Particle acceleration in astrophysical jets

James H. Matthews¹, Anthony R. Bell² & Katherine M. Blundell³

¹Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA
²Clarendon Laboratory, Parks Road, University of Oxford, Oxford, OX1 3PU
³Astrophysics, Keble Road, University of Oxford, Oxford, OX1 3RH

Abstract

In this chapter, we review some features of particle acceleration in astrophysical jets. We begin by describing four observational results relating to the topic, with particular emphasis on jets in active galactic nuclei and parallels between different sources. We then discuss the ways in which particles can be accelerated to high energies in magnetised plasmas, focusing mainly on shock acceleration, second-order Fermi and magnetic reconnection; in the process, we attempt to shed some light on the basic conditions that must be met by any mechanism for the various observational constraints to be satisfied. We describe the limiting factors for the maximum particle energy and briefly discuss multimessenger signals from neutrinos and ultrahigh energy cosmic rays, before describing the journey of jet plasma from jet launch to cocoon with reference to the different acceleration mechanisms. We conclude with some general comments on the future outlook.

Keywords: particle acceleration, jets, magnetic fields, plasma physics, cosmic rays.

1. Introduction

High-energy, superthermal particles are present in a wide variety of astrophysical systems, particularly those that produce collimated jets of plasma. Indeed, most of what we know about jets comes from observations in the radio, X-ray and gamma-ray bands where relativistic, nonthermal particles are undergoing synchrotron and inverse Compton radiative losses via interactions with either magnetic fields or ambient radiation fields. A significant fraction of the jet’s energy budget can be converted into nonthermal particles and the radiation they emit. Remarkably, power-law spectra with cut-offs at various frequencies are observed in jetted systems that have central masses ranging from a few $M_\odot$ up to ~ $10^{10} M_\odot$. This behaviour hints at, but does not necessitate, universal acceleration physics.

Astrophysical jets are ideal sites for particle acceleration. They are often supersonic, so can produce strong shock waves. They create environments where velocity shear and turbulence are important, and also transport strong magnetic fields. Each of these effects can lead to the acceleration of high-energy particles in magnetic fields, where particles are accelerated by scattering in a turbulent velocity field or across a shock, shear layer or reconnection site. As we shall see, regardless of the mechanism, the energy increase normally results from moving charged particles through a $-\vec{v} \times \vec{B}$ electric field. The consequent energy gain can be balanced with particle escape from the acceleration site in such a way that a power-law spectrum of particle energies is produced, although this requires special circumstances.

This review is principally focused on how particles can be accelerated to high energies in the jets themselves, the shocks they produce and the turbulence they initiate. We discuss Fermi mechanisms in shocks, turbulence and sites of magnetic reconnection, with reference to other processes. The dominant acceleration process can vary both between and within each astrophysical system concerned, and in many cases the detailed acceleration physics is still debated. We generally avoid discussion of detailed observational characteristics except insofar that they inform our physical understanding of particle acceleration in jets. Instead, we refer the reader to the relevant review chapters on respective source types within this volume.

We begin (Section 2) by describing four key observational signatures of high-energy particle acceleration that a successful theory must explain. Section 3 contains the main theoretical discussion, outlining the basic physical principles behind particle acceleration to high energy. In Section 4, we discuss what constrains the maximum energy, before exploring what cosmic ray and neutrino data can tell us about jets and their related phenomena, in Section 5. We discuss our results in Section 6 by considering the journey of plasma along the jet length. Finally, we conclude in Section 7. We use centimetre-gram-second units with particle energies in eV, $u$ and $\Gamma$ to denote bulk velocities and Lorentz factors, $v_{pel}$ and $\gamma_{pel}$ to denote particle/electron velocities and Lorentz factors, $E$ for energies and $\vec{E}$ for electric field. We use the phrase cosmic ray (CR) to refer to nonthermal protons and ions.

2. Observations

We pick four key observational aspects of astrophysical jets: Images, synchrotron spectra, gamma-ray emission and time-resolved properties. We focus here on electromagnetic radiation and discuss CR and neutrino signals in Section 5. We will
use these radiative signatures to inform our discussion of the physics in the next Section.

2.1. Images

Images of jetted sources provide important clues about particle acceleration. In radio galaxies, jets can extend 100s of kpc from the central active galactic nucleus (AGN). The morphologies of the sources are, broadly speaking, split into two Fanaroff & Riley (1974) classes, denoted FRI and FRII. In the FRI sources, the synchrotron emission is brightest near the jet core, with the flux dropping outwards as the jet is disrupted, forming an extended plume of radio-emitting plasma. In the higher power FRII sources, the jet instead remains well-confined for much longer distances and deposits its energy in a termination shock. As with most dichotomies in astrophysics, the FRI-FRII distinction is blurred and a diverse continuum of properties exists that depends on the jet power, composition and environment (e.g. Ming0 et al., 2019); however, in general, the dependence of FRI/FRII morphology on the radio luminosity is a robust one (Fanaroff & Riley, 1974; Owen & Ledlow, 1994).

Radio galaxies produce nonthermal emission at a few distinct locations: compact core regions, the jet beam, lobes, plumes and, in FRIIs, hotspots. Very long baseline interferometry has allowed the jet of M87 (Virgo A) to be imaged in the radio on scales as small as 10 gravitational radii (Asada & Nakamura, 2012; Hada et al., 2016; Walker et al., 2018), revealing an edge-brightened transition with an apparent transition from parabolic to conical streamlines. A similar structure is observed in 3C 84 (Perseus A) at larger distances (100s of gravitational radii) from the central engine (Giovannini et al., 2018). On kiloparsec scales, multi-wavelength images at optical, radio and X-ray wavelengths commonly consist of knotty or localised bright spots that imply acceleration in special sites such as shocks (Bridle et al., 1994; Kraft et al., 2002; Perlman et al., 2011; Cara et al., 2013; Clautice et al., 2016; Perlman et al., 2019; Wykes et al., 2015; Snios et al., 2019), mixed with smoother emission that suggests distributed particle acceleration (Jester et al., 2001; Hardcastle & Croston, 2005; Hardcastle et al., 2007; Clautice et al., 2016; Hardcastle et al., 2016; Perlman et al., 2019). If the X-ray emission is synchrotron in origin, the short cooling time requires an in situ acceleration process (Perlman et al., 2001; Hardcastle et al., 2003; Jester et al., 2006), which is generally in-keeping with optical data (Jester et al., 2001). The hotspots are even more robust particle acceleration sites, based on the comparison between the synchrotron cooling time and light travel time, and are consistent with acceleration at the ‘working surface’ of a strong shock (e.g. Blandford & Rees, 1974; Scheuer, 1974; Meisenheimer & Heavens, 1986; Heavens & Meisenheimer, 1987; Laing, 1989; Meisenheimer et al., 1997; Brunetti et al., 2003; Araudo et al., 2015). In FRIs, there is no bright hotspot, but the particle acceleration may still occur primarily while the jet is supersonic and relativistic (Laing & Bridle, 2014). In most FRIs, the jet is then disrupted and forms a plume-like structure of synchrotron emitting plasma, whereas higher power lobed-FRIs (Laing & Bridle, 2012) and FRIIs create cocoons and backflows surrounding the jet and extending backwards from the hotspot (Hargrave & Ryle, 1974; Alexander & Leahy, 1987; Alexander, 1987; Carilli et al., 1991; Carilli & Barthel, 1996).

Although radio galaxies are perhaps the most spectacular examples of lobes and hotspots, X-ray binary systems produce similar structures (Mirabel et al., 1992; Mirabel & Rodríguez, 1999; Fender, 2001; Corbel et al., 2002; Blundell & Bowler, 2004; Jeffrey et al., 2016), often extending several parsecs into the interstellar medium. X-ray binary jet morphologies change depending on whether the jet is steady or transient (Fender, 2006). Transient jets consist of distinct ejections of blobs or plasmoids that can show superluminal motion (e.g. Mirabel & Rodríguez, 1994; Mirabel & Rodríguez, 1999; Miller-Jones et al., 2009; Bright et al., 2020), while corkscrew-like precession is observed in the slower-moving SS433 jet (e.g. Blundell & Bowler, 2004; Blundell et al., 2018). Protostellar systems such as the Herbig-Haro (HH) objects (Herbig, 1951; Haro, 1952) also produce lobes, hotspots and jet emission. Much of this morphology is actually seen in the optical through, e.g., Hz or [Oiii] emission from shock-heated regions (e.g. Lada, 1985; Reipurth & Heathcote, 1997; Reipurth & Bally, 2001), but radio observations of protostellar jets and HH objects do also reveal signatures of nonthermal emission from high-energy electrons (e.g. Carrasco-Gonzalez et al., 2010; Ainsworth et al., 2014; Masque et al., 2015; Anglada et al., 2018; Rodriguez-Kamenetzky et al., 2019). In some cases, these jets show striking similarities to those in AGN (Carrasco-Gonzalez et al., 2010).

Overall, images of jets tell us that in some sources there are particles being accelerated at the ‘working surface’ of a strong shock, but we must also be able to explain emission from (i) compact regions close to the jet launching point; (ii) both smooth and knotty emission along the jet beam and (iii) more extended plumes, lobes and cocoons clearly associated with jet activity. However, the emission from lobes and cocoons can often be explained by hydrodynamic effects, whereas hotspots, core regions and certain portions of the jet beam appear to be privileged sites for in-situ particle acceleration.1

2.2. Synchrotron power-laws

Perhaps the most universal observational signature of astrophysical jets is synchrotron radiation produced by relativistic electrons spiralling in magnetic fields. The exact spectrum observed depends on the distribution of electron energies, the observing waveband, the number of different radiating electron populations, the bulk Lorentz factor of the radiating plasma and whether the emission is subject to absorption effects. There is also significant diversity in spectral indices and degree of curvature; however, in general, power-law synchrotron spectra with spectral indices \( \alpha \approx 0 - 1 \) (where \( \alpha \) is defined such that \( F_\nu \propto \nu^{-\alpha} \), where \( F_\nu \) is the flux density in erg s\(^{-1}\) cm\(^{-2}\) Hz\(^{-1}\)) are common in jetted systems.

To illustrate this behaviour, broadband spectra for three systems with jets are shown in Fig. 1. In the GRB afterglow

\[ 1 \text{A more complete summary of the observational characteristics of radio galaxy, X-ray binary and protostellar jets can be found in Chapters 3, 5 and 10 of this anthology, respectively.} \]
from GW170817, an optically synchrotron power-law with $\alpha \approx 0.55$ extends from radio through to X-ray (Margutti et al., 2018). In the low-hard state observations of X-ray binary XTE J1118+480, the spectrum is characterised by a broken power-law and is only optically thin at IR/optical through X-ray frequencies (Hynes et al., 2000; Markoff et al., 2001). In the radio galaxy Cygnus A, the hotspot shows an optically thin power-law (Meisenheimer et al., 1997; Carilli et al., 1999; Brunetti et al., 2003), with hints of a low-frequency cutoff that has been confirmed with LOFAR observations (McKean et al., 2016). The spectrum exhibits curvature and deviates from a pure power-law which may be due to synchrotron cooling or nonlinear acceleration effects (the latter is discussed in Section 3.2.1). We also show a broadband spectral energy distribution (SED) for the BL Lac type blazar, Mrk 421, in Fig. 2, in which a synchrotron spectrum is observed from radio to X-ray frequencies (the gamma-rays are discussed in the next subsection).

One way of parameterising an electron population subject to synchrotron cooling losses is via a broken power-law (e.g. Heavens & Meisenheimer, 1987; Rybicki & Lightman, 1979; Longair, 1994), in which the spectrum steepens at high frequencies due to synchrotron losses and the low-frequency emission is self-absorbed. In this case, the spectral shape is given by

$$F_\nu \propto \begin{cases} \nu^{-\alpha_{\text{thick}}}, & \nu < \nu_0, \text{Optically thick} \\ \nu^{-\alpha_{\text{thin}}}, & \nu_0 < \nu < \nu_{\text{cool}}, \text{Optically thin} \\ \nu^{-\alpha_{\text{thin}}-1/2}, & \nu > \nu_{\text{cool}}, \text{Cooled} \end{cases}$$

where $\nu_0$ and $\nu_{\text{cool}}$ denote the frequencies above which the emission becomes optically thin and the emitting electrons’ synchrotron cooling timescale is less than the time since the electron was accelerated, respectively. In reality, model spectra show a gradual transition rather than sharp breaks (e.g Jaffe & Perola, 1973; Heavens & Meisenheimer, 1987; Brunetti et al., 2003). At low frequencies self-absorption produces either an inverted spectrum ($\alpha_{\text{thick}} \approx -5/2$; in the case of a single self-absorbed population of electrons) or a flat-spectrum with $\alpha_{\text{thick}} \approx 0$ (in the case of a conical self-absorbed jet; Blandford & Kland, 1979). At intermediate frequencies ($\nu_0 < \nu < \nu_c$), we typically observe $\alpha_{\text{thin}} \approx 0.5 - 0.8$. If this optically thin emission comes from a single population of electrons with a constant magnetic field $B$ and power-law distribution of particle energies $n(E)\mathrm{d}E \propto E^{-s}\mathrm{d}E$, then the relationship between $s$ and $\alpha$ is given by $\alpha_{\text{thin}} = (s - 1)/2$, implying $s \approx 2 - 2.6$.

The characteristic emission frequency of an electron with Lorentz factor $\gamma_e$ spiralling in a magnetic field is given by $\nu_c = \Gamma\gamma_e^2eB/(2m_e c)$, where $\Gamma$ is the bulk Lorentz factor of the flow. Since astrophysical jets can be highly relativistic, the critical frequencies ($\nu_0$, $\nu_{\text{cool}}$,...) can be boosted by large factors if the jet is beamed towards us. The low-frequency cutoff is determined by either the minimum Lorentz factor of the electrons ($\gamma_m$) multiplied by $\Gamma$, or absorption effects (e.g. McKean et al., 2016). In some cases, such as at early times in GRBs, the ‘fast-cooling’ limit is realised and $\nu_c(\gamma_m) > \nu_{\text{cool}}$ (Sari et al., 1998; Granot et al., 2000). Steepening at high-frequency can be caused by synchrotron or inverse Compton cooling and, in GRBs in particular, the time evolution of this break provides useful information about particle populations; however, it has been shown that in radio galaxies the high-frequency cutoff can reflect the maximum particle energy rather than cooling effects (see e.g. Section 4, Araudo et al., 2016, 2018). The break frequency between optically thick and optically thin emission ($\nu_0$) can also be used to infer information about jet dynamics and sites of particle acceleration, in, e.g. X-ray binary jets (Markoff et al., 2001; Gandhi et al., 2011; Koljonen et al., 2015).

2.3. Gamma-ray emission

Gamma-ray emission has been observed in a wide variety of jetted systems, with examples including GRB prompt emission (e.g. Klebesadel et al., 1973; Gehrels et al., 2009), microquasars such as SS433 and Cygnus X-3 (Araudo et al., 2011; Corbel et al., 2012; Bordas et al., 2015; Abeysekara et al.,

Figure 1: a) Broadband spectrum at two different epochs of the afterglow associated with neutron-star merger and short GRB event GW170817, from Margutti et al. (2018). b) Broadband spectrum of the X-ray binary XTE J1118+480 in its low-hard state, compared with an illustrative broken power-law spectrum with $\alpha_{\text{thick}} = 0.8$, $\alpha_{\text{thick}} = -0.1$, synchrotron cooling break at $1.2 \times 10^{20}$ Hz and a low-frequency exponential cut-off at $1.2 \times 10^8$ Hz (Hynes et al., 2000; Markoff et al., 2001, M01). c) Radio spectrum of the southwestern hotspot of Cygnus A (Carilli et al., 1999), with the DSA model from Brunetti et al. (2003, B03) marked in grey. The dashed line shows a power-law with $\alpha = 0.7$. The data are digitised from the respective publications.
Gamma-rays can be produced by both leptonic and hadronic interactions. In the leptonic case, the radiation is due to the inverse Compton (IC) interactions of energetic electrons with a radiation field such as the CMB, thermal emission from the accretion process, or the electrons’ own synchrotron radiation (e.g. Dermer & Schlickeiser, 1993). In the latter case, this is referred to as synchrotron self-Compton (SSC). An illustrative synchrotron plus SSC model is shown in Fig. 2. One-zone formulations of synchrotron-Compton models are commonly invoked to fit blazar spectra (e.g. Maraschi et al., 1992; Kirk et al., 1998; Ghisellini et al., 2010), but these models can encounter problems when fitting the spectral shape (e.g. Bednarek & Protheroe, 1997), or by requiring that the synchrotron cooling time is longer than the variability timescale (e.g. Baloković et al., 2016). However, more complete treatments in which the electron population is self-consistently evolved along a model jet can provide good fits to blazar spectra (Potter & Cotter, 2012, 2013a,b; Morris et al., 2018). Other leptonic models have been proposed for gamma-ray production, such as $e^+e^-$ pair cascades (Blandford & Levinson, 1995; Bednarek, 1997). In hadronic models, gamma-rays are produced when high-energy protons interact with either other protons, photons or magnetic fields, with the relative importance of each depending on the density, radiation energy density and magnetic field energy density (e.g. Mannheim, 1993; Mücke et al., 2003; De Angelis & Mallamaci, 2018). In $p\gamma$ interactions, neutral pions are produced via $p\gamma \rightarrow p\pi^0$, and pairs are produced via the Bethe-Heitler process, $p\gamma \rightarrow pe^+e^-$; subsequent pion decay ($\pi_0 \rightarrow \gamma\gamma$) or annihilation then produces gamma-rays. Inelastic $pp$ collisions also create neutral pions which decay in the same manner. Each of the pion production channels also leads to charged pions, which ultimately create neutrinos, meaning hadronic gamma-rays are intrinsically linked to high-energy neutrinos (see Section 5). Finally, protons can produce synchrotron radiation when the magnetic field is strong enough (Aharonian, 2000; Aharonian et al., 2002; Biteau et al., 2020).

Distinguishing between hadronic and leptonic gamma-ray emission is challenging, but, at least in principle, gamma-rays offer one of the few possible ways to study the composition of jets on small scales, particularly in GRBs and blazars. Constraining jet composition is important in order to (i) help differentiate between launching mechanisms (for example, whether the angular momentum extracted originates from a BH ergosphere [Blandford & Znajek 1977] or accretion disc [Blandford & Payne 1982]); (ii) understand multimessenger signals from UHECRs/neutrinos (see Section 5) and (iii) accurately estimate the energy partitioning in the jet and resulting effect on surroundings.

2.4. Variable sources and transients

We can also use the time domain to learn about particle acceleration. In GRBs or X-ray binaries, where dynamical timescales are short compared to those in AGN, time-resolved observations can be used to produce light-curves over the full lifetime of the outburst or explosion. However, short-timescale changes can also occur in AGN, examples being blazar flares (Aharonian et al., 2007), jet variability in e.g. M87 (Giannios et al., 2010) and Pictor A (Marshall et al., 2010), and even variability in radio galaxy hotspots (Hardcastle & Croston, 2005).

Blazars are split into two main classes, BL Lacs and flat-spectrum radio quasars (FSRQs). Similarly to FRI/FRII radio galaxies, these classes can be divided approximately by their jet power (Celotti & Ghisellini, 2008), with FSRQs the more powerful population. Extremely fast variability has been detected in both classes of blazar, with ~ 10 minute flares observed at TeV energies in, e.g., the BL Lac PKS 2155-304 (Aharonian et al., 2007) and at GeV energies in, e.g., the FSRQ 3C 279 (Ackermann et al., 2016b). This behaviour requires that any complete model for particle acceleration can produce rapid variability in some parts of the blazar jet.

Intriguing timing properties are also seen in X-ray binaries. During a recent outburst of V404 Cygni, Gandhi et al. (2017) found 0.1 second lags between the X-ray and optical variability, associating this time lag with the distance from the black hole.
to the optical-emitting region in the jet. In GRS 1915+105, de Gouveia dal Pino & Lazarian (2005) suggested that the supertlimal motion seen in the jet (Mirabel & Rodríguez, 1994; Mirabel & Rodríguez, 1999) can be explained by fast reconnection occurring just above the black hole, which can cause rapid X-ray variability. In both V404 Cygni and GRS 1915+105, the models invoked draw upon blazar phenomenology and theory in an attempt to unify some of the physical mechanisms at work in both X-ray binaries and AGN.

Time-resolved observations of GRBs\(^3\) reveal two main populations: short-duration GRBs, which were recently confirmed to be associated with a binary neutron star merger (Abbott et al., 2017); and long-duration GRBs, thought to be caused by collapsars (Galama et al., 1998; Iwamoto et al., 1998; MacFadyen & Woosley, 1999). The classes are split by the length of the prompt emission, with the dividing line at \( \approx 2s \) (Kouveliotou et al., 1993). Following the prompt emission, there is a longer-lived ‘afterglow’ phase of synchrotron emission from a shock front (e.g. Waxman, 1997; Sari et al., 1998; Piran, 1999; Yost et al., 2003; Kumar & Zhang, 2015). The evolution of the spectrum can be used to infer information about the particle distribution function and magnetic field (Sari et al., 1998; Wijers & Galama, 1999), while breaks in the power-law slopes of the light curves provide useful constraints on the jet dynamics and geometry (Rhoads, 1997; Sari et al., 1999; Kumar & Panaitescu, 2000; De Colle et al., 2012). X-ray binary sources also show characteristic rise and decay phases (e.g. Fender et al., 1999; Fender et al., 2006; Corbel et al., 2012; Kim & Kim, 2016), with the morphology and spectrum of the radio emission depending on the accretion state. In hard or quiescent states, optically thick emission from a compact jet is observed (Fender, 2001; Fender et al., 2004; Gallo et al., 2005; Russell et al., 2013), with \( \alpha \lesssim 0 \), and the transient ejections described in section 2.1 are then produced as the source crosses into the soft state (Mirabel & Rodríguez, 1999; Fender et al., 2004) often in combination with an optically thin radio flare. Discrete ejection events can be modelled using the ‘synchrotron bubble’ model (van der Laan, 1966), or variants thereof (Hjellming & Johnston, 1988; Tetarenko et al., 2017; Tetarenko et al., 2019), but it can be necessary to account for more complex time-dependent spectral evolution (e.g. Atoyan & Aharonian, 1999; Miller-Jones et al., 2004). While the variability is often primarily driven by accretion physics and jet dynamics, good models for the acceleration and subsequent evolution of the synchrotron electrons are important.

3. Theory of Particle Acceleration

There already exist a number of reviews covering the theory of particle acceleration; some relevant examples include those focusing on shock acceleration (Drury, 1983; Blandford & Eichler, 1987; Schure et al., 2012; Bell, 2014; Marcowith et al., 2016) and magnetic reconnection (Hoshino & Lyubarsky, 2012; de Gouveia Dal Pino & Kowal, 2015; Kagan et al., 2015), as well as more general discussions of astrophysical jets (Begelman et al., 1984; Mirabel & Rodríguez, 1999; Fender, 2006; Bykov et al., 2012; Romero et al., 2017; Blandford et al., 2018), particle acceleration simulations (Marcowith et al., 2020) or high-energy astrophysics (Longair, 1994). For this reason, we only cover some of the essential theory and refer the enthusiastic reader to these efforts. Much of the physics is quite general and not necessarily specific to jets.

We begin by writing the equation for the rate of change in momentum of a particle with Lorentz factor \( \gamma \) and velocity \( \vec{v}_p \):

\[
\frac{d}{dt}(\gamma m \vec{v}_p) = e(\vec{E} + \vec{v}_p \times \vec{B}),
\]

(1)

where \( \vec{E} \) is the electric field and \( \vec{B} \) the magnetic field. The motion of free charges means that large electrostatic fields are difficult to maintain in the mostly ionized media that are common in astrophysical systems, so the electric field is often just \(-\vec{v} \times \vec{B}\), where \( \vec{u} \) is now the bulk velocity of the plasma. It is this \(-\vec{v} \times \vec{B}\) term that leads to particle acceleration. Given the magnetic nature of the acceleration, it is also useful to define the Larmor radius (or gyroradius) of the particle, \( r_g = p/(ZeB) \), where \( Ze \) is the charge on the particle and \( p \) the momentum. For relativistic particles with energy \( E = pc \) this becomes

\[
r_g = \frac{E}{Ze c B},
\]

(2)

This gives the first constraint on acceleration mechanisms: the acceleration site must have a size \( R \) larger than \( r_g \). We call this the confinement condition.

Let us now consider a particle undergoing some form of scattering process in which it gains energy. We assume we have \( N_0 \) particles in the process, with initial energy \( E_0 \). Each time the particles undergo a collision, they gain a fractional increase in energy of \( \beta_i \) such that after \( m \) collisions they have energy \( E = E_0 \prod_{i=1}^{m} \beta_i \). Similarly, if they have a probability \( P_i \) of remaining in the acceleration region after each collision, then the number of particles left after \( m \) collisions, with energy \( E \), is \( N = N_0 \prod_{i=1}^{m} P_i \). We can therefore write

\[
\frac{\ln(N/N_0)}{\ln(E/E_0)} = \frac{\sum_{i=1}^{m} \ln P_i}{\sum_{i=1}^{m} \ln \beta_i}.
\]

(3)

Observed synchrotron spectra (Section 2.2) and the cosmic ray spectrum (Section 5) require a power-law distribution for the differential number of particles \( n(E) = dN/dE \). A power-law distribution for \( n(E) \) is only possible when \((\sum_{i=1}^{m} \ln P_i)/(\sum_{i=1}^{m} \ln \beta_i))\) is constant regardless of the number of collisions. One way of achieving this is if \( P_i \) and \( \beta_i \) are constant values with increasing particle energy, which we label \( P \) and \( \beta \). In this special case, one obtains a power-law of the form

\[
n(E) \ dE \propto E^{(\ln P/\ln \beta)-1} \ dE.
\]

(4)

This equation shows that particles undergoing stochastic collisions can produce a power-law distribution of particle energies whose spectral index depends on the energy gain per collision and escape probability, but only if \( P \) and \( \beta \) remain constant.

---

\(^3\)See Chapter 7.
with increasing energy. The fractional energy gain might feasibly be roughly constant with increasing energy, but arranging for \( P_i \) to be independent of energy is difficult. The escape probability per collision \((1 - P_i)\) is the ratio of the scattering time (the interval between scattering events), \(\Delta t\), to the escape time, \(\tau_{\text{esc}}\). The scattering time is \(\Delta t = \lambda/c \propto r_s/c\), where \(\lambda\) is the mean free path. If the escape is diffusive, then \(\tau_{\text{esc}} \propto r_s^{-1}\), which then implies \((1 - P_i) \propto r_s^2 \propto E^2\). As we shall see, one way of getting around this apparent difficulty is via shock acceleration, which provides a physically motivated reason for \(P_i\) to be constant with increasing energy. Acceleration can also come from other electric fields and non-stochastic processes (see in particular Section 3.3). Nonetheless, this simple model with escape balancing energy gain provides a useful framework that aids discussion of a variety of different acceleration mechanisms.

### 3.1. Second-order Fermi Acceleration

In second-order Fermi acceleration (Fermi II; Fermi, 1949), particles scatter off magnetised clouds or in turbulent velocity field with random velocities. This process can be understood by considering a series of ‘magnetic mirrors’ that either reflect the particle or isotropise the particle distribution function in the rest frame of the scattering centre. In Fermi’s approach, the scatterers are ‘clouds’ moving with velocity \(u_c\), and relativistic particles gain or lose a fraction \(-u_i/c\) of their energy per collision. The energy increase comes from the fact that head-on collisions are more likely. For relativistic test particles, the average energy gain per collision is found by transforming from the lab frame to the cloud frame and back, then integrating over the pitch angle distribution (e.g. Longair, 1994), and is given by

\[
\frac{\Delta E}{E} = \frac{8}{3} \left( \frac{u_i}{c} \right)^2.
\]

Alternatively, if the energy gain per collision is small, Fermi acceleration can be treated using the Fokker-Planck equation (e.g. Blandford & Eichler, 1987). Although Fermi’s original discussion considers clouds, subsequent elaborations (e.g. Kul- srud & Ferrari, 1971; Melrose, 1980) have instead considered scattering off MHD waves, in which case the relevant velocity becomes the Alfvén speed, \(v_A = B/\sqrt{4\pi \rho}\). The acceleration rate for Fermi II is second-order in \((u_i/c)\) or \((v_A/c)\). As a result, Fermi II is slow and a particle must be confined for a long time if the energy is to reach a large value, except in special conditions (e.g. Jones, 1994). Furthermore, as discussed above, there is no obvious reason why the escape time from a volume of turbulent plasma should scale with the scattering time so as to achieve a constant escape probability – in fact, the opposite behaviour \((\Delta t \propto 1/\tau_{\text{esc}})\) is expected if the escape is diffusive. Even if this can be accomplished, there is then no a priori reason why a given spectral index should be produced. As a result, a Fermi II model requires significant fine-tuning in order to produce the observed power law spectra and their spectral index values.

### 3.2. Shock Acceleration

Particles can be accelerated by a few different mechanisms at collisionless shock waves: shock drift acceleration (SDA; Begelman & Kirk, 1990; Decker & Vlahos, 1985; Chalov, 2001), shock-surfing acceleration (SSA; Sagdeev, 1966; Katsooulas & Dawson, 1983; Lee et al., 1996; Zank et al., 2001; Shapiro & er, 2003) and diffusive shock acceleration (DSA, described below). These mechanisms and the relevant microphysics is summarised by Marcowith et al. (2016). SDA occurs in quasi-perpendicular shocks and involves the drifting of particles along the shock front; the drifting is effectively a \(\mathbf{vB}\) drift introduced by the compression of the magnetic field across the shock front that leads to tighter gyrations in the downstream field with respect to the upstream. SSA is similar, and occurs when ions are ‘reflected’ by the potential of the shock and then return from upstream. SSA can only accelerate particles to mildly suprathermal energies since the particle quickly becomes energetic enough to overcome the shock potential. A single episode of SDA also normally leads to only a small energy gain because the particle drifts away from the shock. To reach highly superthermal energies in a quasi-perpendicular shock a particle must return to the shock multiple times and undergo repeated episodes of SDA. This version of first-order Fermi acceleration is thus similar to conventional DSA; the difference between SDA and DSA is in many ways a superficial and semantic one, and first-order Fermi acceleration at shocks does not need to involve only drifts or only diffusion (e.g. Jokipii, 1982).

We discuss drifts and SSA further in the context of injection in Section 3.2.3. In this section, we discuss the general features of first-order Fermi acceleration at shocks using a DSA framework.

The theory of DSA was laid out in four landmark papers (Axford et al., 1977; Krymskii, 1977; Bell, 1978a; Blandford & Ostriker, 1978), while similar mechanisms had earlier been described by Jokipii (1966) and Fisk (1971). The presence of a shock allows for a situation in which ‘collisions’ are essentially always head-on, although the actual scattering process is deflection in a turbulent magnetic field. To illustrate this, let us consider an already relativistic test particle, such that we can set \(v_p = c\) and \(E = pc\), encountering a non-relativistic, collisionless shock with upstream and downstream velocities \(u_1\) and \(u_2\) (see Fig. 3). We follow Bell (1978a) in considering a steady state solution where the number of particles being injected at the shock is equal to the number being carried away downstream, and we assume that all particles from upstream are scattered back towards the shock by self-generated turbulence.

In this scenario, the particle’s energy as it crosses from the upstream to downstream regions is given by a Lorentz transformation such that \(E' = \gamma(E + 3pu_2\cos \theta)\), since for a strong shock the downstream plasma has a velocity of \(3u_2\) relative to the upstream plasma. Here \(\theta\) is the ‘pitch angle’, the angle with respect to the shock normal. An integration over pitch angles reveals that the fractional energy gain per half cycle is \(2\eta_2/c\), and thus the quantity \(\ln(\beta)\), which includes a factor of two for the full cycle there and back across the shock, is given by

\[
\ln \beta = \ln \left(1 + \frac{4\eta_2}{c}\right) \sim \frac{u_2}{c}.
\]
state, the number of particles being removed from the shock by being swept downstream is \( N u_2 \), whereas the number of particles crossing the shock is \( N c/4 \). Thus the fraction of particles lost from the acceleration site per unit time \( 1 - P \) is \( 4Nu_2/Nc \) and we can write
\[
\ln P = \ln \left( 1 - \frac{4u_2}{c} \right) \sim -\frac{u_1}{c}.
\] (7)

Crucially, neither the rate at which particles leave the acceleration region nor the fractional energy gain they experience depends on energy. Both are instead constants that depend only on the shock velocity. Combining these two expressions with equation 4 then results in a spectral index of \( s = 2 \). To obtain this spectral index, we have assumed a strong shock (Mach number, \( M \rightarrow \infty \)) and an adiabatic index of \( 5/3 \) such that the shock compression ratio \( \chi = 4 \), as well as fully relativistic particles, but the more general result is a power-law in momentum in the particle distribution function such that
\[
f(p) \propto p^{-3\chi/(\chi - 1)},
\] (8)
which gives \( f(p) \propto p^{-4} \) for \( \chi = 4 \). The relationship to the differential number density is \( n(p) = 4\pi p^2 f(p) \) for an isotropic particle distribution, so \( n(p) \propto p^{-2} \).

DSA is attractive as an acceleration mechanism because the energy gain is first-order in shock velocity, and because the balance between escape and fractional energy gain is hardwired by the shock jump conditions; thus, the index \( \chi \) in the power-law energy distribution does not require fine-tuning and there is a good reason why a power-law with a fairly universal spectral index might exist. Since the theory was first developed in the late 1970s, the theoretical understanding of DSA has developed considerably. We therefore discuss considerations that go beyond the simple treatment we have outlined, by considering non-linearity, injection and relativistic shock velocities with reference in particular to recent progress using numerical simulations.

### 3.2.1. Non-linearity and self-generated turbulence

Non-linearity in shock acceleration is introduced via the back-reaction of the high-energy particles on the conditions of the background plasma. This is inevitable if particle acceleration is efficient since a relatively high fraction of the energy density will be contained in superthermal particles. In this case the CR pressure must be included in the fluid equations and this modifies the shock structure (Axford et al., 1977; Drury & Voelk, 1981; Drury, 1983; Drury & Falle, 1986). Applying a two-fluid model reveals two main effects: the smoothing of the shock on the diffusion length, and an increase of the compression ratio \( \chi \) due to the relativistic CR equation of state (adiabatic index of \( 4/3 \)). The CRs can also carry away internal energy from the shock region, effectively acting as a coolant and increasing the compression ratio further. An increased compression ratio acts to flatten the spectrum at high energies (see e.g. Ellison, 1995).

Non-linearity is also introduced via waves or plasma instabilities excited by CRs. This is a crucial ingredient of DSA theory, since turbulence is needed to scatter the particles back across the shock front, as noticed by Bell (1978a) and Blandford & Ostriker (1978). Both originally invoked Alfvén waves excited by CRs streaming with super-Alfvén velocities ahead of the shock. This effect and its derivatives are often referred to as the resonant, streaming or Alfvén instability, so-called because it grows at wavenumbers resonant with the Larmor radius of the streaming CRs (Lerche, 1967; Kulsrud & Pearce, 1969; Wentzel, 1974; Skilling, 1975a,b,c). The resonant instability...

---

Figure 3: Diagram showing, schematically, how particles are accelerated in DSA (left) and in Fermi II (right). Bulk velocities are shown with orange dashed lines, crude particle trajectories in red dashed lines and magnetic field lines as black arrows. In the DSA case, the velocities are in the shock rest frame.
saturates at $\delta B/B \sim 1$ (McKenzie & Voelk, 1982; Amato & Blasi, 2009).

A key physical limit of diffusive particle transport is that of Bohm diffusion, when the scattering mean free path is roughly equal to the Larmor radius of the particle. In this case, the diffusion coefficient is given by $D_B = \tau_c c/3$. In supernova remnants (SNRs), there is observational evidence that Bohm diffusion applies (Stage et al., 2006; Uchiyama et al., 2007), but this can only be the case if the turbulent field is comparable to the uniform field, i.e. $\delta B/B \sim 1$. As a result, turbulent magnetic field amplification is a general feature of DSA theory and can affect the maximum attainable energy (see Section 4) as well as the spectrum of the accelerated particles (Caprioli et al., 2009; Bell et al., 2019a; Malkov & Aharonian, 2019).

In high Mach number shocks, there is an instability that grows faster than the resonant instability at wavenumbers nonresonant with the Larmor radius, known as the non-resonant hybrid (NRH) or Bell instability (Lucek & Bell, 2000; Bell, 2004, 2005). The basic physics of the instability is that a return current $-j_{r\delta}$ produced in reaction to streaming CRs stretches and distorts the field via $-\delta j_{r\delta} \times \vec{B}$ force. Numerical simulations have shown that the instability grows exponentially even in the non-linear regime and amplifies the magnetic field to $\delta B/B \sim 10 – 100$ (e.g. Zirakashvili et al., 2008; Reville et al., 2008; Riquelme & Spitkovsky, 2009, 2010; Matthews et al., 2017), although the saturated field value depends on the particle energy regime considered. Caprioli & Spitkovsky (2014b) find, using PIC simulations, a transition from resonant to non-resonant instability at an Alfvénic Mach number, $M_A$, of 30; in their study, the magnetic field is amplified in both regimes with an amplification factor that scales as $\sqrt{M_A}$. One important general property of the NRH instability is that the scale-size of the turbulence grows until it reaches the Larmor radius of the particles driving the instability, thus providing a natural way for the Bohm limit to be realised. Although most of the work on the NRH instability has focussed on SNRs, it has also been applied to jets in radio galaxies (Araudo et al., 2018) and GRBs (Milosavljević & Nakar, 2006).

### 3.2.2. Relativistic shocks

Given that astrophysical jets are often relativistic, we must consider how the original analysis changes as the shock velocity enters the relativistic regime. One of the main differences is that the energy spectrum is steeper than the non-relativistic case; Kirk et al. (2000) find $s = 2.23$ and Achterberg et al. (2001) find $s = 2.2 – 2.3$, while Sironi et al. (2013) find an even steeper spectrum from particle-in-cell (PIC) simulations. The reason for the steeper spectral index compared to non-relativistic shocks is due to a combination of anisotropy in the particle distribution function and the time available for scattering (see e.g. Bykov et al., 2012, for a summary). Although the simple diffusion approximation employed above breaks down, the spectral steepening can be naively understood as due to a smaller value of $\beta$ compared to non-relativistic shocks (Achterberg et al., 2001). In addition, particles in the upstream region do not diffuse far upstream and instead are rapidly overtaken by the shock, before the direction of the particle has been significantly altered, while particles in the downstream region are rapidly advected downstream. Thus, in both the upstream and downstream regions CRs have limited time available to generate a turbulent magnetic field.

The ability of CRs to generate turbulence is crucial and at very high bulk Lorentz factors this effect can prohibit DSA occurring at all (Sironi & Spitkovsky, 2014). At an ultra-relativistic shock ($u_i \rightarrow c, \Gamma \gg 1$), the combination of quasi-parallel magnetic field, steep CR spectrum and reduced time available before the shock overtakes, or advects away, the CRs means that both the scale-size and magnetic field amplification associated with the turbulence generated by streaming or drifting CRs are severely limited. As a result, diffusion occurs well above the Bohm regime and the maximum particle energy is lowered accordingly (Lemoine & Pelletier, 2010; Reville & Bell, 2014; Bell et al., 2018, see also Section 4).

#### 3.2.3. Injection

The DSA theory outlined above assumes that the particle being injected in the system has $v_p \gg u_i$ and is already relativistic—that is, the injection of the particle from the thermal pool onto the superthermal power-law tail is not considered. This “injection problem” is a quite general issue for Fermi-type processes, although much work has been devoted to understanding the injection into the DSA process in particular (e.g. Bell, 1978b; Giacalone et al., 1993; Levinson, 1996; Gruzinov & Waxman, 1999; Zank et al., 2001; Nishikawa et al., 2003; Spitkovsky, 2008; Caprioli & Spitkovsky, 2014a; Marcowith et al., 2016). The injection process is quite different to DSA at the highest-energies and requires a full kinetic treatment that includes the backreaction on the fluid. Injection is therefore a highly-nonlinear problem that is often studied using PIC simulations. Over the past decade or so, PIC and hybrid MHD-PIC simulations have been hugely successful in modelling the DSA process at low energies and self-consistently producing the turbulence that is so critical for its operation (e.g. Spitkovsky, 2008; Sironi & Spitkovsky, 2009b,a, 2011; Riquelme & Spitkovsky, 2011; Caprioli & Spitkovsky, 2014a,b; Caprioli et al., 2015; Bai et al., 2015; Crumley et al., 2019).

A variety of injection mechanisms have been proposed, often involving microinstabilities excited at the shock. For example, the Weibel instability (Weibel, 1959) can mediate collisionless shocks and lift particles from thermal to suprathermal or mildly superthermal energies (Spitkovsky, 2008). Other possibilities are that particles are injected by interaction with whistler waves (e.g. Baring, 1991; Riquelme & Spitkovsky, 2011), or, if the shock is locally quasi-perpendicular, via SDA (e.g. Ball & Melrose, 2001; Park et al., 2015). Injection efficiencies can be expected to vary between electrons and ions for non-relativistic shocks since at non-relativistic temperatures the electrons have gyroradii a factor $\sqrt{m_e/m_p}$ smaller than protons and it is harder to scatter them across the shock (e.g. Bell, 1978b). PIC simulations do show that protons are injected more efficiently than ions (Sironi & Spitkovsky, 2011; Park et al., 2015; Crumley et al., 2019), with the electron efficiency increasing as the electron temperature $k_B T_e$ approaches $m_e c^2$ (Sironi & Spitkovsky, 2011; Crumley et al., 2019). Possible observational evidence
for this phenomenon can be found by comparing the energy fraction in electrons at non-relativistic supernova shocks, where it is low (Morlino & Caprioli, 2012), to relativistic GRB shocks, where it is higher (Wijers & Galama, 1999).

Particles may also be injected by being reflected from the shock, due to either the electrostatic shock potential or magnetic mirror effects at oblique shocks (Hudson, 1965). These ions subsequently return from upstream with a fractional energy gain on the order of a few (Paschmann et al., 1980; Tsurutani & Rodriguez, 1981), a process that can be repeated until the particle can overcome the shock potential, which is often referred to as shock-surfing (e.g. Shapiro & er, 2003). Processes involving reflection preferentially inject ions since their Larmor radii are larger with respect to the shock width compared to that of the electrons, which instead get trapped in the shock transition region. Counter-streaming ions can induce microinstabilities, for example the modified two-stream instability (McBride et al., 1972; Matsukio & Scholer, 2003, 2006) and the Whistler modes mentioned above, which help inject electrons and further ions. While PIC simulations have provided many important recent advances regarding injection, the detailed injection physics for shock acceleration of both electrons and ions is extremely complicated and depends on the detailed shock characteristics (see Marcowith et al., 2016, for a summary).

3.3. Magnetic Reconnection

Magnetic reconnection has recently gained a lot of attention as a particle acceleration mechanism in jets, partly because it may help to explain how magnetic energy is converted into radiation in the bases of jets that are initially Poynting flux dominated – a phenomenon dubbed “magnetoluminescence” by Blandford et al. (2017). The magnetic energy is liberated and converted into bulk motion, heat, and energetic particles. A simple schematic of a reconnection site is shown in Fig. 4, showing how converging field lines of opposite polarity approach each other and undergo resistive dissipation. The characteristic shape of the field lines means that the central reconnection region is often referred to as an ‘X-point’ or ‘X-line’. Magnetic reconnection is often described using two main non-relativistic frameworks: Sweet-Parker reconnection (Parker, 1957; Sweet, 1958), where a long current sheet forms between the converging field lines, or Petschek reconnection (Petschek, 1964), where slow shocks form just outside the X-point region.

Sweet-Parker reconnection is extremely slow, even when kinetic effects are accounted for (Comisso & Asenjo, 2014), and it is well known that the solar flares in the sun require reconnection rates orders of magnitude faster than the Sweet-Parker value (e.g. Isobe & Shibata, 2009; Kagan et al., 2015; Galtier, 2016). Therefore, one of the fundamental questions relating to magnetic reconnection is how fast reconnection can take place, which is generally thought to happen via tearing/plasmoid instabilities that cause the current sheet to fragment into a series of magnetic islands (e.g. Loureiro et al., 2007; Bhattacharjee et al., 2009; Uzdensky et al., 2010; Huang & Bhattacharjee, 2012; Loureiro et al., 2012; Ni et al., 2015) and/or turbulent reconnection (e.g. Lazarian & Vishniac, 1999; Lazarian, 2005; Onofri et al., 2006; Kowal et al., 2009; Loureiro et al., 2009; Lazarian et al., 2012; Kowal et al., 2017; Isliker et al., 2017). Some of the studies discussed here focus on relativistic reconnection, defined as the regime where the magnetic energy density exceeds the rest mass energy density of the plasma, a regime that is relevant to jets. Fortunately, the non-relativistic models can be generalised to the relativistic case (Lyubarsky, 2005).

Particles can be accelerated in a number of different ways close to a reconnection site, as summarised by e.g. Blandford et al. (2017); Drake et al. (2018). The first is by ‘direct’ mechanisms in an electric field associated with the current sheet (Litvinenko, 1996, 1999; Kirk, 2004; Guo et al., 2016). Acceleration in this electric field is generally thought to be fairly inefficient (Dahlin et al., 2016, 2017), but it is probably an important injection mechanism (Lyutikov et al., 2016; Li et al., 2016; Comisso & Sironi, 2018). The second is by betatron acceleration, where the first adiabatic invariant (p²/B) of a particle is conserved in a slowly-increasing magnetic field such that the particle energy increases, but this is generally unimportant in this case since B is generally decreasing. The third is by Fermi-type mechanisms that can be first-order in (u/c), where, in the case of relativistic reconnection, u is a significant fraction of c.

Magnetic reconnection creates a situation where two regions of opposing magnetic polarity approach each other at a velocity u₀. This is a convergent flow that is closely analogous to that in a shock. If particles cross from the lower to upper regions they acquire an energy gain in much the same way as in DSA. The relative velocity in the rest frame of either of the converging two regions is 2u₀, so if we average over an isotropic pitch angle distribution (factor of 1/3) and consider the round trip (factor of 2) then we have (AE/E) ~ 8/3u₀/c, and β = 1 + (8/3)(u₀/c), identical to the case of Fermi II except that the energy change is first-order in (u₀/c). Using a simple approximation for particle escape, de Gouveia dal Pino & Lazarian (2005); de Gouveia Dal Pino & Kowal (2015) derive P = 1 − 4u₀/c. Combining these values with equation 4 gives a predicted spectrum of s = 2.5, close to the inferred index from some synchrotron spectra. However, Drury (2012) notes that these assumptions about escape do not really apply and reformulates the problem in terms of a compression ratio, ρout/ρin, where the subscripts denote densities in the outflowing and inflowing plasma. This leads to a spectral index that tends towards s = 1 for the ‘maximally compressive’ case, i.e. when u₀/utₐₐ is maximised. Reconnection is most compressive when it is efficient in converting magnetic energy to kinetic energy, which may be exactly what is needed to transition from a Poynting flux dominated jet to a kinetically dominated one. This is an important general point; a Fermi process in a reconnection site creates an intrinsic link between the release of magnetic energy and the acceleration of particles. In both these models, there is a similarity to shock acceleration in that there are reasonable grounds for expecting the particle escape probability to be constant with increasing energy.

Although these simple models are instructive, many of the advances in understanding Fermi-type processes close to reconnection sites have come through detailed PIC simulations. For
example, Drake et al. (2006) show that particles can undergo a Fermi process by following field lines around a contracting magnetic island, experience an energy boost of \( \sim v_A/c \) each time the particle reaches the ends of the islands and reverses its direction (see e.g. our Fig. 4, Drake et al.’s fig. 2). Sironi & Spitkovsky (2014) have shown that the plasmoid instability fragments the current sheet into a series of magnetic islands of varying sizes; they find a variety of first-order Fermi-type mechanisms at work, with particles scattering between merging magnetic islands and the largest islands generally acting as reservoirs for the highest-energy particles (see also Lyutikov et al., 2016; Ball et al., 2018). An important difference to Drury’s picture is that the velocity in the energy gain is proportional to \( u_{\text{out}} \) which can be highly relativistic, whereas generally \( u_{\text{in}} \sim v_A \sim 0.1c \) (Lyubarsky, 2005; Sironi & Spitkovsky, 2014). Guo et al. (2015) show using test particle MHD simulations that particles undergo a curvature drift along an electric field induced by the reconnection flows (see also Li et al., 2017). These Fermi mechanisms differ clearly from the DSA case, since the presence of self-generated turbulence is not necessarily needed to facilitate a first-order Fermi process.

Simulations have therefore been relatively successful in producing superthermal particles with power-law spectra from reconnection sites, but the energy regime probed is relatively low, up to around 100 to 1000 times the thermal energy, and the maximum scale is a similar factor of the ion inertial length (e.g. Li et al., 2017). Ignoring losses, the maximum energy in a Fermi-type process in an individual island or X-line region will (optimistically) be set by the confinement condition, but if multiple magnetic islands exist then this also facilitates a stochastic process similar to Fermi II in which particles interact with multiple reconnection sites. This could be thought of as a ‘Fermi 1.5’ process, since the particle gains energy in a first-order process until its gyroradius becomes comparable to the size of an individual reconnection site, before escaping and scattering off other islands. Hoshino (2012) argues that this process is first-order in \( v_A/c \) since the particles preferentially interact with the outflowing plasma from the reconnection sites. Comisso & Sironi (2018) have conducted simulations of a turbulent plasma in which particles are injected at X-points before undergoing second-order Fermi acceleration in the turbulent plasma; they find a spectral index of \( s = 2.9 \) which is not universal and softer than in the more ordered situation considered by Sironi & Spitkovsky (2014). Given that fast reconnection is likely to be inherently turbulent, it is important to consider the overall, 3D turbulent environment, since the details of the process clearly depend on the scales of the various reconnection sites/magnetic islands. PIC simulations and MHD simulations with test particles are therefore critical tools for understanding particle acceleration at reconnection sites. The particle spectral index from these simulations varies, but values for \( s \) tend to lie in the range 1 to 2 for both \( e^+ e^- \) and \( e^- \)-ion plasmas (e.g. Romanova & Lovelace, 1992; Hoshino et al., 2001; Zenitani & Hoshino, 2001; Larrabee et al., 2003; Lyubarsky & Liverts, 2008; Sironi & Spitkovsky, 2014; Guo et al., 2015).

3.4. Other mechanisms

The mechanisms we have described are the most commonly invoked to explain high-energy particles, but there are of course other ways in which particles can acquire highly superthermal energies. Shear acceleration is a form of Fermi acceleration - the flow is not converging, but the particle scatters across a shear layer such as that at the edge of a jet (e.g. Rieger & Duffy, 2004). The particle distribution function can once again take the form of a power-law in momentum whose index is determined by how the time between scattering events depends on the momentum of the particle (Berezhko & Krymskii, 1981, 1982). The non-relativistic particle transport equation was derived by Earl et al. (1988), and the problem has been since been revisited (in the case of relativistic flows) by Webb et al. (2018, 2019) and Rieger & Duffy (2019). Recent reviews are provided by Rieger (2019a,b). Shear acceleration may also be able to provide a ‘one-shot’ energy boost by a factor of \( \sim \Gamma^2 \) to an existing population of high-energy particles encountered by the jet (Caprioli 2015; see also Kimura et al. 2018). Magnetic kink instabilities can also produce nonthermal distributions of particles (e.g. Alves et al., 2018) and are interesting since the kink-mode operates in jets under certain conditions (Begelman et al., 2018).
1984; Mignone et al., 2010; Tchekhovskoy & Bromberg, 2016; Barniol Duran et al., 2017).

3.5. General Considerations

Although the microscopic physics changes depending on which mechanism is in operation, the key underlying principles are similar. The dominant mechanism at a given point in a jet is set by the local physical conditions; we discuss this in Section 6. First, we make a few general comments.

The first relates to the availability of different Fourier modes in magnetic field structure at the particle acceleration site. To accelerate a particle from, say, GeV to TeV energies requires structure in the magnetic field on the scale length of the respective Larmor radii - i.e. over the same dynamic range as the energy range in question. One great advantage of a shock in this regard is that it is a discontinuity, so, provided the shock is not too smoothed out, it is possible for every Fourier mode of turbulent magnetic field to be present at the shock, driven by the self-regulated processes we described in Section 3.2.1. It is of course possible that this condition is fulfilled by other mechanisms, but it depends critically on the spectrum of magnetic irregularities. In Kolmogorov turbulence, with $E(k) \propto k^{-5/3}$, where $k$ is the wavenumber of the turbulence and $E(k)dk$ is the energy density contained in the wavenumber range $(k, k + dk)$, the turbulence covers a wide range of scales, but one has to account for the fact that the energy contained at small scales (large $k$) is small compared to the energy density at the driving scale.

Another related point concerns the spectral index of inferred particle energy distributions. We have implied throughout this section that DSA has an advantage over some of the other acceleration mechanisms not only in its first-order velocity dependence, but also in that the spectral index does not require fine-tuning and has a ‘universal’ value. It is worth noting that a particle energy index of $s = 2$ spreads the energy in superthermal particles out such that each decade in energy contributes an equal amount to the energy density. It is therefore possible that the implied value of $s = 2 - 2.6$ inferred from synchrotron power-laws is a reflection of some self-regulating process that tries not to put the majority of energy in the high-energy particles, and may not provide deep physical insight into the specific mechanism.

4. Timescales and the maximum energy

The characteristic maximum energy achievable by a particle with charge $Ze$ being moved a distance $R$ through a $-\mathbf{u} \times \mathbf{B}$ electric field is set by the Hillas energy (Hillas, 1984), given by $E_H = ZeuBR$, which can be written as

$$E_H = 10^{18} \text{ eV} \left( \frac{R}{1 \text{ kpc}} \right) \left( \frac{B}{1 \mu \text{G}} \right) \left( \frac{u}{c} \right) Z. \quad (9)$$

This can also be obtained by taking the time derivative of a magnetic flux $BR^2$, which gives a potential drop $uBR$ (e.g. Lemoine & Waxman, 2009; Waxman, 2011). If the flow is relativistic, an additional factor of $\Gamma$ is included in estimates for the maximum energy (e.g. Achterberg et al., 2001) depending on the details of the acceleration process, the geometry, the frame in which $B$ is defined, and the escape of the particle after acceleration. Equation 9 becomes the confinement condition (Equation 2) from Section 3 if we set $u = c$, but the distinction is important. In DSA, $u$ is the shock velocity, whereas in magnetic reconnection it is some flow velocity comparable to $v_k$. Since $u < c$, it is, unsurprisingly, harder to accelerate a particle than to confine it and the accelerator must therefore be larger than the Larmor radius by a factor $c/u$.

4.1. Acceleration timescale

The Hillas criterion is necessary but not sufficient for acceleration to high energy. In reality the crucial aspect is a balance between the acceleration timescale and the dominant timescale of escape, adiabatic losses or radiative cooling. The general expression for the acceleration time is

$$t_{\text{acc}} \sim \frac{E}{E_{\text{acc}}} \sim \left( \frac{\Delta E}{E} \right) t_{\text{cycle}}, \quad (10)$$

where $t_{\text{cycle}}$ is either the collision time (for Fermi II) or crossing time (for DSA). A more specific timescale for DSA was derived in the SNR context by Lagage & Cesarsky (1983a,b) and Drury (1983), who find

$$t_{\text{acc,DSA}} = \frac{3}{u^2} \left( \frac{\chi D_\text{u} + \chi^2 D_d}{\chi - 1} \right). \quad (11)$$

where $D_u$ and $D_d$ are the upstream and downstream diffusion coefficients and $\chi$ is the shock compression ratio. As noted by Blasi et al. (2007), the above formula is not appropriate for a CR-modified shock and the expression becomes more complicated, but it is suitable for our purposes. Except for in special situations (e.g. Bell et al., 2019b), the Hillas energy is only reached in DSA at a parallel or oblique shock if Bohm diffusion applies. This can be demonstrated using equation 11. If we assume the upstream and downstream dwell times ($\sim D/u$) are identical, and set $t_{\text{acc}} = R/u_c$, then for a strong, non-relativistic shock with $\chi = 4$ we recover $E_{\text{max}} = 3/8(D_{\text{Bohm}}/D_u)uBR$, where the exact preceding numerical factor is determined by the relationship between $D_u$ and $D_d$ as well as the value of the denominator in $D_{\text{Bohm}}$. At a perpendicular shock, the particle can in principle drift along the shock for its entire length if the turbulence either side of the shock is just sufficient to counteract the $\mathbf{E} \times \mathbf{B}$ drift, in which case the Hillas energy can be reached. However, in practice this is unlikely, so the particle must return to the shock multiple times to reach the Hillas limit, which is harder to achieve than for a parallel shock. Similar expressions to the non-relativistic case can be derived for the maximum energy at ultra-relativistic shocks (e.g. Achterberg et al., 2001; Bell et al., 2018), where $u_c \to c$ and $\chi \to 3$, but the generation of the turbulence becomes a severe limitation (see sections 3.2.2 and 4.2).

The acceleration time for Fermi II actually has a very similar form to that for DSA; although the energy gain is smaller by a factor $(u/c)$, the collision time is also shorter than the shock crossing time and the acceleration time is given by

$$t_{\text{acc,F2}} \propto \frac{D}{u^2_{\text{turb}}} \quad (12)$$
4.2. Instability Growth Times

The acceleration time for particle acceleration in reconnection sites depends on the details of the acceleration process. We might, however, expect similar arguments to DSA/Fermi II to apply depending on how ordered the process is. Detailed studies using test-particle simulations find $t_{\text{acc,rec}} \propto (u_{\text{rec}}/c)^{-\kappa}$ with $\kappa$ ranging from 2.1 to 2.4 (del Valle et al., 2016). This is close to the $u^{-2}$ dependence quoted above for DSA and Fermi II.

4.2. Instability Growth Times

Acceleration timescales based on diffusion coefficients involve an implicit assumption that there is turbulent magnetic field on the scale of the Larmor radius of the particle. In the case of self-generated turbulence from streaming instabilities we must also consider the growth rate of the turbulence. For NRH/Bell turbulence, the process is self-regulating and the CRs drive turbulence on the scale of their respective Larmor radii which in turn scatters particles of that energy. The growth rate of the NRH instability is given by $\gamma$ of the NRH instability is given by $\gamma$ which in turn scatters particles of that energy. In supersonic jets, the turbulent velocity cannot be comparable to the bulk velocity, highlighting again why shocks and reconnection sites are much more efficient at channeling magnetic or kinetic energy into superthermal particle distributions. Fermi II might however be important disrupted, turbulent jets (e.g. Manolakou et al., 1999) or in the jet cocoon or lobe (e.g. Begelman et al., 1984; Hardcastle et al., 2009).

The acceleration time for particle acceleration in reconnection sites depends on the details of the acceleration process. We might, however, expect similar arguments to DSA/Fermi II to apply depending on how ordered the process is. Detailed studies using test-particle simulations find $t_{\text{acc,rec}} \propto (u_{\text{rec}}/c)^{-\kappa}$ with $\kappa$ ranging from 2.1 to 2.4 (del Valle et al., 2016). This is close to the $u^{-2}$ dependence quoted above for DSA and Fermi II.

4.3. Losses

Radiative cooling limits the maximum energy if the radiative cooling timescale for the particles is short enough. The generic expression for any radiative loss timescale is

$$t_{\text{cool,rad}} = \frac{E}{E_{\text{rad}}}$$

where $E_{\text{rad}}$ describes the rate at which a particle radiates due to synchrotron, inverse Compton or synchrotron self-Compton processes. The synchrotron power of an electron in a magnetic field with energy density $U_B = B^2/8\pi$ is

$$E_{\text{sync}} = \frac{4}{3} \sigma_T c^3 \beta_e^2 U_B$$

where $\beta_e = v_e/c$ and $\sigma_T$ is the Thomson cross-section. Similarly, the inverse Compton cooling power is

$$E_{\text{IC}} = \frac{4}{3} \sigma_T c^3 \beta_e^2 U_{\text{rad}}$$

giving us the simple result that the maximum energy loss rate is just determined by the energy density of the dominant field interacting with the electron. The value of $U_{\text{rad}}$ is often uncertain and will vary with height in the jet. For example, close to an accretion disc the disc luminosity may be dominant, but further from the jet base the electrons’ own synchrotron radiation or the CMB dominates. Similarly, we might expect $U_B$ to vary dramatically along a jet if it transitions from Poynting flux dominated to kinetically dominated.

Protons and ions also cool. The proton synchrotron and inverse Compton cooling times can be derived by considering the mass ratio $m_p/m_e \approx 1837$, which, for a given energy, leads to a cooling time a factor $(m_p/m_e)^4 \approx 10^{13}$ longer than for electrons (e.g. Begelman et al., 1990; Aharonian, 2000). Thus, proton synchrotron and inverse Compton radiation are often negligible, although may be important in strong $\gtrsim 100$ MG magnetic fields in the bases of AGN jets (Aharonian, 2000, 2002; Biteau et al., 2020) or in GRBs (Waxman, 2001; Gupta & Zhang, 2007). Nonthermal protons also lose energy via the same $pp$ and $py$ channels that produce gamma-ray emission. The cross-section for $py$ processes depends strongly on the particle energy and shape of the radiation field; for low temperature photon fields such as the CMB, radiation losses are only significant at ultrahigh energies (e.g. Alves Batista et al., 2015), whereas close to an accretion disc the strong UV radiation field can cause fast $py$ losses at $\sim$PeV energies (Begelman et al., 1990, see their fig. 2). Additionally, ions can photodisintegrate leading to a reduction in energy (Stecker, 1968; Stecker & Salamon, 1999). Inelastic collisions with thermal protons also result in energy losses, but the associated timescale is quite long: $t_{pp} \approx (1/\sigma_T m_p c)^{-1} \sim 600$ Gyr for a density of $n_p = 10^{-4}$ cm$^{-3}$ (Sikora et al., 1987), where $\sigma_T \approx 4 \times 10^{-28}$ cm$^2$ is the cross-section. Losses from $pp$ collisions can nevertheless be the dominant proton cooling mechanism for $\lesssim 10$ TeV energies. For protons, the regimes in which the different types of losses are expected to dominate are summarised by Begelman et al. (1990), showing a strong dependence on $U_{\text{rad}}$, $n_p$, $U_B$ and particle energy.
4.4. Power requirement

The Hillas energy can be used to derive a maximum energy based on the power of a source (Lovelace, 1976; Waxman, 1995; Blandford, 2000; Lemoine & Pelletier, 2010). This power requirement is sometimes referred to as a Hillas-Waxman-Lovelace limit. It can be derived by considering the ‘magnetic power’ passing through a surface of cross-sectional area $R^2$, which is $Q_B = (B^2/8\pi) R^2 u$, where $u$ is the bulk velocity through the surface. Combining this with equation 9 then gives

$$Q_{B,\text{min}} = 10^{44} \text{ erg s}^{-1} \left( \frac{E}{10 \text{EeV}} \right)^2 \left( \frac{u}{0.1c} \right)^{-1} Z^2.$$  \hspace{1cm} (16)

This limit is due to the maximum electric field produced by a source with a certain magnetic power, assuming that we arrange for optimal acceleration conditions. A less restrictive power requirement is sometimes derived by setting $u = c$, but equation 16 applies quite generally irrespective of acceleration mechanism. For example, following e.g. Potter (2017), in a reconnection site $u = u_{in}$ and $R^2$ is the surface area of the current sheet, whereas in shock acceleration $u = u_s$ and $R$ is roughly the shock size. We can convert this minimum magnetic power into a kinetic power using a parameter $\eta$ such that $Q_{k,\text{min}} = \eta^{-1} Q_{B,\text{min}}$, where $\eta$ represents the partitioning between kinetic and magnetic energy densities; in shock acceleration, $\eta$ can be thought of as an efficiency of magnetic field amplification at the shock.

For most of the observations described in Section 2, the maximum particle energy is significantly lower than that associated with equation 16, with the important consequence that either the acceleration process or loss processes are limiting the maximum energy. However, for UHECRs, where the particle acceleration process is presumably stretched to its limits, this limit becomes important and already rules out a large number of different sources such as starburst winds, AGN disc winds or microquasar jets. In fact, jets in AGN and GRBs are two of the few known sources that, for reasonable efficiency parameters, satisfy $Q_k > Q_{k,\text{min}}$ for 10 EeV particles.

4.5. A maximum energy ladder

Having explored what factor limits the maximum particle energy we are able to build a ‘maximum energy ladder’: a set of gradually more restrictive criteria that determine the maximum energy for a given mechanism for a specific set of physical conditions. An example diagram is shown in Fig. 5. The first four criteria are based purely on the size, speed and energetics of the accelerator and at this point one can remain agnostic about the detailed acceleration process. The bottom half of the flowchart refers to more involved physics, and in this case the limiting factor will depend on the physical parameters in the plasma, for example $t_{acc}$, $B$ and $U_{rad}$. The diagram illustrates why it is crucial not only to have reasonable constraints on the physical conditions in a given source, but also a good understanding of the detailed plasma instability growth rates (if applicable) and acceleration timescales for the most likely acceleration mechanism.

5. Cosmic Rays and Neutrinos

In Fig. 6, we show differential particle fluxes ($J$) arriving at Earth, multiplied by $E^2$, for three types of particles: gamma-ray photons, neutrinos and hadrons (CRs). The figure is based on similar plots produced by, e.g., Fang & Murase (2018) and Ahlers (2018). The data are from the Fermi Large Area Telescope (Ackermann et al., 2015), the IceCube Neutrino Ob-
servation (Aartsen et al., 2015) and the Pierre Auger Observatory (PAO) (Pierre Auger Collaboration et al., 2017a), respectively, and the units are chosen so that a flat line represents equal amounts of energy per logarithmic energy bin. The contributions of UHECRs (at >EeV energies), high-energy IceCube neutrinos (at TeV to PeV energies) and the extragalactic gamma-ray background (at GeV to TeV energies) are comparable. This behaviour might initially appear to suggest universal physics; however, at least in the UHECR case the power-law is quite steep and so will inevitably cross a given value of $E^2/J$. However, as has been pointed out by numerous authors (e.g. Mannheim, 1995; Waxman & Bahcall, 1997; Bhatthacharjee, 2000; Fang & Murase, 2018; Ahlers, 2018; Alves Batista et al., 2019) the three sets of high-energy particles may have common origins and extragalactic jets are plausible acceleration sites for protons, nuclei and their decay products.

5.1. Neutrinos

High-energy protons produce pions via interactions with photons or proton-proton collisions, which then decay to produce gamma-ray photons (in the case of $\pi^0$) or leptons and neutrinos (in the case of $\pi^\pm$). Neutrinos are therefore sometimes considered to be a “smoking gun” of hadronic acceleration, making them useful messengers for studying jet composition and connections to UHECRs. The IceCube neutrino detector has proved to be a game-changer in the field of high-energy neutrino astrophysics. IceCube has detected PeV neutrinos (Aartsen et al., 2013, 2014; Collaboration, 2013), for which jets from GRBs or AGN are possible sources. Tentative coincidences have been investigated between PeV neutrino events and AGN sources (Kadler et al., 2016; Fraija et al., 2018). Jet sources may also produce the diffuse TeV to PeV flux observed by IceCube and shown in Fig. 6, although the maximum contribution to this diffuse flux from AGN in the 2nd Fermi-LAT AGN catalog is 27% (Aartsen et al., 2017).

Recently, IceCube detected neutrinos spatially coincident with the BL Lac-type blazar TXS 0506+566 as it was undergoing a gamma-ray flare (IceCube Collaboration, 2018b,a). They infer a neutrino energy of $\sim$30 TeV and maximum proton energies in the range $10^{14} - 10^{18}$ eV. While this is some way off explaining the highest energy CRs, the IceCube result is one of the clearest signals yet that jets accelerate hadrons to very high energies, and demonstrates the potential of multimessenger observations for probing conditions in extragalactic jets. However, this detection does not necessitate that the hadrons are significant in terms of the overall jet energy budget or that the gamma-rays from blazars are hadronic in origin.

5.2. UHECRs

CRs were first detected over 100 years ago (Hess, 1912), even before astrophysical jets, but it is more recent progress that is relevant to this review. The CR spectrum extends smoothly across many decades in energy and takes the form of a power-law with various breaks (e.g. Hillas, 2006). The CR power-law provides one of the motivations for seeking a power-law distribution for $n(E)$ in section 3. Galactic CRs, with energies up to and possibly beyond the ‘knee’ in the CR spectrum at PeV energies, are thought to be accelerated by SNRs (e.g. Bell, 2014), although microquasars may also contribute (Heinz & Sunyaev, 2002; Fender et al., 2005; Cooper et al., 2020). However, the origin of ultrahigh energy CRs (UHECRs), with energies extending beyond $10^{20}$ eV, is still unknown. Jets have been suggested as possible acceleration sites, both in various classes of AGN (Hillas, 1984; Norman et al., 1995; Fang & Murase, 2018) as well as GRBs (Waxman, 1995, 2001). Deflections by intervening magnetic fields (e.g. Sigl et al., 2003; Taylor et al., 2011; Farrar & Sutherland, 2017) and the attenuating effect of photo-pion, pair-production and photodisintegration interactions (e.g. Greisen, 1966; Zatsepin & Kuz’min, 1966; Stecker & Salamon, 1999; Alves Batista et al., 2016) make robust directional associations difficult. Simple energetic arguments, such as those discussed in the previous Section, favour extreme sources (Hillas, 1984; Waxman, 1995; Blandford, 2000), and thus systems with jets remain natural candidates due to their enormous kinetic powers.

Recent advances have come from the combined power of the UHECR observatories, in particular the Telescope Array (TA) and the Pierre Auger Observatory (PAO). TA reported the detection of a ‘hotspot’ in the Northern sky with a spread of $\sim 20^\circ$ (Abbasi et al., 2014) while PAO have found a clear dipole above 8 EeV (Pierre Auger Collaboration et al., 2017b, hereafter PAO17) and more tentative indications of anisotropy above $39\text{EeV}$ and $60\text{EeV}$ (Pierre Auger Collaboration et al., 2018, hereafter PAO18). Fig. 7 shows Mollweide projections of the data from PAO17 and PAO18, in Galactic coordinates, with positions of some radio galaxies, blazars and starburst galaxies marked. PAO18 found that the anisotropic component could be associated with extragalactic gamma-ray sources, reporting correlations with starburst galaxies and AGN. Starburst galaxy ‘superwinds’ have been proposed as UHECR sources (e.g. Anchordoqui, 2018), but these winds are generally too slow and low-power to meet the power requirement (equation 16, see also Romero et al., 2018; Matthews et al., 2018). However, starburst galaxies are expected to host large numbers of GRBs (Chary et al., 2002; Christensen et al., 2004; Floch et al., 2006; Becker et al., 2009). GRBs generally offer better prospects for UHECR acceleration than starburst winds, although the required efficiencies are high (Waxman, 1995, 2011; Eichler & Pohl, 2011; Lemoine, 2018; Alves Batista et al., 2019).

An alternative is that UHECRs are accelerated by AGN jets. Radio galaxies have long been suggested as UHECR sources and recent work suggests they may be able to explain the data from PAO; for example, Eichmann et al. (2018) suggest a combination of Cygnus A and Centaurus A, while Matthews et al. (2018) propose Fornax A and Centaurus A. One of the problems with accelerating UHECRs in AGN and GRB jets is that both internal and termination shocks are relativistic and this poses severe problems for acceleration to EeV energies (Lemoine & Pelletier, 2010; Reville & Bell, 2014; Bell et al., 2018). One possibility in AGN is that non-relativistic shocks form in the backflows in the jet lobe/coocoon, offering more conducive conditions for UHECR acceleration (Matthews et al., 2019b). Other proposed mechanisms for UHECR acceleration in AGN
include magnetic reconnection in blazars (Giannios, 2010) and Fermi II acceleration in the lobes of radio galaxies (Hardcastle et al., 2009; Hardcastle, 2010).

UHECRs are interesting for our purposes because they stretch the theory of particle acceleration to its limits. However, it is extremely challenging to link the physics of UHECR acceleration to that of, e.g., GeV synchrotron electrons due to the different factors limiting the maximum energy and the vast range of Larmor radii involved. Progress is likely to come from multimessenger approaches that combine neutrino, gamma-ray and hadronic signals as well as the combined power of future TA and PAO datasets. Further constraints on UHECR composition and the intervening magnetic field will also prove crucial for identifying UHECR sources.

6. Discussion: A Journey Along A Jet

Let us now consider the journey made by jet plasma, from a launching point near the compact object, through, e.g., a termination shock and into a lobe or cocoon. The plasma can be described at any point along this journey by a few key variables: the bulk Lorentz factor, $\Gamma$, the sonic and Alfvénic Mach numbers, the plasma-$\beta$, the magnetic field $B$ and the velocity field $u$. However, for the purposes of this discussion, we focus particularly on three key parameters that are essentially ratios of the various energy densities:

- $M_{ms}$, the magnetosonic Mach number of the flow;
- $\sigma = B^2/(4\pi \Gamma \rho c^2)$, the magnetisation parameter, or the ratio of Poynting flux to mass-energy flux (in the observer frame);
- $u_{turb}/u$, the ratio of the turbulent and bulk velocities.

The magnetosonic Mach number is defined as $M_{ms} = (\Gamma u)/(\Gamma ms c_{ms})$ where $c_{ms} = \sqrt{v_A^2 + c^2(1 - v_A^2/c^2)}$ is the magnetosonic speed in a relativistic flow (Cohen et al., 2014), and $\Gamma ms$ is the Lorentz factor associated with $c_{ms}$. When the flow is magnetically dominated, $c_{ms} \approx v_A$, whereas at low magnetisations $c_{ms} \approx c_r$. We use $M_{ms}$ because it takes into account both the magnetic and thermal energy densities and governs the nature of the shocks that form (Martí et al., 2016). These three parameters determine which particle acceleration processes are likely to operate most efficiently. A schematic showing how they might vary along a jet length and which processes are expected to dominate in each regime is shown in Fig. 8.

We assume a powerful jet similar to those associated with FRII radio galaxies and powerful blazars that is also applicable to other situations where a persistent jet is produced. We assume that the jet is initially highly magnetised, as expected from simulations of jet launch (McKinney, 2006; Komissarov et al., 2009) and constraints from radiative losses and variability in blazars (Potter, 2017; Morris et al., 2018). The magnetisation parameter can change dramatically along a jet if it undergoes a transition from magnetically dominated near the base to kinetically dominated on large scales. Studies have suggested that this transition to $\sigma \lesssim 1$ occurs at around $10^5 r_g$ ($\approx 0.5 \text{pc}$ for a $10^8 M_\odot$ black hole) in blazar jets (Potter, 2017; Chatterjee et al., 2019). In the magnetically dominated regime, reconnection is likely to be important. Fast, relativistic reconnection is a good candidate mechanism for explaining a few key observational constraints from Section 2, that is: the fast variability observed from compact jet regions, the apparent need to dissipate magnetic energy in these regions and the requirement for in situ particle acceleration along the jet length. The region of parameter space where reconnection is most likely to accelerate particles efficiently is shaded red in Fig. 8.
Figure 7: Top: Mollweide projection of UHECR fluxes above 8EeV in Galactic coordinates from PAO (Pierre Auger Collaboration et al., 2017b). A dipole anisotropy is observed in the data. Bottom: Mollweide projection of the anisotropic excess events per beam above 60 EeV in Galactic coordinates from PAO (Pierre Auger Collaboration et al., 2018). In both plots, the PAO exclusion zone in the northern hemisphere is marked. We also plot the gamma-ray AGN and SBG samples from (Pierre Auger Collaboration et al., 2018) and luminous ($\nu L_\nu > 2 \times 10^{40}$ erg s$^{-1}$) radio galaxies within 100 Mpc from the van Velzen et al. (2012) catalogue.
Not all jets start off highly magnetised, in particular those in protostellar systems and some microquasars, while even the Doppler-boosted emission from blazars may originate outside the high-$\sigma$ region (e.g. Sikora et al., 2005). Thus, the regime where $\sigma < 1$ in Fig. 8 can either correspond to regions further along the jet than the magnetised base, or to jets that are not initially Poynting flux dominated. In this lower magnetisation regime, acceleration at reconfinement shocks and internal shocks becomes important. We define internal shocks as those that are roughly planar and created by time-varying jet activity in which shells of different density and velocity in the jet collide with one another (e.g. Rees & Meszaros, 1994; Spada et al., 2001). Reconfi nement shocks are caused by the confining pressure of the cocoon/ambient medium, which drives an oblique shock into the jet (e.g. Falle, 1991; Komissarov & Falle, 2003), and may disrupt the jet (Gourgouliatos & Komissarov, 2018). Both of these classes of shock appear to accelerate superthermal particles. Internal shock models are invoked in GRBs (Rees & Meszaros, 1994; Kobayashi et al., 1997; Piran, 2004), microquasars (Jamal et al., 2010; Malzac, 2013; Drappeau et al., 2015) and AGN jets (Spada et al., 2001; Ghisellini et al., 2002; Bai & Lee, 2003) to explain some of the observed temporal behaviour. Reconfi nement shocks may explain the appearance of knots and quasi-periodic brightenings along the jet length in radio galaxies (Stawarz et al., 2006; Nalewajko, 2012; Hardcastle et al., 2016; Levinson & Globus, 2017). Knots and spatially intermittent emission may also be explained by shocks from the jet interacting with stellar winds (Wykes et al., 2015), a process that may also produce gamma-ray emission (Araudo et al., 2013).

Jets can be disrupted and decelerated by entrainment and mass-loading (Komissarov, 1990; Bowman et al., 1996; Laing & Bridle, 2002; Perucchi & Marti, 2007) or instabilities (Hardee et al., 1995; Tchekhovskoy & Bromberg, 2016; Gourgouliatos & Komissarov, 2018). A low power or highly mass-loaded jet resembles a subsonic or transonic plume of radiating material rather than a focused supersonic beam. However, if the jet avoids disruption and remains at least moderately supersonic, the majority of its remaining kinetic power is dissipated through a strong shock. In the case of steady supersonic jets, the jet terminates in a reverse “termination” shock responsible for the bright hotspots observed in radio galaxies. In transient events, the jet instead sweeps up surrounding material into a relativistic fireball (Blandford & McKee, 1976; Rees & Meszaros, 1992; Waxman, 2001) and the particle acceleration occurs at a forward shock thought to be responsible for the afterglows seen in GRBs (Waxman, 1995; Yost et al., 2003; van Eerten, 2013), although emission from reverse shocks can be observed at early times in long-duration GRBs (Laskar et al., 2016, 2018; Bright et al., 2019). Both forward and termination shocks will generally produce efficient particle acceleration as the shock velocity can be large, so the acceleration time is short, and a large fraction of the jet energy can be transferred to superthermal particles. The termination shock is labelled in our schematic in Fig. 8, and coincides with a sharp drop in $M_{\text{ms}}$, as well as an increase in magnetic field due to amplification and compression at the shock.

Inside a supersonic jet, $u_{\text{turb}}/u$ must be small, else the turbulence would also be supersonic and would rapidly heat the plasma and decrease the Mach number. Beyond the termination shock, or in the plumes of disrupted jets, the flow becomes subsonic or transonic, and less ordered. Thus, in jet lobes and cocoons $u_{\text{turb}}/u$ might feasibly be close to unity. However, fast bulk flows can still occur, a prominent example being the back-
flows of radio galaxies, which can produce shocks and acceler- 
ate particles (Saxton et al., 2002; Matthews et al., 2019b). The 
shocks in backflows may be ideal candidates for accelerat- 
ing UHECRs (Matthews et al., 2019a,b). More generally, however, 
the cocoons and lobes around jets are probably fairly slow ac- 
celerators of superthermal particles since the characteristic val- 
ues of \( u \), \( \nu_{\text{turb}} \) and \( \nu \) are low. The low value of the velocities in 
this regime is in contrast to a) the velocity of the jet, which is the 
characteristic velocity associated with DSA at the termination 
shock, and b) the Alfven speed in a Poynting flux dominated jet, 
which is the characteristic velocity associated with relativistic 
reconnection near the jet base.

7. Concluding Remarks

We have described how particles can be lifted out of the ther- 
al pool and accelerated to high energy, usually via Fermi- 
type processes that allow them to move stochastically through 
a \( -i \times \vec{B} \) electric field. We have shown that there are a number of 
mechanisms, of which shock acceleration and recon- 
nection are the most notable, that produce power-law momentum dis- 
tributions of superthermal particles. We have discussed what 
limits the maximum energy of protons and electrons and ex- 
plored how jets may produce multimessenger signals of particle 
acceleration.

The overall future outlook for understanding particle accel- 
eration in jets is bright; a wealth of observational data is be- 
coming available and these data are truly ‘multimessenger’. In 
the last two years alone, there have been robust detections of 
particle acceleration in the afterglow from a gravitational wave 
source (Abbott et al., 2017) and more tentative associations be- 
tween jetted sources and the production of UHECRs (Pierre 
Auger Collaboration et al., 2018) and neutrinos (IceCube Col- 
aboration, 2018b,a). The Event Horizon Telescope has recently 
observed stunning images of the horizon-scale environ- 
ment around the central black hole in M87 (Event Horizon Tele- 
scope Collaboration, 2019), an AGN that possesses one of the 
best studied, and first discovered (Curtis, 1918) astrophysical 
jets. Longer term, the Cherenkov Telescope Array will offer un- 
precedented sensitivity in gamma-rays from approximately 0.1 
to 100 TeV. On the computational side, recent advances have al- 
lowed PIC simulations to be run in 3D with impressive results 
(Riquelme & Spitkovsky, 2011; Sironi & Spitkovsky, 2011, 
2014; Guo et al., 2015; Crumley et al., 2019), while (GRM)HD 
simulations of relativistic jets provide insights into the magneti- 
sation, morphology and stability of the jets and lobes on a range of 
scapes (e.g. Krause, 2005; Mignone et al., 2010; van Eerten et al., 
2012; van Eerten, 2013; Mendoza et al., 2013; Hard- 
castle & Krause, 2013, 2014; Yoon & Heinz, 2015; Yoon et al., 
2016; English et al., 2016; Tchekhovskoy & Bromberg, 2016; 
Moscibrodzka et al., 2016; Barniol Duran et al., 2017; Liska 
et al., 2018; Chatterjee et al., 2019).

Despite all this exciting progress, questions remain plentiful: 
Are jets leptonic, or lepto-hadronic? How important is entrain- 
ment? Why are many jets apparently ‘magnetoluminescent’ 
at their bases? What limits the maximum energy in different 
sources? Where do UHECRs and PeV neutrinos come from? 
What creates the high-energy hump in blazars? Which parti- 
cle acceleration mechanisms matter most in GRBs? What is 
the microphysics of particle acceleration at shocks and recon- 
nection sites? How is TeV gamma-ray emission produced in 
microquasars? What mechanism produces the rapid variability 
and flaring in blazars? In many cases where nonthermal emis- 
sion is observed, the origins of the nonthermal particles is not 
fully known and it is difficult to distinguish between the possi- 
ble acceleration processes.

A combination of theoretical and observational advances will 
be needed to fully understand the particle acceleration process 
in jets. Other areas of astrophysics, involving the Sun, solar 
wind, supernova remnants and pulsar wind nebulae provide im- 
portant comparison points for the jets community. For exam- 
ple, we can study magnetic reconnection in the Sun (Priest 
& Forbes, 2002; dalla & Browning, 2005; Zharkova et al., 
McKenzie, 2011) and stripped wind of pulsars (Lyubarsky 
& Kirk, 2001; Ptri & Lyubarsky, 2007; Cerutti et al., 2012), 
or DSA in the bright rims of supernova remnants (Long et al., 
2003; Vink & Laming, 2003; Cassam-Chena et al., 2007; 
Uchiyama et al., 2007; Bell et al., 2013) and in the solar sys- 
tem (Tsurutani & Rodriguez, 1981; Eichler, 1981; Wenzel et al., 
1985; Ellison & Moebius, 1987; Burgess, 2007; Burgess et al., 
2012). Each of these environments provides a laboratory that 
can be used to learn about the same particle acceleration mech- 
anisms thought to operate in astrophysical jets.

Acknowledgements

We would like to thank R. Fender, R. Wijers, A. Araudo, M. 
Hardcastle and J. Bright for helpful discussions. JM acknowl- 
dedges a Herchel Smith Research Fellowship at Cambridge. This 
work was also supported by the Science and Technology Facili- 
ties Council under consolidated grant ST/N000919/1. We grate- 
fully acknowledge the use of matplotlib 2.0.0 (Hunter, 2007), 
NAIMA (Zabalza, 2015) and PLUTO (Mignone et al., 2007). 
This research made use of Astropy, a community-developed core 
Python package for Astronomy (Astropy Collaboration et al., 
2013; Price-Whelan et al., 2018).

References

References

Aartsen, M. G., Abbasi, R., Abdou, Y., et al. 2013, Physical Review Letters, 
111, 021103

Aartsen, M. G., Ackermann, M., Adams, J., et al. 2014, Physical Review Let- 
ters, 113, 101101

Aartsen, M. G., Abraham, K., Ackermann, M., et al. 2015, The Astrophysical 
Journal, 809, 98

Aartsen, M. G., Abraham, K., Ackermann, M., et al. 2017, The Astrophysical 
Journal, Volume 835, Issue 1, article id. 45, 17 pp. (2017), 835, 45

Abbasi, R. U., Abe, M., Abu-Zayyad, T., et al. 2014, The Astrophysical Journal 
Letters, 790, L21

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, The Astrophysical Journal 
Letters, 848, L13

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010a, Science, 328, 725

4http://www.astropy.org
