negative) contribution to the total growth rate σ equal to the ion–neutral collision frequency [130].

b) The Alfvén wave can decay into a sound wave and another Alfvén wave [165]. Three-wave decay is forbidden unless the phase velocity of the Alfvén waves, v_{ph}, exceeds the sound velocity, v_s (i.e., $\beta < 1$). However, under conditions of rapid growth the resonances will be broadened and this selection rule can be relaxed. Moreover, if the Alfvén waves are driven to large amplitude, higher order decay processes and non-random phase correlation may be important [166]. The turbulence at the Earth's foreshock is observed to be highly compressive [167]; this may be a result of non-linear Alfvén wave decay. The generation of acoustic waves, which are Landau damped against electrons when $T_e/T_i \gg 1$, is a conceivable mechanism for electron heating in astrophysical shocks.

c) In general, linear transit-time damping by thermal ions will be very important for waves propagating at a finite angle θ to the magnetic field direction in a high β plasma [168, 169]. (Transit-time damping is Landau damping (i.e., $\omega = \boldsymbol{k} \cdot \boldsymbol{v}$) by the wave *magnetic* field acting on the magnetic moments of the resonant electrons.) A useful approximate formula for the slowest wave damping rate in a high β plasma is

$$\frac{\sigma_{TT}}{\Omega_i} = \frac{-\pi^{1/2} \kappa \theta^2}{2[1 + \kappa^2 \beta/8][1 + 4\pi\theta^4/\kappa^2]} \tag{5.16}$$

where $\kappa = kv\beta^{1/2}/\Omega_i$ [170]. Damping by electrons should be important at even smaller propagation angles. The existence of transit time damping provides a justification for the common approximation of only considering waves with $\boldsymbol{k} \| \boldsymbol{B}_0$, which do not possess a Landau damping ($n = 0$) resonance.

d) The fourth damping process (which we consider in section 5.2) is non-linear Landau damping.

Cosmic rays must also be scattered efficiently behind the shock front if the acceleration mechanism is to operate. If we assume that the wave modes have been generated ahead of the shock and convected by the background fluid across the shock, then they may be efficiently damped in the high β post-shock

fluid. This is because a backward propagating Alfvén wave ahead of the shock (with $\boldsymbol{k}\|\boldsymbol{B}$) will be transmitted in the form of up to seven oblique modes and all of these will be efficiently damped (unless $\boldsymbol{u}\|\boldsymbol{B}$) [170]. The use of non-linear transit time damping rates softens this conclusion and allows *relativistic* particle acceleration to proceed in the presence of downstream wave turbulence of low enough amplitude to justify the application of quasi-linear theory. However, the overall conclusion is that if *non-relativistic* particle acceleration is to occur at a shock front, then either the post-shock wave turbulence must be highly non-linear, or hydromagnetic waves must be freshly generated behind the shock (cf. section 5.2 below).

We have assumed up to now that the Alfvén Mach number M_A is very large and have mostly ignored the velocity difference between the wave frame and the fluid frame. However, the upstream waves are excited by the cosmic ray density gradient and propagate away from the shock along the field with speed $\Delta \boldsymbol{u}_- = -v_{A-}(\hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{B}})\hat{\boldsymbol{z}}$ independent of $\boldsymbol{k}$ relative to the fluid. The transmitted distribution function has a slope q which is determined by the velocity of the wave frame rather than the fluid frame. If the downstream waves are isotropized then the slope will be increased to [130]

$$q = \frac{3(u_- + \Delta u_-)}{u_- + \Delta u_- - u_+} = \left(\frac{3r}{r-1}\right)\left\{1 + \frac{v_{A-}\cos\theta_-}{(r-1)u_-}\right\} \tag{5.17}$$

where we have expanded to $O(M_A^{-1})$. Note that if we wish to use eq. (5.17) to $O(M_A^{-2})$ we must include magnetic corrections to the Rankine–Hugoniot relations [209] when deriving the compression ratio r. It is probably more realistic to assume that the backward propagating waves ahead of the shock are transmitted into backward and forward propagating Alfvén waves behind the shock [172]. The effective scattering velocity behind the shock is the intensity-weighted mean of the backward and forward wave frames (cf. eq. (3.37)), and can be simply computed when the magnetic field is parallel to the shock normal. The result is, somewhat surprisingly, that the spectrum will *flatten*, i.e.,

$$q = \left(\frac{3r}{r-1}\right)\left\{1 - \frac{v_{A-}}{(r+1)u_-}\right\} \tag{5.18}$$

(ref. [170]). When $\theta_{Bn} \neq 0$, the downstream waves will comprise a mixture of magnetosonic and oblique Alfvén waves.

Yet another complication arises because a large cosmic ray density can increase the wave phase velocity. When σ is comparable to ω_R, it becomes significant in the real part of eq. (5.5) [173, 174, 175]. When the influence of the cosmic rays can be treated in the linear approximation, $|v_{ph} - v_A| \ll |v_A|$, the cosmic rays affect the quasi-linear phase velocities of the oppositely polarized modes by equal and opposite amounts and this correction can be ignored. However when the cosmic ray energy density is large enough to cause the phase velocity to depart significantly from the Alfvén speed the effect can be large and the acceleration will presumably be inhibited.

There is an additional subtlety in the interaction between cosmic rays and Alfvén waves. When a cosmic ray in the fluid frame is scattered by creating Alfvén waves, it loses momentum parallel to the field $\Delta p_\|$ and energy $\Delta p_\| v_{ph}$. (See the discussion following eq. (5.8).) Therefore it loses energy to momentum in the ratio v_{ph}. However, if we allow an Alfvén wave, viewed in the frame of the fluid to damp (e.g. by ion-neutral damping) then we find that the ratio of the energy lost by the wave to the momentum imparted to the background fluid during the damping is $2v_{ph}$, twice the ratio required to create the wave.

An easy way to see this is to consider a spatially damped transverse Alfvén wave with magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{x}}\cos(kz - \omega t)\,\mathrm{e}^{-\alpha z}$; $z \geqq 0$. The total force exerted on unit area by the damping wave for $z \geqq 0$ is then $\int \mathrm{d}z\,\langle(\nabla \times \boldsymbol{B}) \times \boldsymbol{B}\rangle/4\pi$, where $\langle\ \rangle$ denotes a local time average. This is given by $B^2/16\pi$. The average Poynting flux at $z = 0$, $B^2 v_{\mathrm{ph}}/8\pi$, is then equal to a factor $2v_{\mathrm{ph}}$ times the force per unit area proving the result. (Yet another way to consider the momentum flux carried by Alfvén waves is to regard the waves as comprising a fluid. The pressure in the wave fluid in the z direction, $P_{\mathrm{w}} = \langle \delta B^2/8\pi \rangle$, is half the sum of the magnetic energy density and the kinetic energy density $W_{\mathrm{w}} = \langle \delta B^2/8\pi \rangle + \frac{1}{2}\langle \rho\, \delta v^2 \rangle$, i.e., the total wave energy density [176, 177].)

We therefore conclude that half of the momentum lost by a cosmic ray when it creates an Alfvén wave is deposited in the background medium and a direct computation of the ponderomotive force confirms that this is the case [178]. So when streaming cosmic rays generate Alfvén waves, they must also exert a direct force on the background fluid.

We can use these ideas to write down the equations that describe energy conservation ahead of the shock. In the frame of the shock in which the fluid moves with speed u, the total wave energy flux towards the shock can be expressed as the sum of an electromagnetic Poynting flux, $W_{\mathrm{w}}(u - v_{\mathrm{ph}})$ plus an extra contribution given by the kinetic energy flux of the background fluid, $\frac{1}{2}\langle \rho\, \delta v^2 \rangle u = \langle \delta B^2/8\pi \rangle\, u = P_{\mathrm{w}} u$. Hence

$$F_{\mathrm{w}} = u\,(W_{\mathrm{w}} + P_{\mathrm{w}}) - v_{\mathrm{ph}} W_{\mathrm{w}} \tag{5.19}$$

(refs. [179, 180]).

The cosmic ray energy flux (eq. (3.42)) must also be modified to take account of the motion of the scattering centers. The cosmic rays can be treated as a fluid moving with speed $(u - v_{\mathrm{ph}})$ and so

$$F_{\mathrm{c}} = (u - v_{\mathrm{ph}})(W_{\mathrm{c}} + P_{\mathrm{c}}) + Q_{\mathrm{c}} \tag{5.20}$$

where Q_{c} is the heat flux (eq. (3.45)). Work can be done on the cosmic ray fluid. Evaluating this in the frame of the shock where the flow is, by assumption, stationary, we obtain

$$\mathrm{d}F_{\mathrm{c}}/\mathrm{d}z = (u - v_{\mathrm{ph}})\,\mathrm{d}P_{\mathrm{c}}/\mathrm{d}z \tag{5.21}$$

(e.g., [70]). Again working in the shock frame, the wave pressure gradient will do work at a rate $-u\,\mathrm{d}P_{\mathrm{w}}/\mathrm{d}z$ on the background fluid. Now the cosmic rays can also do work on the waves. However, the waves grow purely temporally in the frame of the background fluid. So working in this frame we can write the growth rate of wave energy density as

$$\int \mathrm{d}k\, 2\sigma \mathscr{E}_k = v_{\mathrm{ph}}\,\mathrm{d}P_{\mathrm{c}}/\mathrm{d}z\,. \tag{5.22}$$

(Equation (5.22) may be verified directly by substituting eq. (3.17) for the pitch angle scattering rate into eq. (5.10) for the wave growth rate and integrating over k.)

Finally there may be extra wave damping (e.g., due to ion neutral friction) at a rate L. Combining these terms we obtain an equation generalizing equation (5.12) for the energy flux of the waves [177]

$$\mathrm{d}F_{\mathrm{w}}/\mathrm{d}z = u\,\mathrm{d}P_{\mathrm{w}}/\mathrm{d}z + v_{\mathrm{ph}}\,\mathrm{d}P_{\mathrm{c}}/\mathrm{d}z - L\,. \tag{5.23}$$

Equations (5.19)–(5.23) must be solved if we want to describe the combined thermal particle, cosmic ray and Alfvén wave structure of a shock wave.

Integrating eq. (5.23) from $z = -\infty$ to the shock front and ignoring wave damping and any changes in u, v_{ph} we find that the fractional wave amplitude $(\delta B/B_0)$ will grow to $(P_c/\rho v_{A-} u_-)^{1/2}$ at the shock front. Under these conditions, the waves become non-linear before the cosmic rays are able to decelerate the background fluid appreciably. In fact, if a fraction M_A^{-1} of the momentum flux or more is converted into cosmic rays, where M_A is the Alfvén–Mach number of the shock, then the waves must be efficiently damped, if we wish to use quasi-linear theory [10, 180, 181].

In fact, eq. (5.22), which involves integrals over p, is approximately correct for cosmic rays of momentum $\sim p$ (for each p) and their resonant waves, as may be verified using eq. (5.19). We can therefore use it to determine which waves in the spectrum go non-linear. To give an example let us consider the galactic cosmic ray spectrum. We have argued on observational grounds in section 2 that the source spectrum is $f(p) \propto p^{-4.2}$ and that the total cosmic ray pressure may be a large fraction (30%) of the total momentum flux. However this cosmic ray pressure is distributed over many octaves of momentum and the pressure associated with a single octave can still be quite small ($\lesssim 0.03\rho u^2$). It is this pressure that is appropriate for estimating the wave amplitude. On this basis, we estimate that quasi-linear theory should be adequate up to $M_A \sim 30$.

5.2. Non-linear calculations

The foregoing discussion strongly suggests that if supernova shock waves are to be efficient accelerators of galactic and interplanetary cosmic rays, then non-linear Alfvén wave interactions must be included. Two approaches, not so different in their conclusions, have been pursued. Either non-linear Landau damping saturates the waves at a (marginally) linear amplitude with all the work done by the cosmic rays on the scatterers being channeled into the internal energy of the fluid, or the magnetic field is regarded as completely chaotic and the particle transport modeled with an appropriate ansatz. We consider the former possibility here and the latter in section 6.

Non-linear Landau damping coefficients have been calculated by several authors [182, 183, 184, 185, 186]. The Alfvén waves ahead of the shock will have a range of k-vectors and both senses of circular polarization. Pairs of waves propagating with different frequencies ω_1, ω_2 and amplitudes B_1, B_2 [with the waves of right hand (left hand) polarization having positive (negative) frequency] produce a non-linear longitudinal electric field and magnetic field gradient. This is in contrast to a pure circularly polarized mode. This beat wave can be damped by the thermal ions at a rate

$$\sigma_{NLTT} \simeq -\tfrac{1}{2}\sqrt{\pi/2}\, k_1 v_A \beta^{1/2} B_2^2/B_0^2 \tag{5.24}$$

when the ion thermal speed v_i exceeds the Alfvén speed. The net effect is to transfer wave energy from short wavelengths to longer wavelengths. The damping will be inhibited by trapping effects if the oscillation period of an electron in the finite potential well created by the beat wave becomes shorter than the correlation time for waves. Nevertheless, this process should be important for waves at the large wave amplitudes anticipated.

A convenient approximation is that the wave amplitude simply saturates at some dimensionless intensity $8\pi k \mathscr{E}_k/B_0^2 = \alpha \simeq 0.1$ independent of k as a consequence of non-linear Landau damping and higher order processes. In this case, we can just bypass the equation for the wave energy density and modify eq. (5.21) to express the transfer of energy from the cosmic rays to the gas [179], i.e.,

$$\frac{u\rho^{\gamma}}{\gamma-1}\frac{\mathrm{d}}{\mathrm{d}z}(P_{\mathrm{g}}/\rho^{\gamma}) \simeq v_{\mathrm{ph}}\frac{\mathrm{d}P_{\mathrm{c}}}{\mathrm{d}z}\,. \tag{5.25}$$

(Remember that the relative velocity between the scatterers and the fluid, v_{ph}, may well exceed the formal Alfvén speed v_{A}.)

5.3. *Non-resonant growth*

Alfvén waves can be driven unstable by a pressure anisotropy, even in the absence of cyclotron-resonant particles. Consider, for example, the dispersion relationship of Alfvén waves in a completely cold plasma, with half the ions moving parallel to B_0 at speed V and the other half moving at $-V$ relative to the electrons, then in the frame of the electrons, the dispersion relation can be written down straight away as

$$\begin{aligned} -c^2k^2 &= \tfrac{1}{2}\omega_{\mathrm{pi}}^2\left[\frac{\omega+kV}{(\omega+kV+\Omega_{\mathrm{i}})}+\frac{\omega-kV}{(\omega-kV+\Omega_{\mathrm{i}})}\right]+\omega_{\mathrm{pe}}^2\frac{\omega}{\omega-\Omega_{\mathrm{e}}}-\omega^2 \\ &= \omega_{\mathrm{pi}}^2\left[\frac{\omega^2-k^2V^2+\Omega_{\mathrm{i}}\omega}{(\omega+\Omega_{\mathrm{i}})^2-k^2V^2}-\frac{\omega}{\Omega_{\mathrm{i}}-(m_{\mathrm{e}}\omega/m_{\mathrm{i}})}\right] \end{aligned} \tag{5.26}$$

where

$$\omega_{\mathrm{p}\alpha}^2 = 4\pi ne^2/m_{\alpha}\,, \qquad \Omega_{\alpha} = eB_0/m_{\alpha}c \tag{5.27}$$

(e.g. [97]). In the long-wavelength limit, $kV \ll \Omega_{\mathrm{i}}$, and neglecting various small terms, (5.26) reduces to

$$\omega^2 = k^2(v_{\mathrm{A}}^2 - V^2) \tag{5.28}$$

implying instability when $V > v_{\mathrm{A}}$.

More generally, it can be shown (e.g. [187]) that the dispersion relationship in the long wavelength limit is

$$\rho\omega^2/k^2 = B_0^2/4\pi + P_{\perp} - P_{\parallel} \tag{5.29}$$

so that instability results when the parallel pressure $P_{\parallel}$ exceeds the sum of the perpendicular pressure $P_{\perp}$ and $B_0^2/4\pi$. This condition is always satisfied in a high Alfvén Mach number viscous subshock (see section 6.3).

In the opposite limit, $kv_{\mathrm{A}} \gg \Omega_{\mathrm{i}}$, the ion contribution to the dielectric becomes small compared to the electrons, and the dispersion relationship is the usual one for stable whistler waves

$$\omega = \frac{\Omega_{\mathrm{e}}c^2k^2}{\omega_{\mathrm{pe}}^2} = \Omega_{\mathrm{i}}\frac{c^2}{\omega_{\mathrm{pi}}^2}k^2\,. \tag{5.30}$$

Thus, the fastest growing unstable modes satisfy $\Omega_{\mathrm{i}}/V \lesssim k \lesssim \Omega_{\mathrm{i}}/v_{\mathrm{A}}$, and have wavelengths similar to the gyroradius of the thermal ions, or possibly as small as the ion inertial length c/ω_{pi}. The firehose

instability together with the scattering due to the hydromagnetic turbulence it generates may well provide the viscosity for low Mach number quasi-parallel-shocks [188]. We discuss this further in the following section.

6. Non-linear theory, structure of collisionless shock waves, and injection

The question of how particles qualify to become cosmic rays, historically termed the "injection" problem, is crucial to understanding the conditions that are necessary to produce cosmic rays, the efficiency with which they are produced, and their composition. The problem, in fact, consists of two distinct questions: a) On which particles does a particular acceleration mechanism operate? and b) Of those particles, how is the energy distributed among them? Only a very small fraction of particles can be elevated to energies greatly in excess of thermal. The fact that cosmic rays are routinely produced in nature so efficiently suggests that the acceleration mechanism generally has more than enough particles to work with, but manages to concentrate much energy into very few of them. This happens automatically if the high energy particles absorb enough energy to control the injection plutocratically and prevent too much energy being democratically shared by the thermal particles. In this case, the acceleration mechanism need not be intrinsically selective; even if it operates on most or all of the particles, a cosmic ray population (as opposed to bulk heating) results.

We have argued in section 2 on observational grounds and in section 4 on theoretical grounds that the cosmic ray pressure is probably sufficient to influence the background gas ahead of the subshock. This suggests a simple realization of the general picture just described [189]. Suppose that a large number of particles in the shocked plasma are injected at a thin subshock and are subject to shock acceleration. Given some spread in particle energies, higher energy particles, having a greater diffusion coefficient diffuse further upstream, mediate the compression felt by less energetic ones, and thereby absorb more energy per particle. In the resulting equilibrium, a dynamically significant population of cosmic rays mediates the flow, allowing just enough acceleration of thermal particles (i.e., injection) to replace the cosmic rays that are eventually transmitted downstream. This equilibrium should be stable. If the injection rate is increased, the Mach number, i.e., shock strength, felt by low energy particles at the subshock will be reduced, thereby reducing the injection rate and vice versa. Now there are two processes that decrease the Mach number in the precursor and both must be considered. One is the slowing down of the incoming gas by the action of the cosmic ray pressure gradient. The second is the entropy increase of the gas through damping of cosmic-ray-generated waves. These processes are considered in the following section. In subsections 6.1 and 6.2, we describe various models of strong shock transitions that are mediated by cosmic rays. Models of the subshock are described in subsection 6.3 and in subsection 6.4 we discuss the constraints on injection scenarios imposed by observations of cosmic ray composition.

6.1. Shock mediation by energetic particles

When we described test particle acceleration in section 4, we ignored the possible exchange of momentum between the cosmic rays and the background fluid. We can allow for this by supplementing eq. (3.36) with equations expressing the conservation of mass and momentum,

$$\rho u = C_1 \tag{6.1}$$

$$P + \rho u^2 = C_2 \tag{6.2}$$

$$\frac{-\partial}{\partial x} u_s f + \frac{\partial}{\partial x} D \frac{\partial f}{\partial x} + \frac{1}{3} \frac{\mathrm{d}u_s}{\mathrm{d}x} \frac{1}{p^2} \frac{\partial}{\partial p} p^3 f = 0 . \tag{6.3}$$

Equation (6.3) is often written in terms of $F \equiv 4\pi p^3 f(\mathrm{d} \ln p / \mathrm{d} \ln E)$, with E replacing p as the independent variable.

The last equation here is the steady state version of (3.36), but with the velocity of the scattering centers, u_s, explicitly distinguished from the fluid velocity that appears in the first two. The total pressure P can include the pressure of energetic particles, P_c, the pressure in gas, P_g, and the pressure due to waves P_w. In all treatments, P_c is assumed to be continuous across the viscous subshock, while P_g will have a discontinuity there. Thermal conduction in the background gas is ignored as the thermal particles are assumed to be much less mobile than the energetic ones. In early treatments, u was equated with u_s, and the wave pressure P_w was neglected, but recent treatments have succeeded in including the waves self-consistently given their phase velocity and damping rate.

We can learn quite a lot about mediation from a two fluid model treating the cosmic rays as a distinct fluid and imposing energy and momentum conservation across both the gas subshock and the total shock incorporating both the cosmic ray precursor [16, 120, 128, 191, 192] and an oblique magnetic field [193]. Now this is only a self-consistent procedure when the shock can be regarded as stationary and the cosmic rays as distinct from the background particles. A clear distinction requires that the slope of the non-relativistic cosmic ray spectrum satisfies $\partial \ln f / \partial \ln p > -5$ so that the partial pressure of the particles whose energy lies in an octave of energy, E, increases with E. However, the slope of the distribution function cannot be too flat, because the acceleration time will generally also increase with energy and must ultimately become comparable with either the age of the shock or at the time for gas to flow past it. We cannot expect energy and momentum to be conserved across the shock unless the cosmic ray pressure is mainly contributed by particles of some intermediate energy. If the high energy particles are relativistic, then we also require $\partial \ln f / \partial \ln p < -4$ as $p \to \infty$.

With these restrictions in mind, the equations of momentum and energy conservation across the total shock can be written down using eqs. (4.1) supplemented with the cosmic ray presure and enthalpy density. We can write down a separate pair of conservation equations (4.3), (4.4) for the gas crossing the viscous subshock. The cosmic rays do not enter here as their pressure and energy density are continuous on the scale of the subshock. Now if the gas in the precursor is presumed to be adiabatic ahead of the shock, then

$$P_g \propto u^{-\gamma} \tag{6.4}$$

where the specific heat ratio γ presumably equals 5/3.

Equations (6.2)–(6.4) can be solved algebraically for the post shock conditions (gas and cosmic ray pressure, velocity, etc.) in terms of the pre-shock conditions provided that we specify the ratio of the cosmic ray pressure to the energy density both ahead of and behind the shock. This ratio decreases from 2/3 to 1/3 as the cosmic rays become increasingly relativistic. The correct value to use depends upon the microscopic details of the diffusion coefficient and the particle injection rate and it turns out that the post shock conditions are extremely sensitive to the ratio that is assumed [194, 195]. Therefore we cannot use fluid models alone to estimate the post shock cosmic ray pressure. Nevertheless these fluid models do demonstrate that it is possible for shocks to accelerate cosmic rays with high efficiency (see fig. 6).

Time-dependent numerical calculations of a fluid shock model have been carried out [196] with a

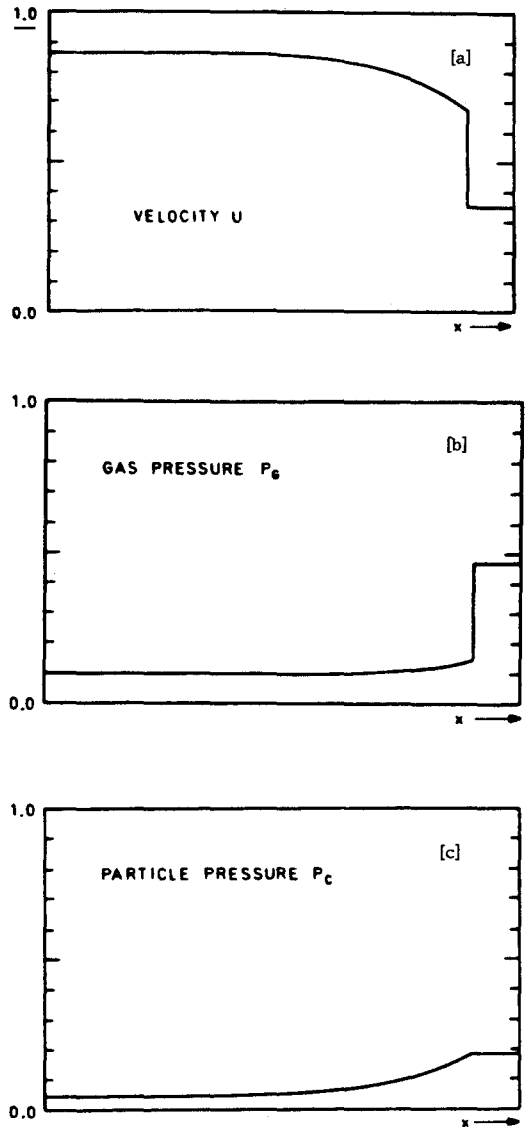

Fig. 6. Shock structure for a simple fluid model of a cosmic ray mediated shock, adapted from [192]. The Mach number is 2 and the cosmic rays are assumed to have an upstream pressure 0.3 times the total pressure. a) The speed of the background fluid is reduced by the pressure gradient of the cosmic rays ahead of the shock. It changes discontinuously at the subshock. b) The cosmic ray pressure builds up as the subshock is approached. c) The gas pressure also increases due to adiabatic compression in the decelerating flow.

view to clarifying their role in supernova remnant dynamics. They also seem to demonstrate the presence of instabilities as is also suggested by a WKB analysis of a planar shock [197]. There appear to be growing modes present when the cosmic ray energy density exceeds $\sim \rho_- u_- v_{A-}$. The physical nature of these instabilities is as yet unclear and requires further analysis.

It is possible for the cosmic rays to mediate the shock entirely. If the post shock cosmic ray pressure is large enough, then there is no necessity for a subshock and indeed for some choices of diffusion coefficient, the subshock must vanish. This state of affairs is analogous to that found in interstellar

shock waves propagating through a partially ionized gas endowed with a transverse magnetic field. Draine [266] has introduced the notation J-type to describe a shock in which the neutral fluid undergoes a discontinuous compression and C-type to describe a shock in which such a jump is absent. Furthermore, there can be more than one self-consistent solution for given upstream conditions although these are generally not all stable [192]. If a subshock is absent, we require that the downstream flow be supersonic with respect to the gas Mach number M_g and subsonic with respect to the total Mach number M. The first condition follows directly from the momentum conservation equation which can be written in the form $dP_c/du = C_1(1/M_g^2 - 1)$. If the cosmic ray pressure P_c is to increase monotonically through the shock as the flow decelerates, then $M_g > 1$. The second condition follows from the requirement that the shock front be stable to small perturbations in the upstream quantities [198].

An analytical approach to the non-linear problem is possible if we use an energy-independent diffusion coefficient D. This allows us to integrate the fluid equations to solve for the shock structure as long as we also specify the ratio of the cosmic ray pressure to energy density and keep this constant through the shock [191, 192, 199]. In the simple case of vanishing gas pressure, when the velocity transition is totally mediated by the cosmic rays, the velocity has a hyperbolic tangent profile

$$u(x) = \tfrac{1}{2}(u_- + u_+) - \tfrac{1}{2}(u_- - u_+)\tanh(x/L) \tag{6.5}$$

where the total compression u_-/u_+ is given by the shock jump conditions and

$$L = \frac{2DM_c^2}{u_-(M_c^2 - 1)} \tag{6.6}$$

is a measure of the shock thickness. (M_c is the Mach number of the shock with respect to the cosmic ray pressure.)

We can now solve eq. (3.36) by separation of variables for the cosmic ray distribution function in the converging flow given by eq. (6.5). The solution can be expressed in terms of hypergeometric functions. The important feature is that it is a power law in momentum at high energy (i.e., relativistic particles) with logarithmic slope given by

$$q = 4M_c^2(M_c^2 - \tfrac{1}{8})/(M_c^2 - 1^2)^2 . \tag{6.7}$$

This tends to limit $q \to 4$ as the shock becomes stronger and the particles are accelerated to greater and greater relativistic energies. The physical interpretation of this limit is that if there is no cut-off in the particle spectrum, the energy density carried by relativistic particles diverges as $q \to 4$ enabling the shock to become stronger without requiring the thermal particles to carry most of the internal energy behind the shock. This calculation verifies that, for an energy-independent D, a gas subshock is not essential for transmitting a power law distribution function at high energy. The existence of a steady state solution depends on the diffusion coefficient being bounded from above, so that particles of even the highest energy feel the mediation due to lower energy ones. Otherwise, the spectrum hardens indefinitely with time. The transmitted distribution function can generally be calculated by separation of variables when the diffusion coefficient is energy-independent and when the velocity field is externally specified [200].

When a high energy cutoff is present, so that the integral over momentum extends only to a finite

upper limit, completely separable solutions are obtainable [201]. We introduce the integral energy spectrum $F(E)$

$$F(E) = 4\pi p^2 f(p)\, \mathrm{d}p/\mathrm{d}\ln E\,. \tag{6.8}$$

$F(E)$ is a measure of the energy contained in a decade of energy and is everywhere proportional to $E^{-\alpha}$, and the velocity profile is as in (6.5) but with the compression ratio increased to

$$u_+/u_- = 1 - (1 + j\alpha/6)^{-1} \tag{6.9}$$

where $j = \mathrm{d}\ln E/\mathrm{d}\ln p$, which is assumed to be constant for this calculation. The calculation shows that the compression ratio can be anywhere between the test particle value and infinity if we allow the high energy particles to escape. This result is obtained no matter how high the cutoff is, as long as it is at finite energy, for, inasmuch as energy is lost from the flow by accelerating particles through the cutoff, the compression ratio is increased and the spectrum is flattened into a divergent energy spectrum. This means that the energy loss through the cutoff is dynamically significant.

Physically speaking, there must always be a high energy cutoff, because sufficiently energetic particles will escape from the shock by virtue of their large mean free paths before being convected to downstream infinity. The above result implies that it must always be included in the calculation before the compression ratio can be found. Once found, its value may yield a convergent cosmic ray spectrum but this is not known *a priori*. This is all the more true when the diffusion coefficient in nature increases with energy.

The treatment of cosmic rays as a second fluid has been refined to include wave heating of the upstream fluid by McKenzie and Völk [179], using similar techniques to those used in the treatment of adiabatic upstream flow [192]. They find yet another type of solution in which the gas is heated to its post shock temperature by the wave heating. This may provide a physical model of a hydrodynamically turbulent shock.

Another approach to the non-linear problem is to regard the deceleration as small and treat it as a perturbation [202, 203, 204]. Working to first order in the cosmic ray pressure and for simplicity ignoring cosmic rays ahead of the shock, the velocity in the precursor is given by

$$u(x) = u_-(1 - \varepsilon(x))$$

where

$$\varepsilon(x) = \frac{M_\mathrm{g}^2}{M_\mathrm{g}^2 - 1}\left(\frac{P_\mathrm{c}}{C_1 u_-}\right) \ll 1 \tag{6.10}$$

and the subscripts c, g refer to the cosmic rays and the gas respectively. We can also calculate the perturbation to the compression ratio r at the subshock,

$$1 - 1/r = (1 - 1/r_0)(1 - (\gamma + 1)\,\varepsilon(0_-)/(M_\mathrm{g}^2 - 1)) \tag{6.11}$$

where r_0 is the unperturbed compression ratio. This perturbed value of the compression ratio must be used in the boundary condition (4.10). Next we can obtain an expression for the slope of the perturbed

distribution function in terms of an integration over the test particle solution by integrating eq. (3.39). The result is

$$q = \frac{-\partial \ln f}{\partial \ln p} = q_0\left[1 + \varepsilon(0_-) + \frac{(\gamma+1)\,\varepsilon(0_-)}{(M^2-1)} - \frac{1}{f_0(0_-)}\int_{-\infty}^{0_-}\left\{\varepsilon\,\frac{\partial f_0}{\partial x} - \tfrac{1}{3}p\,\frac{\mathrm{d}\varepsilon}{\mathrm{d}x}\,\frac{\partial f_0}{\partial p}\right\}\mathrm{d}x\right]. \tag{6.12}$$

Making different assumptions about the form of the diffusion coefficient, we can confirm that if the diffusion coefficient increases with energy, then the spectrum has positive curvature and that the influence of the mildly relativistic cosmic rays is generally to increase the compression ratio through decreasing the effective value of γ in the overall Rankine–Hugoniot conditions and thereby to enhance the acceleration rate. This perturbation scheme can be generalized to higher order though this is not particularly useful; the main point of these fluid approaches being to highlight the physics rather than to produce accurate results.

6.2. Non-perturbative model of shock mediation

The preceding calculations assumed either that the cosmic rays are a small perturbation to the flow, or that the pressure that mediates the flow is in cosmic rays whose diffusion coefficient is more or less energy-independent. The considerations expressed at the beginning of this section, however, suggest that the cosmic rays in real astrophysical shocks are both dynamically significant and that D increases significantly with p. This forces us to a non-separable non-perturbative approach.

Such an approach has been developed [189, 201, 205, 206], by exploiting the large number of decades of energy and spatial scales in the problem, which generally implies that the total pressure in fast particles generally varies much more slowly ahead of the shock than does $F(E)$ at any given E. To understand this separation of scales, note that the pressure in particles having $D \gg D(E)$, which varies over a scale $\sim D/u \gg D(E)/u$, is roughly equal to its post-shock value over the region where $F(E)$ varies, and, similarly, pressure in particles with $D \ll D(E)$ is very small. (Here, for simplicity, we have taken $F(E)$ to vanish as $x \to -\infty$.) Since the total energetic particle pressure is distributed over many decades in energy, very little of the pressure at any point is in particles failing to obey $D \gg D(E)$ or $D \ll D(E)$, for a given E, particularly if D increases rapidly with E. Thus, $F(E)$ executes most of its change over a region where P_c, and hence u, are nearly constant. Assuming $D(E)$ is monotonically increasing (rapidly) with E, we can "map" u and E into each other so that $\bar{u}(E)$ is the velocity over which $F(E)$ increases from zero to nearly its post shock value ahead of the shock, and $\bar{E}(u)$ is the inverse function, i.e.,

$$\bar{E}(\bar{u}(E)) = E\,. \tag{6.13}$$

Using u as the independent spatial variable instead of x, we can approximate $F(E)$ as

$$F(E,u) \cong F_+(E)\,\theta[u(\bar{E}) - u] \cong F_+(E)\,\theta[E - \bar{E}(u)] \tag{6.14}$$

where θ is the Heaviside step function, and where $F_+(E)$ is the value of $F(E)$ at $x = +\infty$, i.e., $u = u_+$. The value of $\bar{u}(E)$ is determined implicitly as follows. Multiplying (6.3) by E and integrating from E to $E_{\max}$, and using definition (6.7), we obtain

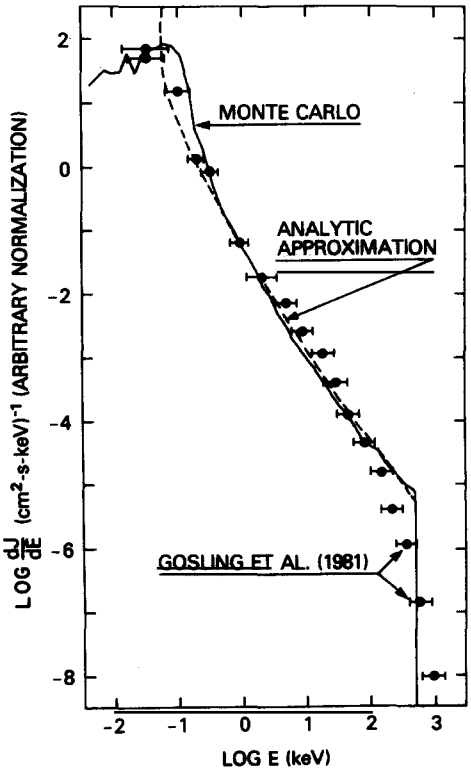

Fig. 7. The entire particle spectrum at a moderate interplanetary shock, as recorded and assembled by [265] is displayed. It is seen that the energetic particle spectrum blends into the thermal population, to within observational uncertainty, about as smoothly as the theoretical model of Ellison and Eichler predicts. The theoretical fit, which neglects the effects of the magnetic field, is taken from ref. [208].

$$-\frac{3}{j}\left(\frac{\partial}{\partial x}u_s\mathcal{P}\right)+\frac{\mathrm{d}u_s}{\mathrm{d}x}\int_E^{E_{\max}}\left(\frac{\mathrm{d}\zeta}{\mathrm{d}\ln E'}-\zeta\right)\mathrm{d}\ln E'+\frac{\partial}{\partial x}\int_E^{E_{\max}}E'D\,\frac{\partial}{\partial x}F\,\mathrm{d}\ln E'=0 \tag{6.15}$$

where

$$j=\mathrm{d}\ln E/\mathrm{d}\ln p\,,\qquad \zeta=(j/3)EF \tag{6.16}$$

and

$$\mathcal{P}(E,u)=\int_E^{E_{\max}}\zeta(E',u)\,\mathrm{d}\ln E'\,.$$

The approximation (6.14) allows several (approximate) identities to be established

$$P_c(u)\equiv\int_0^{E_{\max}}\zeta(E',u)\,\mathrm{d}\ln E'\cong\int_{\bar{E}(u)}^{E_{\max}}\xi(E',u)\,\mathrm{d}\ln E'=\mathcal{P}[(\bar{E}(u),u)]\cong\int_{\bar{E}(u)}^{\infty}\zeta_+(E)\,\mathrm{d}\ln E=\mathcal{P}_+[\bar{E}(u)] \tag{6.17}$$

where the subscript + denotes the value at $x = +\infty$. Also,

$$\mathscr{P}(E, u) \cong \mathscr{P}[\bar{E}(u), u] = \mathscr{P}_+[\bar{E}(u)] \qquad \text{at } u \geqq \bar{u}(E)$$

and (6.18)

$$\mathscr{P}(E, u) = \mathscr{P}[E, \bar{u}(E)] \qquad \text{at } u \leqq \bar{u}(E)\,.$$

Integrating (6.15) over x from $-\infty$ to $+\infty$ and taking F and $\partial F/\partial x$ to vanish at $x = -\infty$ yields

$$-\frac{3}{j}\,u_{s+}\mathscr{P}_+ + \int_{u_-}^{u_+} \zeta(E_{\max})\,\mathrm{d}u_s - \int_{u_-}^{u_+} \zeta\,\mathrm{d}u_s - u_{s+}\mathscr{P}_+ + \int_{u_-}^{\bar{u}(E)} u_s\,\mathrm{d}\mathscr{P} + \int_{\bar{u}(E)}^{u_+} u_s\,\mathrm{d}\mathscr{P} = 0 \tag{6.19}$$

where $u_- \equiv u(-\infty)$, and the limits of the integrals refer to values of u, of which u_s is a function.

By eqs. (6.17) and (6.18), $\mathscr{P}(E, u)$ is just P_c at $u > \bar{u}(E)$ so the next to last term in (6.19) is just $\int_{u_-}^{u(E)} u_s\,\mathrm{d}P_c$.

The integral $\int u_s\,\mathrm{d}P_c$ can be reduced to an expression for u and v_{ph} using the equation for wave energy flux F_w, (5.22), the first law of thermodynamics, and the momentum equation

$$P_c + P_g + P_w + \rho u^2 = C_2\,. \tag{6.20}$$

Assuming that the damped wave energy goes into heat, the damping term in eq. (5.23) is

$$L = C_1 T\,\mathrm{d}s/\mathrm{d}x\,, \tag{6.21}$$

where C_1 is the mass flux and $T\,\mathrm{d}s$ is the heat increment per unit mass.

The first law of thermodynamics is written for the present purpose as $T\,\mathrm{d}s = \mathrm{d}h - \mathrm{d}P_g/\rho$ where h is the enthalpy per unit mass, $\frac{5}{2}P_g/\rho$.

Combining eqs. (5.23), (6.20) and (6.21) yields

$$u_s\,\mathrm{d}P_c = -\mathrm{d}F_w - C_1\,\mathrm{d}h - C_1\,\mathrm{d}u^2/2\,. \tag{6.22}$$

Finally, using all of the above results, one can manipulate eq. (6.19) into the form

$$-\left(1 + \frac{3}{j}\right) u_{s+}[P_g + P_w + C_1 u]_{u_-}^{\bar{u}(E)} + \zeta_+(E)[-\bar{u}_s(E) + u_{s+}] + \left[F_w + \tfrac{5}{2}P_g u + C_1\frac{u^2}{2}\right]_{u_-}^{\bar{u}(E)} + Q = 0 \tag{6.23}$$

where

$$Q = \zeta_+(E_{\max})\,(u_{s-} - u_{s+})\,. \tag{6.24}$$

The quantity Q is the value of the quantity

$$\int_{u_-}^{u_+} (-E_{\max})\,(\mathrm{d}u_s/\mathrm{d}x)\,[(j/3)\,F(E_{\max})]\,\mathrm{d}x \tag{6.25}$$

under the approximation (6.14) [201] and represents the energy loss due to acceleration of particles through $E_{\max}$, as discussed in connection with eq. (6.9). Thus, Q fixes the compression ratio and hence u_+. It is chosen by a procedure that becomes clear after the family of solutions that it parametrizes is generated. (See the following discussion and fig. 8.)

The unknown parameters are the damping rate σ, and the phase velocity $(u - u_s)$, which, as discussed in the previous section, may differ from v_A. These parameters depend on the still unresolved plasma physics discussed in section 5. But we can perform the calculation in the two extreme limits of $L \to \infty$ and $L = 0$.

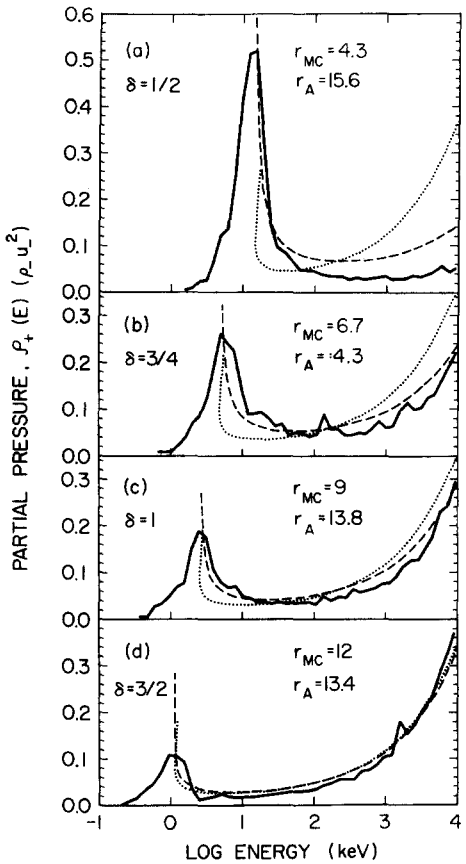

Fig. 8. The approximate analytic solutions of the non-linear shock acceleration equations [201] are compared with the Monte Carlo numerical solutions of the Boltzmann equation with a BGK collision operator [the large angle scattering version of eq. (3.19) of the text]. The true analytic solution is represented in the bottom frame. The numerical solutions are for diffusion coefficients proportional to E^δ. The dotted line is the analytic solution with the compression ratio r_A that follows from eq. (4.3b) but with Q chosen such that the thermal peak of the analytic solution coincides with that of the numerical solution. The dashed line uses the value of Q that gives the same location of the thermal peak, but with the compression ratio r_{MC} of the numerical solution. It is seen that as δ, and (hence) the dynamic range of spatial scales in the energetic particle precursor increases, cosmic ray production becomes more efficient and the analytic approach becomes more accurate. As seen in the last frame, where the range of scales spans six decades, the energetic particle spectrum predicted by the analytic solution is nearly exact, and insensitive to the uncertainty in the compression ratio. Adapted from ref. [208].

i) $L \to \infty$: Rapid wave damping.

Here it is assumed that the wave pressure P_w and energy flux F_w are infinitesimal, so that formally, they make no contribution to eq. (6.23).

To eliminate P_g in favor of $\bar{u}$, we need to derive the thermodynamic trajectory of the preshock fluid as it approaches the shock in the presence of the entropy generation term (6.21).

For simplicity, we first take v_{ph} to be a constant upstream of the shock. Assuming the gas to be monoatomic, we write the First Law of Thermodynamics as

$$\tfrac{3}{2}\, \mathrm{d}(P_g/\rho) = T\, \mathrm{d}s - P_g\, \mathrm{d}(1/\rho)\,; \tag{6.26}$$

using eqs. (5.23) and (6.21) for $T\,\mathrm{d}s$ with $P_w = 0$, $F_w = 0$ we obtain $T\,\mathrm{d}s = v_{ph}\,\mathrm{d}P_c$, and, using the momentum equation to eliminate P_c, we finally obtain

$$P_g + \tfrac{2}{5} v_{ph} = S(u + \tfrac{2}{3} v_{ph})^{-5/3}\,, \tag{6.27}$$

where S is a constant of integration. This is the adiabatic law when $v_{ph} = 0$. The enthalpy flux $C_1 h$ is $\frac{5}{2} P_g u$. Thus eq. (6.23) specifies $\zeta_+(E)$ in terms of $\bar{u}(E)$ (the independent variable) with Q, v_{ph}, u_+ and the fluid parameters at $-\infty$ all as parameters. We can now recover E as the independent variable on which ζ explicitly depends by writing

$$\zeta_+(E)\, \mathrm{d}\ln E = -\mathrm{d}\mathscr{P}_+(E) = -\mathrm{d}\mathscr{P}_+\{\bar{E}[\bar{u}(E)]\} = -\frac{\mathrm{d}P_c[\bar{u}(E)]}{\mathrm{d}\bar{u}(E)}\, \mathrm{d}\bar{u}(E) \tag{6.28}$$

with the last step following from (6.17). From here on we drop the argument in $\bar{u}(E)$.

Specifying P_g in terms of u and using the momentum equation fixes $\mathrm{d}P_c(\bar{u})/\mathrm{d}\bar{u}$ in terms of $\bar{u}$, so that

$$\ln E_2 - \ln E_1 = \int_{\bar{u}_1}^{\bar{u}_2} -\frac{(\mathrm{d}P_c/\mathrm{d}\bar{u})}{\zeta_+(\bar{u})}\, \mathrm{d}\bar{u}\,, \tag{6.29}$$

is a well defined integral over $\bar{u}$. $\zeta_+(E)$ is now specified in terms of E. The analytic results are plotted in fig. 8. They are compared to the results of Monte Carlo numerical solutions to the kinetic Boltzmann equation (3.19) generated by Ellison [207, 208]. Each show that two more or less distinct populations of particles make significant contributions to the post-shock pressure; the low energy pressure peak corresponds to thermal particles and the high energy peak to cosmic rays.

At this point, we actually have a family of solutions parametrized by Q and u_+, given all of the physical parameters, i.e., the upstream parameters and E_{max}. The analytic procedure for choosing Q and u_+, which picks the "true" analytic solution from the rest, is to satisfy eq. (4.3b) and the requirement that the thermal peak implicit in eqs. (6.23) and (6.29) occur at $m_p u_+^2$ to within a factor of order unity. This is discussed in greater detail in ref. [208]. (The residual ambiguity in this factor, which ultimately introduces a slight uncertainty in the compression ratio r, is shown to be insignificant in fig. 8d, where the "true" analytic solution is plotted.) As illustrated in this figure, the analytic procedure is extremely accurate when the dynamic range of spatial scales in the problem exceeds six decades. For galactic cosmic rays, this range exceeds ten decades.

Also illustrated is the fact that cosmic ray production depends on the diffusion coefficient increasing

with energy. In the examples shown, the efficiency of cosmic ray production clearly increases significantly with $\mathrm{d}\ln D/\mathrm{d}\ln E \equiv \delta$.

The appearance of two distinct pressure peaks, one corresponding to thermal particles and the other to cosmic rays, can be understood by differentiating eq. (6.23) with respect to E. In the limit where v_{ph}, P_{w} and F_{w} all vanish (the most transparent case), one obtains

$$\mathrm{d}[(\bar{u}-u_+)\zeta]/\mathrm{d}\bar{u} = [\bar{u} - \tfrac{5}{2}u_+][1 - M^{-2}] \tag{6.30}$$

where M is the Mach number of the flow at $\bar{u}(E)$. Since M must be greater than unity upstream of the viscous subshock, it follows that the quantity $(u-u_+)\zeta$ has a local minimum at $u = \frac{5}{2}u_+$; this implies a two peak structure in $[(u-u_+)\zeta]$.

ii) $L = 0$: No wave damping.

If the waves are assumed not to be damped, a similar procedure exists for computing the energetic particle spectrum. Here, in addition to a thermodynamic trajectory for the gas, which should be taken as adiabatic in the absence of heating by cosmic rays, one needs an "equation of state" for the waves themselves. This has been done using two fluid formalism, by McKenzie and Völk [179], who in fact derive it for various assumptions about the wave damping. The cosmic ray spectra obtained are similar to those obtained for the cases of rapid wave damping.

The buildup of magnetic field energy ahead of the shock can also be computed. As discussed in section 5, the field attains an energy density ahead of the shock of the order of M_{A}^{-1} of the energetic particle pressure. If there is a further compression of the field at the viscous subshock, the post-shock field energy may in fact attain rough equipartition with the cosmic rays for moderate M_{A} [206]. In any case, the post-shock magnetic field energy is likely to exceed its value far upstream of the shock by a large factor.

6.3. *Models of the subshock*

We now turn to the question of particle injection. As discussed in section 2, there is strong observational evidence from both the earth's bow shock and galactic cosmic rays that this occurs at the shock front and that there is not a large seed population of incipient cosmic rays. Furthermore, there is good evidence from X-ray observations of supernova remnants (cf. section 2.5) that high Mach number shocks contain gas subshocks and are not mediated by the pressure of the cosmic rays alone (cf. section 5.1).

To answer the question of which, and how many, thermal particles at the shock are subject to shock acceleration therefore depends on understanding the collisionless shock structure itself. There has been impressive progress in understanding laminar, low Mach number shocks. The shocks that are most efficient at accelerating cosmic rays are of high Mach number, probably are turbulent, and are not well understood. A complete review of the existing theory of collisionless shocks and its confrontation with spacecraft observations is beyond the scope of this article. Relevant reviews include refs. [171, 210, 211, 212] and several articles in the proceedings of the Chapman Conference on Collisionless Shocks in the Heliosphere [213] and the proceedings of Symposium 619 of the 25th Plenary meeting of Cospar [214]. Nevertheless, some of the issues that connect shocks to shock acceleration can be discussed at a rather basic physical level.

A shock is an interface between hot and cold fluid. Though the hot fluid flows away from the shock, it does so subsonically, and some of the shock-heated particles tend to backstream into the pre-shock

(cold) fluid. If they then couple to the cold fluid, the motion of the latter then sweeps them back downstream. This coupling, whatever it happens to be, is what sustains the shock. In its absence, particles reflected off the piston would stream indefinitely through the cold fluid. A major difficulty is that this coupling both determines and depends on the details of the shock structure.

To make these remarks more specific, consider a few simple limiting cases, all of which lead to the same basic conclusion that backstreaming takes place in supercritical shocks. Suppose, without specifying how, that all of the heat goes into the electrons, and neglect magnetic fields that are transverse to the shock normal. In trying to evaporate back upstream, the hot electrons create a thermoelectric potential that confines them to the region of the shock, and simultaneously slows down the oppositely charged ions flowing through the shock. Self-consistent models can be constructed for moderate Mach numbers ($M_s \equiv u_-/c_s \lesssim 3$, $T_e \gg T_i$, where c_s is the sound velocity [215]) in which energy removed from the ions heats the electrons, creating a temperature gradient that supports the potential. (The potential is not necessarily monotonic.) However, at a higher Mach number, the potential is found to be large enough to reflect some of the ions, and solving the shock structure problem then entails figuring out how these reflected ions couple to the incoming fluid. For intermediate Mach numbers, $3 \lesssim M_s \lesssim 7$, self-consistent shock models [216, 217] can be constructed in which the reflected ions continue upstream indefinitely, but in our context, the fate of these ions is the central question.

Similarly, other shock models [218, 219, 220] that are based on a laminar disturbance of some kind break down at moderate Mach numbers. (Here the Mach number refers to the shock velocity relative to the propagation velocity of the disturbance at low amplitude.) The point is that a large amplitude disturbance can propagate at perhaps a few times its phase velocity at low amplitude, but no faster. For shock velocities exceeding that maximum the problem appears to reduce to determining how counter-streaming ion beams which represent an unstable distribution relax into a single, hot fluid.

Now if we reverse the above assumption and assume that all of the energy dissipated into the shock goes into ions, we arrive at a similar conclusion. Although there is no significant thermoelectric potential to reflect ions, the hot ions nevertheless backstream through the preshock fluid. The situation is in some sense the same as that of cosmic rays undergoing shock acceleration: they try to stream along their density gradient, i.e., upstream of the shock, but in doing so interact with the pre-shock fluid, and are continually swept back downstream.

The fate of hot ions backstreaming into the cooler plasma depends on the magnetic field geometry, and it is now useful to distinguish between the parallel and perpendicular cases introduced in section 2. If the magnetic field has a "quasi-perpendicular" geometry, the ions can couple to the pre-shock fluid by the Lorentz force of the background magnetic field. Within a gyroradius, they are turned back downstream by the field. Leroy et al. [221, 222, 223, 224] have simulated high Mach number perpendicular shocks, and find excellent agreement with observations of perpendicular regions of the earth's bow shock. They find that there is a narrow spike of enhanced field strength with a large electric potential that is much thinner than an ion gyroradius. Ions may be reflected off this potential barrier back upstream, or they may be turned upstream by the Larmor gyration behind it. Fully kinetic simulations of high Mach number perpendicular shocks exhibit large electron heating [268] even when the shock speed exceeds the upstream electron thermal speed. The relative importance of each of these processes depends on the Alfvén Mach number and the anomalous resistivity. In any case, these ions are forced downstream of this reflection plane within a few gyroperiods by the convective effects of the perpendicular magnetic field. For a field geometry that is sufficiently close to being perpendicular, there is probably no need for the ions to feel any randomizing turbulence in order for a shock to form. (Similarly, cosmic rays may follow more or less scatter-free trajectories when encountering a quasi-

perpendicular shock, and their acceleration may be limited to the scatter-free type. Near co-rotating interplanetary shocks, which in the inner heliosphere are generally quasi-perpendicular, there are strong anisotropies associated with the fast particles at the shock, which is a sign that they interact with the shock without much stochastic scattering.) When θ_{Bn} is significantly less than 90°, Leroy et al. find that some of the reflected ions obtain a sufficient velocity component along the field to escape upstream of the shock, and the fraction that do so increases as the shock becomes closer to being quasi-parallel.

Now consider hot, post-shock ions in a quasi-parallel geometry. Backstreaming particles, if they are to be eventually forced back across the shock, must interact with the pre-shock fluid via instabilities since they can slip along magnetic field lines in a direction that is close to that of the shock normal. Kinematic considerations [225, 226, 227] suggest that this effect should become important when the upstream field vector is within 45° of the shock normal. Several different approaches (see also [227]) arrive at this conclusion, so it is probably not very sensitive to the history of the backstreaming ions; the basic assumption is that there is a dispersion in the post-shock ion speeds and that backstreaming ions begin a backstreaming trajectory with a velocity vector that points upstream with a magnitude that is a factor of two or so higher than the average downstream thermal velocity. Observations of the earth's bow shock and interplanetary shocks support the qualitative distinction between quasi-parallel and quasi-perpendicular shocks (cf. sections 2.1, 2.2), though clearly the two categories should join smoothly. When the angle between the upstream field and the shock normal θ_{Bn} exceeds 45°, the shocks are clearly defined by the magnetic field data and have structure on a scale much less than an ion gyroradius. When θ_{Bn} is less than 45°, on the other hand, the shocks are not well defined and are surrounded by thick regions of strong magnetic turbulence, which is presumably generated by backstreaming particles.

Which instabilities couple backstreaming ions to the cold fluid? In the case of suprathermal ions it is almost universally accepted that the dominant instability is the resonant ion cyclotron instability. This hypothesis is in good agreement with observations of waves and particle distributions ahead of the bow shock [228, 229, 230, 231] and it appears that the ion-cyclotron instability dominates particles down to speeds of roughly twice the upstream fluid velocity, i.e., just above thermal velocities. In the case of backstreaming thermal particles, there is less certainty as to what instabilities dominate deep within the shock. Leaving aside historical differences, we now consider possible differences between the cosmic ray streaming problem and the high Mach number collisionless shock problem. In the latter case, the "beam" density is comparable to the "background fluid" density. This suggests that non-resonant instabilities may play a more significant role. Moreover, a large ion beam density drags a larger electron current through the background ions. Since anomalous resistivity is generally an increasing function of current, the power dissipated could, *prima facie*, be a rapidly increasing function of beam density. Another possibly important distinction between cosmic rays and thermal particles is that the former travel with speeds in excess of the whistler phase velocity, and shocked thermal particles do not necessarily do so.

Obliquely counterstreaming cosmic rays ahead of the shock may be able to excite lower hybrid waves in the background fluid. This is an efficient way of accelerating suprathermal electrons [232].

The non-resonant ion two stream instability is shorted out by electrons unless their temperature is already close to $m_p u^2$ [233]. Thus, if it is to be important, some additional mechanism must be invoked in order to heat the electrons. Simulations [233, 235, 236] appear to demonstrate that the electron–ion two stream instability (i.e., electrostatic anomalous resistivity) cannot accomplish this and instead tends to increase T_i/T_e which shuts off the instability.

In our view, electromagnetic instabilities are more likely to produce a high Mach number shock.

They have the advantage of being able to feed purely off anisotropy, without requiring a deep minimum in the one dimensional ion velocity distribution function [96], and may therefore be more potent at the middle of a shock, where the distinction between counterstreaming ion beams may be washed out.

The firehose instability, the non-resonant counterpart of the ion-cyclotron instability (cf. section 5), is driven by pressure anisotropy and may therefore support a quasi-parallel shock [237, 238, 239]. It may be described in physical terms as follows: A transverse perturbation in the magnetic field causes the ions moving along the field line to exert a centrifugal force in the direction that would enhance the perturbation, while the tension in the field and the mirror force of the cross field motion of the ions resist the perturbation. When the excess of parallel over perpendicular pressure exceeds the magnetic pressure, the centrifugal term dominates and the perturbation grows in the center of mass frame of the ion population with the dispersion relationship (5.28). As derived in (5.3) the fastest growing modes have a wavelength of the order of the ion gyroradius. If the particle pressure dominates the magnetic pressure, mode coupling is slower than wave growth and the magnetic field probably becomes strongly perturbed. The mean free path of thermal particles at the shock and hence the thickness of the shock itself are of the order of a few ion gyroradii. When both ion streams interact nonresonantly with the growing Alfvén wave, the wave can be shown to grow in the center of mass of the fluid [237]. If one of the beams makes a resonant contribution, it may cause the wave to propagate through the center of momentum frame with a substantial speed. In the limit of slow wave growth, the wave propagates at about the Alfvén velocity relative to the center of momentum frame. When growth is rapid, this phase velocity can be larger but under such circumstances the growth rate is comparable to the real part of the wave frequency, and in strong turbulence, nonlinear effects should set in within a time of the order of one wave period (cf. section 6). In the original firehose shock model of Parker it was assumed that the scattering centers were essentially locked into the center of mass frame of the fluid. The fluid was not treated at the kinetic level, so the question of cosmic ray production was suppressed. Later treatments of the firehose shock [238, 239, 240] assumed a relatively weak shock so that the firehose instability saturated in the quasi-linear regime, i.e., the field could absorb the parallel pressure excess without heating the fluid. For high Alfvén Mach numbers, however, the magnetic field line would be sharply bent long before quasi-linear saturation could set in, and ions would probably be scattered by the strong magnetic turbulence. It does not seem possible to mediate a high Alfvén Mach number shock via firehose instability merely by adding wave energy; unless the original field energy can be amplified by an enormous factor through some unknown effect, the pressure of the scattered particles must be dynamically important.

Note that if the particles dominate the total pressure, then the condition for the firehose instability ($P_{\|} > P_{\perp} + B^2/4\pi$, cf. eq. (5.29)) must be formally satisfied within the viscous subshock. This follows from the definition of viscosity μ which can be cast in the momentum equation as

$$P + \rho u^2 - \tfrac{4}{3}\mu \frac{\mathrm{d}}{\mathrm{d}x} u = C_2 \tag{6.31}$$

[120]. If the particles dominate the total pressure, then it is easy to show that

$$P_{xx} + \rho u^2 = C_2 \tag{6.32}$$

where P_{xx} is defined here to be $\int fm\,(v_x - u)^2\,\mathrm{d}^3p$. These two equations identify the viscosity term as being proportional to the pressure anisotropy, $P_{xx} - P$. Given that the magnetic pressure is negligible, the existence of viscosity implies firehose instability.

Recent hybrid simulations, which, by treating the electrons as a massless fluid can run for many ion gyroperiods, show a parallel shock forming on scales larger than an ion gyroradius [240, 241, 242] (fig. 9). In previous particle simulations of parallel shocks, which ran for only a fraction of a gyroperiod [243], ion counterstreaming was not disrupted and a shock did not form. The magnetic whistler turbulence appears in these simulations mainly at scales of order $\sim c/\omega_{pi}$, which, given the lower Alfvén Mach number, is consistent with the hypothesis that it is excited by the firehose instability [244].

Ellison has made a Monte Carlo kinetic model of a shock [207, 208, 245] in which the particle pressure at the shock is assumed to be large compared to the wave pressure, which is neglected. It is a theoretical model of a firehose-generated shock in the limit of very high Alfvén Mach number. The scattering of particles off the magnetic turbulence at the shock is modeled as BGK (i.e., large angle) scattering, with a scattering rate ν^* that can in principle be any function of particle energy in the frame of the scattering centers and of particle rigidity. This frame, assumed to be a function only of x, is chosen such that the scatterers absorb no net momentum or energy. As is expected of the hydromagnetic waves at a high M_A firehose shock, they merely act as agents for transmitting these conserved quantities between particles. The model is essentially a shock-like solution to eq. (3.19) with the angle average of f made in the frame of the scatterers. The results make a variety of quantitative predictions given a choice of $\nu^*(E)$. Most of the velocity transition within the viscous subshock is found to occur within a single mean free path. Post-shock particle spectra for the principle ion species are shown in fig. 8 for various choices of $\nu^*(E)$ (or equivalently, the diffusion coefficient or mean free path as a function of energy). The principle result is that *when the mean free path increases with energy, a strong shock always puts much of its energy into very energetic particles*. The reason, as discussed in the introduction to section 6, is that the energetic particles are more mobile than less energetic ones. This is illustrated by the fact that the superthermal acceleration becomes increasingly efficient as the diffusion coefficient

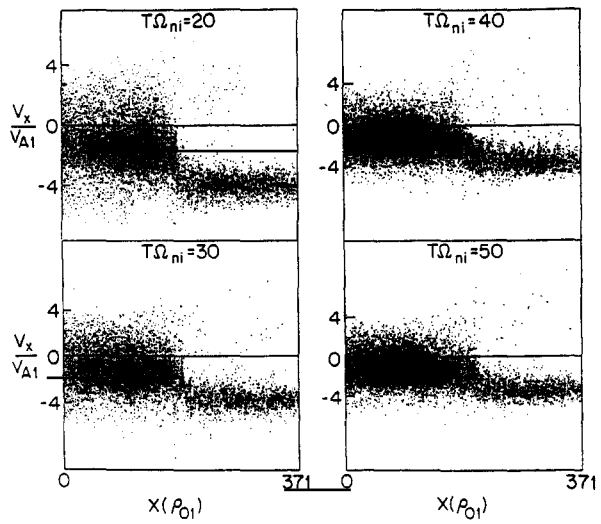

Fig. 9. The temporal evolution of the ions in phase space in a simulated parallel shock [241]. The Alfvén Mach number of the shock, relative to the Alfvén velocity v_A, far upstream, is 4. The units of length are given in terms of the upstream thermal ion gyroradius. The unit of time, Ω_{ni}, is the ion gyroperiod taking only the shock normal component of the magnetic field. The simulation begins with an artificially narrow shock which then relaxes to the configurations at later times shown in the respective frames. There is particle acceleration observed; however, the effects of a finite Alfvén velocity and an upstream free escape boundary are expected to restrict shock acceleration severely for the parameters of the simulation. The history of the individual particles that attained high energies is not shown in the original paper.

becomes an increasingly sensitive function of energy. In the limit where the scattering centers move with the center of mass of the fluid (scattering rate independent of velocity), the scattering of individual particles is elastic and does not heat the rest of the fluid, it merely changes the energy of the scattered particle in the shock frame. As there is very little heating, the bulk kinetic energy is dissipated into non-thermal particles.

With regard to cosmic ray injection, then, a crucial question is whether a high Mach number shock is governed by waves which are fairly well frozen into the fluid (as is likely to be the case for Alfvén waves in a high M_A shock), providing quasi-elastic scattering in the fluid frame, or whether the waves move through the fluid at a hydrodynamically significant velocity. (In the latter case, the waves driven by some particles take their energy and distribute it among the rest, thereby inhibiting the formation of a non-thermal distribution.)

With this question in mind, we consider alternative electromagnetic instabilities that could conceivably make a shock. Another instability that can draw purely on anisotropy is the ion Weibel instability [96, 246] which generates a magnetic field that is transverse to the direction of pressure excess. In this instability, a perturbative magnetic field drives the ions into current filaments of alternating sign which have the correct phase relation to amplify the field. No net current is needed; ions moving in opposite directions at the same point are driven in opposite directions into opposing filaments. The growing field, by Lenz's Law, drives an electron reverse current that can sharply limit its rate of growth, but if the electrons are warm, only those moving nearly perpendicular to the k-vector in the frame of the wave, i.e., with small $\omega - k \cdot v$, yield an appreciable reverse current. The phase velocity of the Weibel mode through the fluid depends on the relative directions of the k-vector and the anisotropy [246]. There is little that can be said from general considerations. In the simulation of Davidson et al. [247] significant electron heating persists after the turbulence saturates, which implies that the ion scattering is not elastic in the frame of the fluid.

The growth rate for the instability for a mode parallel to the ion anisotropy (shock normal) with T_i and $T_e = 0$ is

$$\sigma = ku\omega_{pi}/(\omega_{pe}^2 + k^2c^2)^{1/2} . \tag{6.33}$$

Small scale modes are stabilized when the temperature anisotropy is reduced because ions with sufficiently large $v_{\|}$ cross many wavelengths in one growth period and cannot be herded into current filaments. The Weibel instability has the advantage that it does not need a background magnetic field to operate. Thus, for sufficiently high Alfvén Mach number, one suspects that it would compete favorably with the firehose instability.

The difficulty in judging the extent to which the Weibel and other short wavelength instabilities compete with the firehose instability is that, although they grow more quickly, they also saturate more quickly, and it is very difficult to determine at what level saturation sets in. Simulations of various dimensionalities have revealed that beams of reflected ions can propagate through the incoming fluid over time scales that are far longer than those of the instabilities. Simulations of high Mach number, unmagnetized or quasi-parallel geometries, however, have not lasted long enough to determine the eventual fate of the reflected beam. In proposing a Weibel mode shock, Moiseev and Sagdeev [216] also proposed that saturation of the instability sets in when the field is just strong enough to magnetize the electrons and that the dominant length scale of the field fluctuation is c/ω_{pi}. Thus the scattering centers that form the shock are regions of magnetic field approximately one electron gyroradius across. In each of these regions, ions are scattered through a small angle

$$\delta\theta \approx r_e/r_i = m_e v_e/m_i u \,. \tag{6.34}$$

The thickness of the shock is presumably of order the distance an ion travels before being scattered by an angle of order $\pi/2$. If the scale length of the magnetic turbulence is of the order of c/ω_{pi}, then the thickness of the shock, according to the rough arguments here, is $-(2\delta\theta/\pi)^{-2}c/\omega_{pi}$. Thus, even if the electrons manage to get heated to a temperature of $\sim m_i u^2$, the shock thickness is less than an ion gyroradius only for Alfvén Mach numbers greater than m_i/m_e, according to the model assumptions of Moiseev and Sagdeev. In the simulation performed by Davidson et al. [247] the Weibel instability was found to saturate at a magnetic field strength of

$$B^2/8\pi \sim 10^{-2}\rho u^2 \tag{6.35}$$

where most of the energy was at wavelengths of order c/ω_{pe}. The simulation ran for a time $\sim 15c/\omega_{pi}u$ with little deterioration of the counterstreaming ion beams. An artificial mass ratio of $m_i/m_e \approx 16$ was used which may have influenced this result. If, however, field fluctuations of the size and strength found by the simulation are assumed to generate a shock, the amount they scatter a proton is given by $\delta\theta \approx 10^{-1}\sqrt{m_e/m_i}$ so that the resulting shock thickness would be of the order of $10^2 m_i c/m_e\omega_{pe}$. This is smaller than an ion gyroradius in a background magnetic field for Alfvén Mach numbers in excess of $10^2\sqrt{m_i/m_e}$.

Davidson et al. [247] argue that the saturation is due to ion trapping. This is to say that the magnetic field becomes strong enough to change the phase of an ion in the wave substantially in one linear growth period, presumably destroying the coherence of the current filaments. According to this saturation hypothesis, an ion that moves along the shock normal over a distance of u/σ is deflected sideways by the saturated field a distance of order k^{-1}, where k is the wave number of the dominant modes. For $k \approx \omega_{pe}/c$, and σ given by eq. (6.33), an ion moving a distance of u/σ is deflected through an angle $(m_e/m_i)^{1/2}$. If u/σ is also chosen to be the correlation length, then, following the procedure in the above paragraphs, one estimates that the shock thickness is at least of the order of $(c/\omega_{pi})(m_i/m_e)$. If the saturation is attributed to electron trapping, the saturated field amplitudes (at the real m_i/m_e) are smaller, and the shock mechanism slower.

These estimates are meant to illustrate that transverse magnetic fluctuations on scales that are much smaller than the ion gyroradius are relatively inefficient at scattering ions, and are probably important only when the background magnetic field is very weak. However, the electrons have a stabilising effect and their immobilization by a finite magnetic field might allow Weibel-like instabilities to be more potent than we have estimated.

"High Mach number" shocks in astrophysics are often comparable in velocity to whistler waves, and we now discuss the possible role of these waves in shocks. When $k^{-1} \ll c/\omega_{pi}$, ion inertia becomes relatively unimportant, and the phase velocity of the waves can, to within a factor of unity, approach the "electron-Alfvén" velocity $v_A\sqrt{m_i/m_e}$. It has been conjectured that this mode could be the basis for shocks with $u_- \lesssim v_A\sqrt{m_i/m_e}$. However, the condition $k \gg c/\omega_{pi}$ implies that the ion inertia does not significantly load the waves – that is, the ion velocity is not significantly perturbed by the wave – and this casts doubt on the role of such short whistler waves in shocks. Moreover, as illustrated by eq. (5.30), counterstreaming ions do not drive such waves unstable non-resonantly, as they do at longer wavelengths.

Golden et al. [249] note that counterstreaming ion beams are resonantly cyclotron unstable to high frequency whistler waves, in addition to the longer wavelength, right-handed circularly polarized Alfvén

wave. The whistler moves nearly at the velocity of the beam so that although its lab frame frequency is much greater than Ω_i, the frequency in beam frame is reduced to Ω_i. The phase velocity of the whistler must be finely tuned, relative to the beam velocity so there is only a very narrow wavelength band that is unstable, $\Delta k/k \ll 1$. It is not clear that the whistlers would build up to a significant amplitude in the noisy environment of a shock, since a small amount of mode coupling would efficiently remove energy from this narrow waveband.

Whistler waves are most likely to be generated in oblique shocks, where the global geometry dictates that the field orientation must change. They would have to form on the ion inertial length scale, if they are to influence the motion of the ions. Since the phase velocity of such waves at the ion inertial length scale exceeds v_A only by a factor of order unity, one still expects supercritical behavior for shocks having large Alfvén Mach numbers. Simulations of oblique, intermediate Mach number shocks, $\theta_{Bn} = 45°$, $M_A \approx 5$, by Biskamp and Welter [248] and their subsequent analysis appears consistent with this view of oblique shocks.

Simulations of oblique, supercritical shocks have, at the time of this writing, been extended down to θ_{Bn} of less than 45°, and, due to better techniques, have running times of many gyroradii [240]. They have proven capable of describing the ion reflection at the shock, and the relative importance of the electrostatic potential barrier associated with the large amplitude whistler. For $M_A = 4$, $\theta_{Bn} = 37°$, the thickness of the viscous subshock is on the order of $5(c/\omega_{pi})$, or about one convective ion gyroradius (CIG, the gyroradius of an ion moving at the shock velocity in the zero electric field frame). They have not run long enough to resolve the fate of the reflected ions (e.g., to what extent they undergo first order Fermi acceleration), which become a larger fraction of the total as the shocks are made more parallel. Moreover, they have been thus far limited to moderate Alfvén Mach numbers ($M_A \lesssim 4$), as simulation at higher M_A require more computing time. The characteristic length of the magnetic turbulence is of the order $2(c/\omega_{pi})$. These values for the shock thickness and scale of the magnetic turbulence are somewhat shorter than the more parallel (hybrid, massless electron fluid) simulations of Kan and Swift and Mandt and Kan, which were also performed at an Alfvén Mach number of 4. The difference may be due to the geometries; however, the shock thickness and stationarity in the latter simulations varied greatly with the choice of upstream electron temperature, so we refrain from further interpretation.

Very recent simulations by Quest [269] of parallel shocks indicate that they are established by the firehose instability. Runs over a wide range of M_A indicate clearly that the scale length of the turbulence and shock thickness scale as the convective ion gyroradius, not the ion inertial length. At $M_A \gg 1$, the turbulence is found to be convected by the fluid with little propagation in the fluid frame, as is assumed in Ellison's kinetic model.

Recently reported observations of several quasi-parallel shocks [250, 251] are consistent with a length scale of several ion Larmor radii for the viscous subshock; see, for example, fig. 10. In determining the length scale of a bow shock, we are faced with the uncertainty of not knowing the velocity of the shock relative to the spacecraft. For traveling interplanetary shocks, this problem is greatly reduced. In the one case reported by Scudder et al., the shock thickness was reported to be about 10 CIG. This is somewhat longer than the estimates for the earth's quasi-parallel bow shock thickness, which were estimated to be only 1 or 2 CIG, but the interplanetary shock was considerably weaker.

Translating a CIG into mean free paths introduces an additional uncertainty. Using the energetic particle profile in the earth's foreshock (cf. section 2), we estimate that a mean free path is about 4 ion gyroradii. The turbulence in the viscous subshock appears to be somewhat stronger than in the foreshock (cf. fig. 10) so a somewhat lower value, perhaps 2 or $3r_g$, is appropriate for the former case. The results of simulations and observations, if they are interpreted to give viscous subshocks of about

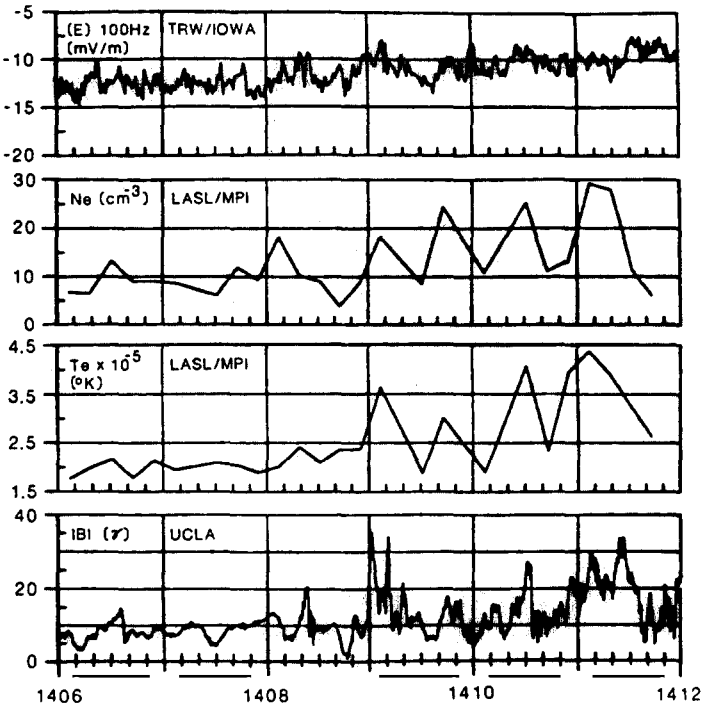

Fig. 10. Data from a parallel bow shock crossing of ISEE-1. The unit of the magnetic field B, "γ", is 10^{-5} gauss. The gyrofrequency in the average field of 10γ is $1\,s^{-1}$. The noisy regions of the magnetic field data, which coincide with sharp increases in the electron density, represent the shock. The intrinsic fluctuations in the solar wind cause the shock to jitter in its location, and the quasi-periodic changes in the density can be attributed to motion of the shock back and forth across the spacecraft. If the shock moves relative to the spacecraft at 10% of the flow speed of the material through the shock, as generally indicated by multiple spacecraft measurements, then the typical transition time of 25 s, as indicated in the plasma data, corresponds to a transition length scale of 2 to 3 convective ion gyroradii. The length scale of the magnetic turbulence, under the assumption that it moves with the fluid, is then about the shock transition length scale. The viscuous subshock appears to be about one mean free path thick (i.e. about 3 gyroradii). The scattering is due to magnetic turbulence on the scale of the order of a gyroradius. Adapted from [250, 251].

three ion gyroradii, thus give a shock thickness of about 1 mean free path, which is in good agreement with the model of Ellison.

The most striking observational support to date for Ellison's model is the quantitative comparison with prediction [270] of the AMPTE/IRM observation of the complete particle spectrum just upstream of the parallel bow shock during a radial field configuration [271]. Because the preshock solar wind particles are included in the measurement, the shock parameters are essentially fixed by it, and with these parameters, the number and spectra of energetic ions are in excellent agreement with the data (fig. 11).

Let us summarize our limited understanding of high Mach number collisionless shocks. Purely electrostatic instabilities do not appear to provide a promising shock mechanism. Purely electromagnetic instabilities, in which a transverse magnetic field is used to scatter incoming ions, can make a shock. The most efficient way to scatter ions is with magnetic field fluctuations distributed on a scale of an ion gyroradius such as would result from a firehose instability. However, if the shock itself can spontaneously produce enough magnetic field, the shock can be thinner than the firehose scale at sufficiently large M_A. Modes which mix electrostatic and electromagnetic effects, could also be as

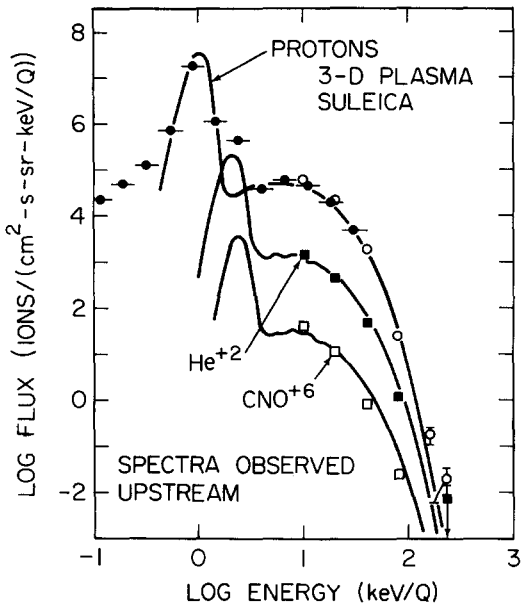

Fig. 11. Comparison of Ellison's shock model [270] with AMPTE/IRM observations of Mobius et al. [271] of diffuse ions just upstream of the earth's bow shock. The observations, made during near ideal radial conditions, give the absolute injection and acceleration efficiencies of the parallel bow shock, and are in excellent quantitative agreement with the model. The calculation assumes a Maxwellian preshock solar wind distribution and makes no attempt to fit the three lowest energy points, which are near the one-count level. The high point in the "proton" spectrum at ~2.5 keV is at the predicted position of the He^{+2} thermal peak and contains a significant contribution of alpha particles. Due to the lack of mass resolution in the 3D plasma instrument, quantitative measurements of thermal He^{+2} flux cannot be made.

important but are complicated. If the shock is created purely by waves whose phase velocities are small compared to the shock velocity, cosmic ray production seems to be a necessary consequence. The firehose-ion cyclotron instability, which creates Alfvén waves, could provide a situation somewhat like this for $M_A \gg 1$. Though even here, neglect of the motion of scattering centers relative to the fluid is an idealization which possibly overestimates the production of non-thermal particles. The ion-cyclotron instability, and the particle acceleration it implies, seem to dominate the coupling of backscattered ions to the fluid upstream of the earth's bow shock for all ions that are just above thermal energies. The coupling of truly thermal ions to each other, which determines the shock structure, may be more complicated because there is no clear distinction between "beam" and "background". Nevertheless, such a distinction could be made at the leading edge of the shock; if the ion-cyclotron instability governs this leading edge, it may be there that particles are injected into the shock acceleration mechanism regardless of the particular structure deep within the shock. Similarly, the acceleration efficiency, which is nearly independent of the injection rate as long as the latter is enough to drive the acceleration mechanism non-linear, may therefore not depend on the structure of the gas subshock. The composition of accelerated particles, on the other hand, surely depends on shock structure, and this may ultimately set strong constraints on theories.

6.4. Composition of accelerated particles

Many authors have studied the elemental and isotopic composition of cosmic rays in the hope of finding clues about their origin. As described in section 2.5, the abundances of galactic cosmic rays and

solar cosmic rays differ from standard local galactic and photospheric abundances. Hydrogen and helium are underabundant relative to much heavier species ($Z > 10$) by respective factors of about 10 and 20 at a given energy per nucleon, with C, N, O and Ne somewhere in between, suggesting some dependence on mass. In addition, elements with high first ionization potentials appear to be underabundant relative to those that are more easily ionized. For example, Ne, while enhanced relative to protons, is underabundant relative to Na, which has nearly the same mass. There is also some evidence that neutron-rich isotopes of Ne and Mg are enhanced in cosmic rays. Measurements at the earth's bow shock do not resolve elements heavier than carbon, but taken as a group, the heaviest (mainly C, N, O) seem to be five to six times more abundant than protons at a given energy per charge, consistent with a weak enhancement at a given energy per nucleon. The difficulties in tracing the origin of cosmic rays by their abundances include the uncertainties in the composition of the thermal gas at any particular astrophysical location and the fact that the relative abundances of a trace element in cosmic rays could be altered significantly by a relatively small amount of acceleration being carried out in an environment where that element has a very high abundance. The apparent deficiency in elements with a high ionization potential led to speculation that cosmic rays originate out of a gas with temperatures of the order of 10^4 K, where some but not all elements are ionized. However, this pattern of elemental abundances is also observed in solar cosmic rays. In the latter case, the charge states can be measured and they indicate that the energetic particles originate largely out of a 10^6 K gas, presumably the solar corona. This has prompted the recent hypothesis (e.g., Meyer [252]) that the solar corona itself is deficient in high first ionization potential elements. The best fits to measurements of composition of the solar wind and the solar corona support this hypothesis [252].

Similar difficulties exist for the hypothesis that galactic cosmic rays originate out of a 10^4 K thermal plasma. Under such conditions, many refractory elements would be largely in grains. Moreover, the presence of neutrals implies damping of the waves that confine cosmic rays to shocks [130, 131] and shock acceleration beyond a few GeV per nucleon in the presence of neutral hydrogen is difficult under such conditions. This suggests that cosmic rays originate out of the hot (10^6 K) phase of the interstellar medium. It may be that this phase has the same composition that is being attributed to the solar corona. Concerning isotopic anomalies in cosmic rays, the overabundant isotope could be coming from astrophysical sites that are extremely rich in them such as the reverse shocks in supernova ejecta, termination shocks in the winds of Wolf-Rayet stars [63], and possibly the interstellar material in sites of young star formation where a supernova shock could encounter material recently enriched by a previous supernova. It has also been suggested [253] that the solar system itself is anomalous in some isotopic abundances as a result of forming in a young stellar association. It would then be deficient in elements that are released by stars that live longer than the association, such as ^{22}Ne. Given all of the factors that could affect cosmic ray composition, its similarity to standard cosmic composition is perhaps more noteworthy than the differences. (In contrast, consider ^{3}He-rich solar flares in which ^{3}He, believed to undergo pre-injection [254], is enhanced relative to ^{4}He and other elements by many orders of magnitude.) That the injection process leaves standard cosmic rays so similar in composition to the thermal plasma is a strong constraint on the physics of the injection process, and possible variations in the thermal composition due to environmental factors only strengthens this argument. It is unlikely, for example, that cosmic rays are selected from the rest of the thermal population by some reflection process off an electrostatic potential barrier. Such a barrier would reflect light ions far more readily than heavy ones. Strong selective reflection is seen clearly in laboratory experiments. In a detailed compositional analysis of energetic particles at the earth's bow shock, Ipavich et al. [255] report that shocks with the three lowest values of M_A yielded energetic particle populations that were deficient in

He. For shocks with $M_A > 8$, He is enhanced at a given energy per charge by an average factor ranging from 1 to 3 with a mean of 1.7 and with no significant correlation with M_A. They suggest that, as is the case in the perpendicular shock simulations of Leroy et al. [221], the electrostatic potential is important at low Alfvén–Mach numbers in reflecting particles, and unimportant compared to magnetic forces for stronger shocks. The diffuse population of energetic particles that is observed for a given event consists of seed particles taken from various points along the bow shock ranging in shock strength, and this may be a source of scatter still remaining in the data. But similar studies for carefully chosen magnetic field geometries (perhaps at lower energies as well) could further extend this very promising line of study. As theories of collisionless shocks and theories of shock acceleration become more closely intertwined, the composition of cosmic rays and the temperatures of the different ion species behind shocks may serve as strong constraints on collisionless shock theory.

The range of helium enhancement observed by Ipavich et al. of 1 to 3 at a given energy per charge should be compared to the value of 3 that results from Ellison's model [256]. The model also predicts that ionized heavy elements ($Z > 6$) are enhanced by about an order of magnitude, at a given energy per nucleon, for mass to charge ratios typical of heavy elements at ionization temperatures of 10^6 K. Similar enhancements, relative to the lowest reasonable reference energy in the solution are obtained from analytic theory [257]. Physically, the reason for the enhancement is that partially ionized heavy elements, having a larger rigidity and hence a larger mean free path, see the shock as being thinner and feel a greater velocity contrast across it. The degree to which the acceleration prefers high rigidity elements depends on the degree to which fast particles mediate the shock, so factors that diminish the overall efficiency of non-thermal particle production also diminish this preference. Probably the enhancement is less than predicted by the simplest theory, but apparently not by much. The whole issue needs to be explored in greater detail and it is hoped that more realistic simulations of high Mach number collisionless shocks will include heavy ion species.

Cosmic rays at very high energies, $E \approx 10^{15}$ eV, may be mostly heavy elements, notably Fe [49, 50]. The airshower studies that suggest this preponderance of heavy elements at 10^{15} eV measure *total* energy, and a 10^{15} eV iron nucleus has a much lower rigidity than a proton of the same energy. Iron nuclei can therefore be better confined to the vicinity of a shock and indeed to our Galaxy than protons.

We have so far emphasised the acceleration of cosmic ray ions. However in several of the synchrotron sources described in section 2, it is the relativistic electrons that we observe [131, 258]. Now, at a given energy, the Larmor radius of an electron is $\sim(m_i/m_e)^{-1/2}$ smaller than the energy of a corresponding ion and so we expect that electron injection is controlled by waves other than Alfvén modes [259]. However at energies $\gg m_i^2 u_-^2/m_e$ relativistic electrons should be accelerated by scattering off just the same Alfvén waves as the ions and should exhibit similar spectra to the ions when compared at a given rigidity. The electron–proton ratio which is controlled by the electron injection rate may depend quite sensitively on the detailed shock structure as observations of supernova remnants and extragalactic radio sources seem to indicate.

7. Summary

In this review, we have attempted to summarize both the incentive for understanding high Mach number collisionless shock waves and recent theoretical progress directed towards this end. As we have described, suprathermal ions are abundant at interplanetary quasi-parallel shock waves. They seem, for the strongest of such shocks, to carry away a substantial fraction (>10%) of the incident energy flux in

the frame of the shock, and they are coupled to the bulk fluid flow by low frequency hydromagnetic waves. Observations seem to be semi-quantitatively consistent with quasi-linear theory, the earth's bow shock having proved itself a valuable laboratory for this and other tests of shock models.

Quite independent arguments have led to the conclusion that the bulk of the galactic cosmic rays are accelerated by middle-aged supernova shock waves. Physical conditions in the interstellar medium are not so dissimilar to those encountered in the interplanetary medium and there is well-founded optimism that understanding of interplanetary shocks can be exported to the interstellar medium and to the even larger scale shock waves associated with double radio sources and the intergalactic medium. With relatively few assumptions, the shock wave theory of cosmic ray origin appears to be in rough quantitative agreement with the observed intensity, spectrum, and composition of high energy particles within these different environments.

More detailed calculations of the cosmic ray distribution function should lead to more precise tests of the theory. It is now possible, in principle at least, to evaluate the spectrum, and the post-shock energy distribution among thermal particles, and the cosmic rays, and the magnetic field strength given the phase velocity and the energy density of the Alfvén waves at all points near the shock (fig. 12). We can make reasonable guesses at the approximate form of these last two quantities, and the answers in some cases are insensitive to them, but in the context of precise quantitative predictions, they must be regarded as crucial, largely unresolved issues.

However, the distribution of energy among the different ions and especially the electrons cannot be properly understood until we understand the structure of collisionless shocks. This is particularly important for the electrons because it is the electrons in astrophysical shocks that are usually the

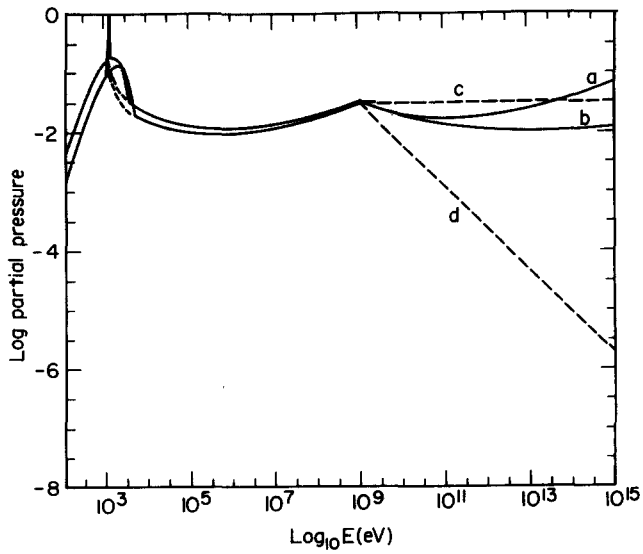

Fig. 12. The steady state, post-shock particle spectra that are predicted by the full non-linear theory of shock acceleration are plotted for parameters that are typical of young to middle-aged supernova remnants expanding in the hot phase of the interstellar medium. The harder spectrum (a) is for an acoustic Mach number M of 47 and the upstream wave velocity of $0.04u_-$. The softer spectrum (b) follows for $M = 16$ and $v_{ph} = 0.04u_-$, and, to within visual accuracy, the same spectrum follows from $M = 33$ and $v_{ph} = 0.105u_-$. In all cases, the high energy cutoff is at $E = 10^{15}$ eV. The upper dotted line (c) represents the spectrum predicted by the test particle theory with a shock compression ratio of 4. The lower dotted line (d) shows the power law spectrum $[N(E)\,dE \propto E^{-2.7}\,dE]$ that best represents the observed galactic cosmic ray spectrum. The y-axis displays partial pressure $[\propto E^2 \times N(E)]$ in arbitrary units.

observable species at astronomical distances; thermal electrons yielding X-rays, and relativistic ones, synchrotron emission.

Undoubtedly, the highest priority is to develop a self-consistent theory or an accurate simulation of a quasi-parallel shock at the plasma physical level. We have several relevant observations to guide us in this task. Interplanetary studies assure us that at low or moderate Mach number, the shock structure depends sensitively both on the Mach number M, the angle between the field and the shock normal θ_{Bn}, and the ratio of the gas pressure to the magnetic pressure β. In particular shocks with M less than some critical value ranging from 2 to 3 are essentially laminar, whereas higher Mach number shocks are turbulent and can freely inject suprathermal ions into the foreshock. This injection is especially pronounced for the quasi-parallel ($\theta_{Bn} \leqq 45°$) shocks. By contrast, the shocks associated with young supernova remnants have large Mach numbers and radio observations give no indication of the difference between quasi-parallel and quasi-perpendicular shocks, such as might be expected to do if the interstellar magnetic has a preferred direction in the neighborhood of the remnant. X-ray observations of these same shocks also tell us that in spite of the existence of solutions in which the background fluid is just compressed adiabatically, its entropy does in fact increase by roughly the factor predicted by the Rankine–Hugoniot relations. There must either be a strong subshock or strong electron heating by the precursor ions and Alfvén waves.

We conclude by listing some specific issues that are being addressed now and are immediately relevant to the theory of shock wave acceleration.

(i) Can we solve the combined cosmic ray transport-shock mediation problem for a specified energy-dependent diffusion coefficient over a sufficient range of energy to model astrophysical shocks? Are approaches based on the kinetic equation (3.36) or those using Monte Carlo methods preferable? Can we make a stationary, numerical simulation of a high Mach number, quasi-parallel subshock making due allowance for the heat flux carried off by suprathermal ions and electrons in the precursor?

(ii) How does Alfvén wave turbulence develop and damp when driven by a strong pressure gradient?

(iii) Given a model of the subshock and pre-existing cosmic rays in the ambient medium, under what circumstances do there exist more than one shock solution for fixed upstream conditions? Which solutions are unstable; and, if there is more than one stable solution, which is likely to be preferred?

(iv) How are suprathermal ion populations in interplanetary shocks observed to vary with β and θ_{Bn} for supercritical Mach numbers?

(v) What is the composition of cosmic rays of energy $T \gtrsim 10^6$ GeV?

It has been realized comparatively recently that understanding the acceleration of cosmic rays and the structure of high Mach number shock waves are intimately related tasks. Future progress depends upon a combined program of numerical simulations and thorough analysis of detailed spacecraft observations in the solar wind. The widespread occurrence of particle acceleration and hypersonic flows in astrophysical environments, strengthens the case for enlarging this program over the coming decade [260].

Acknowledgements

We thank Dr. W. Kohn, Director of the Institute for Theoretical Physics, Santa Barbara (RB, DE) and D. Lynden Bell (RB), Director of the Institute of Astronomy, Cambridge for hospitality during the writing of much of this review. We are particularly grateful to other participants in the Space and

Astrophyical Plasmas program for extensive discussions of collisionless shock theory. We thank Drs. A. Achterberg and D. Ellison for further vital assistance. Financial support by the National Science Foundation under grants AST80-17752, 82-13001, 80-02673, 84-15355 and 83-17755, the NASA Solar-Terrestrial Theory Program and the Alfred P. Sloan Foundation is gratefully acknowledged.

References

[1] Selected Papers on Cosmic Ray Origin Theory, ed. S. Rosen (Dover Publications Inc., New York, 1969).
[2] V.N. Ginzburg and S.I. Syrovatski, Origin of Cosmic Rays (Pergamon Press, New York, 1964).
[3] S. Hayakawa, Cosmic Ray Physics (Wiley, New York, 1969).
[4] G.R. Burbidge, in: Composition and Origin of Cosmic Rays, ed. M. Shapiro (Reidel, Dordrecht, Holland, 1983).
[5] D.B. Melrose, Plasma Astrophysics, vol. 1 (Gordon and Breach, New York, 1980).
[6] D.B. Melrose, Plasma Astrophysics, vol. 2 (Gordon and Breach, New York, 1980).
[7] J. Arons, C.E. Max and C.F. McKee (eds.), AIP Conf. Proc. No. 56 Particle Acceleration Mechanisms in Astrophysics (American Institute of Physics, New York, 1979).
[8] L. Koch-Miramond and M.A. Lee (eds.), Particle Acceleration Processes, Shock Waves, Nucleosynthesis, and Cosmic Rays, Adv. Sp. Res. 4 (1984).
[9] D. Eichler, Proc. VI Moriond Astrophys. Meeting, eds. J. Adouze and Tran Then Van (1986) in press.
[10] R.D. Blandford, in: Particle Acceleration Mechanisms in Astrophysics, eds. Arons, Max and McKee (American Institute of Physics, New York, 1979).
[11] C. Cesarsky, Ann. Rev. Astron. Astrophys. 18 (1980) 289.
[12] I.N. Toptygin, Sp. Sci. Rev. 26 (1980) 157.
[13] W.I. Axford, Origin of Cosmic Rays, eds. Setti, Spada and Wolfendale, IAU Symp. No. 94 (Reidel, Dordrecht, Holland, 1980) p. 339.
[14] W.I. Axford, Proc. NY Acad. Sci. 375 (1981) 297.
[15] W.I. Axford, Plasma Astrophysics, eds. Guyenne and Long (ESA, Paris, 1981) p. 425.
[16] L.O'C. Drury, Rep. Prog. Phys. 46 (1983) 973.
[17] C.F. Kennel, J.P. Edmiston and T. Hada, Proc. AGU Chapman Conf. on Collisionless Shocks in the Heliosphere, eds. Stone and Tsurutani, J. Geophys. Res. 90 (1985) A1.
[18] J.R. Asbridge, S.J. Baume and I.B. Strong, J. Geophys. Res. 73 (1987) 5777.
[19] J.T. Gosling, J.R. Asbridge, S.J. Baume, G. Paschmann and N. Sckopke, Geophys. Res. Lett. 5 (1978) 957.
[20] M. Scholer, F.M. Ipavich, G. Glockler and D. Movestadt, J. Geophys. Res. 85 (1980) 4602.
[22] E.W. Greenstadt and R.W. Fredericks, Solar System Plasma Physics III, eds. Parker, Kennel and Lanzerotti (North-Holland, Amsterdam, 1979).
[23] F.M. Ipavich, A.B. Galwin, G. Gloecker, M. Scholer and D. Havestadt, J. Geophys. Res. 86 (1981) 4337.
[24] M. Scholer, F.M. Ipavich and G. Gloecker, J. Geophys. Res. 86 (1981) 4374.
[25] D.G. Mitchell and E.C. Roelof, J. Geophys. Res. 88 (1983) 5623.
[26] F.M. Ipavich, M. Scholer and G. Gloecker, J. Geophys. Res. 86 (1981) 11153.
[27] M.M. Hoppe, C.T. Russell, L.A. Frank, T.E. Eastman and E.W. Greenstadt, J. Geophys. Res. 86 (1981) 4471.
[28] G.C. Anagostopoulos, E.T. Sarris and S.M. Krimigis, J. Geophys. Res. 91 (1986) 3020.
[29] F.B. McDonald, B.J. Teegarden, J.H. Trainer, T.T. von Rosenvinger and W.R. Webber, Astrophys. J. Lett. 203 (1976) L149.
[30] C.W. Barnes and J.A. Simpson, Astrophys. J. Lett. 210 (1976) L91.
[31] R.G. Decker, M.E. Pesses and S.M. Krimigis, J. Geophys. Res. 86 (1979) 8819.
[32] C.F. Kennel, F.L. Scarf, F.V. Coroniti, E.J. Smith and D.A. Gurnett, J. Geophys. Res. 87 (1982) 17.
[33] M.A. Lee, J. Geophys. R. 88 (1983) 6109.
[34] C.F. Kennel, F.L. Scarf, F.V. Coroniti, C.T. Russell, K.P. Wenzel, T.R. Sanderson, P. Van Ness, W.C. Feldman, C.K. Parks, E.J. Smith, B.T. Tsurutani, F.S. Mozer, M. Temerin, R.R. Anderson, J.D. Scudder and M. Scholer, J. Geophys. Res. 89 (1984) 5419, 5436.
[35] W.C. Feldman, R.D. Anderson, S.J. Bame, J.T. Gosling, R.D. Zwickl and E.J. Smith, J. Geophys. Res. 88 (1983) 9949.
[36] D. Eichler, Ap. J. 247 (1981) 1089.
[37] M.A. Pomerantz, Proc. Nat. Acad. Sci. 58 (1967) 2136.
[38] D.E. Page, A. Domingo and D.G. Wentzel, Correlated Interplanetary and Magnetospheric Observations, ed. D.E. Page (Reidel, Dordrecht, 1974) p. 573.
[39] A. Achterberg and C.A. Norman, Astron. Astrophys. 89 (1980) 353.
[40] D.C. Ellison and R. Ramaty, Astrophys. J. 298 (1985) 400.
[41] M.M. Breneman and E.C. Stone, Astrophys. J. Lett. 299 (1985) L57.
[42] W.S. Kurth, D.A. Gurnett, F.L. Scarf and R.L. Poynter, Nature 312 (1984) 27.

[43] M.E. Pesses, J.R. Jokipii and D. Eichler, Astrophys. J. Lett. 246 (1981) L85.
[44] M. Cassé and J. Paul, Ap. J. 237 (1980) 236.
[45] H.J. Völk and M. Forman, Astrophys. J. 253 (1982) 188.
[46] J.A. Simpson, in: Composition and Origin of Cosmic Rays, ed. M. Shapiro (Reidel, Dordrecht, 1983) p. 1.
[47] A.A. Watson, Adv. Sp. Res. 4 (1984) 35.
[48] A.M. Hillas, Ann. Rev. Astr. Astrophys. 22 (1984) 425.
[49] G.B. Yodh, Cosmology and Particles, ed. Audouze (Editions Frontières, Dreux, France, 1982).
[50] Rm.M. Baltrusaitis, R. Cady, G.-L. Cassiday, R. Cooper, J.W. Elbert, P.R. Gerhardy, S. Ko, E.C. Loh, Y. Mizumoto, M. Salamor, P. Sokolsky and D. Steck, Phys. Rev. Lett. 54 (1985) 1875.
[51] L. Spitzer, Physical Processes in the Interstellar Medium (Wiley, New York, 1978).
[52] E.N. Parker, Cosmical Magnetic Fields (Oxford University Press, Oxford, 1979).
[53] A.J. Owens and J.R. Jokipii, Astrophys. J. 215 (1977) 677.
[54] J.D. Simpson, Ann. Rev. Nucl. Part. Sci. 33 (1983) 323.
[55] N.R. Brewster, P.S. Freier and C.J. Waddington, Astrophys. J. 264 (1983) 324.
[56] M. Cassé and P. Goret, Astrophys. J. 221 (1978) 703.
[57] M.E. Wiedenbeck, Adv. Sp. Res. 4 (1984) 15.
[58] W.R. Binns, R.K. Fickle, T.L. Garrard, M.H. Israel, J. Larmann, K.E. Krombel, E.C. Stone and C.J. Waddington, Astrophys. J. Lett. 267 (1983) L93.
[59] N. Lund, Adv. Sp. Res. 4 (1984) 5.
[60] R. Mewaldt, Rev. Geophys. Sp. Sci. 21 (1983) 295.
[61] G.A. Tamman, in: Supernovae: A Survey of Current Research, eds. Rees and Stoneham (Reidel, Holland, 1982).
[62] N. Duric, Adv. Sp. Res. 3 (1984) 91.
[63] M. Cassé and J.A. Paul, Astrophys. J. 237 (1980) 236.
[64] H.J. Völk and M. Forman, Astrophys. J. 253 (1982) 188.
[65] D. Eichler, Ap. J. 237 (1980) 809.
[66] J.P. Meyer, Ap. J. Suppl. 57 (1985) 173.
[67] J.B.G.M. Bloemen, K. Bennett, G.F. Bignami, L. Blitz, P.A. Caraveo, M. Gottwald, W. Hermsen, F. Lebrun, H.A. Mayer-Hasselwander and A.W. Strong, Astron. Astrophys. 135 (1984) 12.
[68] M. Garcia-Munoz, G.M. Mason and J.A. Simpson, Astrophys. J. 217 (1977) 859.
[69] T.A. Prince, Astrophys. J. 191 (1979) 331.
[70] L.D. Landau and E.M. Lifshitz, Fluid Mechanics (Pergamon Press, Oxford, 1969).
[71] C.F. McKee and D.J. Hollenbach, Ann. Rev. Astron. Astrophys. 18 (1980) 219.
[72] D.P. Cox and B.W. Smith, Astrophys. J. Lett. 189 (1974) L105.
[73] C.F. McKee and J.P. Ostriker, Astrophys. J. 218 (1977) 148.
[74] R.D. Blandford and J.P. Ostriker, Astrophys. J. 237 (1980) 793.
[75] C. Franson and R.I. Epstein, Astrophys. J. 242 (1980) 411.
[76] R. Silberg, C.H. Tsao, J.R. Letau and M.M. Shapiro, Phys. Rev. Lett. 51 (1983) 1217.
[77] I. Lerche and R. Schlickeiser, Astron. Astrophys. 151 (1985) 408.
[78] R.D. Blandford, in: Supernovae: A Survey of Current Research, eds. Rees and Stoneham (Reidel, Holland, 1982).
[79] P.A.G. Scheuer, Adv. Sp. Res. 4 (1984) 337.
[80] Proc. IAU Symp. No. 101, Supernova Remnants and their X-ray Emission, eds. Darziger and Goverstein (Reidel, Dordrecht, Holland, 1983).
[81] J.R. Jokipii and G.E. Morfill, Ap. J. 290 (1985) L1.
[82] D.S. Heeschen and C.M. Wade (eds.), Extragalactic Radio Sources Proc. IAU Symp. No. 97 (Reidel, Dordrecht, Holland, 1982).
[83] K.I. Kellermann and I.I.K. Pauliny-Toth, Ann. Rev. Astron. Astrophys. 19 (1981) 373.
[84] G.K. Miley, Ann. Rev. Astron. Astrophys. 18 (1980) 165.
[85] A.H. Bridle and R.A. Perley, Ann. Rev. Astron. Astrophys. 22 (1984) 319.
[86] M.C. Begelman, R.D. Blandford and M.J. Rees, Rev. Mod. Phys. 56 (1984) 255.
[87] J. Biretta, R.N. Owen and P.E. Hardee, Ap. J. Lett. 274 (1983) L27.
[88] J. Miller (ed.), Astrophysics of Active Galaxies and Quasi-Stellar Objects (Univ. Science Books, Mill Valley, California, 1985).
[89] J.E. Dyson (ed.), Active Galactic Nuclei (Manchester University Press, 1985).
[90] R.D. Blandford and C.F. McKee, Mon. Not. R. Ast. Soc. 180 (1977) 343.
[91] D. Kazanas and D.C. Ellison, Astrophys. J. 304 (1986) 178.
[92] C. Jones and W. Forman, in: Extragalactic Radio Sources, eds. Heeschen and Wade (Reidel, Holland, 1982).
[93] P.J. Young, A. Boksenberg and W.L.W. Sargent, Astrophys. J. Suppl. 48 (1982) 455.
[94] J.P. Ostriker and L. Cowie, Astrophys. J. Lett. 243 (1981) L127.
[95] G. Cavallo, Astron. Astrophys. 111 (1982) 368.
[96] P.C. Clemmow and J.P. Dougherty, Electrodynamics of Charged Particles and Plasmas (Addison Wesley, Chicago, 1969).

[97] E.M. Lifshitz and L.P. Pitaevski, Physical Kinetics (Pergamon, Oxford, 1982).
[98] E. Fermi, Phys. Rev. 75 (1949) 1169.
[99] L. Davies, Phys. Rev. 101 (1956) 351.
[100] R.M. Kulsrud and A. Ferrari, Astrophys. and Sp. Sci. 12 (1971) 302.
[101] A. Achterberg, Astron. Astrophys. 97 (1981) 161.
[102] C. Lacombe, Astron. Astrophys. 54 (1977) 1.
[103] J.R. Jokipii, Astrophys. J. 146 (1966) L80.
[104] J.R. Jokipii, Rev. Geophys. Sp. Sci. 9 (1971) 27.
[105] R. Kulsrud and W.F. Pearce, Astrophys. J. 156 (1969) 445.
[106] D. Bhatnager, E. Gross and M. Krook, Phys. Rev. 94 (1954) 511.
[107] G. Holman, P.J. Morrison, J. Scott and J. Ionson, in: Particle Acceleration Mechanisms in Astrophysics, eds. Arons, Maxard and McKee (American Institute of Physics, New York, 1979).
[108] D.G. Wentzel, in: Unstable Current Systems and Plasma Instabilities in Astrophysics, Proc. IAU Symp. No. 108, eds. Kundu and Holman (Reidel, Dordrecht, Holland, 1985).
[109] M.A. Lee and H.J. Völk, Astrophys. Sp. Sci. 24 (1973) 31.
[110] M.L. Goldstein, Astrophys. J. 204 (1976) 900.
[111] F.C. Jones, T.J. Birmingham and T.B. Kaiser, Phys. Fluids 21 (1978) 347.
[112] A. Achterberg, Astron. Astrophys. 98 (1981) 161.
[113] W.I. Axford, Planet. Sp. Sci. 13 (1965) 115.
[114] P.A. Isenberg and J.R. Jokipii, Ap. J. 234 (1979) 746.
[115] J.R. Jokipii and E.N. Parker, Phys. Rev. Lett. 21 (1968) 44.
[116] J. Skilling, I. McIvor and J.A. Holmes, Mon. Not. R. Astr. Soc. 167 (1974) 87.
[117] J. Skilling, Mon. Not. R. Astr. Soc. 172 (1975) 557.
[118] G.M. Webb, Ap. J. 296 (1985) 319.
[119] G.F. Krymsky, Izv. ANSSR, Ser. Fiz. 45 (1981) 461.
[120] Ya.B. Zel'dovich and Yu.P. Raizer, Physics of Shock Waves and High Temperature Phenomena (Academic Press, New York, 1966).
[121] S. Chandrasekhar, Plasma Physics (University of Chicago Press, Chicago, 1965).
[122] P.D. Hudson, Mon. Not. R. Astr. Soc. 131 (1965) 23.
[123] M. Pesses, Unpublished thesis, University of Iowa (1979).
[124] G.M. Webb, W.I. Axford and T. Terasawa, Astrophys. J. 270 (1983) 537.
[125] J.A. Peacock, Mon. Not. R. Astr. Soc. 196 (1981) 13.
[126] G.M. Webb, Astron. Astrophys. 124 (1983) 163.
[127] J.R. Jokipii, Ap. J. 255 (1982) 716.
[128] W.I. Axford, E. Leer and A. Skadron, Proc. 15th Intern. Cosmic Ray Conf. (Plovdiv) 11 (1977) 132.
[129] G.F. Krimsky, Dok. Akad. Nauk. SSSR 234 (1977) 1306.
[130] A.R. Bell, Mon. Not. R. Astr. Soc. 182 (1978) 147.
[131] A.R. Bell, Mon. Not. R. Astr. Soc. 182 (1978) 443.
[132] R.D. Blandford and J.P. Ostriker, Astrophys. J. Lett. 227 (1978) L49.
[133] E. Schatzman, Ann. d'Ap. 137 (1963) 135.
[134] P.D. Hudson, Mon. Not. R. Astr. Soc. 131 (1965) 23.
[135] D.G. Wentzel, Astrophys. J. 140 (1964) 1013.
[136] L.R. Fisk, J. Geophys. Res. 76 (1971) 1662.
[137] F.C. Michel, Astrophys. J. 247 (1981) 664.
[138] D. Lynden-Bell, MNRAS 136 (1967) 101.
[139] W.C. Saslaw, Gravitation Physics of Stellar and Galactic Systems (Cambridge University Press, Cambridge, 1985).
[140] R.D. Blandford and M.J. Rees, Mon. Not. R. Astr. Soc. 169 (1974) 395.
[141] B.J. Burn, Astron. Astrophys. 45 (1976) 435, 395.
[142] R.D. Blandford and J.P. Ostriker, Astrophys. J. 237 (1980) 793.
[143] A.M. Bykov and I.N. Toptyghin, J. Geophys. 50 (1982) 221.
[144] J.A. Peacock, Mon. Not. R. Astr. Soc. 196 (1981) 135.
[145] D. Eichler, Proc. VIth Moriond Astrophysics Meeting, eds. J. Audouze and Tran Thanh Van (1986).
[146] L.A. Fisk, J. Geophys. Res. 76 (1971) 1662.
[147] P.O. Lagage and C.J. Cesarsky, Astron. Astrophys. 118 (1983) 223.
[148] V.N. Fedorenko, Astrophys. Sp. Sci. 96 (1983) 25.
[149] A.F. Heavens, Mon. Not. R. Astr. Soc. 207 (1984) 1.
[150] I.N. Toptyghin, Sp. Sci. Rev. 26 (1980) 157.
[151] T.J. Bogdan and H.J. Völk, Astron. Astrophys. 122 (1983) 129.
[152] H. Moraal and W.I. Axford, Astron. Astrophys. 125 (1983) 204.

[153] R.A. Chevalier, Astrophys. J. 272 (1983) 765.
[154] M.A. Forman, 18th ICRC 3 (1983) 153.
[155] D. Eichler, Astrophys. J. 244 (1981) 711.
[156] T. Terasawa, J. Geophys. Res. 86 (1981) 7595.
[157] M.A. Lee, J. Geophys. Res. 87 (1982) 5063.
[158] G.M. Webb, T.J. Bogdan, M.A. Lee and I. Lerche, Mon. Not. R. Astr. Soc. 215 (1985) 341.
[159] L. Fisk and M.A. Lee, Astrophys. J. 237 (1980) 620.
[160] R.I. Epstein, Mon. Not. R. Astr. Soc. 193 (1980) 723.
[161] R. Cowsik and M.A. Lee, Proc. Roy. Soc. A. 383 (1982) 409.
[162] R.D. Blandford and D.G. Payne, Mon. Not. R. Astr. Soc. 194 (1981) 1041.
[163] Yu.Eh. Lyarbarski and R.A. Sunyaev, Sov. Astron. Lett. 8 (1982) 612.
[164] I. Lerche, Astrophys. J. 147 (1967) 689.
[165] Y-C. Chin and D.G. Wentzel, Astrophys. Space Sci. 16 (1972) 465.
[166] S. Ghosh, Ph.D. thesis, University of Maryland (1985).
[167] M.M. Hoppe, C.T. Russell, L.A. Frank, T.E. Eastman and F.U. Greenstadt, J. Geophys. Res. 86 (1981) 4471.
[168] E.A. Foote and R.M. Kulsrud, Astrophys. J. 233 (1979) 302.
[169] A. Achterberg, Astron. Astrophys. 119 (1983) 274.
[170] A. Achterberg and R.D. Blandford, Mon. Not. R. Astr. Soc. 218 (1986) 551.
[171] D.A. Tidman and N.A. Krall, Shock Waves in Collisionless Plasmas (John Wiley and Sons, Canada, 1971).
[172] J.F. McKenzie and K.O. Westphal, Phys. Fluids 11 (1970) 2350.
[173] E. Zweibel, Proc. AIP Conf. No. 56 (American Institute of Physics, New York, 1979).
[174] P.J. Morrison, J. Scott, G.D. Holman and J.A. Ionson, Unpublished preprint (1981).
[175] A. Achterberg, Astron. Astrophys. 119 (1983) 274.
[176] D.G. Wentzel, J. Geophys. Res. 82 (1977) 714.
[177] R.L. Dewar, Phys. Fluids 13 (1970) 2710.
[178] A. Achterberg, Astron. Astrophys. 98 (1982) 195.
[179] J.F. McKenzie and M.J. Völk, Astr. Astrophysics. 116 (1982) 191.
[180] M.J. Völk, L.O'C. Drury and J.F. McKenzie, Astron. Astrophys. 130 (1984) 19.
[181] V.N. Federenko, Astrophys. Sp. Sci. 96 (1983) 25.
[182] M.A. Lee and H.J. Völk, Astrophys. Sp. Sci. 24 (1973) 31.
[183] C.J. Cesarsky and R.M. Kulsrud, Proc. IAU Symp. 94. Origin of Cosmic Rays, eds. Setti Speda and Wolfendale (Reidel, Dordrecht, 1982).
[184] H.J. Völk and C.J. Cesarsky, Z. Naturforsch 37 (1982) 809.
[185] J.H. McKenzie and R.A.B. Bond, Astron. Astrophys. 123 (1983) 111.
[186] A. Achterberg and R.D. Blandford, Mon. Not. R. Astr. Soc. 218 (1986) 551.
[187] D.C. Montgomery and D.A. Tidman, Plasma Kinetic Theory (McGraw Hill, New York, 1964).
[188] C.F. Kennel and R.Z. Sagdeev, J. Geophys. Res. 72 (1967) 3303, 3327.
[189] D. Eichler, Astrophys. J. 229 (1979) 419.
[190] F. Hoyle, Mon. Not. R. Astr. Soc. 120 (1960) 338.
[191] W.I. Axford, E. Leer and A.G. Skabon, Astron. Astrophys. 111 (1982) 317.
[192] L.O'C. Drury and H. Völk, Astrophys. J. A248 (1981) 344.
[193] G.M. Webb, Astron. Astrophys. 127 (1983) 97.
[194] A. Achterberg, R.D. Blandford and V. Periwal, Astron. Astrophys. 132 (1984) 97.
[195] A.F. Heavens, Mon. Not. R. Astr. Soc. 210 (1984) 813.
[196] E. Dorfi, Adv. Sp. Res. 4 (1984) 205.
[197] L.O'C. Drury, Adv. Sp. Res. 4 (1984) 185.
[198] F.V. Coroniti, J. Plasma Phys. 4 (1970) 265.
[199] R.D. Blandford and D.G. Payne, MN 194 (1981) 1041.
[200] L.O'C. Drury, W.I. Axford and D. Summers, Mon. Not. R. Astr. Soc. 198 (1982) 833.
[201] D. Eichler, Astrophys. J. 277 (1984) 429.
[202] R.D. Blandford, Astrophys. J. 238 (1980) 410.
[203] A.F. Heavens, Mon. Not. R. Astr. Soc. 204 (1983) 699.
[204] G.M. Webb, Astrophys. J. 270 (1983) 319.
[205] D.C. Ellison, F.C. Jones and D. Eichler, J. Geophys. 50 (1981) 110.
[206] D. Eichler, Ap. J. 294 (1985) 40.
[207] D.C. Ellison, unpublished Ph.D. thesis, Catholic University (1981).
[208] D.C. Ellison and D. Eichler, Ap. J. 286 (1984) 691.
[209] M. Matsumoto and T. Sato (eds.), Computer Simulations of Space Plasmas (Reidel, Boston, 1985).

[210] D. Biskamp, Nuclear Fusion 13 (1973) 719.
[211] V. Formisano, J. de Phys., Colloque C5 Suppl. 12:38 (1977) C6-65.
[212] C.F. McKee and D. Hollenbach, Ann. Rev. Astron. Astrophys. 18 (1980) 219.
[213] E. Stone and B.T. Tsurutani (eds.), Proc. Chapman Conf. on Collisionless Shocks in the Heliosphere, J. Geophys. Res. 90 (1985) A1.
[214] L. Koch-Miramond and M.A. Lee, Adv. Sp. Res. 4 (1984).
[215] D.C. Montgomery and G. Joyce, J. Plasma Phys. 3, Pt. 1 (1969) 1.
[216] S.S. Moiseev and R.Z. Sagdeev, J. Nucl. Energy, Pt. C (1963) 43.
[217] D.W. Forslund and J.P. Friedberg, Phys. Rev. Lett. 27 (1971) 1189.
[218] T. Adam and J. Allen, Phil. Mag. 3 (1958) 449.
[219] L. David, R. Lust and A. Schluta, Z. Naturforsch. 13a (1958) 916.
[220] R.Z. Sagdeev, in: Reviews of Plasma Physics, Vol. 4, ed. M.A. Leontovich (Consultants Bureau, New York, 1966).
[221] M.M. Leroy, D. Winske, C.C. Goodrich and K. Papadopoulos, J. Geophys. Res. 87 (1982) 5081.
[222] M.M. Leroy, Adv. Sp. Res. 4 (1984) 231.
[223] D.W. Forslund, K.B. Quest, J.U. Brackhill and K. Lee, J. Geophys. Rev. 84 (1984) 2142.
[224] K.B. Quest, Phys. Rev. Lett. 54 (1985) 1872.
[225] J.P. Edmiston, C.F. Kennel and D. Eichler, Geophys. Res. Lett. 9 (1982) 531.
[226] J.T. Gosling, M.R. Thomsen, S.J. Bame, W.C. Feldman, G. Paschman and N. Skopke, Geophys. Res. Lett. 9 (1983) 1333.
[227] M. Tanaka, C.C. Goodrich, D. Winske and K. Papadopoulos, J. Geophys. Res. 88 (1983) 3049.
[228] S.P. Gary, J. Geophys. Res. 86 (1981) 4331.
[229] D.D. Sentman, C.F. Kennel and L.A. Frank, J. Geophys. Res. 86 (1981) 4365.
[230] D.D. Sentman, J.P. Edmiston and L.A. Frank, J. Geophys. Res. 86 (1981) 7487.
[231] M.A. Lee, J. Geophys. Res. 87 (1982) 5095.
[232] O.L. Vaisberg, A.A. Galeev, G.N. Zastenka, S.I. Limow, M.N. Nordrachev, R.Z. Sagdeev, A.Yu. Sokolov and V.D. Shapiro, Sov. Phys. JETP 85 (1983) 1232.
[233] T.E. Stringer, Plasma Phys. 6 (1964) 267.
[234] C.F. McKee, Phys. Rev. Lett. 24 (1970) 990.
[235] C.F. McKee, Phys. Fluids 14 (1971) 2164.
[236] D.W. Forslund and C.R. Shonk, Phys. Rev. Lett. 25 (1970) 1699.
[237] E.N. Parker, J. Nucl. Fus. Part C2 (1961) 146.
[238] C.F. Kennel and R.Z. Sagdeev, J. Geophys. Res. 72 (1967) 3302.
[239] R.D. Auer and H.J. Völk, Astrophys. Sp. Sci. 22 (1973) 243.
[240] K.B. Quest, D.W. Forslund, J.U. Brockhill and K. Lee, Geophys. Res. Lett. 10 (1983) 471.
[241] J.R. Kan and D.W. Swift, J. Geophys. Res. 88 (1983) 6919.
[242] M.E. Mandt and J.R. Kan, J. Geophys. Res. 90 (1985) 115.
[243] Papadopoulos, unpublished.
[244] K.B. Quest, Proc. Chapman Conf. on Collisionless Shocks, J. Geophys. Res. 90 (1985) A1.
[245] D.C. Ellison, Adv. Sp. Res. 4 (1984) 139.
[246] F.D. Kahn, J. Fluid Mech. 19 (1964) 210.
[247] R.C. Davidson, D.A. Hammer, I. Haber and C.E. Wagner, Phys. Fluids 15 (1972) 317.
[248] D. Biskamp and M. Welter, Nuclear Fusion 12 (1972) 663.
[249] K.I. Golden, L.M. Linson and S.A. Mani, Physics of Fluids 16 (1973) 2319.
[250] J.D. Scudder, L.F. Burlaza and E.W. Greenstadt, J. Geophys. Res. 89 (1984) A9.
[251] E.W. Greenstadt, Proc. Chapman Conf. on Collisionless Shocks, J. Geophys. Res. 90 (1985) A1.
[252] J.P. Meyer, Proc. 17th Intern. Conf. on Cosmic Rays (1981).
[253] K. Olive and D.N. Schramm, Ap. J. 257 (1982) 276.
[254] K. Hsieh and J.A. Simpson, Ap. J. Lett. 162 (1970) 491.
[255] F.M. Ipavich, J.T. Gosling and M. Scholer, J. Geophys. Res. 89 (1984) 1501.
[256] D. Ellison, J. Geophys. Res. 90 (1985) 29.
[257] D. Eichler and K. Hainebach, Phys. Rev. Lett. 47 (1981) 1560.
[258] L. Drury, Sp. Sci. Rev. 36 (1983) 57.
[259] A.A. Galeev, Adv. Sp. Res. 4 (1984) 255.
[260] C.F. Kennel, J. Arons, R.D. Blandford, F.U. Coroniti, M. Israel, L. Lanzerotti, A. Lightman, K. Papadopoulos, R. Rosner and F. Scarf, Unstable Current Systems and Plasma Instabilities in Astrophysics, Proc. IAU Symp. No. 108, B537, eds. Kundu and Holman (Reidel, Dordrecht, Holland, 1985).
[261] B.T. Tsurutani and P. Rodriguez, J. Geophys. Res. 86 (1981) 4319.
[262] M. Scholer, Adv. Sp. Res. 4 (1984) 419.
[263] J.A. Biretta, F.N. Owen and P.E. Hardee, Ap. J. Lett. 274 (1983) L27.

[264] A.M. Hillas, Ann. Rev. Astron. Astrophys. 22 (1984) 425.
[265] J.T. Gosling, J.R. Asbridge, S.J. Bame, W.C. Feldman, R.D. Zwickl, G. Paschmann, N. Sckopke and R.J. Hynds, J. Geophys. Res. 86 (1981) 547.
[266] B.T. Draine, Astrophys. J. 241 (1980) 1021.
[267] R.D. Blandford and D.G. Payne, Mon. Not. R. Astr. Soc. 194 (1981) 1033.
[268] R.L. Tokar, C.H. Aldrich, D.W. Forslund and K.B. Quest, Phys. Rev. Lett. 56 (1986) 1059.
[269] K. Quest, Proc. XXVIth COSPAR Symp., Toulouse (1986).
[270] D. Ellison and E. Möbius, Ap. J. (1987) in press.
[271] E. Möbius, D. Hovestadt, B. Klecker, M. Scholer, G. Gloeckler, F.M. Ipavish, C.W. Carlson and R.P. Lin, Geophys. Res. Lett. (1987) in press.