Cambridge University Press
978-0-521-51459-0 - Collisionless Shocks in Space Plasmas: Structure and Accelerated Particles
David Burgess and Manfred Scholer
Excerpt
More information

Chapter 1

Shocks

1.1 Introduction

A shock is an abrupt transition between supersonic and subsonic flows. The best-known example is that formed by an obstacle, such as an aircraft, travelling through air faster than the speed of sound. For the aircraft to move forward the air ahead of it must be diverted around it, and there has to be a layer of subsonic flow adjacent to the obstacle. If this were not the case the influence of the pressure force, which propagates away from the obstacle at the sound speed, would be swept downstream and could not affect the flow ahead of the obstacle. Relative to the obstacle the distant flow is supersonic, so there has to be a transition to the subsonic flow close to the obstacle.

Considering such transitions within the framework of gas dynamics leads to the study of discontinuous, or near-discontinuous, solutions which must satisfy governing equations such as conservation of mass, momentum and energy. These solutions represent abrupt, well-defined changes in flow state, with so called 'jump conditions' which describe the supersonic to subsonic transition. Two defining qualities can be extracted from this framework. Firstly, at a shock the flow speed changes, but also the temperature increases due to dissipation, so that the shock mediates a transfer of bulk kinetic energy in the upstream flow into thermal energy downstream. The presence of dissipation means that the change of state at a shock corresponds to an entropy increase, and it is irreversible. Secondly, by their nature, shock solutions are fundamentally nonlinear, since not only do they accomplish a change of state, but they can also be thought of as a 'wave' whose propagation speed is determined by the supersonic flow speed, i.e., faster than any small-amplitude linear wave. It is nonlinearity which leads to wave steepening at the shock, and the importance of discontinuous solutions.

Shocks have been explored in laboratory plasmas since the 1950s, but the discovery of shocks in interplanetary space, and the confirmation that they are relatively stable structures, has been one of the major advances of space physics. Nevertheless, the study of shocks in plasmas which are essentially collisionless poses some fundamental problems. The conceptual framework provided by gas dynamics has to be stretched to account for the complex behaviour possible in a collisionless plasma such as the presence of multiple wave modes, non-Maxwellian particle distributions leading to wave growth via instabilities, and the presence of energetic particle populations. But, more importantly, the basic process of transferring upstream kinetic flow energy to downstream 'thermal' energy has to be accomplished by the complex interaction of plasma fields and particles, and without collisions.

A gas dynamic shock has a width of the order of the mean free path between collisions, whereas in a collisionless plasma the scale length of a shock might be associated with one or more plasma kinetic processes operating in a nonlinear transition region. We shall see later that, unlike in gas dynamic shocks, there is no single, unique mechanism which controls the shock dissipation, and hence length scales. Different types of shock can be dominated by very different processes, so that physically they behave differently, while still providing a transition between supersonic and subsonic flow states. Nevertheless, shocks in collisionless plasmas share one important property, which is that they are always associated with the acceleration of particles. Not only can they be a source of energetic particles, but their structure can be modified by, and sometimes even dominated by, the effects of such particles.

In this chapter we outline the basic ideas of shocks in gas dynamics, particularly the importance of steepening and dissipation. We then discuss in general terms the issues which have to be confronted when studying the physics of shocks in collisionless plasmas. Finally we give an introduction to shocks as found in the heliosphere, since they provide the motivation for the more theoretical studies described in later chapters.

1.2 Shocks in gases

The fundamental framework for the study of shocks is based on the behaviour of nonlinear waves in fluid or gas continuum systems. We shall see that there are two essential factors for the existence of shock solutions: nonlinearity in order to have steepening of a perturbation, and dissipation which constrains that steepening. Other, even discrete, systems can have shock-like solutions, but in all cases shocks arise as a balance, or competition, between steepening and relaxation.

Consider first a linear system with wave propagation. As an example, we take a gas which supports sound waves, so that small-amplitude, linear pressure perturbations

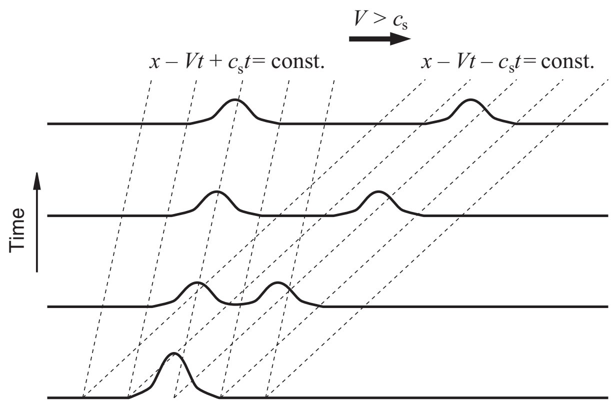

Figure 1.1 Time evolution of wave equation from d'Alembert's solution for the case of a background flow $V > c_s$, showing the two families of characteristics (dashed).

are described by the wave equation

$$\frac{\partial^2 p}{\partial t^2} - c_s^2 \frac{\partial^2 p}{\partial x^2} = 0, \tag{1.1}$$

with sound speed c_s and assuming dependence on just time t, and one spatial coordinate x. The general solution to this partial differential equation is given by the well-known d'Alembert's solution

$$p(x, t) = f(x - c_s t) + g(x + c_s t), \tag{1.2}$$

where f and g are arbitrary functions of a single variable. If the initial value of the solution is $p(x, t = 0) = p_0(x)$, then the solution becomes

$$p(x, t) = \frac{1}{2}(p_0(x - c_s t) + p_0(x + c_s t)). \tag{1.3}$$

The solution consists of two form-conserving parts: the first is constant along lines $x - c_s t = k_+$ and the second is constant along lines $x + c_s t = k_-$, for constants k_+ and k_-. The two parts of the solution retain constant shape, and travel to right and left at the sound speed. If there is some background flow speed V in the system, then there is an additional convection of the solution, changing the slopes of the lines of constant solution value. The case for $V > c_s$ is shown in Fig. 1.1, illustrating how, in a supersonic flow, both parts of the wave solution are convected in the direction of the flow, i.e., downstream.

The linear wave equation is a second-order, hyperbolic partial differential equation (PDE), and the families of lines $x - c_s t = k_+$ and $x + c_s t = k_-$ are its characteristics. A partial differential equation (or system of equations) is termed hyperbolic if it has real characteristics, i.e., its characteristics are rays in (x, t) space. As in the case of the wave equation, the solution of a hyperbolic PDE system can be expressed in terms of an evolution along its characteristics. It is often convenient to think of the characteristics as paths conveying information about the solution forward in time, or that information propagates along the characteristics.

For the linear wave equation above, the characteristics are two families of parallel lines, so that the left and right-going components of the solution preserve their shape. For the case of a gas, the wave equation arises for small-amplitude fluctuations by a linearization of the full, nonlinear governing equations. Following the discussion in LeVeque (1992), we now show how nonlinearity can naturally lead to the steepening of a wave. We start with a simpler PDE, namely the linear advection equation

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0, \tag{1.4}$$

for a function $u(x, t)$ where the advection speed c is constant. We consider the initial value problem for this equation, seeking a solution for $t \geq 0$, with initial condition $u(x, 0) = u_0(x)$. The solution is

$$u = u_0(x - ct). \tag{1.5}$$

At any time $t > 0$ the profile of the solution is exactly the same as at $t = 0$ but shifted to the right by a distance ct. The solution is constant along the characteristics of the PDE, which are the family of lines $x - ct = x_0$, where x_0 is a parameter (given by the x position of the characteristic at $t = 0$). The linear advection equation behaves like the linear wave equation, but has only one family of straight-line characteristics. As before, the solution can be thought of as propagating along the characteristics without change of shape.

The simplest form of nonlinearity is to let the wave speed depend linearly on the solution value, as in Burgers' equation with zero viscosity:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0. \tag{1.6}$$

This is similar to the situation for sound waves in a gas where $c_s^2 = \gamma p/\rho$, with c_s increasing with pressure. The wave speed corresponds to the slope of the characteristic, which is no longer constant as for the linear advection equation:

$$\frac{dx(t)}{dt} = u(x(t), t). \tag{1.7}$$

The solution is again constant along the characteristics since

$$\begin{aligned}\frac{d}{dt}u(x(t), t) &= \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}\frac{dx}{dt} \\ &= \frac{\partial u}{\partial t} + u\,\frac{\partial u}{\partial x} = 0.\end{aligned} \tag{1.8}$$

Since the slope of the characteristics is not constant, there is the possibility that they will intersect immediately for $t > 0$, given an arbitrary initial solution. Intersection of the characteristics would correspond to a multi-valued solution. But, by choosing a suitable initial function, the development of the solution can be followed until such an intersection occurs. An example is shown in Fig. 1.2. The

Cambridge University Press
978-0-521-51459-0 - Collisionless Shocks in Space Plasmas: Structure and Accelerated Particles
David Burgess and Manfred Scholer
Excerpt
More information

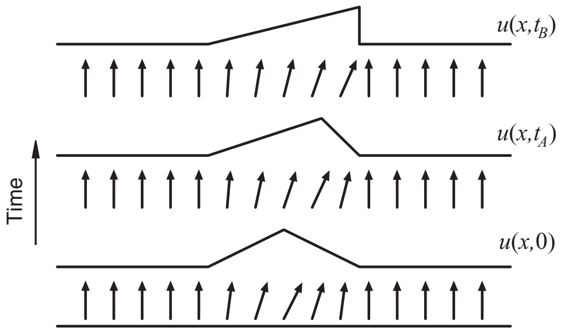

Figure 1.2 Development of a discontinuous solution of Burgers' equation. Arrows indicate the slopes of the characteristics.

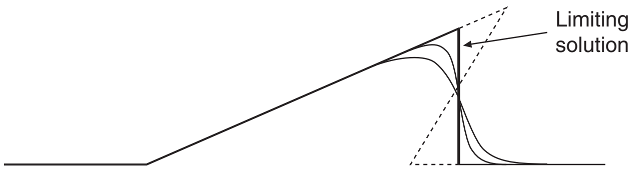

Figure 1.3 Multi-valued solution to the inviscid Burgers' equation (dashed) and sequence of solutions to the viscous Burgers' equation showing the limiting solution as $\epsilon \to 0$.

initial $t = 0$ solution has a triangular form that rises to some maximum. The effect of the nonlinearity is that the profile evolves, so that the points on the leading (right-most) edge move to the right, but the points at high solution value move faster. The effect is that the leading edge becomes steeper, and the trailing edge becomes less steep (as at $t = t_A$). Eventually at $t = t_B$ the characteristics from different points on the leading edge intersect for the first time, and at this time the solution develops an infinite slope. The solution develops a discontinuity at the time when the crest of the wave catches up with the leading edge.

Taking the evolution beyond this time (i.e., $t > t_B$) would predict multi-valued solutions, which would be unphysical (Fig. 1.3). In a realistic physical system one possibility is that the multi-valued solution corresponds to the transition to a turbulent regime, as a water wave which breaks as it comes into shore. However, in the case of Burgers' equation, a steady, shock-like solution can be recovered by retaining the viscous term with parameter ϵ:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \epsilon \frac{\partial^2 u}{\partial x^2}. \tag{1.9}$$

For small values of ϵ, and a smooth initial solution, the viscous term is unimportant and the solution evolves as before. But, as the wave steepens and the gradient increases, the second derivative becomes more important and the viscous term begins to play a role in keeping the function smooth and single valued. For a smaller value of viscosity ϵ the same process operates, but the gradient has to be larger (and the scale length of the transition smaller) for the viscous term to be effective.

Cambridge University Press
978-0-521-51459-0 - Collisionless Shocks in Space Plasmas: Structure and Accelerated Particles
David Burgess and Manfred Scholer
Excerpt
More information

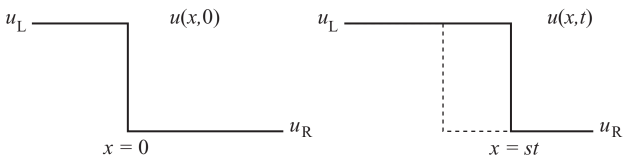

Figure 1.4 Discontinuous solution to Burgers' equation for a change of state $u_L \rightarrow u_R$ propagating at a speed s, for $t = 0$ (left) and $t > 0$ (right).

Thus, as $\epsilon \rightarrow 0$ the solution corresponding to the intersection of the characteristics approaches a discontinuity in the small viscosity limit, as illustrated in Fig. 1.3.

For what follows later we can summarize this in general terms: a nonlinear system will naturally exhibit wave steepening, with solutions evolving to large gradients at short length scales. If there is an additional dissipation process, such as viscosity, which depends on those gradients and acts to limit them, then wave steepening can be halted when the solution reaches a scale length which depends on the nature of the dissipation process. The ideal discontinuous shock solution can be obtained as the scale length of the transition tends to zero.

Since Burgers' equation supports solutions which develop a discontinuity (at least in the small viscosity limit), we can ask what can be learnt about such solutions. We start from a Riemann problem for Burgers' equation with an initial piecewise constant solution with a single discontinuity, with values to the left and right of the discontinuity being u_L and u_R (Fig. 1.4). The initial solution can be written as

$$u(x, 0) = \begin{cases} u_L & x < 0 \\ u_R & x > 0. \end{cases} \tag{1.10}$$

If $u_L > u_R$, then the situation is similar to the discussion above, with the characteristics sloping towards the discontinuity. The solution in this case is that the discontinuity stays at the same level and moves to the right at a speed s:

$$u(x, t) = \begin{cases} u_L & x < st \\ u_R & x > st, \end{cases} \tag{1.11}$$

where the shock speed s is given by

$$s = \frac{1}{2}(u_L + u_R). \tag{1.12}$$

This can be shown to be a solution as follows. The inviscid Burgers' equation can be written as

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2\right) = 0. \tag{1.13}$$

This is a conservation law with a flux function $\frac{1}{2}u^2$. Integrating this equation over a region $-L < x < +L$, with L large enough to contain the shock as it propagates, gives

$$\frac{d}{dt}\int_{-\mathrm{L}}^{+L} u(x,t)dx = \frac{1}{2}u_{\mathrm{L}}^2 - \frac{1}{2}u_{\mathrm{R}}^2 = \frac{1}{2}(u_{\mathrm{L}}+u_{\mathrm{R}})(u_{\mathrm{L}}-u_{\mathrm{R}}). \tag{1.14}$$

From the shock solution for $u(x,t)$ we also have

$$\int_{-\mathrm{L}}^{+L} u(x,t)dx = (L+st)u_{\mathrm{L}} + (L-st)u_{\mathrm{R}} = L(u_{\mathrm{L}}+u_{\mathrm{R}}) + st(u_{\mathrm{L}}-u_{\mathrm{R}}), \tag{1.15}$$

so that

$$\frac{d}{dt}\int_{-\mathrm{L}}^{+L} u(x,t)dx = s(u_{\mathrm{L}}-u_{\mathrm{R}}). \tag{1.16}$$

Comparing with (1.14) one has

$$s = \frac{1}{2}(u_{\mathrm{L}}+u_{\mathrm{R}}). \tag{1.17}$$

Thus we arrive at the important conclusion that a discontinuity is only a solution if it propagates at this speed, i.e., at the shock speed.

It is important to note from Eq. (1.17) that u_{L}, u_{R} and s are in a unique relationship. In other words, *any* size of discontinuity $u_{\mathrm{L}} \to u_{\mathrm{R}}$ is allowed provided it propagates at the speed given by Eq. (1.17). Alternatively, if u_{R} and s are given, then there is a unique solution for u_{L}. The same applies in the context of a shock in gas dynamics: given the upstream state and the speed of the shock there is a unique solution for the downstream state. This leads to the Rankine–Hugoniot jump relations, or shock conservation relations, which relate the upstream and downstream states in terms of the shock parameters. The corresponding conservation relations for shocks in collisionless plasmas are derived in Section 2.2.

The characteristics for Burgers' equation are right sloping, and for a discontinuous shock solution with $u_{\mathrm{L}} > u_{\mathrm{R}}$, they are directed towards the shock as it propagates. Fluctuations upstream and downstream of the shock both advect towards the shock, and so the solution is stable. If $u_{\mathrm{L}} < u_{\mathrm{R}}$, other solutions are possible, namely an entropy violating shock (which is unstable) and a rarefaction wave. For more details on discontinuous solutions and their stability, see LeVeque (1992).

All the features of the discontinuous solution for the Burgers' equation can be found for shock solutions in ordinary gas dynamics. One finds a discontinuous solution in the small viscosity limit, where nonlinear steepening is balanced against viscous dissipation. In gas dynamics waves can propagate in both (e.g., left and right) directions, so one can have shock propagation in both directions, but similar conditions as above on the left and right states apply in order to have a stable shock solution. Note, for a rightward propagating shock, in a system with forward and backward sets of characteristics, upstream perturbations associated with both directions of waves are convected towards the shock, but in the downstream region only rightward propagating perturbations can travel towards the shock.

The shock speed is the defining property of the shock solution, and in gas dynamics it is usual to characterize a shock using V_u, the upstream flow speed normal to the shock, in the shock frame where the shock is stationary. Note that only the component of the upstream flow velocity normal to the shock is used. The Mach number of a shock is the ratio of the shock speed to a wave speed associated with the shock. So that the sonic Mach number is

$$M_{cs} = \frac{V_\mathrm{u}}{c_\mathrm{s}}, \tag{1.18}$$

where c_s is the sound speed.

In this section we have discussed a simple model of a continuum system which has demonstrated all the components necessary to understand how shocks, as nonlinear discontinuous solutions, arise in other systems such as the equations of gas dynamics. The nonlinearity is in the sense that wave speed increases with solution value. If there is a large-scale, nonlinear perturbation then shorter-scale fluctuations at the wave crest travel faster, leading to a pile-up of increasingly short-scale fluctuations. This drives the steepening of the large-scale wave. As the wave steepens to shorter scales, viscous effects come into play, keeping the solution smooth at some small scale determined by the value of the viscosity. For an ordinary gas this viscosity is associated with collisions between atoms which damp the shortest-scale fluctuations. The shock solution is inherently nonlinear, and can be viewed as a discontinuity, but at the shortest scales a dissipative process (such as viscosity) ensures that the solution is smooth, and indeed determines the scale size of the transition. The shock transition that results is a balance between nonlinear steepening and viscous dissipation.

The shock solution can have any size jump across it, but must propagate at a commensurate speed, as dictated by conservation properties of the governing system of equations. Finally, the stable shock solution has the characteristics of the system directed towards it. In physical terms, this means that fluctuations both upstream and downstream must propagate or be convected towards the shock. In the upstream region, into which the shock is propagating, this means that the flow relative to the shock is faster than the wave speed, or supersonic. All upstream waves are convected towards the shock. Similarly for the downstream region, in which the shock is propagating away, waves must be able to overcome any effects of convection, and be able to propagate towards the shock. Therefore, in the downstream region the flow speed relative to the shock is slower than the wave speed, or subsonic.

1.2.1 Shocks in MHD plasma

Magnetohydrodynamics (MHD) is a fluid model of a plasma, and has many of the properties of the gas dynamic system. The use of discontinuous shock solutions can be applied to the MHD system of equations, but there are a number of complications which need to be taken account of. The ideal MHD equations are a system of hyperbolic PDEs, with a set of characteristics corresponding to the linear MHD wave modes: Alfvén, and fast and slow magnetosonic waves. The compressive modes have corresponding shock solutions, namely fast-mode and slow-mode shocks. In this book we will concentrate almost exclusively on shocks that correspond to the MHD fast-mode shock, where the density and magnetic field compress across the shock in phase. The MHD shocks and other types of discontinuous solutions are discussed in Section 2.1.

One of the most important differences in MHD, compared with ordinary gas dynamics, is that the wave speeds are anisotropic. Consequently, as well as Mach number, MHD shocks are also characterized by the magnetic field geometry at the shock. This is parameterized by the shock normal angle θ_{Bn} between the shock normal, directed into the upstream region, and the upstream magnetic field direction. The shock normal angle is, as we shall see later, a key parameter for describing different types of plasma shock. It has a strong controlling influence on both the shock jump conditions in MHD and shock structure in collisionless plasmas.

The shock normal angle is used, for convenience but based on observed and theoretically based behaviour, to classify shocks according to the following scheme: perpendicular shocks with $\theta_{Bn} = 90°$, parallel shocks with $\theta_{Bn} = 0°$, quasi-perpendicular with $\theta_{Bn} > 45°$, and quasi-parallel with $\theta_{Bn} < 45°$. The $\theta_{Bn} = 45°$ dividing line between quasi-perpendicular and quasi-parallel is a conventional one, rather than indicating any radical and abrupt transition between shock types (although see the discussion in Section 2.6). The term ‘oblique’ (as in: ‘oblique shock’) is used to refer to an ill-defined range of θ_{Bn} in the quasi-perpendicular regime but not close to perpendicular.

The wave speeds of the MHD modes in the region upstream of the shock will depend on the shock geometry, so that, as well as the sonic Mach number M_{cs}, it is possible to relate the upstream normal flow velocity to the Alfvén speed v_{A} and the magnetosonic mode speeds. The Alfvén Mach number is given by

$$M_{\mathrm{A}} = \frac{V_{\mathrm{u}}}{v_{\mathrm{A}}}, \tag{1.19}$$

and the fast and slow magnetosonic Mach numbers by

$$M_{\mathrm{f}} = \frac{V_{\mathrm{u}}}{v_{\mathrm{f}}} \quad \text{and} \quad M_{\mathrm{s}} = \frac{V_{\mathrm{u}}}{v_{\mathrm{s}}}, \tag{1.20}$$

where the fast- and slow-mode wave speeds v_{f} and v_{s} in the direction normal to the shock surface are

$$\left.\begin{matrix} v_{\mathrm{f}} \\ v_{\mathrm{s}} \end{matrix}\right\} = \left\{\frac{1}{2}\left[v_{\mathrm{A}}^2 + c_{\mathrm{s}}^2 \pm \sqrt{\left(v_{\mathrm{A}}^2 + c_{\mathrm{s}}^2\right)^2 - 4v_{\mathrm{A}}^2 c_{\mathrm{s}}^2 \cos\theta_{Bn}}\right]\right\}^{1/2}. \tag{1.21}$$

1.3 Shocks in collisionless plasmas

The previous section has outlined how a shock solution can arise in a gas dynamic or MHD system. Dissipation is required at some small scale length (the shock transition scale) to limit steepening, but at larger scales the solution can be treated as discontinuous. In these systems the dissipation is due to viscosity or resistivity from collisional scattering in the gradients of the transition. With this dissipation the shock transition connects two states of the gas, i.e., the supersonic upstream and subsonic downstream regions, which have distinct thermodynamic and flow properties.

When we begin to consider whether a collisionless plasma can have shock solutions we immediately encounter some fundamental questions. Can a shock exist in a plasma where there is no dissipation due to Coulomb collisions? Does wave steepening lead to discontinuous solutions as in gas dynamics? We will discuss these two questions to demonstrate that there is a framework in which we can sensibly study shocks in collisionless plasmas.

A two-fluid description of a plasma can be derived from the collisionless Vlasov–Maxwell system, and in the long-wavelength, low-frequency limit one recovers waves with similar properties to the MHD modes. Therefore, one expects a collisionless plasma to behave similarly to an MHD plasma in terms of nonlinear wave steepening, at least until the scale length of the perturbations becomes short enough for the departures from MHD in a two-fluid theory to be important. The key difference to MHD is that wave dispersion appears as the scale lengths in the system approach the ion kinetic scales, such as the ion gyroradius or ion inertial length. An example of this behaviour, which is relevant for collisionless shocks, is the magnetosonic/whistler mode which has positive dispersion as its phase speed increases above the low frequency limit value of the fast magnetosonic speed.

Wave dispersion fundamentally changes the usual framework for understanding shock solutions, since there is the possibility, as a perturbation steepens, of short-wavelength waves being excited which have a speed faster than the (long-wavelength limit) wave speed. In gas dynamics there is a single constant wave speed, and in MHD the different wave modes each have constant speed, while a shock is a flow transition across a characteristic wave speed. Dispersion, however, negates the premise of a constant wave speed. The effect is that short-wavelength fluctuations may propagate upstream away from a steepening shock transition, and in doing so they would carry away (or 'dissipate') energy at the shortest scales of the transition. This leads to the concept of a dispersive shock in which the transition scale is given by the wavelength of a dispersive wave which can stand in the flow at the shock. Nonlinear steepening is thus limited, since the presence of any shorter scales generates perturbations which cannot remain in the rest frame of the shock. This behaviour may seem similar to the role played by viscosity in gas dynamic shocks, but with the difference that there is no associated dissipation. If dispersion is to play a major role in the shock transition, then there must also be an another process which provides dissipation, for