4 Relativistic Shocks

GRBs invoke relativistically moving fluids, whose behavior is delineated by relativistic fluid dynamics. *Shocks* (both the *external forward and reverse shocks* and *internal shocks*) are believed to play an important role to interpret the observed GRB afterglow and prompt emission. This chapter describes the physics of *relativistic shocks*. Section 4.1 briefly summarizes the key equations of fluid dynamics, in both non-relativistic and relativistic regimes. *Shock jump conditions* (for both non-relativistic and relativistic cases) are introduced in §4.2, with a treatment of a generalized problem invoking an arbitrarily magnetized outflow. In astrophysical problems, usually a pair of shocks develop when the relative Lorentz factor between the two fluids exceeds the sound speeds in both fluids, so in §4.3 the physics of a system of a pair of (forward vs. reverse) shocks is treated. The next two sections deal with two important functions of relativistic shocks. One is their ability to accelerate particles (§4.4) through the *first-order Fermi acceleration mechanism*; while the other is their ability in amplifying magnetic fields through *plasma or fluid instabilities* (§4.5). These two properties of relativistic shocks lay the foundation for non-thermal synchrotron radiation from the shocks, which holds the key to interpreting the GRB afterglow, and probably prompt emission as well. Finally, §4.6 introduces a parameterization method widely used in the GRB community to describe relativistic shocks. Some constraints on the parameters derived from the observational data are also summarized.

4.1 Relativistic Fluid Dynamics

4.1.1 Non-Relativistic Hydrodynamics

Fluid dynamics (*hydrodynamics*) handles the dynamical evolution of a fluid system. The fundamental equations are three conservation laws regarding mass, momentum, and energy.

In the *non-relativistic* regime, the rest mass energy density, ρc^2, is $\gg \epsilon = (1/2)\rho v^2 + e$, the sum of the kinetic energy density $(1/2)\rho v^2$ and internal energy density e. When gravity is neglected,[1] the three conservation equations (in Eulerian form) can be written as

[1] Since the GRB luminosity ($\sim 10^{51}$–10^{52} erg s^{-1}) is more than 10 orders of magnitude higher than the "Eddington" luminosity of the central engine ($L_{\rm Edd} \simeq 1.3\times 10^{39}$ erg s$^{-1}(M/10M_\odot)$, which is the maximum luminosity achievable in a steady, isotropic, gravitationally bound system), the GRB ejecta is gravitationally unbound. Therefore, gravity can be safely neglected in GRB emission regions.

- Mass conservation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0; \tag{4.1}$$

- Momentum conservation:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{1}{\rho}\nabla p; \tag{4.2}$$

- Energy conservation:

$$\frac{\partial \epsilon}{\partial t} + \nabla \cdot [(\epsilon + p)\mathbf{v}] = 0. \tag{4.3}$$

Here $\mathbf{v}$ is the three-dimensional velocity vector, and ρ, p, and ϵ are the mass density, pressure, and non-relativistic energy density of the fluid element at location (x, y, z) and time t, respectively.

4.1.2 Relativistic Hydrodynamics

For a relativistic fluid, the velocities approach the speed of light, and various energy density terms approach or exceed the rest mass energy density, ρc^2.

Let us define a four-dimensional event vector $x^\mu = (ct, \mathbf{x})$ and the proper time $d\tau = dt/\gamma$. One can then define a normalized 4-velocity

$$u^\mu = dx^\mu/(cd\tau) = \gamma(1, \mathbf{v}/c), \tag{4.4}$$

and a 4-mass-current

$$j^\mu = (j^0, \mathbf{j}) = \rho u^\mu, \tag{4.5}$$

with $j^2 = \eta_{\mu\nu} j^\mu j^\nu = -\rho^2$. Mass conservation can be expressed as

$$\nabla_\mu j^\mu = 0. \tag{4.6}$$

Let us recall the energy–momentum tensor of a relativistic fluid (Eq. (3.106)):

$$T^{\mu\nu} = (\rho c^2 + e + p)u^\mu u^\nu + p\eta^{\mu\nu} = hu^\mu u^\nu + p\eta^{\mu\nu},$$

where

$$h = \rho c^2 + e + p \tag{4.7}$$

is the relativistic *enthalpy density* of the gas. It can be reduced to Eq. (3.105) in the fluid rest frame given the definition of $\eta^{\mu\nu}$ (Eq. (3.112)). Energy and momentum conservation can be cast in the simple form

$$\nabla_\mu T^{\mu\nu} = 0. \tag{4.8}$$

More explicitly, the three equations can be written as (Exercise 4.1):

- Mass conservation:

$$\frac{\partial(\gamma\rho)}{\partial t} + \nabla \cdot (\gamma\rho\mathbf{v}) = 0; \tag{4.9}$$

- Momentum conservation (the space component ($\nu = i$) of Eq. (4.8)):

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = -\frac{1}{\gamma^2 h}\left(c^2\nabla p + \mathbf{v}\frac{\partial p}{\partial t}\right); \tag{4.10}$$

- Energy conservation (the time component ($\nu = 0$) of Eq. (4.8)):

$$\frac{\partial(\gamma^2 h)}{\partial t} - \frac{\partial p}{\partial t} + \nabla\cdot(\gamma^2 h\mathbf{v}) = 0. \tag{4.11}$$

4.1.3 Equation of State

Equations (4.9)–(4.11) include five differential equations solving for six unknown parameters (ρ, e, and p, and three vector components of $\mathbf{v}$). Another equation is needed to close the problem, which is the equation of state.

An *equation of state* describes how gas pressure is related to other thermodynamic properties, e.g. $p = p(\rho, T, \ldots)$. Statistically, it connects the macroscopically measured quantity p to microscopically defined quantities such as particle number density n, internal energy density e, etc., i.e. $p = p(n, e, \ldots)$.

Microscopically, pressure is a momentum flux, i.e. momentum per unit area per unit time. Considering an ideal gas (with the kinetic energy much greater than the interaction energy of the particles) and recalling in three-dimensional space, one can write the *gas pressure* as[2]

$$p = \frac{1}{3}\int_0^\infty n(\epsilon)\mathrm{p}v d\epsilon, \tag{4.12}$$

where $\epsilon = (\gamma - 1)mc^2$ is the kinetic energy of a particle, $\mathrm{p} = \gamma m v$ is the one-dimensional momentum of the particle, and $n(\epsilon) = dn/d\epsilon$ is the number density of the particles in the energy interval $(\epsilon, \epsilon + d\epsilon)$.

For a non-relativistic gas, one has $\mathrm{p} \simeq (2m\epsilon)^{1/2}$ and $v \simeq (2\epsilon/m)^{1/2}$, so that

$$p \simeq \frac{2}{3}\int_0^\infty n(\epsilon)\epsilon d\epsilon = \frac{2}{3}n\langle\epsilon\rangle = \frac{2}{3}e, \tag{4.13}$$

where $n = \int_0^\infty n(\epsilon)d\epsilon$ is the number density of the particles, $\langle\epsilon\rangle = \int n(\epsilon)\epsilon d\epsilon/n$ is the mean random kinetic energy of the particles, and $e = n\langle\epsilon\rangle$ is the internal energy density of the system. Similarly, in the relativistic regime ($\gamma \gg 1$), one has $\mathrm{p} \simeq \gamma mc$ and $v \simeq c$, so that

$$p \simeq \frac{1}{3}\int_0^\infty n(\epsilon)\epsilon d\epsilon = \frac{1}{3}n\langle\epsilon\rangle = \frac{1}{3}e. \tag{4.14}$$

More generally, noticing $\gamma = \epsilon/mc^2 + 1$, Eq. (4.12) can be expressed as

$$p = \frac{1}{3}\int_0^\infty n(\epsilon)\left(1 + \frac{\epsilon}{mc^2}\right)^{-1}\left(2 + \frac{\epsilon}{mc^2}\right)\epsilon d\epsilon, \tag{4.15}$$

[2] Notice that, in this subsection, γ, v, and ϵ are the Lorentz factor, velocity, and kinetic energy of the microscopic particles, in contrast with the definitions in the other sections of this chapter, where they are defined for fluid elements.

or in terms of the Lorentz factor,

$$p = \frac{1}{3}\int_1^\infty n(\gamma)\frac{\gamma^2-1}{\gamma}mc^2 d\gamma. \tag{4.16}$$

The results depend on the distribution function $n(\epsilon)$ or $n(\gamma)$. In any case, it is insightful to investigate the simplest *mono-energetic* case ($\gamma = \text{const}$). Noticing $n = \int_0^\infty n(\gamma)d\gamma$, one has (e.g. Uhm, 2011)

$$p = \frac{1}{3}\frac{\gamma+1}{\gamma}n\epsilon = \frac{\gamma+1}{3\gamma}e. \tag{4.17}$$

In general, one may write

$$p = \kappa e, \tag{4.18}$$

with

$$\kappa \simeq \frac{\bar{\gamma}+1}{3\bar{\gamma}}, \tag{4.19}$$

where $\bar{\gamma}$ is the average Lorentz factor of the gas particles. Such an approximation is usually good enough (Uhm, 2011).

The so-called *adiabatic index* is defined as (e.g. Kumar and Granot, 2003; Uhm, 2011)

$$\hat{\gamma} = \frac{c_{\rm P}}{c_{\rm V}} = \frac{e+p}{e} = \kappa + 1 \simeq \frac{4\bar{\gamma}+1}{3\bar{\gamma}} \simeq \begin{cases} \frac{5}{3}, & \bar{\gamma} \sim 1 \text{ (non-relativistic)}, \\ \frac{4}{3}, & \bar{\gamma} \gg 1 \text{ (relativistic)}, \end{cases} \tag{4.20}$$

where $c_{\rm P}$ and $c_{\rm V}$ are specific heat capacity at constant pressure and constant volume, respectively.

With $\hat{\gamma}$, the equation of state of an ideal gas can be also written as

$$p \propto \rho^{\hat{\gamma}}. \tag{4.21}$$

This is proven below.

The first law of themodynamics for an adiabatic system ($dQ = 0$) can be written

$$dU + pdV = d(eV) + pdV = 0, \tag{4.22}$$

where $U = eV$ is the total internal energy of the system. Noting Eq. (4.18), one can write

$$(1+\kappa)\frac{dV}{V} + \frac{dp}{p} = 0, \tag{4.23}$$

or

$$p \propto V^{-(1+\kappa)} \propto \rho^{1+\kappa} \propto \rho^{\hat{\gamma}}. \tag{4.24}$$

4.1.4 Relativistic Magnetohydrodynamics

When electromagnetic (EM) fields are present, the description of a plasma becomes more complicated. A macroscopic description treats the system as a single fluid, in which EM fields fluctuate with the fluid on the same time and length scales. Such a theory is called *magnetohydrodynamics (MHD)*.

The simplest MHD assumes zero resistivity so that the fluid can be treated as a perfect conductor. The relativistic version of this *ideal MHD* can be described by the following equations:

- Mass conservation:

$$\nabla_\mu(\rho u^\mu) = 0; \tag{4.25}$$

- Energy–momentum conservation:

$$\nabla_\mu T^{\mu\nu} = 0; \tag{4.26}$$

- Perfect MHD condition:

$$F^{\mu\nu}u_\nu = 0; \tag{4.27}$$

- Maxwell's equations:

$$\nabla_\nu F^{\mu\nu} = 4\pi J^\mu; \tag{4.28}$$

$$\nabla_\lambda F_{\mu\nu} + \nabla_\mu F_{\nu\lambda} + \nabla_\nu F_{\lambda\mu} = 0. \tag{4.29}$$

Here $J^\mu = (c\rho_e, \mathbf{J})$ is the 4-current (ρ_e is the charge density), and

$$T^{\mu\nu} = T^{\mu\nu}_{\rm FL} + T^{\mu\nu}_{\rm EM} \tag{4.30}$$

is the energy–momentum tensor, which includes two components: the fluid component

$$T^{\mu\nu}_{\rm FL} = (\rho c^2 + e + p)u^\mu u^\nu + p\eta^{\mu\nu} = hu^\mu u^\nu + p\eta^{\mu\nu}, \tag{4.31}$$

and the EM component

$$T^{\mu\nu}_{\rm EM} = \frac{1}{4\pi}\left(F^\mu_\lambda F^{\lambda\nu} - \frac{1}{4}\eta^{\mu\nu}F^{\lambda\delta}F_{\lambda\delta}\right), \tag{4.32}$$

where

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix} \tag{4.33}$$

is the electromagnetic field tensor.

Explicitly splitting these equations in 3+1 space-time, one gets the following ideal MHD equations:

$$\frac{\partial(\gamma\rho)}{\partial t} + \nabla\cdot(\gamma\rho\mathbf{v}) = 0; \tag{4.34}$$

$$\frac{\partial}{\partial t}\left(\frac{\gamma^2 h}{c^2}\mathbf{v} + \frac{\mathbf{E}\times\mathbf{B}}{4\pi c}\right) + \nabla\cdot\left[\frac{\gamma^2 h}{c^2}\mathbf{v}\otimes\mathbf{v} + \left(p + \frac{E^2+B^2}{8\pi}\right)\mathbf{I} - \frac{\mathbf{E}\otimes\mathbf{E} + \mathbf{B}\otimes\mathbf{B}}{4\pi}\right] = 0; \tag{4.35}$$

$$\frac{\partial}{\partial t}\left(\gamma^2 h - p - \gamma\rho c^2 + \frac{B^2+E^2}{8\pi}\right) + \nabla\cdot\left[(\gamma^2 h - \gamma\rho c^2)\mathbf{v} + \frac{c}{4\pi}\mathbf{E}\times\mathbf{B}\right] = 0; \tag{4.36}$$

$$\frac{\partial\mathbf{B}}{\partial t} + c\nabla\times\mathbf{E} = 0; \tag{4.37}$$

$$\nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{J}; \tag{4.38}$$

$$\nabla \cdot \mathbf{B} = 0; \tag{4.39}$$

$$\nabla \cdot \mathbf{E} = 4\pi\rho_e; \tag{4.40}$$

$$\mathbf{E} = -\frac{\mathbf{v}}{c} \times \mathbf{B}. \tag{4.41}$$

Note that the symbol $\otimes$ denotes a tensor product. These equations can be reduced to the relativistic hydrodynamic equations by setting $E = B = 0$ (Exercise 4.2).

4.2 Relativistic Shock Jump Conditions

A perturbation in a non-magnetized fluid propagates in the form of sound waves. A *sound wave* is a longitudinal pressure wave. Namely, the direction of wave propagation direction ($\mathbf{k}$) is always parallel to the direction of the displacement ($\boldsymbol{\xi}$) that perturbs the medium. The wave equation can be written

$$\nabla^2 p - \frac{1}{c_s^2}\frac{\partial^2 p}{\partial t^2} = 0, \tag{4.42}$$

with the *sound speed*

$$c_s = \sqrt{\frac{\partial p}{\partial \rho}} = \sqrt{\frac{\hat{\gamma} p}{\rho}}, \tag{4.43}$$

where $\hat{\gamma}$ is the adiabatic index (Eq. (4.20)).

If an abrupt disturbance occurs so that the speed of a fluid changes by more than the speed of sound in the fluid (i.e. *supersonic*), a shock wave, characterized by a discontinuity, the *shock front*, in the properties of the fluid (density, pressure, temperature, velocity, etc.) would develop. These properties across the shock front are connected through a series of conditions related to the strength of the shock. These are called *shock jump conditions* or *Rankine–Hugoniot conditions*.

4.2.1 Non-Relativistic Shocks

When a fluid moves supersonically ($v > c_s$), the information does not have time to propagate via sound waves, and a shock is inevitable. The shock front separates the un-perturbed region (upstream) and the perturbed region (dowstream). Across the shock front, one again needs to satisfy three conservation laws. Written in *the rest frame of the shock front*, these conditions read:

- Mass conservation:

$$\rho_1 v_1 = \rho_2 v_2; \tag{4.44}$$

- Momentum conservation:

$$\rho_1 v_1^2 + p_1 = \rho_2 v_2^2 + p_2; \tag{4.45}$$

- Energy conservation:

$$\left(\frac{1}{2}\rho_1 v_1^2 + e_1 + p_1\right) v_1 = \left(\frac{1}{2}\rho_2 v_2^2 + e_2 + p_2\right) v_2. \tag{4.46}$$

Here v_i is the speed of the region i with respect to the shock, ρ_i, p_i, and e_i are the mass density, pressure, and internal energy density of the region i in their respective rest frames, and $i = 1, 2$ stands for the upstream and downstream regions, respectively. These equations can be understood straightforwardly. The v_i factor in both sides of Eqs. (4.44) and (4.46) is introduced to convert the relevant densities (mass density and enthalpy density) to the relevant conservative quantities, since for a flowing fluid the volume is proportional to speed. Equation (4.45) states the pressure balance, where ρv_i^2 denotes the *ram pressure* and p_i denotes the gas pressure in stream i.

These conditions can be translated to the ratios of the parameters between the upstream and the downstream, which more directly reflect the strength of the shock. Defining the *Mach number*

$$M = \frac{v}{c_s} = \frac{v}{\left(\hat{\gamma} p/\rho\right)^{1/2}} = \left(\frac{\rho v^2}{\hat{\gamma} p}\right)^{1/2}, \tag{4.47}$$

which delineates how "supersonic" the upstream is, one can write three Rankine–Hugoniot relations for an ideal gas (Exercise 4.3):

$$\frac{\rho_2}{\rho_1} = \frac{v_1}{v_2} = \frac{(\hat{\gamma}+1)M_1^2}{(\hat{\gamma}-1)M_1^2 + 2}, \tag{4.48}$$

$$\frac{p_2}{p_1} = \frac{2\hat{\gamma}M_1^2 - \hat{\gamma} + 1}{\hat{\gamma}+1}, \tag{4.49}$$

$$\frac{T_2}{T_1} = \frac{p_2\rho_1}{p_1\rho_2} = \frac{(2\hat{\gamma}M_1^2 - \hat{\gamma} + 1)[(\hat{\gamma}-1)M_1^2 + 2]}{(\hat{\gamma}+1)^2 M_1^2}. \tag{4.50}$$

Here, both the upstream and downstream regions are assumed to be in thermal equilibrium, and T_i stands for the temperature in the region i. As will be discussed in §4.4 below, particles are accelerated at the shock front, so that both the upstream and downstream are not in strict thermal equilibrium. The temperatures in Eq. (4.50) can be understood as effective temperatures that delineate the mean internal specific energy density in the two streams.

For strong non-relativistic shocks, $M \gg 1$, $\hat{\gamma} = 5/3$, one has

$$\frac{\rho_2}{\rho_1} = \frac{v_1}{v_2} \simeq \frac{\hat{\gamma}+1}{\hat{\gamma}-1} \simeq 4, \tag{4.51}$$

$$\frac{p_2}{p_1} \simeq \frac{2\hat{\gamma}}{\hat{\gamma}+1}M_1^2 \simeq \frac{5}{4}M_1^2 = \frac{3}{4}\frac{p_{1,\rm ram}}{p_1}, \tag{4.52}$$

$$\frac{T_2}{T_1} \simeq \frac{2\hat{\gamma}(\hat{\gamma}-1)}{(\hat{\gamma}+1)^2}M_1^2 \simeq \frac{5}{16}M_1^2 = \frac{3}{16}\frac{p_{1,\rm ram}}{p_1}. \tag{4.53}$$

One can see that the downstream is compressed by a factor of 4, the pressure and temperature are increased by a factor $\sim M_1^2$ (with a prefactor of 5/4 and 5/16, respectively), or the ratio between the upstream ram pressure ($p_{1,\rm ram} = \rho_1 v_1^2$) and gas pressure ($p_1$) (with a prefactor of 3/4 and 3/16, respectively).

4.2.2 Relativistic Shocks: Preparation

The general expression for sound speed in the relativistic regime can still be expressed as Eq. (4.43), except that the "effective" density now includes all the internal energy and pressure, i.e.

$$\rho = \frac{h}{c^2} = \rho_0 + \frac{e+p}{c^2} = \rho_0 + \frac{1}{c^2}\frac{\hat{\gamma}}{\hat{\gamma}-1}p. \tag{4.54}$$

Plugging Eq. (4.54) into Eq. (4.43), one gets

$$c_s = c\sqrt{\frac{\hat{\gamma}p}{\rho_0 c^2 + \frac{\hat{\gamma}}{\hat{\gamma}-1}p}} = \begin{cases} \sqrt{\hat{\gamma}p/\rho_0}, & p \ll \rho_0 c^2 \text{ (non-relativistic)}, \\ \sqrt{\hat{\gamma}-1}\cdot c \simeq \frac{c}{\sqrt{3}}, & p \gg \rho_0 c^2 \text{ (relativistic)}, \end{cases} \tag{4.55}$$

where $\hat{\gamma} = 4/3$ has been used in the relativistic regime. If the fluid speed exceeds this relativistic sound speed, i.e. $\Gamma > \sqrt{3/2} \simeq 1.225$, a relativistic shock will develop.

In GRB problems, the shocks are usually relativistic,[3] i.e. the upstream is moving with a relativistic speed with respect to the shock front. The Rankine–Hugoniot conditions are revised accordingly.

From the normalized 4-velocity definition (Eq. (4.4)), one can define a dimensionless 4-speed:

$$u = \gamma\beta. \tag{4.56}$$

For a shock problem, the velocities can be defined with respect to one of the three references, the upstream frame "1", the downstream frame "2", and the shock front frame "s". To reduce confusion, we will explicitly introduce two subscripts to describe the relative velocities/Lorentz factors. For example, γ_{12} stands for the relative Lorentz factor between the upstream and downstream, β_{1s} stands for the relative dimensionless speed between the upstream and the shock front, and u_{2s} stands for the relative dimensionless 4-speed between the downstream and the shock front.

There are several transformation relations among these quantities, which are useful for deriving relativistic shock equations. We list them below (Zhang and Kobayashi, 2005; Exercise 4.4):

$$\beta_{2s} = \frac{\beta_{1s} - \beta_{21}}{1 - \beta_{1s}\beta_{21}}, \tag{4.57}$$

$$\beta_{1s} = \frac{\beta_{2s} + \beta_{21}}{1 + \beta_{2s}\beta_{21}}, \tag{4.58}$$

[3] The external forward shock is relativistic before the blastwave is decelerated to a non-relativistic speed at late times (typically months after the explosion). The external reverse shock can be non-relativistic early on, but typically would reach at least "trans-relativistic" (say, $\gamma \sim 2$), or even relativistic (if the engine duration is long enough). Internal shocks are believed to be at least mildly relativistic in order to have high enough efficiency to interpret the prompt γ-ray emission.

$$\beta_{21} = \frac{\beta_{1s} - \beta_{2s}}{1 - \beta_{1s}\beta_{2s}}, \tag{4.59}$$

$$\gamma_{2s} = \gamma_{1s}\gamma_{21}(1 - \beta_{1s}\beta_{21}), \tag{4.60}$$

$$\gamma_{1s} = \gamma_{2s}\gamma_{21}(1 + \beta_{2s}\beta_{21}), \tag{4.61}$$

$$\gamma_{21} = \gamma_{1s}\gamma_{2s}(1 - \beta_{1s}\beta_{2s}), \tag{4.62}$$

$$u_{2s} = \gamma_{1s}\gamma_{21}(\beta_{1s} - \beta_{21}), \tag{4.63}$$

$$u_{1s} = \gamma_{2s}\gamma_{21}(\beta_{2s} + \beta_{21}), \tag{4.64}$$

$$u_{21} = \gamma_{1s}\gamma_{2s}(\beta_{1s} - \beta_{2s}), \tag{4.65}$$

$$\beta_{1s} - \beta_{2s} = \frac{u_{21}}{\gamma_{1s}\gamma_{2s}}. \tag{4.66}$$

4.2.3 Relativistic Hydrodynamic Shock Jump Conditions

Let us define the specific enthalpy density (enthalpy density per particle):

$$\mu = \frac{h}{n} = \frac{nm_pc^2 + e + p}{n} = m_pc^2 + \frac{\hat{\gamma}}{\hat{\gamma} - 1}\frac{p}{n}. \tag{4.67}$$

The three Rankine–Hugoniot conditions (mass, energy, momentum conservation) for a hydrodynamic relativistic shock can be written in *the rest frame of the shock* as

$$n_1u_{1s} = n_2u_{2s}, \tag{4.68}$$

$$\gamma_{1s}\mu_1 = \gamma_{2s}\mu_2, \tag{4.69}$$

$$\mu_1u_{1s} + \frac{p_1}{n_1u_{1s}} = \mu_2u_{2s} + \frac{p_2}{n_2u_{2s}}. \tag{4.70}$$

Cold Upstream

In many problems (e.g. the afterglow problem), the unshocked upstream is cold. One has $e_1 = p_1 = 0$, $\mu_1 = m_pc^2$. The jump conditions lead to the following solution (Blandford and McKee, 1976; Exercise 4.5):

$$u_{2s}^2 = \frac{(\gamma_{21} - 1)(\hat{\gamma} - 1)^2}{\hat{\gamma}(2 - \hat{\gamma})(\gamma_{21} - 1) + 2}, \tag{4.71}$$

$$u_{1s}^2 = \frac{(\gamma_{21} - 1)(\hat{\gamma}\gamma_{21} + 1)^2}{\hat{\gamma}(2 - \hat{\gamma})(\gamma_{21} - 1) + 2}, \tag{4.72}$$

$$\frac{e_2}{n_2} = (\gamma_{21} - 1)m_pc^2, \tag{4.73}$$

$$\frac{n_2}{n_1} = \frac{\hat{\gamma}\gamma_{21} + 1}{\hat{\gamma} - 1}, \tag{4.74}$$

$$\gamma_{1s}^2 = \frac{(\gamma_{21} + 1)[\hat{\gamma}(\gamma_{21} - 1) + 1]^2}{\hat{\gamma}(2 - \hat{\gamma})(\gamma_{21} - 1) + 2}. \tag{4.75}$$

For strong relativistic shocks, $\gamma_{21} \gg 1$, $\hat{\gamma} = (4\gamma_{21}+1)/3\gamma_{21}$ (Eq. (4.20)), one has[4]

$$\frac{e_2}{n_2} = (\gamma_{21}-1)m_p c^2 \simeq \gamma_{21} m_p c^2, \tag{4.76}$$

$$\frac{n_2}{n_1} = 4\gamma_{21}, \tag{4.77}$$

$$u_{1s} \simeq \gamma_{1s} \simeq \sqrt{2}\gamma_{21}, \quad \beta_{1s} \simeq 1, \tag{4.78}$$

$$u_{2s} \simeq \frac{\sqrt{2}}{4}, \quad \beta_{2s} \simeq \frac{1}{3}, \quad \gamma_{2s} \simeq \frac{3}{4}\sqrt{2}. \tag{4.79}$$

So a relativistic shock is much stronger than a non-relativistic one, with downstream relativistic "temperature" of the order of $\gamma_{21} m_p c^2$, and a compression ratio of $4\gamma_{21}$. The physical understanding of this formula is the following (Kumar and Zhang, 2015): a downstream (region 2) observer sees a cold upstream (region 1) moving towards the observer with a bulk Lorentz factor γ_{21}. After passing the shock, this bulk motion is converted to *random* motion of the particles in the downstream rest frame with a Lorentz factor of the same order.

Combining Eqs. (4.76) and (4.77), the downstream internal energy density can be written in the form

$$e_2 = 4\gamma_{21}(\gamma_{21}-1)n_1 m_p c^2 \simeq 4\gamma_{21}^2 n_1 m_p c^2, \tag{4.80}$$

which is about $4\gamma_{21}^2$ times the upstream rest mass energy density.

Hot Upstream

In some problems (e.g. energy injection into the shocked blastwave region, Kumar and Piran 2000b; Zhang and Mészáros 2002c), the upstream fluid is hot (relativistic). The $m_p c^2$ term in the expression of μ (Eq. (4.67)) can be dropped. Equations (4.68)–(4.70) can be revised to

$$n_1 \gamma_{1s} \beta_{1s} = n_2 \gamma_{2s} \beta_{2s}, \tag{4.81}$$

$$(e_1 + p_1)\gamma_{1s}^2 \beta_{1s} = (e_2 + p_2)\gamma_{2s}^2 \beta_{2s}, \tag{4.82}$$

$$(e_1 + p_1)\gamma_{1s}^2 \beta_{1s}^2 + p_1 = (e_2 + p_2)\gamma_{2s}^2 \beta_{2s}^2 + p_2. \tag{4.83}$$

This leads to the solution (Kumar and Piran, 2000b; Zhang and Mészáros, 2002c)

$$\beta_{1s} = \sqrt{\frac{e_2(\hat{\gamma}-1) + e_1(\hat{\gamma}-1)^2}{e_1 + e_2(\hat{\gamma}-1)}}, \tag{4.84}$$

$$\beta_{2s} = \sqrt{\frac{e_1(\hat{\gamma}-1) + e_1(\hat{\gamma}-1)^2}{e_2 + e_1(\hat{\gamma}-1)}}, \tag{4.85}$$

$$\gamma_{1s} = \sqrt{\frac{e_1 + e_2(\hat{\gamma}-1)}{\hat{\gamma}(2-\hat{\gamma})e_1}}, \tag{4.86}$$

[4] Sari and Piran (1995) adopted $\hat{\gamma} = 4/3$, so that Eq. (4.77) becomes $n_2/n_1 = 4\gamma_{21} + 3$, which cannot be reduced to $n_2/n_1 = 4$ in the non-relativistic regime.

$$\gamma_{2s} = \sqrt{\frac{e_2 + e_1(\hat{\gamma} - 1)}{\hat{\gamma}(2 - \hat{\gamma})e_2}}, \tag{4.87}$$

$$\gamma_{21}^2 = \frac{[e_1 + e_2(\hat{\gamma} - 1)][e_2 + e_1(\hat{\gamma} - 1)]}{\hat{\gamma}^2 e_1 e_2}, \tag{4.88}$$

$$\left(\frac{n_2}{n_1}\right)^2 = \frac{e_2[e_2 + e_1(\hat{\gamma} - 1)]}{e_1[e_1 + e_2\hat{\gamma} - 1]}. \tag{4.89}$$

For $\hat{\gamma} = 4/3$, these equations become

$$\beta_{1s} = \sqrt{\frac{3e_2 + e_1}{3(3e_1 + e_2)}}, \tag{4.90}$$

$$\beta_{2s} = \sqrt{\frac{3e_1 + e_2}{3(3e_2 + e_1)}}, \tag{4.91}$$

$$\gamma_{1s} = \sqrt{\frac{3(3e_1 + e_2)}{8e_1}}, \tag{4.92}$$

$$\gamma_{2s} = \sqrt{\frac{3(3e_2 + e_1)}{8e_2}}, \tag{4.93}$$

$$\gamma_{21}^2 = \frac{(3e_1 + e_2)(e_1 + 3e_2)}{16e_1 e_2}, \tag{4.94}$$

$$\left(\frac{n_2}{n_1}\right)^2 = \frac{e_2(e_1 + 3e_2)}{e_1(3e_1 + e_2)}. \tag{4.95}$$

4.2.4 MHD Waves

When a magnetic field exists in a fluid, the propagation of waves in the fluid becomes much more complicated. This is because the direction of $\mathbf{B}$ breaks the isotropy in the fluid. One must consider the interplay among three vectors: the magnetic field vector $\mathbf{B} = B\hat{\mathbf{b}}$, the wave propagation vector $\mathbf{k} = k\hat{\mathbf{k}}$, and the perturbation displacement vector $\boldsymbol{\xi}$.

Waves in a Non-relativistic, Low-σ Fluid

A non-relativistic MHD fluid is defined by $\beta \ll 1$, $p \ll \rho_0 c^2$, and the magnetization parameter $\sigma \ll 1$. The σ parameter is defined as

$$\sigma \equiv \frac{B_0^2}{4\pi\rho_0 c^2} = \frac{B^2}{4\pi\Gamma\rho c^2}, \tag{4.96}$$

which is the ratio between the total internal magnetic energy density (including magnetic energy density $B_0^2/8\pi$ plus magnetic pressure $B_0^2/8\pi$) and the rest mass energy density ($\rho_0 c^2$) in the rest frame of the fluid. In the lab frame where the magnetic field is B and density is ρ, the expression has a bulk Lorentz factor Γ in the denominator. The meaning of σ is the ratio between the Poynting flux energy density $|\mathbf{E} \times \mathbf{B}|/(4\pi) = B^2/(4\pi)$ (since $\mathbf{E} = -\boldsymbol{\beta} \times \mathbf{B}$, so that $|\mathbf{E}| \simeq |\mathbf{B}|$) and the matter flux $\Gamma\rho c^2$.

Let us consider a uniform fluid with a uniform magnetic field $\mathbf{B_0} = B_0\hat{\mathbf{b}}$ and a plane wave solution for displacement ($\boldsymbol{\xi}(\mathbf{r}, t) = \Sigma\boldsymbol{\xi}_k e^{i(\mathbf{k}\cdot\mathbf{r}+w_k t)}$). With the non-relativistic ideal MHD equations, one gets a general equation for MHD wave propagation (e.g. Schnack, 2009):

$$[\omega^2 - (\hat{\mathbf{k}}\cdot\hat{\mathbf{b}})^2 v_{\rm A}^2]\boldsymbol{\xi} = [(c_s^2 + v_{\rm A}^2)(\hat{\mathbf{k}}\cdot\boldsymbol{\xi}) - v_{\rm A}^2(\boldsymbol{\xi}\cdot\hat{\mathbf{b}})(\hat{\mathbf{k}}\cdot\hat{\mathbf{b}})]\hat{\mathbf{k}} - v_{\rm A}^2(\hat{\mathbf{k}}\cdot\boldsymbol{\xi})(\hat{\mathbf{k}}\cdot\hat{\mathbf{b}})\hat{\mathbf{b}}, \tag{4.97}$$

where $c_s = \sqrt{\hat{\gamma}p/\rho_0}$ is the non-relativistic sound speed, and

$$v_{\rm A} \equiv \frac{B_0}{\sqrt{4\pi\rho_0}} = \sqrt{\sigma}c \tag{4.98}$$

is the *non-relativistic Alfvén speed*. The vector equation (4.97) includes three equations, which would yield three roots for ω^2 and six possible waves:

- two *shear Alfvén waves*, with $\omega = \pm\omega_0$;
- two *magneto-sonic (or magneto-acoustic (MA)) waves*, with $\omega = \pm\omega_1$;
- two *sound waves*, with $\omega = \pm\omega_2$.

To disentangle these wave modes, let us define that the direction of the magnetic field is the z-direction, i.e. $\hat{\mathbf{b}} = \hat{\mathbf{e}}_z$, that the direction of wave propagation is in the x–z plane, i.e. $\mathbf{k} = k_\perp\hat{\mathbf{e}}_\mathbf{x} + k_\parallel\hat{\mathbf{e}}_\mathbf{z}$ with $k_\perp = k\sin\theta$ and $k_\parallel = k\cos\theta$ (angle θ defined as the angle between vector $\hat{\mathbf{b}}$ and $\hat{\mathbf{k}}$), and that the direction of displacement is arbitrary, i.e. $\boldsymbol{\xi} = \xi_x\hat{\mathbf{e}}_x + \xi_y\hat{\mathbf{e}}_y + \xi_z\hat{\mathbf{e}}_z$. The vector equation (4.97) can then be disentangled into three equations:

$$x\text{-component}: (v_{\rm A}^2k^2 + c_s^2k_\perp^2)\xi_x + c_s^2k_\parallel k_\perp\xi_z = \omega^2\xi_x, \tag{4.99}$$

$$y\text{-component}: v_{\rm A}^2k_\parallel^2\xi_y = \omega^2\xi_y, \tag{4.100}$$

$$z\text{-component}: c_s^2k_\parallel k_\perp\xi_x + c_s^2k_\parallel^2\xi_z = \omega^2\xi_z. \tag{4.101}$$

The y-component decouples from the other two components. Solving Eq. (4.100) directly gives the *phase velocity*, ω/k, of the well-known transverse wave, i.e. the *shear Alfvén* mode:

$$\left(\frac{\omega}{k}\right)_0^2 = v_{\rm A}^2\cos^2\theta. \tag{4.102}$$

The phase speed is $(\omega/k)_0 = v_{\rm A}$ at $\theta = 0$, and 0 at $\theta = \pi/2$.

Jointly solving Eqs. (4.99) and (4.101) gives a quadratic equation for ω^2, i.e.

$$\omega^4 - (v_{\rm A}^2 + c_s^2)k^2\omega^2 + c_s^4k_\parallel^2k_\perp^2 = 0. \tag{4.103}$$

This gives solutions of the phase velocities of two *longitudinal* waves:

$$\left(\frac{\omega}{k}\right)_{1,2}^2 = v_{\rm F,S}^2 = \frac{1}{2}(c_s^2 + v_{\rm A}^2)\left[1 \pm \sqrt{1 - \frac{4c_s^2v_{\rm A}^2\cos^2\theta}{(c_s^2 + v_{\rm A}^2)^2}}\right]. \tag{4.104}$$

The mode 1 with "+" sign is the *fast magneto-sonic* (or *fast MA*) wave mode, whereas mode 2 with "−" sign is the *slow magneto-sonic* (or *slow MA*) wave mode. For the fast MA mode, one has $v_{\rm F} = (\omega/k)_1 = v_{\rm A}$ at $\theta = 0$, and $= \sqrt{v_{\rm A}^2 + c_s^2}$ at $\theta = \pi/2$; whereas for

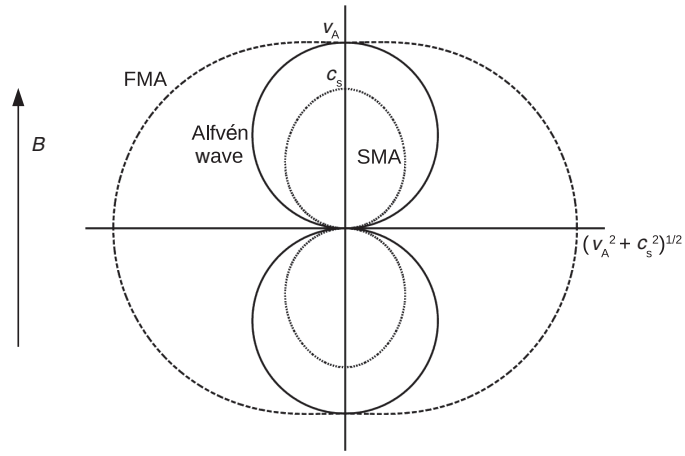

Figure 4.1 Phase velocities of the three MHD wave modes as a function of angle from the magnetic field direction in a uniform medium. Figure courtesy Wei Deng.

the slow MA mode, one has $v_S = (\omega/k)_2 = c_s$ at $\theta = 0$, and $= 0$ at $\theta = \pi/2$. The latter is essentially a sound wave.

Figure 4.1 shows the phase velocity of the three modes at an arbitrary angle. One can see that the largest speeds are v_A, c_s, and $\sqrt{v_A^2 + c_s^2}$ for Alfvén, slow MA, and fast MA waves, respectively.

In a magnetized fluid, a MHD shock is excited if the relative speed of another fluid with respect to this fluid exceeds the maximum value of the fast MA speed, i.e. $v > v_{F,max} = \sqrt{v_A^2 + c_s^2}$.

Waves in a Relativistic, High-σ Fluid

In the relativistic regime, characterized by $\beta \sim 1$, $p \geq \rho_0 c^2$, or $\sigma \geq 1$, the phase velocities of the three modes are modified.

The general expression of the *Alfvén speed in the relativistic regime* can be derived by replacing $\rho_0 c^2$ in Eq. (4.98) by $\rho c^2 = \rho_0 c^2 + e + p + B_0^2/4\pi$, i.e.

$$v_A = c \frac{B_0}{\sqrt{4\pi \left(\rho_0 c^2 + \frac{\hat{\gamma}}{\hat{\gamma}-1} p + \frac{B_0^2}{4\pi} \right)}}, \tag{4.105}$$

which can be reduced to Eq. (4.98) when $p \ll \rho_0 c^2$ and $B_0^2/(4\pi\rho_0 c^2) \ll 1$.

Along the same lines, the relativistic phase speed of the fast/slow MA waves can be written

$$\left(\frac{\omega}{k}\right)_{1,2}^2 = v_{F,S} = \frac{1}{2}\left[v_A^2 + c_s^2 \left(1 - \frac{v_A^2}{c^2}\right) + \frac{v_A^2}{c^2} c_s^2 \cos^2\theta \right] \pm \sqrt{\left[v_A^2 + c_s^2 \left(1 - \frac{v_A^2}{c^2}\right) + \frac{v_A^2}{c^2} c_s^2 \cos^2\theta \right]^2 - 4 c_s^2 v_A^2 \cos^2\theta}. \tag{4.106}$$

The maximum speed of the fast MA waves is

$$v_{\rm F,max}=\sqrt{v_{\rm A}^2+c_s^2\left(1-\frac{v_{\rm A}^2}{c^2}\right)}=c\sqrt{\frac{\hat{\gamma}p+\frac{B_0^2}{4\pi}}{\rho_0c^2+\frac{\hat{\gamma}}{\hat{\gamma}-1}p+\frac{B_0^2}{4\pi}}}. \tag{4.107}$$

A relativistic MHD fluid can be "cold", namely, $p\ll\rho_0c^2$, but $\sigma\gg1$. For such a case, the Alfvén speed reads

$$v_{\rm A}=\frac{v_{\rm A,NR}}{\sqrt{1+v_{\rm A,NR}^2/c^2}}=\sqrt{\frac{\sigma}{1+\sigma}}c, \tag{4.108}$$

where $v_{\rm A,NR}$ follows Eq. (4.98). The corresponding Alfvén Lorentz factor reads

$$\gamma_{\rm A}=\sqrt{1+\sigma}. \tag{4.109}$$

The fast MA wave speed can be written as

$$v_{\rm F,max}=\sqrt{v_{\rm A}^2+\frac{c_s^2}{\gamma_{\rm A}^2}}=c\sqrt{\frac{\sigma}{1+\sigma}+\frac{c_s^2}{c^2}\frac{1}{1+\sigma}}, \tag{4.110}$$

so that

$$\gamma_{\rm F,max}=\sqrt{1+\sigma}\cdot\sqrt{\frac{c^2}{c^2-c_s^2}}. \tag{4.111}$$

For $c_s^2=(1/3)c^2$ (relativistic fluid), one has $\gamma_{\rm F,max}=\sqrt{3/2}\cdot\sqrt{1+\sigma}\simeq1.22\gamma_{\rm A}$.

4.2.5 MHD Shocks

For a relativistic shock propagating into a magnetized medium characterized by σ (see definition in Eq. (4.96)), the electromagnetic terms would explicitly enter the shock jump condition conservation equations. The equations become more complicated. Below, we discuss a simple case with ordered magnetic fields lying in the plane of shock, i.e. $\langle\mathbf{B}\cdot\hat{\mathbf{n}}\rangle=90^{\rm o}$, where $\hat{\mathbf{n}}$ is the unit vector of the shock plane normal (called a $90^{\rm o}$ shock). Such a configuration may be a good approximation for GRBs. The GRB central engine is believed to be strongly magnetized, so that the ejecta wind launched from the GRB central engine may carry a globally ordered magnetic field with both a poloidal and a toroidal component. For a conical jet (the jet opening angle θ_j remains constant during the propagation), due to magnetic flux conservation, the poloidal component decays with radius more rapidly ($B_p\propto r^{-2}$) than the toroidal component ($B_t\propto r^{-1}$). At the GRB emission radius, where $R_{\rm GRB}\gg R_0$ (R_0 is the size of the GRB central engine), one usually has $B_t\gg B_p$, so that the field lines are approximately in the shock plane.

The Rankine–Hugoniot jump conditions for such a MHD shock are modified from those of a hydrodynamic shock. First, the magnetic field in the rest frame of upstream 1, when viewed in the shock frame, should have an associated electric field E, which should equal the E field in the downstream. One therefore has a new equation (Eq. (4.113) below) to connect the magnetic fields between the upstream and the downstream. Next, the internal energy e and pressure p should be modified to include the magnetic contribution, i.e. e_i

should be replaced by $e_i + B_i^2/(8\pi)$, and p_i should be replaced by $p_i + B_i^2/(8\pi)$. Plugging them into the relativistic shock jump conditions Eqs. (4.68)–(4.70), one finally gets the following set of equations (Kennel and Coroniti, 1984; Zhang and Kobayashi, 2005) (Exercise 4.6):

$$n_1 u_{1s} = n_2 u_{2s}, \tag{4.112}$$

$$E = \beta_{1s} B_{1s} = \beta_{2s} B_{2s}, \tag{4.113}$$

$$\gamma_{1s}\mu_1 + \frac{EB_{1s}}{4\pi n_1 u_{1s}} = \gamma_{2s}\mu_2 + \frac{EB_{2s}}{4\pi n_2 u_{2s}}, \tag{4.114}$$

$$\mu_1 u_{1s} + \frac{p_1}{n_1 u_{1s}} + \frac{B_{1s}^2}{8\pi n_1 u_{1s}} = \mu_2 u_{2s} + \frac{p_2}{n_2 u_{2s}} + \frac{B_{2s}^2}{8\pi n_2 u_{2s}}, \tag{4.115}$$

where B_{1s} and B_{2s} are the strengths of the magnetic fields in the upstream and downstream in the rest frame of the shock front, respectively.

Let us consider a cold upstream $e_1 = p_1 = 0$ (so that $\mu_1 = m_p c^2$), and define the σ parameter of the cold upstream (noticing $B_1 = B_{1s}/\gamma_{1s}$):

$$\sigma \equiv \sigma_1 = \frac{B_1^2}{4\pi n_1 \mu_1} = \frac{B_{1s}^2}{4\pi n_1 \mu_1 \gamma_{1s}^2}. \tag{4.116}$$

The Rankine–Hugoniot relations can be cast in the following form (Zhang and Kobayashi, 2005):

$$\frac{e_2}{n_2 m_p c^2} = (\gamma_{21} - 1)\left[1 - \frac{\gamma_{21} + 1}{2u_{1s}(\gamma_{21}, \sigma) u_{2s}(\gamma_{21}, \sigma)}\sigma\right], \tag{4.117}$$

$$\frac{n_2}{n_1} = \frac{u_{1s}(\gamma_{21}, \sigma)}{u_{2s}(\gamma_{21}, \sigma)} = \gamma_{21} + \frac{[u_{2s}^2(\gamma_{21}, \sigma) + 1]^{1/2}}{u_{2s}(\gamma_{21}, \sigma)}(\gamma_{21}^2 - 1)^{1/2}, \tag{4.118}$$

where

$$u_{1s}(\gamma_{21}, \sigma) = u_{2s}(\gamma_{21}, \sigma)\gamma_{21} + [u_{2s}^2(\gamma_{21}, \sigma) + 1]^{1/2}(\gamma_{21}^2 - 1)^{1/2}, \tag{4.119}$$

and u_{2s}^2 is the root of the following equation:

$$Ax^3 + Bx^2 + Cx + D = 0, \tag{4.120}$$

where

$$A = \hat{\gamma}(2 - \hat{\gamma})(\gamma_{21} - 1) + 2, \tag{4.121}$$

$$\begin{aligned} B = &-(\gamma_{21} + 1)\left[(2 - \hat{\gamma})(\hat{\gamma}\gamma_{21}^2 + 1) + \hat{\gamma}(\hat{\gamma} - 1)\gamma_{21}\right]\sigma \\ &-(\gamma_{21} - 1)\left[\hat{\gamma}(2 - \hat{\gamma})(\gamma_{21}^2 - 2) + (2\gamma_{21} + 3)\right], \end{aligned} \tag{4.122}$$

$$\begin{aligned} C = &(\gamma_{21} + 1)\left[\hat{\gamma}\left(1 - \frac{\hat{\gamma}}{4}\right)(\gamma_{21}^2 - 1) + 1\right]\sigma^2 \\ &+(\gamma_{21}^2 - 1)\left[2\gamma_{21} - (2 - \hat{\gamma})(\hat{\gamma}\gamma_{21} - 1)\right]\sigma \\ &+(\gamma_{21} + 1)(\gamma_{21} - 1)^2(\hat{\gamma} - 1)^2, \end{aligned} \tag{4.123}$$

$$D = -(\gamma_{21} - 1)(\gamma_{21} + 1)^2(2 - \hat{\gamma})^2\frac{\sigma^2}{4}. \tag{4.124}$$

For $\hat{\gamma} = 4/3$, the four coefficients can equivalently be written as

$$A = 8\gamma_{21} + 10, \tag{4.125}$$

$$B = -(\gamma_{21}+1)(8\gamma_{21}^2 + 4\gamma_{21} + 6)\sigma - (\gamma_{21}-1)(8\gamma_{21}^2 + 18\gamma_{21} + 11), \tag{4.126}$$

$$C = (\gamma_{21}+1)(8\gamma_{21}^2+1)\sigma^2 + (\gamma_{21}^2-1)(10\gamma_{21}+6)\sigma + (\gamma_{21}+1)(\gamma_{21}-1)^2, \tag{4.127}$$

$$D = -(\gamma_{21}-1)(\gamma_{21}+1)^2\sigma^2. \tag{4.128}$$

No analytical expression of u_{2s} is available.

This general solution can be reduced to simpler equations under certain conditions:

- When $\sigma = 0$, the solution can be reduced to the standard hydrodynamic equations (4.71)–(4.75).

- In the $\gamma_{21} \gg 1$ limit, an analytical solution of u_{2s} is available:

$$u_{2s}^2 = \frac{\hat{\gamma}(1-\frac{\hat{\gamma}}{4})\sigma^2 + (\hat{\gamma}^2 - 2\hat{\gamma} + 2)\sigma + (\hat{\gamma}-1)^2 + \sqrt{X}}{2\hat{\gamma}(2-\hat{\gamma})(\sigma+1)}, \tag{4.129}$$

  where

$$X = \hat{\gamma}^2\left(1-\frac{\hat{\gamma}}{4}\right)^2\sigma^4 + \hat{\gamma}\left(\frac{\hat{\gamma}^3}{2} - 3\hat{\gamma}^2 + 7\hat{\gamma} - 4\right)\sigma^3 + \left(\frac{3}{2}\hat{\gamma}^4 - 7\hat{\gamma}^3 + \frac{31}{2}\hat{\gamma}^2 - 14\hat{\gamma} + 4\right)\sigma^2 + 2(\hat{\gamma}-1)^2(\hat{\gamma}^2 - 2\hat{\gamma} + 2)\sigma + (\hat{\gamma}-1)^4. \tag{4.130}$$

  Notice that the quadratic equation of u_{2s}^2 has another solution invoking $-\sqrt{X}$. However, the solution gives rise to a negative pressure, which is unphysical.

- For a relativistic downstream region, i.e. $\hat{\gamma} = 4/3$, the solution is reduced to

$$u_{2s}^2 = \frac{8\sigma^2 + 10\sigma + 1 + \sqrt{64\sigma^2(\sigma+1)^2 + 20\sigma(\sigma+1) + 1}}{16(\sigma+1)} = \frac{8\sigma^2 + 10\sigma + 1 + (2\sigma+1)\sqrt{16\sigma^2 + 16\sigma + 1}}{16(\sigma+1)}. \tag{4.131}$$

  This is Eq. (4.11) of Kennel and Coroniti (1984), which has been used to treat pulsar wind nebula problems.

4.3 Forward–Reverse Shock System

A disturbance that induces a shock wave usually invokes a "supersonic" collision between two fluids. If the relative speed between the two fluids exceeds the sound speed (or the fast MA speed for a MHD fluid) of one fluid, it usually also exceeds that of the second one. So

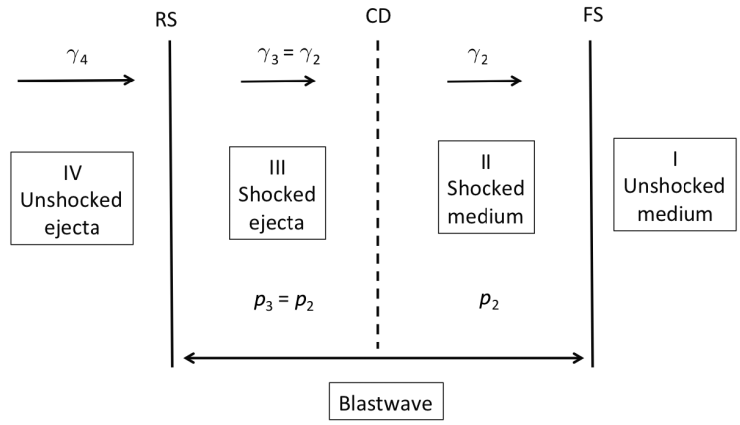

Figure 4.2 A system of colliding two fluids separated by two (forward and reverse) shocks and one contact discontinuity. The region between the two shocks is defined as the blastwave.

a pair of shocks will develop. For a jet running into a medium at rest in the cosmic proper frame (such as the GRB problem), the shock propagating into the medium is called the *forward shock* (FS),[5] while that propagating into the jet itself is called the *reverse shock* (RS). The region between the FS and the RS is called the *blastwave* (Fig. 4.2).

In a FS–RS system, there are four regions separated by three discontinuities (Fig. 4.2): I. the unshocked medium; II. the shocked medium; III. the shocked ejecta; IV. the unshocked ejecta. The FS separates Regions I and II, whereas the RS separates Regions III and IV. The surface between Region II and Region III is called the *contact discontinuity* (CD).

4.3.1 Cold Fluids

The simplest case is that both Regions I and IV are cold (internal energy negligible) and hydrodynamic (no dynamically important magnetic fields).

For an unmagnetized ejecta, the FS/RS system can be solved with the following conditions:

$$\frac{e_2}{n_2 m_p c^2} = \gamma_{21} - 1, \tag{4.132}$$

$$\frac{n_2}{n_1} = \frac{\hat{\gamma}_2 \gamma_{21} + 1}{\hat{\gamma}_2 - 1}, \tag{4.133}$$

$$\frac{e_3}{n_3 m_p c^2} = \gamma_{34} - 1, \tag{4.134}$$

$$\frac{n_3}{n_4} = \frac{\hat{\gamma}_3 \gamma_{34} + 1}{\hat{\gamma}_3 - 1}, \tag{4.135}$$

where $\hat{\gamma}_2$ and $\hat{\gamma}_3$ are the adiabatic indices in Regions II and III, respectively.

[5] For the jet systems whose images are available, the forward shock usually displays a "bow" shape, so sometimes the FS is also called the *"bow" shock* in other fields. GRBs are too far away to have the jets directly imaged, so usually this term is not used in GRB problems.

In such problems, the properties of Regions I and IV (e.g. n_1, n_4, γ_4) are specified. There are six unknowns in the problem: γ_{12}, γ_{34}, e_2, e_3, n_2, and n_3. In order to close the problem, besides the above four equations, two more equations are needed. The first is that the pressure across the CD should be the same in order to achieve a balance, i.e. $p_2 = p_3$. This can be written as

$$(\hat{\gamma}_2 - 1)e_2 = (\hat{\gamma}_3 - 1)e_3. \tag{4.136}$$

The second condition is that Regions II and III should move with the same Lorentz factor, i.e.

$$\gamma_2 = \gamma_3, \tag{4.137}$$

where $\gamma_2 = \gamma_{21}$ and $\gamma_3 = \gamma_{31}$. Otherwise the blastwave region would detach ($\gamma_3 < \gamma_2$) or squeeze ($\gamma_3 > \gamma_2$).

Reducing Eqs. (4.132)–(4.137) and noting Eq. (4.20), one gets the following simple formula relation:[6]

$$\frac{n_4}{n_1} = \frac{\gamma_{21}^2 - 1}{\gamma_{34}^2 - 1}. \tag{4.138}$$

Assuming $\gamma_{21} \gg 1$ and $\gamma_{34} \gg 1$, and noting $\gamma_{34} = \gamma_{43} \simeq (1/2)(\gamma_4/\gamma_3 + \gamma_3/\gamma_4)$, one gets (Sari and Piran, 1995)

$$\gamma_2 = \gamma_3 \simeq \frac{\gamma_4^{1/2}(n_4/n_1)^{1/4}}{\sqrt{2}}, \tag{4.139}$$

and

$$\gamma_{34} = \frac{\gamma_4^{1/2}}{\sqrt{2}(n_4/n_1)^{1/4}}. \tag{4.140}$$

Notice that in the above treatment, one has assumed a uniform pressure and Lorentz factor in Regions II and III. This condition is not met in reality, since it violates energy conservation. Nonetheless, for a short-lived reverse shock, this approximation is reasonably good. We will come back to this issue in §4.3.4.

4.3.2 Magnetized Fluid(s)

In the GRB problem, it is possible that the jet contains a dynamically important magnetic field. For the external shocks (FS and RS), a more general problem is to treat an arbitrarily magnetized jet interacting with a non-magnetized medium (Zhang and Kobayashi, 2005).

Let us consider an arbitrarily magnetized ejecta characterized by a magnetization parameter $\sigma = \sigma_4$. Equations (4.134) and (4.135) are modified to

$$\frac{e_3}{n_3 m_p c^2} = (\gamma_{34} - 1)f_a, \tag{4.141}$$

[6] Assuming that both Regions I and IV are extremely relativistic fluids, so that $\hat{\gamma}_2 = \hat{\gamma}_3 = 4/3$, Sari and Piran (1995) derived $n_4/n_1 = [(\gamma_2 - 1)(4\gamma_2 + 3)]/[(\gamma_{34} - 1)(4\gamma_{34} + 3)]$. By applying a more general expression of $\hat{\gamma}$ (Eq. (4.20)), we note that the factor $(4\gamma + 3)$ becomes $(4\gamma + 4)$, so that the ratio n_4/n_1 can be reduced to the elegant form in Eq. (4.138), which is valid for any bulk Lorentz factor γ_{12} and γ_{34}.

$$\frac{n_3}{n_4} = \frac{\hat{\gamma}_3\gamma_{34}+1}{\hat{\gamma}_3-1} f_b, \tag{4.142}$$

where

$$f_a = f_a(\sigma,\gamma_{34}) = 1 - \frac{\gamma_{34}+1}{2[u_{3s}^2\gamma_{34} + u_{3s}(u_{3s}^2+1)^{1/2}(\gamma_{34}^2-1)^{1/2}]}\sigma, \tag{4.143}$$

$$f_b = f_b(\sigma,\gamma_{34}) = \frac{\gamma_{34} + \frac{(u_{3s}^2+1)^{1/2}}{u_{3s}}(\gamma_{34}^2-1)^{1/2}}{\frac{\hat{\gamma}_3\gamma_{34}+1}{\hat{\gamma}_3-1}}. \tag{4.144}$$

The pressure equilibrium equation (4.136) is modified to

$$(\hat{\gamma}_2-1)e_2 = (\hat{\gamma}_3-1)e_3 f_c, \tag{4.145}$$

with

$$f_c = 1 + \frac{p_{b,3}}{p_3}, \tag{4.146}$$

where $p_{b,3} = B_3^2/8\pi$, and $p_3 = (\hat{\gamma}_3-1)e_3$.

Equation (4.138) is correspondingly modified to[7]

$$F\frac{n_4}{n_1} = \frac{\gamma_{21}^2-1}{\gamma_{34}^2-1}, \tag{4.147}$$

where

$$F = f_a f_b f_c. \tag{4.148}$$

In the above derivations, the existence of a reverse shock is already pre-assumed. In reality, a reverse shock cannot exist if σ is high enough, so that the thermal pressure in Region II cannot overcome the magnetic pressure in Region IV. The reverse shock condition can be written as (Zhang and Kobayashi, 2005)

$$\frac{4}{3}\gamma_4^2 n_1 m_p c^2 > \frac{B_4^2}{8\pi}, \tag{4.149}$$

or

$$\sigma < \sigma_c \equiv \frac{8}{3}\gamma_4^2\frac{n_1}{n_4}. \tag{4.150}$$

So for a given ejecta Lorentz factor γ_4 and density radio n_1/n_4, there exists a characteristic magnetization factor σ_c, above which no reverse shock can form.[8] A reverse shock in a high-σ flow is typically weak, since the available (kinetic) energy to tap is only a factor $(1+\sigma)^{-1}$ of the total wind energy. The majority of the energy (a fraction of $\sigma/(1+\sigma)$) is stored in the magnetic fields and cannot be tapped without dissipation.

[7] This is essentially Eq. (32) of Zhang and Kobayashi (2005), with the factor $(4\gamma+3)$ parameter replaced by $(4\gamma+4)$, which is a more precise description for an arbitrary γ value. See Footnote 6.

[8] It is a misunderstanding that a reverse shock does not exist when $\sigma > 1$. Giannios et al. (2008) made an argument for this condition, but soon the same group found that a weak reverse shock does exist in the $\sigma > 1$ regime (Mimica et al., 2009). The correct condition is the pressure balance condition (4.150), as suggested by Zhang and Kobayashi (2005) and proven by the solution to a 1-D Riemann problem of the deceleration of an arbitrarily magnetized relativistic outflow (Mizuno et al., 2009).

In the GRB prompt emission models, collisions between two (highly) magnetized shells are possible. For two identical, magnetized blobs, since the magnetic pressure is already in balance, a violent collision with the relative Lorentz factor exceeding the Alfvén Lorentz factor $\gamma_{\rm A} \simeq (1+\sigma)^{1/2}$ would trigger a pair of shocks propagating into both magnetized shells (e.g. Zhang and Yan, 2011; Narayan et al., 2011). The shocks are weak however, again because of the small fraction ($(1+\sigma)^{-1}$) of the available energy involved.

4.3.3 Hot Fluid(s)

In the case that one or more upstream(s) are hot, one should use the hot jump conditions (Eqs.(4.90)–(4.95)) to solve the problem.

For example, Kumar and Piran (2000b) considered the collision between two hot shells, which may be relevant to some internal shocks (each of the two shells might have just experienced a collision and internal shock heating). Zhang and Mészáros (2002c) considered a late shell catching up with the blastwave (which is itself hot). There are six distinct regions in the problem: I. unshocked ISM; II. shocked ISM; III. hot early ejecta (heated by an early reverse shock); IV. an even hotter shocked ejecta by the late shell; V. reverse-shocked late shell; and VI. unshocked late shell. The hot jump conditions apply at the shock between Region III and Region IV.

For hot upstream problems, whether or not a strong shock will form not only depends on the standard supersonic condition, but also depends on whether the pressure in the shocked region of another fluid can exceed the pressure in the hot fluid. In order to excite a shock into a hot fluid, a strong collision (with large Lorentz factor contrast) is needed (Zhang and Mészáros, 2002c).

4.3.4 Mechanical Model

Even though at the contact discontinuity pressure is balanced, Eqs. (4.136) and (4.137) actually make the assumption that, within the entire blastwave (Regions II and III), both the pressure and the Lorentz factor are constant. In reality, there should exist a pressure gradient and probably a γ-profile as well within the blastwave. However, no simple analytical description is available to determine these profiles.

Beloborodov and Uhm (2006) pointed out that the assumptions of constant p and γ (Eqs. (4.136) and (4.137)) do not conserve energy. An accurate treatment of the pressure gradient and Lorentz factor profile in the blastwave region can only be achieved via numerical simulations. Beloborodov and Uhm (2006) developed a *mechanical model* by introducing a pressure gradient, but still adopting the constant Lorentz factor (Eq. (4.137)) assumption. This treatment guarantees energy conservation (Uhm, 2011; Uhm et al., 2012), and can more accurately model emission of the blastwave system, especially if there exists a long-lasting reverse shock. For systems with a short-lived reverse shock, the treatment invoking the two assumptions (as extensively discussed above) is a reasonably good approximation.

The mechanical model solves the following set of equations: in a spherical coordinate system, the three relativistic hydrodynamical equations (4.9), (4.10), and (4.11) are written as (Beloborodov and Uhm 2006, see Uhm 2011 for derivations)

$$\frac{1}{r^2 c}\frac{d}{dt}(r^2\rho\gamma) = -\rho\gamma\frac{\partial\beta}{\partial r}, \tag{4.151}$$

$$\frac{1}{r^2 c}\frac{d}{dt}(r^2 h\gamma^2\beta) = -\frac{\partial p}{\partial r} - h\gamma^2\beta\frac{\partial\beta}{\partial r}, \tag{4.152}$$

$$\frac{1}{r^2 c}\frac{d}{dt}(r^2 h\gamma) = \frac{\gamma}{c}\frac{dp}{dt} - h\gamma\frac{\partial\beta}{\partial r}, \tag{4.153}$$

where $d/dt \equiv \partial/\partial t + c\beta(\partial/\partial r)$, $\beta = v/c$, and other symbols take their standard definitions. The key assumption of the treatment is

$$\gamma(t, r) = \Gamma(t),\ \partial\beta/\partial r = 0,\ r_r < r < r_f, \tag{4.154}$$

i.e. within the blast between the reverse shock front (at r_r) and forward shock front (at r_f), the Lorentz factor is uniform at a given time ($\Gamma(t)$), so that there is no velocity gradient ($\partial\beta/\partial r = 0$). Dropping out the $\partial\beta/\partial r$ terms in the above equations, one gets the following equations at any instant t:

$$\frac{\Gamma}{r^2}\frac{d}{dr}(r^2\Sigma\Gamma) = \rho_r(\beta - \beta_r)\Gamma^2 + \frac{1}{4}\rho_f, \tag{4.155}$$

$$\frac{1}{r^2}\frac{d}{dr}(r^2 H\Gamma^2) = h_r(\beta - \beta_r)\Gamma^2 + p_r, \tag{4.156}$$

$$\frac{\Gamma}{r^2}\frac{d}{dr}(r^2 H\Gamma) = \Gamma^2\frac{dP}{dr} + (h_r - p_r)(\beta - \beta_r)\Gamma^2 + \frac{3}{4}p_f, \tag{4.157}$$

where the subscripts "f" and "r" denote the FS and RS, respectively, $c\beta_r = dr_r/dt$, $c\beta_f = dr_f/dt$, $\Sigma \equiv \int_{r_r}^{r_f}\rho dr$, $H \equiv \int_{r_r}^{r_f} h dr$, and $P \equiv \int_{r_r}^{r_f} p dr$. Noting that

$$\beta - \beta_r = \frac{\Gamma_4^2 - \Gamma^2}{2\Gamma^2(\Gamma_4^2 + 2\Gamma^2)}, \tag{4.158}$$

$$\rho_r = 2\left(\frac{\Gamma_4}{\Gamma} + \frac{\Gamma}{\Gamma_4}\right)\rho_4, \tag{4.159}$$

$$p_r = \frac{1}{3}\left(\frac{\Gamma_4}{\Gamma} - \frac{\Gamma}{\Gamma_4}\right)^2\rho_4 c^2, \tag{4.160}$$

$$h_r = \frac{4}{3}\left(\frac{\Gamma_4^2}{\Gamma^2} + \frac{\Gamma^2}{\Gamma_4^2} + 1\right), \tag{4.161}$$

where Γ_4 and ρ_4 are the Lorentz factor and density of the unshocked ejecta, one finally has three equations (Eqs. (4.155)–(4.157)) solving for four unknowns, Σ, H, P, and Γ, with known parameters (Γ_4, ρ_4, etc.). In order to close the problem, Beloborodov and Uhm (2006) assumed an approximate "equation of state" relation,

$$H - \Sigma c^2 = 4P, \tag{4.162}$$

which is accurate in the limits of ultra-relativistic and non-relativistic regimes.

The mechanical model has been used to treat problems that invoke a long-lasting reverse shock (e.g. Uhm and Beloborodov, 2007; Uhm, 2011; Uhm et al., 2012).

4.4 Particle Acceleration

Astrophysical shocks are ideal sites for accelerating charged particles (electrons and ions) to high energies. The accelerated particles typically have a power-law distribution in energy. The non-thermal electrons accelerated in the shocks emit photons via synchrotron radiation or inverse Compton scattering. The high-energy protons/ions accelerated in the shocks may produce high-energy photons and neutrinos through hadronic processes (Chapter 6). Some ions escape the shock region and become *cosmic rays*.

The acceleration of particles in shocks proceeds through the so-called *first-order Fermi acceleration mechanism*. In the following, we first discuss a less-efficient *second-order Fermi acceleration* mechanism through *stochastic processes* (§4.4.1) originally proposed by Enrico Fermi. Then we introduce the more efficient *first-order*, *diffusive shock* acceleration processes for both non-relativistic (§4.4.2) and relativistic (§4.4.3) shocks. Sections 4.4.1 and 4.4.2 closely follow Longair (2011).

4.4.1 Second-Order, Stochastic Fermi Acceleration

In his original paper (Fermi, 1949), Fermi considered charged particles gaining energy in interstellar space through stochastic "collisions" against "moving magnetic fields", i.e. moving plasma clouds that carry magnetic fields. Fermi found that on average the energy gain for each collision is proportional to $(V/c)^2$, where V is the average random motion speed of the clouds. Since $V/c \ll 1$ and the energy gain is to the second order of this small number, this mechanism is inefficient.

The origin of the $(V/c)^2$ factor comes from two effects:

1. For random motion of clouds with an average speed V, the chance for a particle with a velocity $v \sim c$ to collide with clouds from different directions is not exactly the same. There is a slightly higher probability for *head-on* collisions (which lead to energy gain) than *tail-on* collisions (which lead to energy loss). This *net energy-gain probability* is of the order (V/c).

To simplify the problem, one may consider a one-dimensional case with a particle colliding with the cloud either head-on or tail-on. For a more realistic three-dimensional problem, an angular average is involved, but the main conclusion remains the same.

Let us assume that the mean distance between clouds is d. The time scales for a head-on and a tail-on collision are

$$\tau_h = \frac{d}{v+V}, \quad \tau_t = \frac{d}{v-V}, \tag{4.163}$$

respectively. So, on average, there are more head-on collisions than tail-on collisions, and the probabilities for the two kinds of collisions are

$$P_h = \frac{\tau_h^{-1}}{\tau_h^{-1}+\tau_t^{-1}} = \frac{v+V}{2v}, \quad P_t = \frac{\tau_t^{-1}}{\tau_h^{-1}+\tau_t^{-1}} = \frac{v-V}{2v}, \tag{4.164}$$

respectively. One can see that the difference between the two probabilities is of the order of V/c for $v \sim c$.

2. For each collision, the energy gain/loss is also of the order (V/c).

For particles interacting with magnetic fields, when energy loss due to radiation is negligible, one may assume an approximately elastic collision. The change of particle velocity for each collision is therefore

$$\Delta v = \pm 2V. \tag{4.165}$$

The energy change is

$$\Delta E_h = \frac{1}{2}m(v+2V)^2 - \frac{1}{2}mv^2 \simeq \frac{1}{2}mv^2 \cdot \left(\frac{4V}{v}\right) = E_k\left(\frac{4V}{v}\right) \tag{4.166}$$

and

$$\Delta E_t = \frac{1}{2}m(v-2V)^2 - \frac{1}{2}mv^2 \simeq -\frac{1}{2}mv^2 \cdot \left(\frac{4V}{v}\right) = -E_k\left(\frac{4V}{v}\right), \tag{4.167}$$

for head-on and tail-on collisions, respectively, where $E_k = (1/2)mv^2$ is the kinetic energy of the particle. One can see that the net change is a factor of (V/c) for $v \sim c$.

Putting these two effects together, one can estimate the average energy gain after many collisions:

$$\langle \Delta E_k \rangle = P_h \Delta E_h + P_t \Delta E_t = \left(\frac{v+V}{2v}\frac{4V}{v} - \frac{v-V}{2v}\frac{4V}{v}\right)E_k = 4\left(\frac{V}{v}\right)^2 E_k. \tag{4.168}$$

One can immediately see that it is to the second order in (V/c) for $v \sim c$.

One can also show that the accelerated particles have a power-law distribution in energy.

On average, a particle gains a small amount of energy after each collision, i.e.

$$E = \zeta E_0 \tag{4.169}$$

with $\zeta \geq 1$, where E_0 and E are the energies of the particle before and after the collision, respectively. During each collision, there is a small probability that the particle, after gaining the energy, will escape the acceleration region. Considering a system with N_0 particles at energy E_0, one can introduce a probability, $P \leq 1$, for the particle to remain in the acceleration region after gaining the energy. The total number of particles at energy $E > E_0$ is therefore

$$N = PN_0. \tag{4.170}$$

Now consider the case of k collisions. After the collisions, the typical particle energy is

$$E = E_0 \zeta^k, \tag{4.171}$$

whereas the total number of particles at energy E is

$$N = N_0 P^k. \tag{4.172}$$

Since $k = \ln(E/E_0)/\ln\zeta = \ln(N/N_0)/\ln P$, one can derive

$$\frac{N}{N_0} = \left(\frac{E}{E_0}\right)^{\ln P/\ln\zeta}, \tag{4.173}$$

or, in terms of $N(E) = dN/dE$,

$$N(E)dE = KE^{\ln P/\ln\zeta - 1}dE = KE^{-p}dE, \tag{4.174}$$

where

$$p = 1 - \ln P / \ln \zeta \tag{4.175}$$

is the power-law index of the particle energy distribution.

Even if this second-order Fermi mechanism catches the essence of particle acceleration, it has two main difficulties in interpreting the astrophysical data. First, it is inefficient (as presented above). Second, there is no available physical process to guide $\ln P$ and $\ln \zeta$ such that the predicted particle power-law index, Eq. (4.175), matches the observed value (typically 2 for supernova remnants and other non-relativistic synchrotron sources).

4.4.2 First-Order, Diffusive Shock Fermi Acceleration

The above-mentioned two difficulties can be overcome when astrophysical shocks are considered. With the existence of a shock, particles more efficiently gain energy by crossing the shock front back and forth. Because of the existence of strong magnetic fields near the shock front, charged particles gyrate around field lines (or are "scattered" by Alfvén waves), so that they can cross the shock front multiple times to gain energy. A particle gains an energy of the order (V/c) whenever it crosses the shock front from either (upstream or downstream) side of the shock front. So, effectively, one gets rid of the "tail-on" collisions that lose energy,[9] so that one factor of (V/c) related to the probability of gaining energy is removed. The acceleration process then becomes "first order" to (V/c). A detailed review on shock acceleration can be found in, e.g., Blandford and Eichler (1987). The physical basis of the mechanism is outlined in the following in the so-called "test particle" approach, with a *non-relativistic shock* taken as an example (Fig. 4.3).

A shock system can be viewed in three different frames: the upstream (region 1 denoted by (p_1, T_1, ρ_1)) frame, the shock front frame, and the downstream (region 2, denoted by (p_2, T_2, ρ_2)) frame. The velocity transformation for a non-relativistic shock is Galilean.

First, let us consider the system in the upstream frame and assume that the shock front is moving with a speed U towards the upstream (upper left Fig. 4.3). Now view the same system in the shock front frame (upper right Fig. 4.3). The upstream moving with a speed $v_1 = U$ is decelerated to a lower speed v_2 in the downstream. Let us define $\rho_2/\rho_1 = v_{1s}/v_{2s} = r$ (the compression ratio); one immediately derives $v_{2s} = v_{1s}/r = U/r$. So the relative speed between the downstream and upstream is $V = \frac{r-1}{r}U$, i.e. in the rest frame of the upstream, the downstream moves with a velocity $V = \frac{r-1}{r}U$ towards the upstream (lower left Fig. 4.3). When the same system is viewed in the downstream frame, the upstream velocity is $\frac{r-1}{r}U$ towards the downstream.

An important ingredient of the argument is that a particle can quickly adjust to the local frame it enters through scattering. The random motion directions of the particles in a certain stream are isotropic. When new particles enter this stream, even though they gain energy, the direction of motion is quickly randomized to become isotropic in the rest frame (lower panels of Fig. 4.3). Since the two streams always move relative to each other with a speed $V = \frac{r-1}{r}U$, on average, a particle will gain a small fractional energy

[9] Crossing the shock front from the downstream is still "tail-on", however, it still gains energy of the same order as "head-on" crossing the shock front from the upstream. See more explanations below.

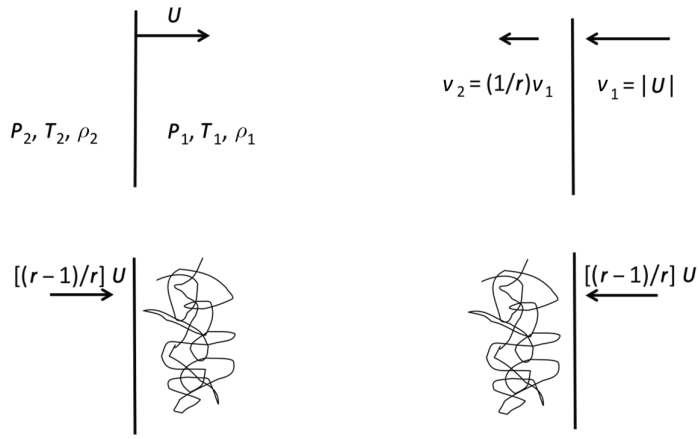

Figure 4.3 Different viewpoints of a non-relativistic shock system. *Upper left*: As viewed in the upstream frame (characterized by p_1, T_1, and ρ_1), the shock front is moving with a speed U towards the right (the upstream). *Upper right*: As viewed in the shock frame, the upstream moves with speed $v_{1s} = U$, which is decelerated to $v_{2s} = U/r$ in the downstream. *Lower left*: Viewed in the upstream, the downstream fluid moves towards the right (the upstream) with a speed $[(r-1)/r]U$, and a particle moves in random directions after crossing the shock and gaining energy. *Lower right*: Similarly, viewed in the downstream, the upstream fluid moves towards the left (the downstream) with a speed $[(r-1)/r]U$, and a particle moves in random directions after crossing the shock and gaining energy. Adapted from Longair (2011).

$$\frac{\Delta E}{E} \sim \frac{V}{c} = \frac{r-1}{r}\frac{U}{c} \tag{4.176}$$

whenever it crosses the shock front from either side. A more precise treatment involving angular integration gives (Bell, 1978)

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{2}{3}\frac{V}{c} = \frac{2}{3}\frac{r-1}{r}\frac{U}{c}. \tag{4.177}$$

For such a first-order Fermi acceleration process, the power-law index p can be derived. According to Eq. (4.177), one can derive the energy gain parameter for one round-trip crossing of the shock, i.e.

$$\zeta = 1 + 2\left\langle \frac{\Delta E}{E} \right\rangle = 1 + \frac{4}{3}\frac{r-1}{r}\frac{U}{c}. \tag{4.178}$$

One may also estimate the probability P of retaining the particle in the acceleration region (i.e. the probability for a downstream particle returning to the upstream) as follows. Let us define the downstream particle number density as n. The average number rate density of particles crossing the shock from the upstream to the downstream (the particle gaining rate) is $\sim nc/4$, where the factor $1/4$ is a result of averaging over the Maxwell–Boltzmann (thermal) distribution of the particles. On the other hand, since the downstream speed is U/r with respect to the shock front, particles that crossed the shock into the downstream would be advected away from the shock front with such a bulk speed. The fraction of particles that are lost (the particle losing rate) is nU/r. The probability of losing the particle

is therefore $(nU/r)/(nc/4) = (4/r)(U/c)$. So the probability of retaining a particle in the acceleration region (those that are *diffused* back into the upstream region) is

$$P = 1 - \frac{4}{r}\frac{U}{c}. \tag{4.179}$$

Again one can derive Eq. (4.174), with

$$\begin{aligned} p &= 1 - \frac{\ln P}{\ln \zeta} = 1 - \frac{\ln(1 - \frac{4}{r}\frac{U}{c})}{\ln(1 + \frac{4}{3}\frac{r-1}{r}\frac{U}{c})} \simeq 1 - \frac{-\frac{4}{r}\frac{U}{c}}{\frac{4}{3}\frac{r-1}{r}\frac{U}{c}} \\ &= \frac{r+2}{r-1} = \frac{v_{1s} + 2v_{2s}}{v_{1s} - v_{2s}} = \frac{\beta_{1s} + 2\beta_{2s}}{\beta_{1s} - \beta_{2s}}, \end{aligned} \tag{4.180}$$

with the condition $U/c \ll 1$ applied for non-relativistic shocks. For strong non-relativistic shocks, one has $r \simeq 4$ (Eq. (4.51)), so that $p \simeq 2$. So for non-relativistic strong shocks, the accelerated particle energy spectrum reads

$$N(E)dE \propto E^{(\beta_{2s}+2\beta_{1s})/(\beta_{2s}-\beta_{1s})}dE \propto E^{-2}dE. \tag{4.181}$$

The expression (4.180) can be derived more rigorously by solving the diffusion equation of particle distribution for shocks, which in a 1-D approximation reads (e.g. Bell, 1978; Blandford and Ostriker, 1978; Jones and Ellison, 1991)

$$\frac{\partial f}{\partial t} + v_2\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left(D(x)\frac{\partial f}{\partial x}\right). \tag{4.182}$$

The physical meaning is the following: whereas the particles that penetrate into the upstream region from the downstream across the shock will inevitably return back to the downstream, those from the upstream streaming into the downstream will not always return to the upstream. The diffusion term in Eq. (4.182) introduces a probability P (Eq. (4.179)) for a particle to recross the shock into the upstream region. Repeating the above arguments, one can derive Eq. (4.180).

4.4.3 Particle Acceleration in Relativistic Shocks

For relativistic shocks, the basic principle of Fermi acceleration also applies. Due to the relativistic nature of the shock, the chance for a downstream particle to diffuse and return to the upstream is lower. On the other hand, for each crossing, the particle gains much more energy. Particles can therefore also be efficiently accelerated.

Semi-analytical studies of relativistic shock kinetic theory suggest that particles can be accelerated with a power-law distribution in energy. For the convention $N(E) \propto E^{-p}$, the derived value of p is quite "universal". For example, Achterberg et al. (2001) derived $p \simeq 2.2$–2.3 for a variety of input parameters. Assuming isotropy in the downstream, Keshet and Waxman (2005) derived[10]

$$p = \frac{\beta_{1s} - 2\beta_{1s}\beta_{2s}^2 + 2\beta_{2s} + \beta_{2s}^3}{\beta_{1s} - \beta_{2s}}. \tag{4.183}$$

[10] The s parameter derived in Keshet and Waxman (2005) is $p + 2$ in our convention.

Notice that, since both β_{1s} and β_{2s} are $\ll 1$ non-relativistic shocks, the two terms with higher order β (i.e. $-2\beta_{1s}\beta_{2s}^2$ and $+\beta_{2s}^3$) can be dropped out in the non-relativistic regime. This expression is reduced to Eq. (4.180). For ultra-relativistic shocks, one has $\beta_{1s} \simeq 1$ and $\beta_{2s} \simeq 1/3$, so that

$$p \simeq \frac{20}{9} \simeq 2.22. \tag{4.184}$$

Such a p value was also derived for *parallel* (where the upstream magnetic field direction is parallel to the shock normal direction) relativistic shocks through Monte Carlo simulations (e.g. Ellison and Double, 2002). This method injects test particles into a background relativistic shock with ordered magnetic fields in both the upstream and the downstream, with the shock properties defined by the shock jump conditions. The test particles are followed kinetically for their evolution in the momentum space. Ellison and Double (2002) showed that the typical spectral index of relativistic protons is $p \sim 2.23$. For *oblique* (where the upstream magnetic field direction is at an angle with respect to the shock normal) and *trans-relativistic* (Lorentz factor of a few) shocks, significant deviation from $p = 2.23$ is possible.

Since significant particle acceleration will modify the shock structure, the test particle Monte Carlo method cannot solve the particle acceleration problem self-consistently. Breakthroughs were made by fully simulating relativistic shock acceleration numerically using the particle-in-cell (PIC) method (e.g. Spitkovsky, 2008; Sironi and Spitkovsky, 2009a, 2011). These simulations solve the Vlasov–Maxwell equations with the particle distribution function $f(\mathbf{x}, \mathbf{v}, t)$ in six-dimensional phase space as a function of time, which self-consistently catches microscopic electromagnetic interactions in the plasmas. With a simple two-spatial-dimension (2-D) simulation for an unmagnetized electron/positron (e^+e^-) pair shock, Spitkovsky (2008) showed that the downstream particle spectrum consists of two components: one quasi-thermal relativistic Maxwellian (with a characteristic temperature defined by the upstream kinetic energy of the flow), and a non-thermal power-law component (Fig. 4.4). With a limited simulation time, he showed that the non-thermal tail extended to 100 times the energy of the thermal peak and had a spectral index of $p = 2.4 \pm 0.1$. The non-thermal population had 1% in number, but $\sim$10% in energy. Spitkovsky (2008) emphasized that the simulation had not reached a steady state, and the non-thermal population would continue to grow as simulation time increased. It is possible that all the particles would be accelerated to the non-thermal component in relativistic shocks.

Sironi and Spitkovsky (2009a) studied the magnetized e^+e^- pair shocks in detail. They found that, similar to the results using the Monte Carlo method, the acceleration properties depend on the oblique angle. In particular, they defined a *critical oblique angle* at which a particle *sliding* along the magnetic field line with the speed of light in the downstream can just catch the motion of the shock. In the upstream frame, it is conveniently defined as

$$\cos\theta_{\rm cr,1} = \beta_{1s}. \tag{4.185}$$

In the downstream frame (in which the simulation of Sironi and Spitkovsky 2009a was performed), the angle can be expressed as (Exercise 4.7)

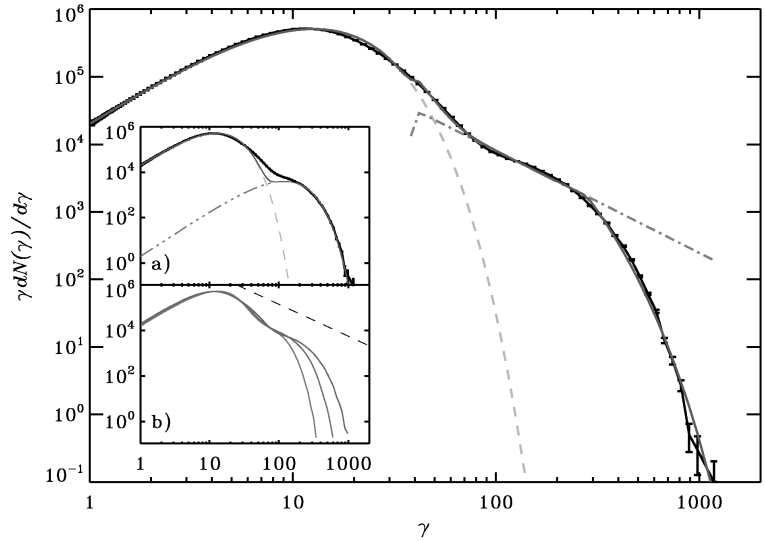

Figure 4.4 The result from a particle-in-cell (PIC) simulation which shows that electrons/positrons are indeed accelerated from a relativistic e^+e^- pair shock. Reproduced from Figure 2 in Spitkovsky (2008) with permission. ©AAS.

$$\theta_{\rm cr,2} = \tan^{-1}\left[\frac{1}{\gamma_{2s}(\beta_{12}+\beta_{2s})}\right]. \tag{4.186}$$

For $\theta > \theta_{\rm cr}$ (in either frame), the downstream particles can never catch up with and cross the shock front even if they travel at the speed of light. These are called *superluminal* shocks. Particles in these shocks are barely accelerated since they cannot complete the cycle of crossing the shock from both (up- and down-) streams. The opposite case ($\theta < \theta_{\rm cr}$) is *subluminal shocks*. Particles in the downstream of these shocks can cross the shock and complete the acceleration cycle. The simulations by Sironi and Spitkovsky (2009a) showed that particles are indeed accelerated, and the energy spectral index p depends on the inclination angle (Fig. 4.5). Interestingly, within the subluminal domain, acceleration progressively becomes significant as θ approaches $\theta_{\rm cr}$, with p ranging from 2.8 to 2.3. Sironi and Spitkovsky (2009a) drew the conclusion that significant particle acceleration is possible only for weak-magnetic-field relativistic shocks ($\sigma < 0.03$).

Sironi and Spitkovsky (2011) extended the e^+e^- pair shock simulations to electron–ion shock simulations. Due to numerical limitations, the ion-to-electron mass ratio (m_i/m_e) is still unrealistic: from 16 to 1000 (recall that the proton-to-electron mass ratio is $m_p/m_1 \simeq 1836$). Nonetheless, some interesting features in contrast to pair shocks are revealed. First, since their gyration radii are much smaller than those of electrons, ions are easier to accelerate. Ions have a larger fraction in energy than electrons deposited in the non-thermal component (30% vs. 1%), and they have a harder power-law spectrum than electrons ($p_p = 2.1 \pm 0.1$ vs. $p_e = 3.5 \pm 0.1$). Second, it is found that electrons are efficiently heated to occupy 15% of the upstream ion energy, but acceleration is

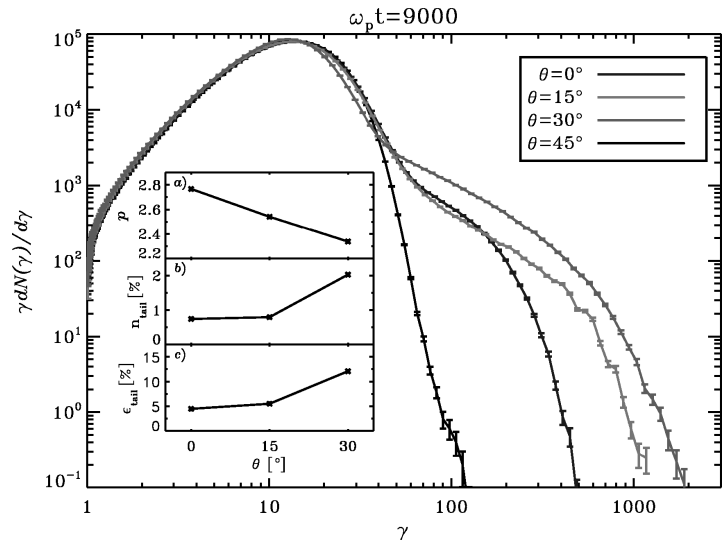

Figure 4.5 Particle-in-cell (PIC) simulations for magnetized pair shocks. The dependence of acceleration on the oblique angle of the magnetic field with respect to the shock normal was investigated. Reproduced from Figure 11 in Sironi and Spitkovsky (2009a) with permission. ©AAS. A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.

inefficient. Finally, the conclusions regarding the role of magnetic fields remain the same as for pair shocks: particle acceleration is suppressed in magnetized ($\sigma > 10^{-3}$) relativistic shocks.

GRB observations demand efficient particle acceleration. The forward shock is initially ultra-relativistic. Since it enters a circumburst medium that is not likely highly magnetized, efficient particle acceleration is warranted, which nicely explains the bright multi-wavelength afterglows. Since the GRB central engine is likely strongly magnetized, the external reverse shock and the internal shocks are likely magnetized. Observational evidence suggests that some GRBs have bright reverse shock emission, which demands moderately strong magnetic fields (Sari and Piran, 1999a; Mészáros and Rees, 1999; Fan et al., 2002; Zhang et al., 2003a; Kumar and Panaitescu, 2003). The inferred σ parameter in the reverse shock region exceeds 10^{-3} (Zhang et al., 2003a; Gao et al., 2015a; Fraija et al., 2016), suggesting that particles can be accelerated in these shocks. On the other hand, the reverse shocks are believed to be mildly relativistic. The superluminal condition (Sironi and Spitkovsky, 2009a) is therefore greatly suppressed, so that significant particle acceleration would become possible. For GRB prompt emission, it is still unclear whether the emission is powered by internal shocks or other processes such as magnetic reconnection (see Chapter 9 for details). If they are related to internal shocks, they must not be too magnetized, and their relative Lorentz factor must not be too strong, in order to allow efficient particle acceleration. On the other hand, in the strong magnetic field regime, magnetic reconnection would offer an alternative mechanism for particle acceleration (see §9.7 and §9.8 for detailed discussion).

4.5 Plasma and Fluid Instabilities and Magnetic Amplification

Another important role of relativistic shocks is to excite instabilities that amplify downstream magnetic fields. The afterglow data of at least some GRBs require a magnetic field strength much greater than the value derived from compressing the upstream magnetic field from a weakly magnetized circumburst medium (e.g. μG level magnetic field in an interstellar medium), suggesting that these instabilities likely play an important role to amplify magnetic fields to the desired level, at least in some GRBs.

4.5.1 Microscopic Weibel Plasma Instability

A well-studied instability on the microscopic plasma scale is the *Weibel instability* (Weibel, 1959; Medvedev and Loeb, 1999). It is a *two-stream* instability invoking two fluids moving relative to each other. The mechanism is best illustrated with the help of the schematic picture in Fig. 4.6 (Medvedev and Loeb, 1999).

Let us consider two streams of plasma moving in opposite directions along the positive and negative x-axis directions. For simplicity, we consider negatively charged electrons only, whose moving directions are marked by open arrows. Repeating the arguments below for protons would lead to exactly the same conclusion.

Suppose initially the system is weakly magnetized, and, for some reason, a weak perturbation of the magnetic field is excited (in the form of a sine wave, solid curve in Fig. 4.6). Due to the Lorentz force $-\frac{e}{c}(\mathbf{v} \times \mathbf{B})$ (negative sign from electron charge), one can see that the electron beams deviate from otherwise straight lines as indicated, and the patterns are different in regions I and II, where different magnetic field configurations are invoked. Let us look at the nearest left-going electron (the lowest left-going arrow solid line) in Fig. 4.6 as an example. The corresponding magnetic field for this electron is in the downward direction. It is immediately seen that the $\mathbf{v} \times \mathbf{B}$ direction is towards the reader, so that the electron Lorentz force direction is away from the reader (noticing the negative

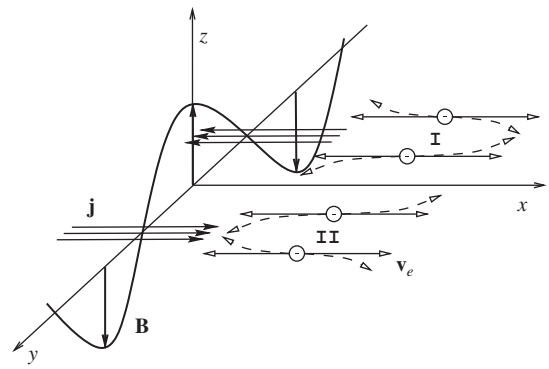

Figure 4.6 A schematic picture that demonstrates the Weibel instability. Reproduced from Figure 1 in Medvedev and Loeb (1999) with permission. ©AAS.

charge of electrons), as is delineated by the curved, dashed, left-going arrow. The moving directions of other electrons can be analyzed similarly, and the results are indicated by the dashed arrows. The new stream pattern will excite a new current pattern (**j**) as illustrated in Fig. 4.6 (again noticing the negative sign of electron charge). This new current pattern will induce a magnetic field, which enhances the original seed magnetic field. This induces a *runaway instability*, which leads to significant amplification of the seed magnetic fields.

This instability has been studied extensively through PIC simulations (e.g. Nishikawa et al., 2005, 2009; Spitkovsky, 2008; Sironi and Spitkovsky, 2009a, 2011). The results show that the instability indeed operates efficiently in a non-magnetized or weakly magnetized plasma right behind the relativistic shock. The magnetic field strength quickly decays far downstream from the shock front. The Weibel instability is suppressed completely if the plasma carries an ordered magnetic field with a relatively large σ (Sironi and Spitkovsky, 2009a).

4.5.2 Macroscopic Turbulence-Driven Instability

Besides the instabilities on the microscopic plasma scale, Sironi and Goodman (2007) suggested that a macroscopic *turbulence-excited fluid instability* can also amplify seed magnetic fields. Invoking an inhomogeneity in the circumburst medium density field, they showed that relativistic vorticity may be produced when the GRB blastwave hits a density clump, and the vorticity excites turbulence, which amplifies magnetic fields to an energy that is comparable to the energy of the turbulence. Such a turbulent dynamo has been extensively studied from numerical simulations (e.g. Schekochihin et al., 2004), and applied to broad astrophysical phenomena (Beresnyak and Lazarian, 2015).

4.6 Shock Parameterization

In general, relativistic shocks are sites for (1) particle acceleration, (2) magnetic field amplification, and (3) photon radiation. For GRBs, strong evidence suggests that the afterglow originates from the *external shocks* (including the forward and the reverse shock) upon interaction between the ejecta and the circumburst medium. The origin of prompt emission is debated. In any case, *internal shocks*, those shocks arising from collisions among shells with different Lorentz factors within the GRB outflow, are believed to play an important role in producing or triggering GRB prompt emission.

In order to describe a relativistic shock in full detail from first principles, very expensive numerical simulations are required. In GRB problems, one usually parameterizes the shocks with some empirical parameters. These parameters reflect one's ignorance of the detailed microscopic physics at the plasma level, but make a direct connection with the observational properties. This greatly reduces the complexity of the problems, and serves as a bridge between the macroscopic and microscopic worlds.

The widely adopted *microphysics parameters* include the following:

- p: the power-law index of the non-thermal electrons, defined by $N(E)dE \propto E^{-p}dE$, or $N(\gamma)d\gamma \propto \gamma^{-p}d\gamma$. Similarly, an index for the non-thermal protons, p_p, may be also defined;
- ϵ_B: the fraction of the shock internal energy that is partitioned to magnetic fields;
- ϵ_e: the fraction of the shock internal energy that is partitioned to electrons;
- ϵ_p: the fraction of the shock internal energy that is partitioned to protons (ions);
- ξ_e: the fraction of electrons that are accelerated to a non-thermal distribution;
- ξ_p: the fraction of protons that are accelerated to a non-thermal distribution.

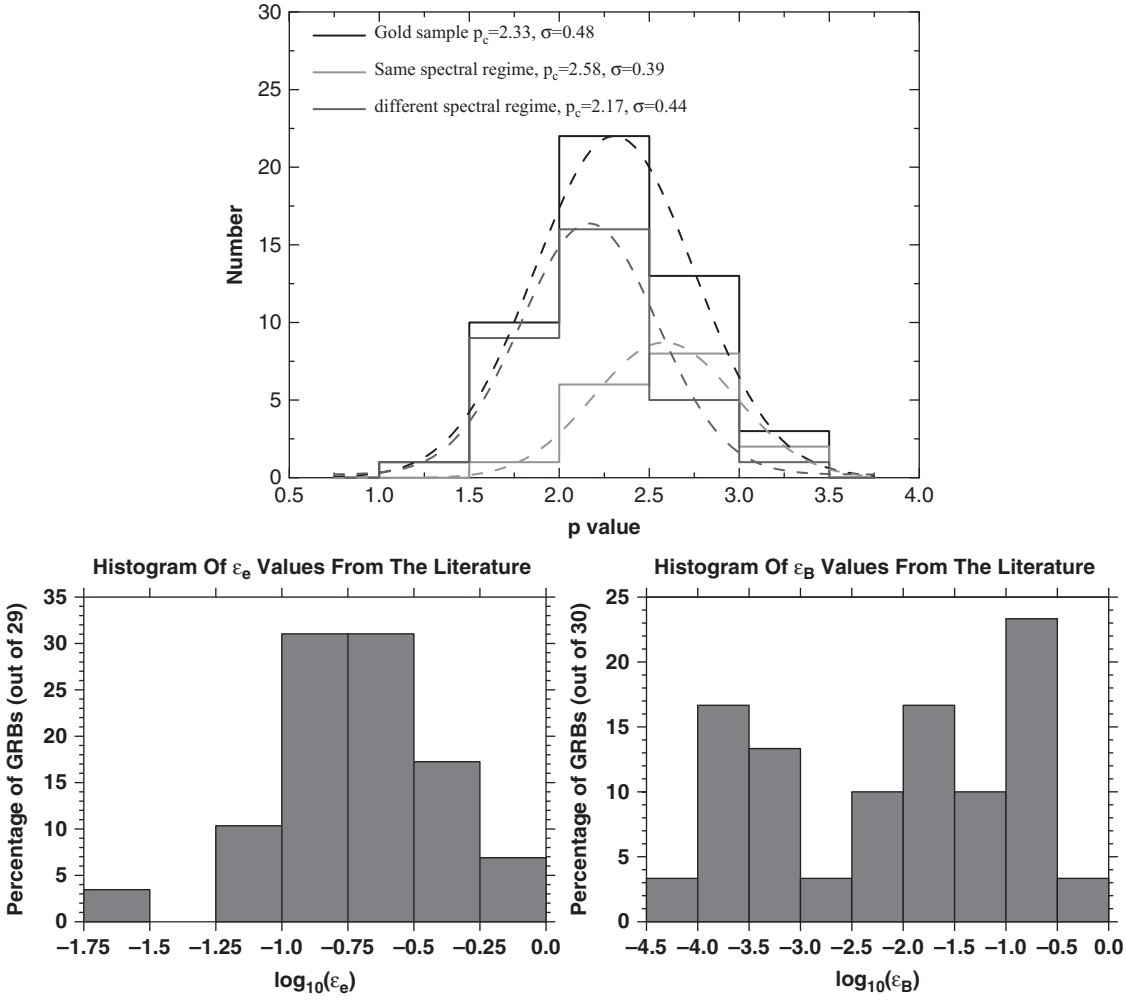

Figure 4.7 Constrained microphysics parameters in the literature. *Upper:* The derived p values from the afterglow data of *Swift* GRBs. From Wang et al. (2015b). *Lower:* The constrained ϵ_e and ϵ_B parameters of the GRB external shocks collected from the literature. Reproduced from Figure 1 in Santana et al. (2014) with permission. ©AAS.

By definition one has

$$\epsilon_p + \epsilon_e + \epsilon_B = 1. \tag{4.187}$$

In the GRB afterglow problem, usually $\xi_e = \xi_p = 1$ is assumed to model the forward shock. Interpreting the GRB prompt emission via the internal shock synchrotron radiation model requires invoking $\xi_e \ll 1$ to get the right E_p as observed (Daigne and Mochkovitch, 1998). Particle-in-cell simulations show the co-existence of a thermal and a non-thermal component, but the thermal bump may be significantly eroded when the simulation time is long enough (e.g. Spitkovsky, 2008). Observationally, there is no evidence of the existence of a thermal electron population in both prompt emission and afterglow data.

These microphysics parameters have been constrained from the GRB data by many authors. Without getting into the details of the methods of constraining them (see details in Chapters 8 and 9), some results are presented in Fig. 4.7. One can see that these microphysics parameters are *not universal*, and seem to be distributed in a wide range. In particular, the electron power-law index p, even though it has a distribution peaking at the theoretically predicted value (2.2–2.3), can vary from being smaller than 2 to above 3. The ϵ_e and ϵ_B parameters, on the other hand, even vary across orders of magnitudes. This suggests that the particle acceleration and magnetic amplification processes must depend on a variety of physical conditions, which could be quite different from burst to burst. An in-depth understanding of the shock microphysics requires the combination of first-principle numerical simulations and insights gained from observational constraints.

Exercises

4.1 Derive Eqs. (4.10) and (4.11) from Eq. (4.8). [Hint: Derive Eq. (4.11) first, and apply it to derive Eq. (4.10).]

4.2 Derive Eqs. (4.10) and (4.11) from the MHD equations (4.35) and (4.36) by assuming $E = B = 0$.

4.3 Use Eqs. (4.44)–(4.46) to derive the non-relativistic Rankine–Hugoniot relations, Eqs. (4.48)–(4.50).

4.4 Derive the relativistic transformation equations (4.57)–(4.66).

4.5 Use Eqs. (4.68)–(4.70) to derive the relativistic Rankine–Hugoniot relations (4.71)–(4.80).

4.6 Derive the MHD shock jump conditions, Eqs. (4.112)–(4.115).

4.7 Derive the critical angle in the downstream frame for a magnetized relativistic shock, Eq. (4.186).