fall midway between the nodes of $|w_1|$, with two nodes per vertical wavelength, and the amplitude would be exactly zero at each node.

Figure 54.7 illustrates refraction of acoustic-gravity waves in response to the continuous variation of properties of the solar atmosphere. Because the sound speed changes little with height, high-frequency waves show little bending of the direction of the phase velocity $\mathbf{v}_p$ or group velocity $\mathbf{v}_g$. Gravity waves, in contrast, show strong refraction from the large changes in ω_{BV}, with $\mathbf{v}_p$ bending away from the vertical as ω_{BV} decreases, and $\mathbf{v}_g$ bending toward the vertical. Though the direction of $\mathbf{v}_g$ tends toward the vertical, the magnitudes of both u_g and w_g decrease to zero as ω_{BV} decreases to ω.

5.3 Shock Waves

The theory developed in §§5.1 and 5.2 applies only to small-amplitude disturbances, which propagate essentially adiabatically and are damped only slowly by dissipative processes. As the wave amplitude increases, this simple picture breaks down because of the effects of the *nonlinear* terms in the equations of hydrodynamics. When nonlinear phenomena become important, the character of the flow alters markedly. In particular, in an acoustic disturbance a region of compression tends to overrun a rarefaction that precedes it; thus as an acoustic wave propagates, the leading part of the profile progressively steepens, eventually becoming a near discontinuity, which we identify as a *shock*.

Once a shock forms it moves through the fluid supersonically and therefore outruns preshock acoustic disturbances by which adjustments in local fluid properties might otherwise take place; it can therefore persist as a distinct entity in the flow until it is damped by dissipative mechanisms. The material behind a shock is hotter, denser, and has a higher pressure and entropy than the material in front of it; the *stronger* the shock (i.e., the higher its velocity) the more pronounced is the change in material properties across the discontinuity. The rise in entropy across a shock front implies that wave energy has been dissipated irreversibly; this process damps, and ultimately destroys, the propagating shock (sometimes rapidly).

In contrast to acoustic waves, internal gravity waves do not develop shocks. Instead in the nonlinear regime they *break* and degenerate into turbulence. We will not discuss these phenomena in this book; see for example, (**M3**) and (**M4**).

Shock phenomena are of tremendous importance in astrophysics. As we saw in §5.2, the growth of waves to finite amplitude occurs naturally and inevitably in an atmosphere having an exponential density falloff. Thus, as Biermann (**B3**), (**B4**) and Schwarzschild (**S8**) first recognized, small-amplitude acoustic disturbances generated by turbulence in a stellar convection zone can propagate outward with ever-increasing amplitude until they steepen into shocks that dissipate their energy, thus heating the

ambient atmosphere. Indeed, this mechanism is thought to provide part of the heating responsible for the outward temperature rise in stellar chromospheres (**A2,** Chaps. 9 and 10), (**B10,** Chap. 7), (**K3**), (**S5**), (**S6**), (**S7**), (**U4**), (**U6**).

Most of the shocks formed by spontaneous growth of randomly generated acoustic waves are rather weak. Much more impressive phenomena are produced in pulsating stars (e.g., Cepheids and RR Lyraes) in which a coherent velocity pulse generated by a radial motion of the entire stellar envelope propagates outward and drives a shock strong enough (1) to alter radically both the thermodynamic properties (e.g., degree of ionization) and the dynamical state (e.g., some layers are lofted outward and subsequently free-fall back) of the atmosphere and (2) to produce interesting spectroscopic phenomena (e.g., emission lines). Even more dramatic are the exceedingly strong shocks, essentially *blast waves*, generated in supernova explosions, which blow away the entire outer envelope of a star.

Similar phenomena also occur in laboratory situations, for example: when a projectile or aircraft moves supersonically through the atmosphere, when a piston is driven rapidly into a tube of gas (a shock tube), in the blast wave produced by a strong explosion, or when rapidly flowing gas encounters a constriction in a flow channel or runs into a wall.

55. The Development of Shocks

Let us now construct a solution of the full nonlinear hydrodynamical equation for a pulse propagating into an infinite homogeneous medium. We assume the flow is one dimensional (along the x axis) and is adiabatic, thus neglecting, for the moment, dissipation. The density, pressure, and velocity of the flow are then completely determined by the continuity and momentum equations and an equation of state $p = p(\rho, s)$ [or $p = p(\rho)$ because s is constant].

In §§48 and 49 we saw that all the physical quantities (ρ, p, T, etc.) and the fluid velocity u in a small-amplitude traveling wave are functions of a single argument $x \pm at$; this implies that any quantity can be expressed as a function of any other [e.g., $p = p(u)$, $\rho = \rho(u)$, etc.] independent of position and time. For finite-amplitude waves, the simple relationships obtained earlier no longer apply. But, as Riemann (**R5,** 157) first showed, it is possible to obtain a general solution of the full nonlinear equations for a traveling wave, in which all physical properties and the fluid velocity are again functions of a single argument $x \pm vt$; but now the propagation speed v of each point on the wave profile is a function of the fluid velocity u at that point in the disturbance. Hence, even in the nonlinear case it is possible to express any physical property of the wave as a function of any other; in particular we can regard the density as a unique function of the fluid velocity.

Thus assuming $\rho = \rho(u)$, we can write the continuity equation as

$$(d\rho/du)(\partial u/\partial t) + [u(d\rho/du) + \rho](\partial u/\partial x) = 0 \tag{55.1a}$$

or

$$(\partial u/\partial t) + [u + \rho(du/d\rho)](\partial u/\partial x) = 0, \tag{55.1b}$$

and the momentum equation as

$$\frac{\partial u}{\partial t} + \left[u + \frac{1}{\rho}\left(\frac{\partial p}{\partial \rho}\right)_s \frac{d\rho}{du}\right]\frac{\partial u}{\partial x} = 0. \tag{55.2}$$

Comparing (55.1b) with (55.2) we see that

$$(du/d\rho) = \pm[(\partial p/\partial \rho)_s]^{1/2}/\rho = \pm a/\rho \tag{55.3}$$

where a is the adiabatic sound speed, regarded here as a function of ρ [recall that $p = p(\rho)$]. Hence, the general relation between the fluid velocity and the density or pressure in the wave is

$$u = \pm\int_{\rho_0}^{\rho} (a/\rho)\, d\rho = \pm\int_{p_0}^{p} dp/\rho a, \tag{55.4}$$

where ρ_0 and p_0 are ambient values in the undisturbed fluid. Note that for a small-amplitude disturbance with $\rho = \rho_0 + \rho_1$, where $|\rho_1|/\rho_0 \ll 1$, (55.4) reduces to $u = \pm a_0\rho_1/\rho_0$, where a_0 is the sound speed in the undisturbed medium, in agreement with (48.18).

Using (55.3) in (55.1b) or (55.2), we obtain

$$(\partial u/\partial t) + (u \pm a)(\partial u/\partial x) = 0. \tag{55.5}$$

Similarly, by inverting the function $\rho(u)$, we can write the continuity equation as

$$(\partial \rho/\partial t) + [\rho(du/dx) + u](\partial \rho/\partial x) = 0, \tag{55.6}$$

which, from (55.3), implies

$$(\partial \rho/\partial t) + (u \pm a)(\partial \rho/\partial x) = 0. \tag{55.7}$$

Equations (55.5) and (55.7) yield general solutions of the form

$$u = F_1[x - (u \pm a)t] \tag{55.8}$$

and

$$\rho = F_2[x - (u \pm a)t] \tag{55.9}$$

where F_1 and F_2 are arbitrary functions that fix the run of u and ρ at $t = 0$.

Equations (55.8) and (55.9) represent traveling waves known as *simple waves*. One sees that a particular value of, say, ρ or u propagates through the ambient medium with phase speed

$$v_p(u) = u \pm a(u) \tag{55.10}$$

where $a(u)$ is given by (55.3) and (55.4). In (55.8) to (55.10), we choose the positive (negative) sign for waves traveling in the positive (negative) x

direction. Because $\rho = \rho(u)$, $p = p[\rho(u)]$, etc., all physical variables in the wave propagate in the same manner as u.

To make the results derived above more concrete, consider a simple wave in a perfect gas. Then $a^2 \propto p/\rho \propto \rho^{\gamma-1}$ implies that $(\gamma-1)(d\rho/\rho) = 2(da/a)$, hence (55.4) yields

$$u = \pm 2(a - a_0)/(\gamma - 1) \tag{55.11}$$

or

$$a = a_0 \pm \tfrac{1}{2}(\gamma - 1)u, \tag{55.12}$$

which implies

$$v(u) = \tfrac{1}{2}(\gamma - 1)u \pm a_0, \tag{55.13}$$

Using the polytropic gas laws we readily find from (55.12) that

$$\rho = \rho_0[1 \pm \tfrac{1}{2}(\gamma - 1)(u/a_0)]^{2/(\gamma-1)}, \tag{55.14}$$

$$p = p_0[1 \pm \tfrac{1}{2}(\gamma - 1)(u/a_0)]^{2\gamma/(\gamma-1)}, \tag{55.15}$$

and

$$T = T_0[1 \pm \tfrac{1}{2}(\gamma - 1)(u/a_0)]^2. \tag{55.16}$$

Consider a finite pulse having an initial sinusoidal shape as sketched in Figure 55.1, moving to the right. In the small-amplitude limit, we recover the acoustic equations (48.18), (48.24a), and (48.24b) from (55.14) to (55.16), and the disturbance propagates to the right with unchanged shape at speed a_0. But in the finite-amplitude regime, (55.10) to (55.16) plainly show that the more compressed parts of the pulse have a larger fluid velocity u, are hotter, have a higher sound speed $a(u)$, and move to the right with a higher velocity than the less compressed regions. Thus the crest of the pulse continuously gains on the pulse front and, as shown in Figure 55.1, the wave front progressively steepens.

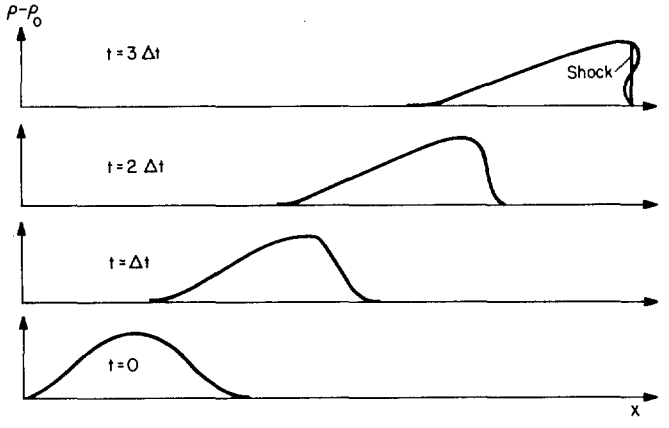

Fig. 55.1 Nonlinear steepening of a simple wave into a shock.

According to (55.12) to (55.16), the wave crest eventually overtakes the pulse front, and at later times the solution becomes multiple valued as sketched in Figure 55.1 for $t = 3\,\Delta t$. This result is unphysical and indicates a break-down of the theory. In reality, the front steepens into a shock, in which all variables change abruptly through a very thin layer, within which, owing to steep gradients, viscosity and thermal conductivity come into play to determine the detailed structure of the front. A continuation of the construction illustrated in Figure 55.1 shows that, as time progresses, the pulse becomes more and more triangular in shape and the velocity amplitude of the shock front decreases monotonically, implying that once a shock forms the wave continuously dissipates energy and is damped. By a similar construction, one can see that in a periodic acoustic wave the crests overrun the troughs, and the wave changes shape from a sinusoid to a train of shocks separated by the wavelength of the original wave. Such a shock train is called a *sawtooth wave* or N *wave*. We discuss the propagation and damping of pulses and sawtooth waves in an exponential atmosphere in §58.

56. Steady Shocks

In laboratory experiments (e.g., flow in a nozzle), it is possible to achieve steady flow, so that if shocks form they are fixed in space and have upstream and downstream properties that are time independent. For such *steady shocks* one can derive analytical expressions relating the values of the physical variables on the upstream and downstream sides of the front. More important, it is usually possible to consider propagating shocks as instantaneously steady because, as shown in §57, the shock thickness is only of the order of a particle mean free path λ, whereas the distance over which the properties of the upstream material can change significantly is some characteristic structural length in the fluid, say a scale height H. Consequently, the ratio of the time required for material to cross the shock front to the time needed for upstream conditions to change appreciably is very small, roughly equal to the Knudsen number $\mathrm{Kn} = \lambda/H$, which is only $\sim 10^{-6}$ in a stellar atmosphere, and even smaller deeper in a star. Therefore, at any instant even a propagating shock is steady to a high degree of approximation.

To simplify the analysis, we transform from the laboratory frame in which the shock moves along the x axis with speed v_s into a frame moving with the shock front itself. Thus if the material into which the shock propagates has a lab-frame speed v_1, and the material behind the shock has a lab-frame speed v_2, then in the shock's frame the upstream material enters the shock with speed

$$u_1 = v_1 - v_s \tag{56.1}$$

and the downstream material leaves the front with a speed

$$u_2 = v_2 - v_s. \tag{56.2}$$

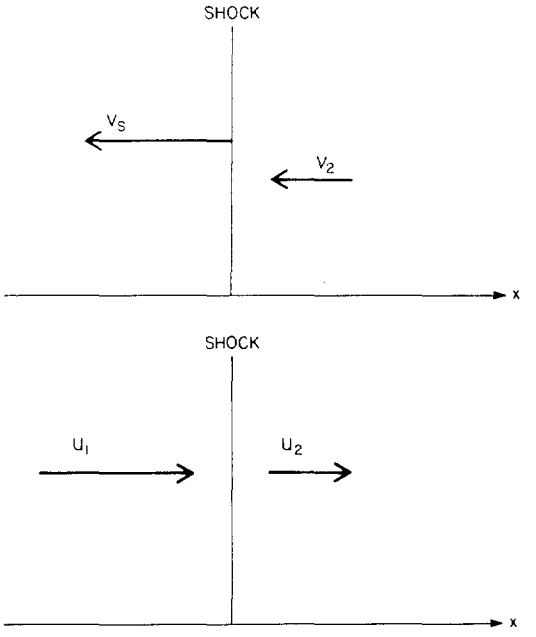

Fig. 56.1 Fluid velocities near shock, measured in lab frame (top), and shock's frame (bottom).

In the case illustrated in Figure 56.1, $v_s < 0$ and $v_1 = 0$, hence $u_1 > 0$, as is u_2.

THE CONSERVATION LAWS

In the frame of the shock the flow is steady, hence the equations of continuity, momentum, and energy (again ignoring viscosity and conduction) reduce to

$$\frac{d(\rho u)}{dx} = 0, \tag{56.3}$$

$$\frac{d}{dx}(\rho u^2 + p) = 0, \tag{56.4}$$

and

$$\frac{d}{dx}[\rho u(h + \tfrac{1}{2}u^2)] = 0, \tag{56.5}$$

which are conservation relations stating that the mass, momentum, and energy fluxes per unit area are constant throughout the flow, and in particular must be constant across the shock front. If we integrate (56.3) to (56.5) across the shock thickness, say from $-\frac{1}{2}\delta$ to $+\frac{1}{2}\delta$, and formally take

the limit as $\delta \to 0$ (because $\mathrm{Kn} \ll 1$), we obtain

$$\rho_1 u_1 = \rho_2 u_2 \equiv \dot{m}, \tag{56.6}$$

$$\rho_1 u_1^2 + p_1 = \rho_2 u_2^2 + p_2, \tag{56.7}$$

and

$$h_1 + \tfrac{1}{2}u_1^2 = h_2 + \tfrac{1}{2}u_2^2; \tag{56.8}$$

here all upstream variables have subscript "1" and all downstream variables have subscript "2". The quantity $\dot{m}$ is the mass flux through the shock. Note that these equations remain valid for curved shock fronts (e.g., in a spherical medium) because the thickness of the front is almost always negligible compared to its radius of curvature.

GENERAL JUMP RELATIONS

The conservation relations can be manipulated into other useful forms. Let $V \equiv 1/\rho$ be the specific volume of the material (this notation is inconsistent with that used in Chapter 1, but is adopted here to avoid confusion between volume and velocity). Then $u_1 = \dot{m}V_1$ and $u_2 = \dot{m}V_2$, hence (56.7) gives

$$p_2 - p_1 = \dot{m}^2(V_1 - V_2), \tag{56.9}$$

which shows that in a (p, V) diagram the initial and final states of the material are connected by a straight line with slope $-\dot{m}^2$ (see Figure 56.2).

Alternatively, using $u_1/u_2 = V_1/V_2$ in (56.7) to eliminate u_1 or u_2 we find

$$u_1^2 = V_1^2(p_2 - p_1)/(V_1 - V_2) \tag{56.10}$$

and

$$u_2^2 = V_2^2(p_2 - p_1)/(V_1 - V_2), \tag{56.11}$$

hence

$$u_1^2 - u_2^2 = (p_2 - p_1)(V_1 + V_2). \tag{56.12}$$

Substituting (56.10) and (56.11) into (56.8) we obtain

$$h_2 - h_1 = \tfrac{1}{2}(V_1 + V_2)(p_2 - p_1) \tag{56.13}$$

or

$$e_2 - e_1 = \tfrac{1}{2}(V_1 - V_2)(p_1 + p_2), \tag{56.14}$$

which are known as the *Rankine-Hugoniot* relations (**R2**), (**H17**).

Suppose we are given the upstream conditions ρ_1, p_1, and u_1, and that the material obeys a caloric equation of state $e(\rho, p)$ or $h(\rho, p)$, which may be quite general, including, for example, excitation and ionization effects. Then, in the (p, V) plane, either (56.13) or (56.14) defines a unique curve $p(V) = f(p_1, V_1, V)$, called the *Hugoniot curve*, passing through (p_1, V_1) and having the general shape sketched in Figure 56.2. Equation (56.9) combined with either (56.13) or (56.14) determines p_2 and V_2, hence u_2; the solution is shown graphically in Figure 56.2 as the intersection of the Hugoniot curve with the straight line given by (56.9). In general one finds that (p_2, V_2) differs substantially from (p_1, V_1), the discrepancy between the

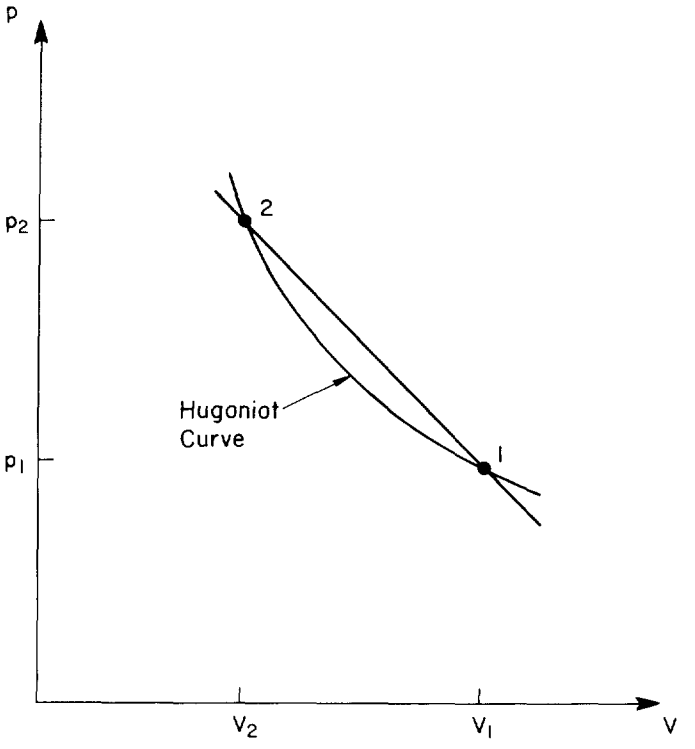

Fig. 56.2 Shock Hugoniot curve in (p, V) diagram.

two growing larger with increasing Mach number. We thus conclude that, to the level of approximation at which we are now working, the equations of hydrodynamics admit a *discontinuous jump* in the physical variables across a shock; such jumps are called *weak solutions* of the fluid equations.

In principle, (56.9) to (56.14) permit two types of solutions: (1) those in which $p_2 > p_1$, $V_2 < V_1$ (or $\rho_2 > \rho_1$), and $u_2 < u_1$, or (2) those in which the inequalities are all reversed. The former are *compression shocks* and the latter are *rarefaction discontinuities*; we see below that while both solutions are permitted mathematically, only compression shocks can exist physically. Rarefactions, when they occur, are always continuous (**C6,** Chap. 3).

If we choose solutions with $p_2 > p_1$, then it follows from (56.13) and (56.14) that $h_2 > h_1$ and $e_2 > e_1$. Further, using the fact that $(u_1 - u_2) = \dot{m}(V_1 - V_2)$ in (56.9), we find

$$u_1 - u_2 = [(p_2 - p_1)(V_1 - V_2)]^{1/2} \tag{56.15}$$

where we chose the positive root on physical grounds. Alternatively from (56.6) and (56.7), we can write

$$u_1 - u_2 = (p_2/\rho_2 u_2) - (p_1/\rho_1 u_1); \tag{56.16}$$

these results will prove useful shortly.

JUMP RELATIONS FOR A PERFECT GAS

If we assume that the fluid is a perfect gas, we can derive a comprehensive set of explicit formulae relating upstream and downstream variables. In this case, (56.8) becomes

$$\tfrac{1}{2}u_1^2+\left(\frac{\gamma}{\gamma-1}\right)\frac{p_1}{\rho_1}=\tfrac{1}{2}u_2^2+\left(\frac{\gamma}{\gamma-1}\right)\frac{p_2}{\rho_2}=\tfrac{1}{2}u_1^2+\frac{a_1^2}{\gamma-1}=\tfrac{1}{2}u_2^2+\frac{a_2^2}{\gamma-1}, \quad (56.17)$$

where a_1 and a_2 are the upstream and downstream sound speeds, respectively. Likewise, (56.13) becomes

$$\gamma(p_2V_2-p_1V_1)=\tfrac{1}{2}(\gamma-1)(V_1+V_2)(p_2-p_1), \quad (56.18)$$

whence we obtain

$$\frac{p_2}{p_1}=\frac{(\gamma+1)V_1-(\gamma-1)V_2}{(\gamma+1)V_2-(\gamma-1)V_1} \quad (56.19)$$

or

$$\frac{V_2}{V_1}=\frac{(\gamma+1)p_1+(\gamma-1)p_2}{(\gamma-1)p_1+(\gamma+1)p_2}=\frac{\rho_1}{\rho_2}=\frac{u_2}{u_1}. \quad (56.20)$$

From the perfect gas law we then have

$$T_2/T_1=p_2V_2/p_1V_1=a_2^2/a_1^2. \quad (56.21)$$

Using (56.20) in (56.10) and (56.11) we find

$$u_1^2=\tfrac{1}{2}V_1[(\gamma-1)p_1+(\gamma+1)p_2] \quad (56.22)$$

and

$$u_2^2=\tfrac{1}{2}V_1[(\gamma+1)p_1+(\gamma-1)p_2]^2/[(\gamma-1)p_1+(\gamma+1)p_2], \quad (56.23)$$

and from (56.22) and $\dot{m}=u_1/V_1$ we obtain

$$\dot{m}^2=[(\gamma-1)p_1+(\gamma+1)p_2]/2V_1. \quad (56.24)$$

In one-dimensional steady flows, it is sometimes convenient to introduce the *critical velocity* u_c at which the flow speed equals the local sound speed. Then, from (56.17),

$$\tfrac{1}{2}u_1^2+\left(\frac{\gamma}{\gamma-1}\right)\frac{p_1}{\rho_1}=\tfrac{1}{2}u_2^2+\left(\frac{\gamma}{\gamma-1}\right)\frac{p_2}{\rho_2}=\frac{(\gamma+1)}{2(\gamma-1)}u_c^2, \quad (56.25)$$

hence

$$u_1+\left(\frac{2\gamma}{\gamma-1}\right)\frac{p_1}{\rho_1u_1}=\left(\frac{\gamma+1}{\gamma-1}\right)\frac{u_c^2}{u_1} \quad (56.26)$$

and

$$u_2+\left(\frac{2\gamma}{\gamma-1}\right)\frac{p_2}{\rho_2u_2}=\left(\frac{\gamma+1}{\gamma-1}\right)\frac{u_c^2}{u_2}, \quad (56.27)$$

whence

$$u_1 - u_2 + \left(\frac{2\gamma}{\gamma-1}\right)\left(\frac{p_1}{\rho_1 u_1} - \frac{p_2}{\rho_2 u_2}\right) = \left(\frac{\gamma+1}{\gamma-1}\right)\frac{(u_2-u_1)u_c^2}{u_1 u_2}. \tag{56.28}$$

In view of (56.16), (56.28) reduces to the *Prandtl relation*

$$u_1 u_2 = u_c^2. \tag{56.29}$$

The relationship between upstream and downstream flow quantities can be expressed concisely in terms of $\not{p} \equiv \Delta p/p_1 = (p_2 - p_1)/p_1$, the *fractional pressure jump* across the shock. Thus from (56.20) we find the *compression ratio*

$$\rho_2/\rho_1 = V_1/V_2 = [2\gamma + (\gamma+1)\not{p}]/[2\gamma + (\gamma-1)\not{p}], \tag{56.30}$$

and from (56.21) we have

$$T_2/T_1 = (1+\not{p})[2\gamma + (\gamma-1)\not{p}]/[2\gamma + (\gamma+1)\not{p}]. \tag{56.31}$$

Furthermore, from (56.22) we find

$$M_1^2 - 1 = \tfrac{1}{2}(\gamma+1)\not{p}/\gamma \tag{56.32}$$

and from (56.23)

$$M_2^2 - 1 = -(\gamma+1)\not{p}/2\gamma(1+\not{p}), \tag{56.33}$$

whence we see that because $\not{p} \geq 0$, $M_1^2 \geq 1$ while $M_2^2 \leq 1$. That is, the upstream flow is always supersonic relative to a shock front, and the downstream flow is always subsonic. Note that if $M_1 = 1$, then $\not{p} = 0$ and the jump in all physical quantities vanishes, that is, there is no shock.

For very strong shocks $\not{p} \to \infty$, which implies that

$$\rho_2/\rho_1 \to (\gamma+1)/(\gamma-1) \tag{56.34}$$

and

$$M_2^2 \to (\gamma-1)/2\gamma, \tag{56.35}$$

while $M_1^2 \to (\gamma+1)\not{p}/2\gamma \to \infty$ and $T_2/T_1 \to (\gamma-1)\not{p}/(\gamma+1) \to \infty$. Hence for a monatomic gas, $\gamma = \frac{5}{3}$, the limiting compression ratio in an extremely strong shock is $(\rho_2/\rho_1)_{\max} = 4$, and the limiting value of the downstream Mach number is $(M_2)_{\min} = 1/\sqrt{5}$.

For weak shocks (i.e., $\not{p} \ll 1$), we have

$$(\rho_2/\rho_1) - 1 \approx \not{p}/\gamma, \tag{56.36}$$

$$(T_2/T_1) - 1 \approx (\gamma-1)\not{p}/\gamma, \tag{56.37}$$

$$M_1^2 - 1 \approx \tfrac{1}{2}(\gamma+1)\not{p}/\gamma, \tag{56.38}$$

and

$$M_2^2 - 1 \approx -\tfrac{1}{2}(\gamma+1)\not{p}/\gamma. \tag{56.39}$$

Equations (56.36) and (56.37) are merely linear expansions of the polytropic relations between ρ, p, and T, and show that to first order in $\not{p}$ weak shocks are essentially adiabatic [but see (56.51) and (56.56)].

Upstream and downstream flow properties can also be related in terms of the upstream Mach number M_1. Thus from (56.32) we find

$$p_2/p_1 = [2\gamma M_1^2 - (\gamma - 1)]/(\gamma + 1), \tag{56.40}$$

hence from (56.20)

$$\rho_2/\rho_1 = (\gamma + 1)M_1^2/[(\gamma - 1)M_1^2 + 2] = u_1/u_2, \tag{56.41}$$

and from (56.21)

$$T_2/T_1 = [2\gamma M_1^2 - (\gamma - 1)][(\gamma - 1)M_1^2 + 2]/(\gamma + 1)^2 M_1^2. \tag{56.42}$$

Then using (56.41) and (56.42) in $M_2^2 = (u_2/a_2)^2 = M_1^2(u_2/u_1)^2(a_1/a_2)^2$ we find

$$M_2^2 = [(\gamma - 1)M_1^2 + 2]/[2\gamma M_1^2 - (\gamma - 1)]. \tag{56.43}$$

For strong shocks $M_1 \to \infty$ and we obtain the same limiting values for ρ_2/ρ_1 and M_2 stated above, while $p_2/p_1 \to 2\gamma M_1^2/(\gamma + 1) \to \infty$ and $T_2/T_1 \to 2\gamma(\gamma - 1)M_1^2/(\gamma + 1)^2 \to \infty$.

For weak shocks with $M_1^2 = 1 + m$, $m \ll 1$, we find

$$(p_2/p_1) - 1 = 2\gamma m/(\gamma + 1) \tag{56.44}$$

$$(\rho_2/\rho_1) - 1 = 2m/(\gamma + 1) \tag{56.45}$$

$$(T_2/T_1) - 1 = 2(\gamma - 1)m/(\gamma + 1) \tag{56.46}$$

and

$$M_2^2 = 1 - m. \tag{56.47}$$

THE ENTROPY JUMP

In the (p, V) diagram, the adiabats form a one-parameter family of curves $p = p(V, s)$ where the specific entropy s is fixed along each curve. In contrast, the Hugoniot curves form a *two*-parameter family, with the curve passing through (p_1, V_1) having the form $p(V, p_1, V_1)$. In general, the Hugoniot through (p_1, V_1) is not identical with the adiabat through (p_1, V_1)—a fact demonstrated below for weak shocks in general materials, and for shocks of arbitrary strengths in a perfect gas. Thus in general Hugoniots *cross* adiabats, which implies that the entropy of the material changes as it passes through a shock; therefore the entropy experiences a discrete jump across the front by an amount determined by the shock strength.

According to the second law of thermodynamics, the entropy of a substance cannot be decreased by internal processes alone (cf. §3); thus the downstream specific entropy in a shock must equal or exceed its upstream value, $s_2 \geq s_1$. This entropy increase, predicted by the mass, momentum, and energy conservation relations alone, implies an irreversible dissipation of energy, even for an ideal fluid, entirely independently of the existence of a dissipation mechanism, which at first sight seems paradoxical. This apparent paradox is easily resolved by studying shock structure for a real

gas (§57). We then find that a shock is not a mathematical discontinuity, but is actually a thin transition layer, a few particle mean free paths λ thick, where dissipative mechanisms (which generate entropy) are strongly operative in response to steep gradients. The ideal fluid is merely a degenerate case obtained when we suppress the internal transport properties of a fluid, which is equivalent to letting $\lambda \to 0$, which in turn implies that the transition layer collapses to a discontinuity. We get the same total entropy jump for given upstream conditions in both cases because the entropy, like any other thermodynamic variable, can be regarded as a function of any two other variables, say (p, V). As we have seen, the downstream values of these variables are uniquely fixed by the hydrodynamical equations alone, regardless of the detailed physical properties of the fluid.

Consider first a weak shock, and examine the implications of (56.13). Take $h = h(p, s)$ and expand in powers of $\Delta p \equiv p_2 - p_1$ and $\Delta s \equiv s_2 - s_1$. Anticipating the result that Δs is $O(\Delta p^3)$, we retain only first-order terms in Δs and terms up to third order in Δp, obtaining

$$h_2 - h_1 = (\partial h/\partial s)_p\, \Delta s + (\partial h/\partial p)_s\, \Delta p + \tfrac{1}{2}(\partial^2 h/\partial p^2)_s\, \Delta p^2 + \tfrac{1}{6}(\partial^3 h/\partial p^3)_s\, \Delta p^3. \tag{56.48}$$

From (2.33) we have $(\partial h/\partial s)_p = T$ and $(\partial h/\partial p)_s = 1/\rho = V$, hence

$$h_2 - h_1 = T_1\, \Delta s + V_1\, \Delta p + \tfrac{1}{2}(\partial V/\partial p)_s\, \Delta p^2 + \tfrac{1}{6}(\partial^2 V/\partial p^2)_s\, \Delta p^3. \tag{56.49}$$

Similarly, take $V = V(p, s)$; inasmuch as $(V_1 + V_2)$ in (56.13) is already multiplied by Δp we can expand V to only second order in Δp, and omit the term in Δs, obtaining

$$V_2 = V_1 + (\partial V/\partial p)_s\, \Delta p + \tfrac{1}{2}(\partial^2 V/\partial p^2)_s\, \Delta p^2. \tag{56.50}$$

Substituting (56.49) and (56.50) into (56.13) we find

$$s_2 - s_1 = \tfrac{1}{12}(\partial^2 V/\partial p^2)_s (p_2 - p_1)^3/T_1. \tag{56.51}$$

From (56.51) we see that Δs is nonzero (unless $\Delta p \equiv 0$), and the requirement that $\Delta s > 0$ fixes the sign of Δp once the sign of $(\partial^2 V/\partial p^2)_s$ is known. For "normal" substances both experiment and theory show that $(\partial^2 V/\partial p^2)_s > 0$, that is, adiabats are concave upward in the (p, V) diagram; for example it follows from (4.16) that for a perfect gas $(\partial^2 V/\partial p^2)_s = (\gamma + 1)V/\gamma^2 p^2$. Thus for normal substances we conclude that $\Delta p > 0$ across a shock front, as asserted earlier. Given that $\Delta p > 0$, (56.9) implies $\Delta V < 0$ $(\Delta \rho > 0)$, (56.12) implies $\Delta u < 0$, (56.13) implies $\Delta h > 0$, and (56.14) implies $\Delta e > 0$.

For a weak shock, (56.9) yields

$$\dot{m} \approx [-(\partial p/\partial V)_s]^{1/2}, \tag{56.52}$$

and to the same order, (56.10) and (56.11) yield

$$u_1 \approx u_2 \approx \dot{m} V = [(\partial p/\partial \rho)_s]^{1/2} = a. \tag{56.53}$$

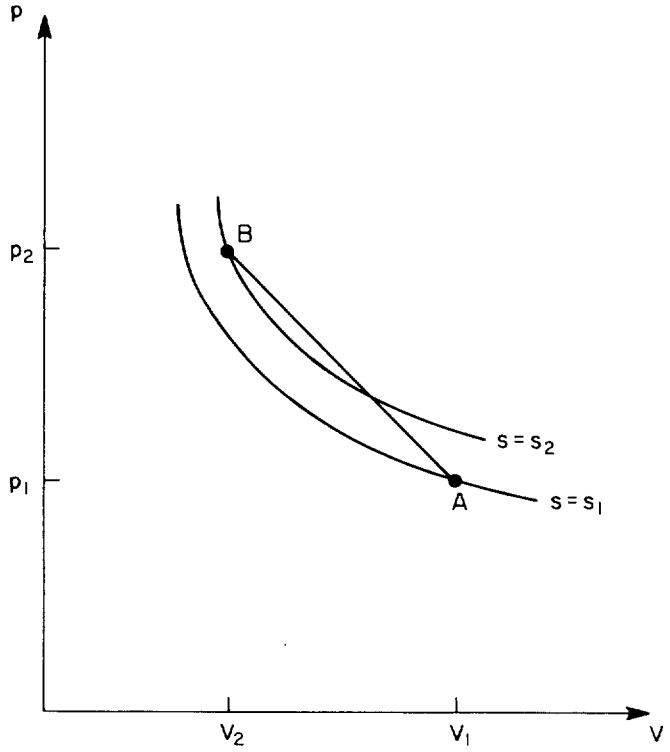

Fig. 56.3 Shock transition connecting two adiabats in (p, V) diagram.

These results can be refined by reference to Figure 56.3. There we see that because the initial and final states are joined by the chord AB whose slope is $-\dot{m}^2$ [cf. (56.9)], we must have $\dot{m}^2 > [-(\partial p/\partial V)_s]_1$. Therefore at point A

$$u_1^2 = V_1^2 \dot{m}^2 > -V_1^2[(\partial p/\partial V)_s]_1 = [(\partial p/\partial \rho)_s]_1 = a_1^2, \tag{56.54}$$

that is, $u_1 > a_1$. By a similar analysis at point B we find $u_2 < a_2$. These general results are consistent with (56.38), (56.39), and (56.47), which, however, apply only for a perfect gas. Indeed it can be shown that all of the inequalities stated above are true for shocks of arbitrary strength provided only that $(\partial^2 V/\partial p^2)_s > 0$; see (**L2,** §84).

For the special case of a perfect gas, we can write explicit formulae for Δs in shocks of arbitrary strength. Thus from (4.10) we have

$$\begin{aligned}
\Delta s &= c_v \ln (p_2 \rho_1^\gamma / p_1 \rho_2^\gamma) \\
&= c_v \left\{ \ln (1 + \not{\epsilon}) + \gamma \ln \left[\frac{2\gamma + (\gamma - 1)\not{\epsilon}}{2\gamma + (\gamma + 1)\not{\epsilon}} \right] \right\} \\
&= c_v \left\{ \ln \left[\frac{2M_1^2 - (\gamma - 1)}{\gamma + 1} \right] + \gamma \ln \left[\frac{(\gamma - 1)M_1^2 + 2}{(\gamma + 1)M_1^2} \right] \right\}.
\end{aligned} \tag{56.55}$$

From (56.55) it is straightforward to show that Δs is a monotone increasing function of $\not{\mu}$ and of M_1, and furthermore that for $\not{\mu} \ll 1$ and $m \ll 1$

$$\Delta s = (\gamma - 1)c_v \not{\mu}^3/12\gamma^2 = 2\gamma(\gamma - 1)c_v m^3/3(\gamma + 1)^2 \geq 0; \qquad (56.56)$$

thus Δs is, in fact, always greater than or equal to zero. Equation (56.56) is, of course, consistent with (56.51) and (56.44).

STABILITY

We have seen that the requirement that entropy not decrease across shock fronts implies that $\Delta p > 0$, hence that rarefaction discontinuities do not exist. There are additional reasons why rarefaction discontinuities cannot exist. If such a discontinuity did exist, it would have $u_1 < a_1$ and $u_2 > a_2$, and would therefore propagate subsonically through the undisturbed medium. But then any small disturbance, which would travel as an acoustic wave at the speed of sound, produced in the flow at the jump could outrun the discontinuity. Therefore the rarefaction region behind the discontinuity would tend to spread into the gas in front of the discontinuity faster than the discontinuity itself could propagate, and in doing so would erode away any initial jump in material properties. That is, a rarefaction discontinuity is immediately smoothed into a continuous transition. Furthermore, because a rarefaction discontinuity would move supersonically with respect to downstream material, it could not be influenced by any process or change in conditions occurring behind the jump. That is, no feedback on the wave is possible, and in that sense the wave is uncontrolled. Both of these properties imply that rarefaction discontinuities are *mechanically unstable*, and disintegrate immediately.

In contrast, in a compression shock the entropy increases. The front outruns acoustic waves that might tend to smear it out, and the upstream material remains "unaware" of the shock until it slams into it; hence the shock can propagate as a sharp discontinuity. Furthermore, the shock propagates subsonically with respect to downstream material, hence the material behind the front can influence the front's behavior; if the downstream gas is strongly compressed and heated, it tends to strengthen the shock; if the downstream material cools rapidly (e.g., by radiation losses) the driving force behind the shock front weakens and eventually the shock dissipates. Thus compression shocks not only satisfy entropy constraints but are, in addition, mechanically stable. Further discussion of these issues can be found in (**C6,** Chap. 3), (**L2,** §84), and (**Z1,** §1.17).

RELATIVISTIC SHOCKS

We can obtain jump conditions across a shock in a relativistic flow by expressing the continuity equation (39.9) and the dynamical equations (42.2) to (42.6) in the frame in which the shock is stationary, and then subjecting them to the same analysis as led to (56.6) to (56.8). Thus writing N for the number of particles per unit proper volume, and $U_x = \gamma u_x$ for the

x component of the four-velocity of the material relative to the front, particle conservation implies continuity of the particle flux:

$$N_1 U_{x1} = N_2 U_{x2} \equiv j. \tag{56.57}$$

Similarly, energy and momentum conservation imply continuity of the energy flux

$$[(\hat{e}+p)U_0 U_x]_1 = [(\hat{e}+p)U_0 U_x]_2 \tag{56.58a}$$

or

$$(\hat{e}+p)_1 \gamma_1 U_{x1} = (\hat{e}+p)_2 \gamma_2 U_{x2}, \tag{56.58b}$$

and of the momentum flux

$$(\hat{e}+p)_1 U_{x1}^2 + p_1 c^2 = (\hat{e}+p)_2 U_{x2}^2 + p_2 c^2. \tag{56.59}$$

Here $\hat{e} = \rho_0(c^2+e) = Nm_0(c^2+e)$ is the total proper energy density of the fluid.

Rewrite (56.57) as

$$U_{x1} = j\tilde{V}_1 \tag{56.60a}$$

and

$$U_{x2} = j\tilde{V}_2, \tag{56.60b}$$

where $\tilde{V} \equiv 1/N$ is the volume per particle. Then (56.59) becomes

$$p_2 - p_1 = j^2(\tilde{h}_1 \tilde{V}_1 - \tilde{h}_2 \tilde{V}_2)/c^2, \tag{56.61}$$

while (56.58) reduces to

$$\gamma_1 \tilde{h}_1 = \gamma_2 \tilde{h}_2. \tag{56.62}$$

Here

$$\tilde{h} \equiv m_0(c^2+e) + (p/N) \tag{56.63}$$

is the total enthalpy per particle. In the nonrelativistic limit, (56.61) reduces to (56.9). Multiplying (56.61) by $(\tilde{h}_1 \tilde{V}_1 + \tilde{h}_2 \tilde{V}_2)$ and using (56.60) we find

$$(\tilde{h}_1 U_{x1}/c)^2 - (\tilde{h}_2 U_{x2}/c)^2 = (p_2 - p_1)(\tilde{h}_1 \tilde{V}_1 + \tilde{h}_2 \tilde{V}_2). \tag{56.64}$$

Then adding the square of (56.62) we obtain, finally,

$$\tilde{h}_2^2 - \tilde{h}_1^2 = (\tilde{h}_1 \tilde{V}_1 + \tilde{h}_2 \tilde{V}_2)(p_2 - p_1), \tag{56.65}$$

which reduces to (56.13) in the nonrelativistic limit. Equations (56.61) and (56.65), first derived by Taub (**T1**), are the relativistic generalizations of the Rankine-Hugoniot jump relations.

It is possible to obtain relativistic generalizations for essentially all of the results derived above for nonrelativistic shocks. Thus one can show (**T2**) that the generalization of (56.51) for the entropy jump across a weak shock is

$$s_2 - s_1 = \tfrac{1}{2}[\partial^2(\tilde{h}\tilde{V})/\partial p^2]_s (p_2 - p_1)^3/\tilde{h}T, \tag{56.66}$$

which, as long as $[\partial^2(\tilde{h}\tilde{V})/\partial p^2]_s > 0$, implies that $p_2 > p_1$, $\tilde{h}_2 > \tilde{h}_1$, $N_2 > N_1$,

$\tilde{V}_2 < \tilde{V}_1$, $U_{x2} < U_{x1}$, $U_{x1}/a_1 > 1$, $U_{x2}/a_2 < 1$ in a shock in which entropy increases. By straightforward manipulation of (56.58) and (56.59), one can also show that

$$\frac{u_{x1}}{c} = \left[\frac{(p_2 - p_1)(\hat{e}_2 + p_1)}{(\hat{e}_2 - \hat{e}_1)(\hat{e}_1 + p_2)}\right]^{1/2} \tag{56.67}$$

and

$$\frac{u_{x2}}{c} = \left[\frac{(p_2 - p_1)(\hat{e}_1 + p_2)}{(\hat{e}_2 - \hat{e}_1)(\hat{e}_2 + p_1)}\right]^{1/2}, \tag{56.68}$$

where u_x is the ordinary velocity (i.e., the three-velocity) of the material relative to the shock. In the nonrelativistic limit, $\hat{e} \to \rho c^2 \gg p$, and (56.67) and (56.68) reduce to (56.10) and (56.11). In the extreme relativistic limit, $p \to \frac{1}{3}\hat{e}$, hence

$$\frac{u_{x1}}{c} \to \left[\frac{3\hat{e}_2 + \hat{e}_1}{3(3\hat{e}_1 + \hat{e}_2)}\right]^{1/2} \tag{56.69}$$

and

$$\frac{u_{x2}}{c} \to \left[\frac{3\hat{e}_1 + \hat{e}_2}{3(3\hat{e}_2 + \hat{e}_1)}\right]^{1/2}. \tag{56.70}$$

For weak shocks, $\hat{e}_2 \approx \hat{e}_1$ and $u_{x1} \approx u_{x2} \approx c/\sqrt{3}$; for strong shocks $\hat{e}_2 \gg \hat{e}_1$ and $u_{x1} \to c$ while $u_{x2} \to c/\sqrt{3}$. Using the relativistic law for the addition of velocities, that is,

$$\Delta u = (u_{x1} - u_{x2})/[1 - (u_{x1}u_{x2}/c^2)], \tag{56.71}$$

(which follows from boosting the four-velocity of a particle moving with velocity u_{x1} in a frame S into a frame S' moving with velocity u_{x2} relative to S), we find that the relative velocity of the gas on the two sides of the shock is

$$\frac{\Delta u}{c} = \left[\frac{(p_2 - p_1)(\hat{e}_2 - \hat{e}_1)}{(\hat{e}_1 + p_2)(\hat{e}_2 + p_1)}\right]^{1/2}. \tag{56.72}$$

In the nonrelativistic limit, (56.72) reduces to (56.15).

More complete discussions of relativistic shocks are given in (**I2**), (**L7**), (**L8**), (**M2**), (**T2**), and (**T4**).

57. Shock Structure

Let us now investigate how dissipative processes—viscosity and thermal conduction—determine the structure and thickness of shock fronts. We expect these processes to play a key role within the front because gradients are very steep there (indeed, infinitely steep according to the idealized analysis of §56).

The conservation relations for the steady flow in the shock's frame now

are [cf. (26.2) and (27.34)]

$$\rho u = \rho_1 u_1 = \dot{m}, \tag{57.1}$$

$$\rho u^2 + p - \mu'(du/dx) = \rho_1 u_1^2 + p_1, \tag{57.2}$$

and

$$\rho u(h + \tfrac{1}{2}u^2) - \mu' u(du/dx) - K(dT/dx) = \rho_1 u_1(h_1 + \tfrac{1}{2}u_1^2), \tag{57.3}$$

and the entropy-generation equation (27.17) in the shock's frame is

$$\rho u T\left(\frac{ds}{dx}\right) = \mu'\left(\frac{du}{dx}\right)^2 + \frac{d}{dx}\left(K\frac{dT}{dx}\right). \tag{57.4}$$

Here $\mu' \equiv \frac{4}{3}\mu + \zeta$ denotes the effective one-dimensional viscosity.

Equations (57.1) to (57.4) apply at all points in the flow and determine the physical properties (ρ, p, u, etc.) as functions of x across the shock front. The constants on the right-hand sides of (57.2) and (57.3) are evaluated in the upstream flow far from the shock, where $(du/dx) = (dT/dx) \equiv 0$. If we evaluate the left-hand sides of (57.2) and (57.3) in the downstream flow far from the shock where (du/dx) and (dT/dx) again vanish, we recover the ideal-fluid jump relations (56.7) and (56.8).

VISCOUS SHOCKS

Consider first a hypothetical fluid having a finite viscosity but zero thermal conductivity. In order to simplify the discussion we assume that μ' is constant and that the fluid is a perfect gas. We can then rewrite (57.2) and (57.3) as

$$p + \dot{m}u - \mu'(du/dx) = p_1 + \dot{m}u_1 \tag{57.5}$$

and

$$u\{\tfrac{1}{2}\dot{m}u + [\gamma p/(\gamma - 1)]\} - \mu' u(du/dx) = \dot{m}\{\tfrac{1}{2}u_1^2 + [\gamma p_1/(\gamma - 1)\rho_1]\}. \tag{57.6}$$

Multiplying (57.5) by $\gamma u/(\gamma - 1)$ and subtracting (57.6) we obtain

$$-\nu u(du/dx) = a_1^2(u - u_1) + u_1^2[\gamma u - \tfrac{1}{2}(\gamma - 1)u_1] - \tfrac{1}{2}(\gamma + 1)u_1 u^2, \tag{57.7}$$

where $\nu = \mu'/\rho$ is the effective kinematic viscosity. Let $w \equiv u_1 - u$; then (57.7) can be rewritten

$$\nu(dw/dx) = w[u_1^2 - a_1^2 - \tfrac{1}{2}(\gamma + 1)u_1 w]/(u_1 - w). \tag{57.8}$$

The *velocity drop* w varies from $w = 0$ far upstream to

$$w_{\max} = u_1 - u_2 = 2(u_1^2 - a_1^2)/(\gamma + 1)u_1 \tag{57.9}$$

far downstream; the second equality in (57.9) follows from the Prandtl relation (56.29).

From (57.8) and (57.9), one finds that $(dw/dx) \geq 0$, with $(dw/dx) = 0$ at $w = 0$ and $w = w_{\max}$; therefore $w(x)$ is monotone increasing and $u(x)$ is

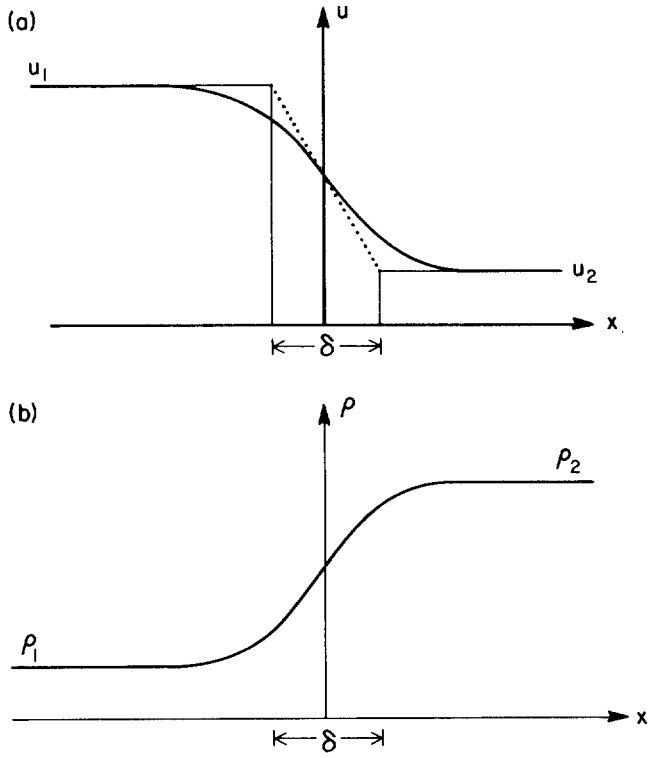

Fig. 57.1 Velocity and density variation in a viscous shock.

monotone decreasing. Furthermore, w has an inflection point because

$$\nu(d^2w/dx^2) = u_1[u_1^2 - a_1^2 - (\gamma+1)w(u_1 - \tfrac{1}{2}w)]/(u_1 - w)^2, \quad (57.10)$$

which shows that $(d^2w/dx^2) > 0$ at $w = 0$, $(d^2w/dx^2) < 0$ at $w = w_{\max}$, and $(d^2w/dx^2) = 0$ at $w = u_1 - \sqrt{u_1 u_2}$. Thus $u(x)$ varies as sketched in Figure 57.1a. From continuity it follows that $\rho(x)$ is monotone increasing, as sketched in Figure 57.1b.

Using (57.8) in (57.5) we find

$$p = p_1 + \rho_1 w[a_1^2 + \tfrac{1}{2}(\gamma-1)u_1 w]/(u_1 - w), \quad (57.11)$$

which shows that $p(x)$ is monotone increasing, like the sketch in Figure 57.1b. Combining (57.1) and (57.11) we find

$$T/T_1 = (p/p_1)(\rho_1/\rho) = 1 + (\gamma-1)(w/u_1) + \tfrac{1}{2}\gamma(\gamma-1)(w/a_1)^2, \quad (57.12)$$

which shows that $T(x)$ increases monotonically, like the sketch in Figure 57.1b.

As an estimate of the shock width δ we take

$$\delta \approx [w/(dw/dx)]_0, \quad (57.13)$$

where x_0 is the point at which $w = \frac{1}{2}w_{\max}$. Substituting from (57.9) into (57.8) one finds

$$(dw/dx)_0 = \tfrac{1}{2}(u_1^2 - a_1^2)^2/\nu(\gamma u_1^2 + a_1^2), \tag{57.14}$$

whence

$$\delta = 2\nu(\gamma u_1^2 + a_1^2)/[(\gamma+1)u_1(u_1^2 - a_1^2)]. \tag{57.15}$$

From mean free path arguments we know that $\nu = \mu_1/\rho_1 \sim a_1\lambda$. Thus in the weak-shock limit, where $u_1 \approx a$, we find, using (56.38),

$$\delta = \frac{2\nu}{a_1^2(M_1^2-1)} = \frac{4\gamma\nu}{(\gamma+1)a_1\not{p}} \sim \left(\frac{4\gamma}{\gamma+1}\right)\left(\frac{\lambda}{\not{p}}\right). \tag{57.16}$$

Thus the shock thickness is of the order of a particle mean free path divided by the fractional pressure jump.

For a strong shock ($M_1 \gg 1$), (57.15) yields

$$\delta \sim [2\gamma/(\gamma+1)](\lambda/M_1), \tag{57.17}$$

which formally predicts that δ becomes much smaller than λ when $M_1 \gg 1$. This result is incompatible with a fluid description, and comes from taking $\nu \sim a_1\lambda$. If instead we assume that the viscous dissipation occurs mainly in the hot material at the back edge of the transition zone and take $\nu \sim \bar{u}\lambda \approx a_2\lambda \sim M_1 a_1 \lambda$, then $\delta \sim C\lambda$ where C is a number of order unity; hence δ remains of the order of λ.

For a purely viscous shock

$$\rho u T(ds/dx) = \dot{m}T(ds/dx) = \mu'(du/dx)^2; \tag{57.18}$$

hence s increases monotonically through the shock like the sketch in Figure 57.1b. We can use (57.18) to estimate the entropy jump across the shock by replacing derivatives with finite differences, writing

$$\dot{m}T_1(\Delta s/\Delta x) \approx \mu'(u_2 - u_1)^2/\Delta x^2. \tag{57.19}$$

Using (56.16) to write $\Delta u = \Delta p/\dot{m} \approx \Delta p/\rho a$ for a weak shock, and adopting $\Delta x \approx \delta$ as given by (57.16), we find that (57.19) gives the same result as (56.56) to within a numerical factor of order unity.

The analysis presented above shows the fundamental role played by viscosity in determining shock structure: it leads to an irreversible conversion of kinetic energy of the inflowing material into heat. Put differently, it transforms *ordered* flow motion of the particles in the gas into *random* motions via the mechanism of dissipation of particle momentum.

CONDUCTING SHOCKS

Now consider a fluid with zero viscosity ($\mu' \equiv 0$), but finite thermal conductivity. These assumptions are of more than hypothetical interest because, as we mentioned in §51, they are realistic for a radiating gas in which radiative energy transport can strongly influence shock structure even when

viscous and thermal-conduction effects are negligible (cf. §104). As we will see, the structure of an inviscid conducting shock can be qualitatively different from that of a pure viscous shock.

When $\mu' \equiv 0$, the momentum and energy conservation relations are

$$p + \dot{m}u = p_1 + \dot{m}u_1 \tag{57.20}$$

and

$$u\{\tfrac{1}{2}\dot{m}u + [\gamma p/(\gamma-1)]\} + q = \dot{m}\{\tfrac{1}{2}u_1^2 + [\gamma p_1/(\gamma-1)\rho_1]\}, \tag{57.21}$$

where $q = -K(dT/dx)$. The entropy-generation equation reduces to

$$(ds/dx) = (K/\dot{m}T)(d^2T/dx^2), \tag{57.22}$$

where K has been assumed to be constant.

Rewrite (57.20) as

$$p = p_1 + \dot{m}u_1(1-\eta) \tag{57.23}$$

where η is the *volume ratio*

$$\eta \equiv V/V_1 = \rho_1/\rho = u/u_1. \tag{57.24}$$

Clearly the pressure is a monotone increasing function through the shock front, rising to p_2 when η equals

$$\eta_2 = V_2/V_1 = [(\gamma-1)M_1^2 + 2]/(\gamma+1)M_1^2, \tag{57.25}$$

as given by (56.41). Using the perfect gas law we then find

$$T/T_1 = \eta[\gamma M_1^2(1-\eta) + 1] \tag{57.26}$$

which shows that $T(\eta)$ is a quadratic function of η, as sketched in Figure 57.2. $T(\eta)$ reaches its maximum value at

$$\eta_{\max} = (\gamma M_1^2 + 1)/2\gamma M_1^2. \tag{57.27}$$

Note in passing that as $M_1 \to \infty$, $\eta_{\max} \to \frac{1}{2}$. Finally, from (57.21) and (57.23) to (57.25), we obtain

$$q = -\dot{m}u_1^2(\gamma+1)(1-\eta)(\eta-\eta_2)/2(\gamma-1), \tag{57.28}$$

which shows that $q \leq 0$ for $\eta_2 \leq \eta \leq 1$. More precisely, $q = 0$ at $\eta = 1$ and at $\eta = \eta_2$, and q reaches an absolute minimum at the midpoint of the compression, that is, at

$$\eta_0 = \tfrac{1}{2}(1+\eta_2). \tag{57.29}$$

Consider first a weak shock for which $\eta_2 > \eta_{\max}$, for example, the transition to point A sketched in Figure 57.2. From (57.26) we see that in this case $T(x)$ increases monotonically from T_1 to T_2, which implies $(dT/dx) \geq 0$, which is consistent with the conclusion that $q \leq 0$. Furthermore, the fact that q achieves a minimum at $\eta = \eta_0$ implies that $(dq/dx)_0 = -K(d^2T/dx^2)_0 = 0$; therefore $T(x)$ has an inflection point at η_0 and varies as sketched in Figure 57.3a. These conclusions further imply that the

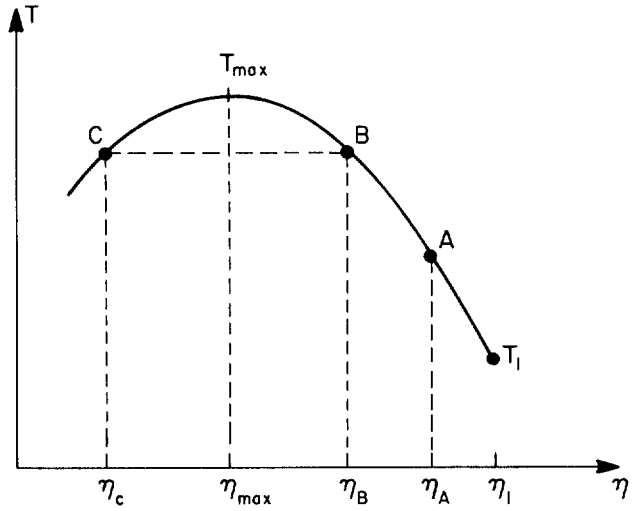

Fig. 57.2 Temperature variation as a function of volume ratio in a conducting shock.

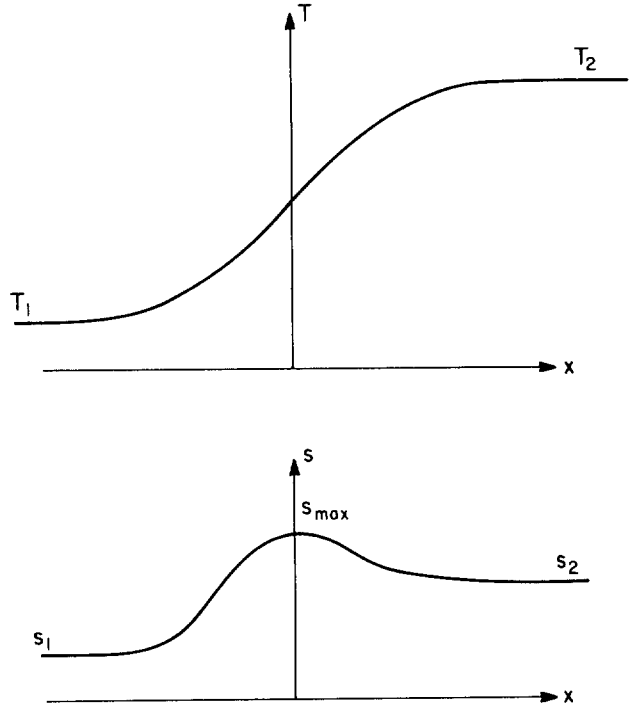

Fig. 57.3 Temperature and entropy variation as a function of spatial position in a weak conducting shock.

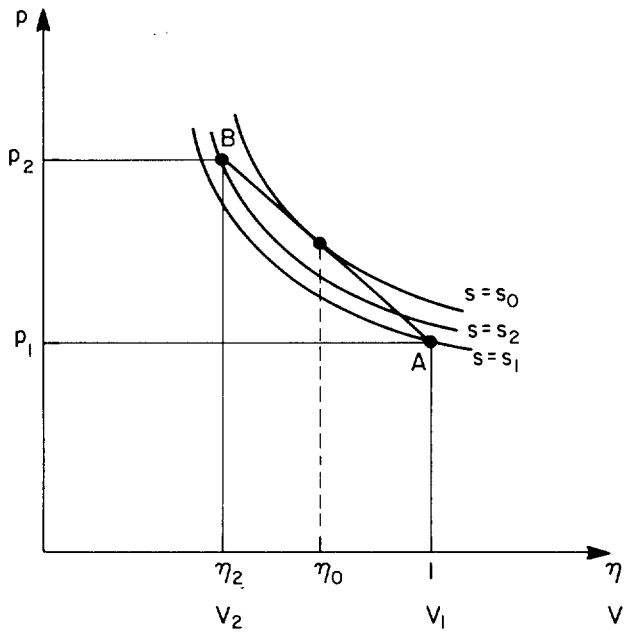

Fig. 57.4 Shock transition and initial, final, and maximum entropy adiabats in (p, V) diagram for a conducting shock.

entropy must achieve a local maximum at $\eta = \eta_0$ because $(d^2T/dx^2) = 0$ implies $(ds/dx) = 0$; thus $s(x)$ varies as sketched in Figure 57.3b.

We can calculate the maximum entropy increase in the shock as follows. According to (57.23), in the (p, V) plane the gas follows the straight line AB shown in Figure 57.4, and the material reaches its maximum entropy where this line is just tangent to an isentrope, say at $s = s_0$. For a weak shock the equation for the straight line is

$$p - p_1 = \frac{(p_2 - p_1)(V - V_1)}{(V_2 - V_1)} \approx \left(\frac{\partial p}{\partial V}\right)_{s_1} (V - V_1) + \frac{1}{2}\left(\frac{\partial^2 p}{\partial V^2}\right)_{s_1} (V_2 - V_1)(V - V_1), \tag{57.30}$$

where we have ignored a term in $(\partial p/\partial s)_V$ because $(s_2 - s_1)$ is third order in Δp or ΔV. Similarly, the equation for the isentrope $s = s_0$ is

$$p - p_1 \approx \left(\frac{\partial p}{\partial V}\right)_{s_1} (V - V_1) + \frac{1}{2}\left(\frac{\partial^2 p}{\partial V^2}\right)_{s_1} (V - V_1)^2 + \left(\frac{\partial p}{\partial s}\right)_{V_1} (s_0 - s_1), \tag{57.31}$$

where again we neglect third-order terms.

To enforce tangency of the two curves, we demand that $[(\partial p/\partial V)_{\text{line}}]_0 = [(\partial p/\partial V)_{\text{isentr}}]_0$ at the point where $(\partial s/\partial V) = 0$ (hence s is a maximum), whence we find that $V_0 = \frac{1}{2}(V_1 + V_2)$, in agreement with (57.29). Then, demanding $(p_{\text{line}})_0 = (p_{\text{isentr}})_0$, we equate the right-hand sides of (57.30) and

(57.31) evaluated at $V = V_0$, obtaining

$$s_0 - s_1 = \tfrac{1}{8}[(\partial^2 p/\partial V^2)_s/(\partial p/\partial s)_V]_1 (V_1 - V_2)^2. \tag{57.32}$$

From (4.16) and (5.15) we find that for a perfect gas $(\partial^2 p/\partial V^2)_s = \gamma(\gamma+1)p/V^2$ and $(\partial p/\partial s)_V = p/c_v$, and from (56.36) we have $\Delta V/V = -\Delta p/\gamma p$, hence

$$s_0 - s_1 = \tfrac{1}{8}\gamma(\gamma+1)c_v(\Delta V/V)^2 = \tfrac{1}{8}[(\gamma+1)/\gamma]c_v \not{s}^2. \tag{57.33}$$

Thus the maximum entropy change *within* the shock front is second order in Δp or ΔV, whereas the total entropy change *across* the front is only third order.

To estimate the thickness of a conducting shock, we calculate the total entropy jump from (57.22), obtaining

$$(\dot{m}/K)\,\Delta s = \int_{-\infty}^{\infty} T^{-1}(d^2T/dx^2)\,dx = \int_{-\infty}^{\infty} [T^{-1}(dT/dx)]^2\,dx, \tag{57.34}$$

where we integrated by parts and noted that $(dT/dx) = 0$ at $x = \pm\infty$. The integral is approximately equal to $(q/KT)_0^2\,\delta$, where q is the heat flux and δ is the shock thickness. Evaluating q_0 from (57.28) and (57.29), using the scaling rule $K \sim a\lambda\rho c_v$, and using (56.56) for Δs, we find

$$\delta \sim \tfrac{16}{3}[(\gamma-1)/(\gamma+1)](\lambda/\not{s}), \tag{57.35}$$

which agrees with (57.15) for a weak viscous shock to within a numerical factor of order unity.

We have thus shown that for shocks below a certain critical strength [i.e., for which the downstream volume ratio η_2 is greater than $\eta_{\max}$ defined by (57.27)] in an inviscid, conducting fluid all physical properties vary continuously through the shock front over a distance of the order of a few particle mean free paths. The shock structure is qualitatively similar to that of a viscous shock, with the velocity decreasing monotonically, while ρ, p, and T rise monotonically; it differs only in that the entropy passes through a local maximum instead of increasing monotonically.

The situation for strong shocks in an inviscid, conducting fluid is quite different; here, as Rayleigh (**R3**) first noted, only the temperature varies continuously, while the other variables experience a discontinuous jump within the shock front [see also (**B2**)]. Thus suppose the material is compressed to some $\eta_2 < \eta_{\max}$, for example, to point C in Figure 57.2; this case arises naturally because, as noted above, as $M_1 \to \infty$, $\eta_{\max} \to \frac{1}{2}$ while $\eta_2 \to (\gamma-1)/(\gamma+1) \le \frac{1}{4}$. Equation (57.28) shows that $q = -K(dT/dx) \le 0$ for $\eta_2 \le \eta \le 1$, regardless of whether η_2 is less than $\eta_{\max}$ or not. We therefore must guarantee that

$$(dT/dx) = (dT/d\eta)(d\eta/dx) \ge 0 \tag{57.36}$$

throughout the shock front. Now $(d\eta/dx)$ is always ≤ 0, hence only those portions of the curve $T(\eta)$ for which $(dT/d\eta) \le 0$ are physically accessible.

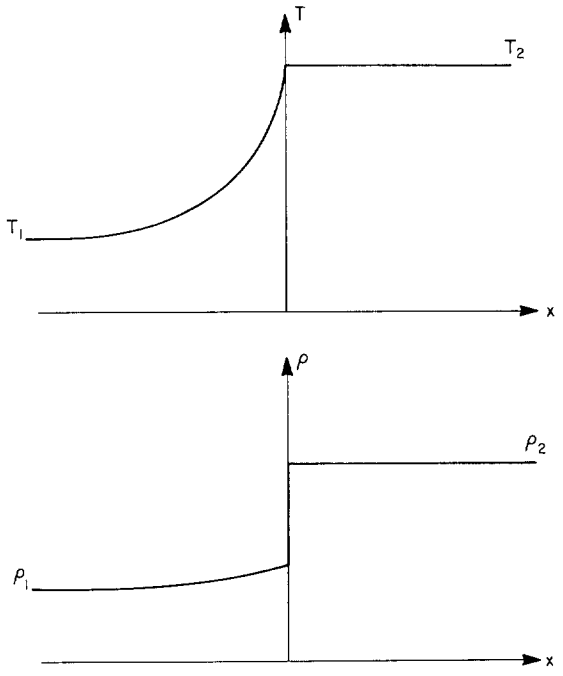

Fig. 57.5 Temperature and density variation as a function of spatial position in an isothermal shock.

Hence as the material is compressed from $\eta = 1$ to $\eta = \eta_C$, it is not possible for the temperature to rise to T_{max} and then track down the descending branch to $T = T_C$. The only way the material can actually make the transition to point C is for the temperature to rise continuously from T_1 to $T_B = T_C$ at $\eta = \eta_B > \eta_{max}$, and then remain constant while the relative volume collapses *discontinuously* from η_B to η_C, giving rise to a density jump like that sketched in Figure 57.5. Because this density discontinuity (which accounts for most of the total density jump across the shock) occurs at a single temperature, this type of solution is called an *isothermal shock*.

Thus in an inviscid, conducting fluid, all shocks above some critical strength will be isothermal. Combining (57.27) with (56.41) we see that the critical Mach number at which η_2 equals η_{max} is

$$(M_1^2)_{crit} = (3\gamma - 1)/\gamma(3 - \gamma), \tag{57.37}$$

which, from (56.40), implies that

$$(p_2/p_1)_{crit} = (\gamma + 1)/(3 - \gamma), \tag{57.38}$$

and, from (57.27),

$$(\eta_2)_{crit} = (\gamma + 1)/(3\gamma - 1). \tag{57.39}$$

For a monatomic gas with $\gamma = \frac{5}{3}$ shocks become isothermal for $M_1 \geq (\frac{9}{5})^{1/2} \approx 1.18$, or $p_2/p_1 \geq 2$, or $\eta_2 \leq \frac{2}{3}$. Conditions for an isothermal shock are met even more easily in polyatomic or ionizing gases, or when radiation is present, for then $\gamma < \frac{5}{3}$.

The notion of an isothermal shock is an idealization because in reality the strong density (hence velocity) jump within the front implies that viscous effects must inevitably come into operation and smooth the discontinuity. When *both* viscosity and conduction act (as they must in any real gas), viscosity converts flow momentum into heat, which is transported by conduction in such a way as to produce a local entropy maximum. All properties vary continuously through the front, though the density and pressure may rise rapidly in a limited region where the temperature changes less swiftly (**B2**). Nevertheless, the analysis presented above is instructive because it shows that we can *guarantee* a continuous solution for arbitrarily strong shocks only through the dissipative effects of viscosity, a point of considerable significance for numerical calculations (cf. §59).

THE RELAXATION LAYER

We have thus far assumed that the gas remains instantaneously in local thermodynamic equilibrium as it flows through a narrow transition zone, the *dissipation zone*, at the shock front, having a thickness δ of only a few particle mean free paths. In reality, the material may not be able to remain in equilibrium because the characteristic flow time $t_f \sim \lambda/u_1$ through the front may be much shorter than the time required for some thermodynamically important process (e.g., ionization of the material) to occur. Thus while some degrees of freedom may equilibrate within the dissipation layer (always true for the translational degrees of freedom of each particle species), others may be far from equilibrium when the material emerges from that layer.

In this event, the dissipation zone is followed downstream by a *relaxation layer* within which internal relaxation processes operate to bring the material to its final equilibrium state. If the characteristic relaxation time for some process is t_{relax}, the thickness of the relaxation layer associated with that process is $\Delta \sim u_2 t_{\text{relax}}$. Clearly $\Delta \gg \delta$ whenever $t_{\text{relax}} \gg t_f$. Several relaxation processes may occur simultaneously (or even sequentially), and the full thickness of the layer (i.e., the distance required to reach the point where the downstream conditions predicted by the Rankine-Hugoniot relations are achieved) is determined by the slowest process.

Relaxation processes can sometimes be described by phenomenological equations of the form

$$(dn/dt) = (n_{\text{equib}} - n)/t_{\text{relax}}, \tag{57.40}$$

where n represents the number of particles in the desired state (e.g., ionized as a result of passing through the shock) and n_{equib} is the number that would be in that state if the material were in equilibrium at the

downstream values of the material properties (i.e., ρ_2, p_2, T_2, etc.). Equation (57.40) implies an exponential relaxation of the form

$$n = n_0 \exp(-t/t_{\text{relax}}) + n_{\text{equib}}[1 - \exp(-t/t_{\text{relax}})] \tag{57.41}$$

where $t \sim x/u_2$, x being the distance downstream from the shock front and u_2 the downstream flow velocity. To obtain a more accurate picture we must specify the *rates* of the relevant relaxation processes, write *kinetic equations* that describe how these processes determine the distribution of particles over various states, and solve these equations (usually numerically) simultaneously with the equations of hydrodynamics.

In general, the problem can be quite complicated because on the one hand the relaxation rates depend on the thermodynamic state of the material, hence the dynamics of the flow, but on the other hand the relaxation processes determine the thermodynamic state of the material (hence the flow dynamics), for example by setting the rate of thermal energy loss into ionization (or the rate of energy gain by recombinations). We will discuss rate coefficients and kinetic equations in §85, and give examples of solutions of the set of coupled equations in §105. For the present, it suffices merely to describe qualitatively some of the basic processes that occur in the absence of radiation in order to get a physical feeling for their relative importance in different regimes.

(*a*) *Molecular Gas* The extent to which any particular process plays a significant role in determining the structure of the relaxation layer depends strongly on the degree of ionization of the gas. Consider first a neutral gas composed of atoms and diatomic molecules. The most rapid of all relaxation processes is the establishment of equilibrium among the translational degrees of freedom (i.e., of a Maxwellian velocity distribution). Typically only a few collisions are required to effect a complete randomization of particle motions and kinetic energy, hence a Maxwellian is usually established within a few mean free paths. Indeed, to a good approximation, the translational relaxation layer is coincident with the dissipation layer, and we can assign a unique kinetic temperature to each particle species at every point in the flow. Similarly, molecular rotation is typically quite easily excited in only a few collisions, and this degree of freedom usually remains in equilibrium with translational motions.

In contrast, molecular vibrational modes, which first become excited at temperatures of the order of 10^3 K, may require hundreds to thousands of collisions to come into equilibrium, and the vibrational relaxation layer in cool material and/or weak shocks may be much thicker than the dissipation layer. However, as temperatures rise to a few thousand kelvins, either because the upstream material is hot or because the shock is strong, vibrational relaxation proceeds much more rapidly, and is displaced from its role as the slowest process by molecular dissociation. When temperatures reach about 10^4 K in the downstream material, molecular dissociation

proceeds very rapidly and the limiting process becomes ionization, which we discuss further below.

To gain insight into the effects of relaxation processes on shock structure in a neutral gas we make the idealization that the shock is composed of two distinct regions: (1) a very thin dissipation zone (also called the *external relaxation zone*) in which viscosity and conduction effects are large, and within which equilibrium of the translational (and perhaps other) degrees of freedom is achieved, followed by (2) a relaxation zone (also called the *internal relaxation zone*) in which viscosity and conduction are unimportant, but some hitherto incompletely excited degree of freedom comes into equilibrium. These zones are assumed to be separated by a definite interface. As before upstream and downstream quantities are denoted by subscripts "1" and "2" respectively; properties at the interface are denoted by a subscript "i". Then the conservation relations are

$$\rho u = \rho_1 u_1 = \rho_i u_i = \rho_2 u_2 \equiv \dot{m}, \tag{57.42}$$

$$p + \rho u^2 = p_1 + \rho_1 u_1^2 = p_i + \rho_i u_i^2 = p_2 + \rho_2 u_2^2, \tag{57.43}$$

and

$$h + \tfrac{1}{2}u^2 = h_1 + \tfrac{1}{2}u_1^2 = h_i + \tfrac{1}{2}u_i^2 = h_2 + \tfrac{1}{2}u_2^2. \tag{57.44}$$

where unsubscripted variables denote quantities measured downstream from the interface. The enthalpy h_i includes only contributions from the translational and other rapidly excited degrees of freedom, other degrees of freedom still being frozen at their upstream values.

In this idealized description the material undergoes the transition sketched in Figure 57.6. Joining the initial state A to the final state C is the straight line (56.9). If all degrees of freedom were excited as rapidly as translational motions, in the dissipation zone the material would jump essentially discontinuously from A to C as defined by the intersection of the equilibrium Hugoniot with the straight line. But if some degrees of freedom are frozen during passage through the dissipation zone, the material has, in effect, a larger γ than it would in equilibrium [recall from kinetic theory that $\gamma = (n+2)/n$ where n = number of available degrees of freedom]. Therefore in the dissipation zone the material jumps essentially discontinuously from A only to point B, defined by the intersection of the straight line with a nonequilibrium Hugoniot which has a steeper slope than the equilibrium curve. Point B corresponds to (p_i, V_i) at the interface. The material then slowly relaxes along the straight line to its downstream equilibrium state C.

From (57.43) one has

$$(p_2 - p_i)/(p_2 - p_1) = (\eta_i - \eta_2)/(1 - \eta_2). \tag{57.45}$$

In a strong shock $\eta_i \approx \frac{1}{4}$ even if only translational motions are excited, and given that $\eta_2 \geq 0$, we see that the fractional pressure rise in the relaxation zone is always small, less than 25 percent of the total pressure jump in the

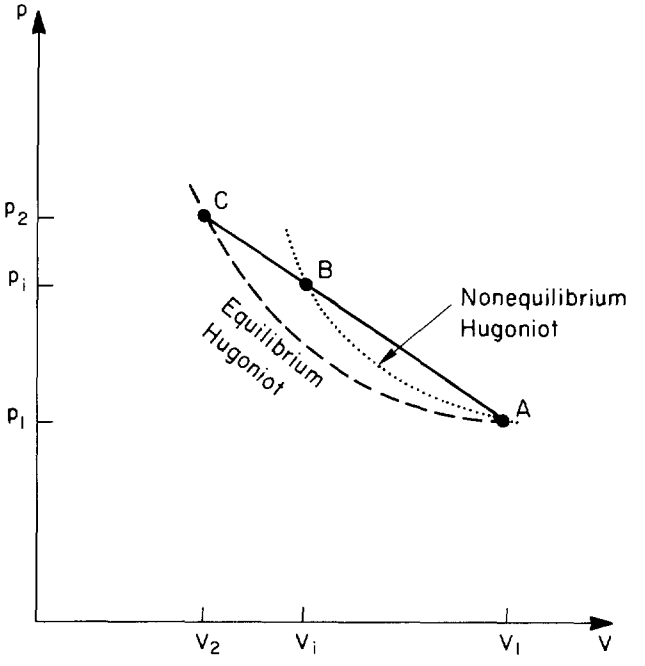

Fig. 57.6 Shock transition in material with a nonequilibrium relaxation layer.

shock. The pressure variation in the relaxation zone is sketched qualitatively in Figure 57.7a. The enthalpy increase behind the interface is even smaller. From (57.44) we have

$$(h_2 - h_i)/(h_2 - h_1) = (\eta_i^2 - \eta_2^2)/(1 - \eta_2^2) \tag{57.46}$$

and by the same reasoning we see that the fractional enthalpy increase in the relaxation zone is less than 6 per cent of the total.

Exploiting the result that h is nearly constant in the relaxation zone we can write

$$T_i/T_2 \approx \gamma_2(\gamma_i - 1)/\gamma_i(\gamma_2 - 1), \tag{57.47}$$

where γ_i and γ_2 are, respectively, the effective adiabatic exponents at the interface and far downstream. Because $\gamma_i \geq \gamma_2$, (57.47) implies that the shock has a significant *temperature overshoot* immediately behind the dissipation zone, as sketched in Figure 57.7b, followed by a long, inviscid, nonconducting *tail* in which the temperature decreases to its equilibrium value. Such shocks are called *partly dispersed*, with part of the total shock dissipation occurring in classical dissipation mechanisms (viscosity, conductivity) and part in a lagging relaxation process. If the shock is sufficiently weak it can become *fully dispersed* by relaxation processes alone, and the solution is continuous even in the absence of viscosity and thermal conductivity.

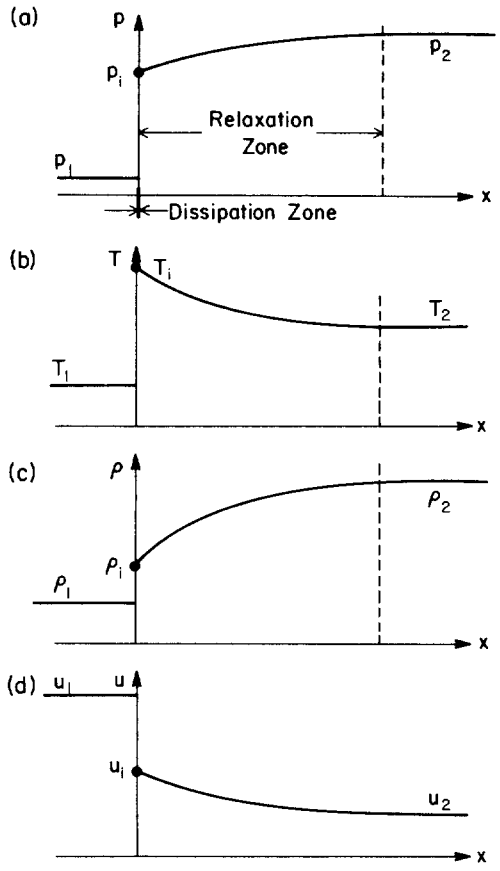

Fig. 57.7 Pressure, temperature, density, and velocity variation as a function of spatial position in a shock with a nonequilibrium relaxation zone.

The postshock temperature overshoot can be quite substantial. For example, if $\gamma_i = \frac{5}{3}$ (only translational motions) and $\gamma_2 = \frac{7}{5}$ (translation plus molecular rotation), $T_i/T_2 = 1.4$; if $\gamma_2 = \frac{9}{7}$ (translation, rotation, and vibration), $T_i/T_2 = 1.8$. Furthermore, noting that $\rho \sim p/T$, we see that the modest rise in p coupled with a significant drop in T leads to a fairly large rise in ρ in the relaxation zone, as sketched in Figure 57.7c. By continuity, u varies inversely as ρ, as sketched in Figure 57.7d.

(*b*) *Fully Ionized Plasma* Suppose now that the gas is sufficiently hot that all molecules have been dissociated, and it is composed of atoms, ions, and electrons. Indeed, consider first the extreme case of a completely ionized hydrogen plasma containing only electrons and protons (**I1**), (**S10**). As a first step in describing the shock structure we suppose that thermal conduction in the plasma can be neglected. Then the only relaxation

phenomenon that occurs is the equilibration of the postshock kinetic temperatures of the two species of particles.

Initially, in the upstream material, $T_e = T_p = T_1$. As the material passes through the shock front, Coulomb interactions among the protons produce viscous forces that dissipate a large fraction of the protons' directed kinetic energy into thermal motions, producing a large proton temperature rise $\Delta T_p \sim m_p u_1^2/k \sim (m_p a_1^2/k) M_1^2$ within a layer of thickness $\delta \sim u_2 t_{pp}$, where t_{pp} is the proton self-collision time [cf. (10.26)]. Because the electron self-collision time $t_{ee} = (m_e/m_p)^{1/2} t_{pp} \approx \frac{1}{43} t_{pp}$, the electrons also have adequate time to convert their own directed motions into thermal energy within the dissipation zone. However, this mechanism leads to an electron temperature increase of only $\Delta T_e \sim m_e u_1^2/k = (m_e/m_p)\Delta T_p \sim \frac{1}{1800}\Delta T_p$, which is clearly negligible. Furthermore, the electron-proton energy-exchange time is much too long $[t_{ep} = (m_p/m_e)^{1/2} t_{pp} = (m_p/m_e) t_{ee}]$ to permit significant energy transfer from the protons to the electrons within the dissipation zone, hence an opportunity for a large discrepancy between T_e and T_p arises.

The strong Coulomb forces coupling the electrons and protons assure that there can be no charge separation over distances much larger than a Debye length. Therefore, as the protons are compressed in the shock, the electrons are also compressed by the same amount, and because the electrons cannot exchange energy with the protons in the time available (and, for the present, we are ignoring thermal conduction), this compression occurs essentially adiabatically. Hence the electron temperature just downstream from the front is $T_{e,i} \approx (\rho_2/\rho_1)^{\gamma-1} T_1$, or, for a strong shock $(\rho_2/\rho_1 = 4)$ in a monatomic gas, $T_{e,i} \approx 2.5 T_1$. This is a large rise, but still much smaller than that experienced by the protons for large Mach numbers.

Within the framework of assumptions made above, we can derive a simple quantitative relation between T_e and T_p in the downstream flow. Behind the dissipation zone the pressure is nearly constant, hence

$$p = n_e k T_e + n_p k T_p \approx p_2 = 2 n_{e2} k T_2, \tag{57.48}$$

which, because the plasma is fully ionized $(n_e = n_p)$ and the postshock density is nearly constant $(n_e \approx n_{e2})$, implies that

$$T_e + T_p \approx 2 T_2 \approx \tfrac{5}{8} M_1^2 T_1. \tag{57.49}$$

For a completely ionized plasma of electrons and ions of charge Z, (57.49) generalizes to

$$Z T_e + T_{ion} = (Z+1) T_{ion,2} = \tfrac{5}{16}(Z+1) M_1^2 T_1. \tag{57.50}$$

For example, if $M_1 = 4$ in hydrogen, $T_2 = 5 T_1$, and given that $T_{e,i} \approx 2.5 T_1$ (adiabatic compression of the electrons), (57.49) shows that $T_{p,i} \approx 7.5 T_1$. As the plasma flows downstream, energy exchange between the electrons and protons finally occurs, producing a large temperature-relaxation layer

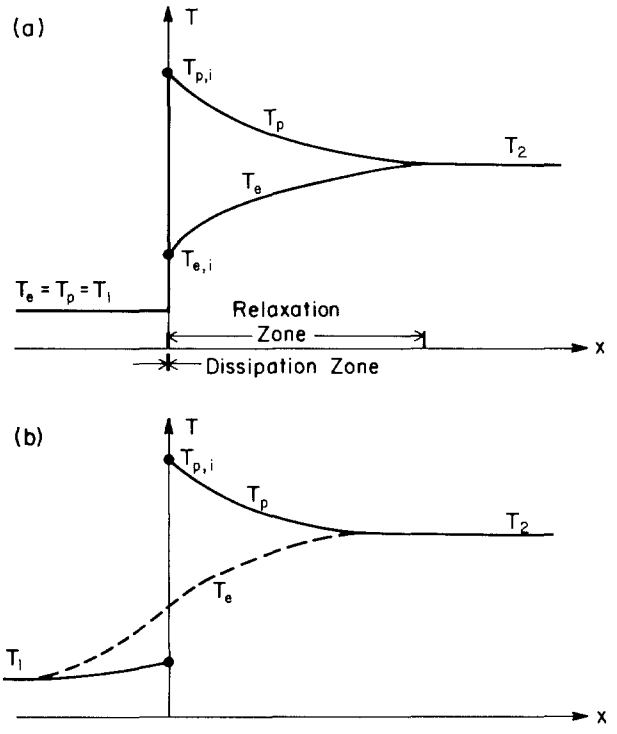

Fig. 57.8 Temperature variation as a function of spatial position in a shock in a two-fluid fully ionized hydrogen plasma. (a) Electron conduction omitted. (b) Electron conduction included.

of thickness

$$\Delta \sim u_2 t_{ep} = (m_p/m_e)^{1/2} u_2 t_{pp} \sim (m_p/m_e)^{1/2} \lambda_{\text{proton}} \sim (m_p/m_e)^{1/2}\, \delta \quad (57.51)$$

in which T_p decreases, and T_e increases, towards T_2, in accordance with (57.49). The resulting temperature profile is sketched in Figure 57.8a.

The picture described above is seriously inadequate, however, because we have ignored conduction effects. This omission is appropriate for protons because the characteristic length scale of the temperature gradient in the relaxation zone greatly exceeds λ_{proton}, the length scale over which viscous and/or conduction effects by protons are important. However the situation is quite different for electrons. From (10.23) or (10.26) we see that the mean free path of particles in a plasma is independent of a particle's mass, hence in ionized hydrogen ($Z = 1$), $\lambda_e = \lambda_{\text{proton}}$, which implies that the thermal diffusivity for electrons $\chi_e \propto \bar{v}_e \lambda_e$ is $\chi_e = (m_p/m_e)^{1/2} \chi_p$ (as noted also in §33). Thus the characteristic length scale over which electron conduction is important is

$$l_e \sim (\chi_e/u_2) \sim (m_p/m_e)^{1/2} (\chi_p/u_2) \sim (m_p/m_e)^{1/2} \lambda_{\text{proton}}, \quad (57.52)$$

that is, l_e is of the same order of thickness as the relaxation zone.

Therefore, thermal conduction by electrons transports energy very efficiently throughout the entire relaxation zone, and thus strongly heats the electrons immediately behind the dissipation zone, while simultaneously promoting a more rapid equilibration of T_2 with T_p because heat is transferred to the postshock electrons by heat conduction at about the same rate as energy is transferred from the protons. Much more important, because the electron velocities behind the shock are roughly a factor of $(m_p/m_e)^{1/2}$ larger than the downstream flow speed, the electrons can overtake the shock front and conduct heat into the upstream material before the shock front arrives. This *conduction precursor* efficiently preheats the upstream electrons, which then transfer some of their excess energy to the upstream ions. The result is a temperature structure like that sketched in Figure 57.8b. (We will find a similar phenomenon, the *radiation precursor*, for shocks in radiating fluids, see §§104 and 105).

Shafranov (**S10**) made detailed numerical calculations of shock structure in ionized hydrogen, including the effects of electron thermal conduction, for a wide range of upstream Mach numbers. In the limit of very strong shocks, $M_1 \gg 1$, he obtains the results listed in Table 57.1, which apply immediately in front of and immediately behind the viscous dissipation zone. Note that the electron temperature is now continuous across the dissipation zone (for which reason such shocks are sometimes called *electron-isothermal* shocks), and has almost achieved its final downstream value already at the shock front. Similarly the protons are preheated to about 15 percent of the downstream temperature, and the postshock proton temperature overshoot is now only about 25 percent of the downstream temperature, the rest of the excess proton energy predicted by (57.49) having been consumed in heating the electrons.

In his numerical work Shafranov found that below a certain critical Mach number electron conduction is sufficient to produce fully dispersed shocks (i.e., all variables continuous across the front). For a plasma of electrons and positive ions of charge Z, Imshennik (**I1**) derived analytical formulae for the critical Mach number

$$(M_1^2)_{\text{crit}} = [\gamma^2 + (3Z+1)\gamma - Z]/\gamma[(3Z+1) - \gamma(Z-1)], \qquad (57.53)$$

Table 57.1. Physical Properties at Shock Front in a Hydrogen Plasma for $M_1 \gg 1$

Quantity	Preshock	Postshock
ρ/ρ_1	1.131	3.526
T_e/T_1	$0.29M_1^2$	$0.29M_1^2$
T_e/T_2	0.928	0.928
T_p/T_1	$1.2+0.05M_1^2$	$0.387M_1^2$
T_p/T_2	$0.16+(3.84/M_1^2)$	1.238

the critical pressure jump

$$(p_2/p_1)_{\text{crit}} = (Z+1)(\gamma+1)/[(Z+1)(\gamma+1)-2Z(\gamma-1)], \qquad (57.54)$$

and the critical volume ratio

$$(\eta_2)_{\text{crit}} = [(Z+1)\gamma(\gamma+1)-Z(\gamma^2-1)]/[(Z+1)\gamma(\gamma+1)-Z(\gamma-1)^2], \qquad (57.55)$$

separating fully dispersed and discontinuous shocks. As $Z \to \infty$, (57.53) to (57.55) reduce to the single-fluid Rayleigh formulae (57.37) to (57.39). For $Z = 1$ shocks are fully dispersed when $M_1^2 \leq \frac{19}{15} \approx 1.125$, $p_2/p_1 \leq \frac{4}{3}$, and $\eta_2 \geq \frac{16}{19} \approx 0.842$. For any $\gamma > 1$ the critical Mach number, or pressure jump, above which we get discontinuous electron-isothermal shocks is a monotone increasing function of Z.

(*c*) *Weakly Ionized Monatomic Gas* If instead of a fully ionized plasma we start with a weakly ionized monatomic gas and generate downstream temperatures of about 10^4 K to 2×10^4 K in the shock, the slowest post-shock relaxation process is ionization of the gas. If the material is originally completely neutral, the first few ionizations behind the shock are produced by atom-atom collisons; this mechanism is slow, and if it were the only ionization mechanism the material would remain neutral for very large distances downstream. However, once a few *seed electrons* have been produced, subsequent ionizations occur efficiently via electron-atom collisions, which are far more effective than atom-atom collisions because the electrons: (1) move much faster than the atoms, hence collide with many more particles per unit time and (2) are charged, hence interact strongly with atoms via the long-range Coulomb potential. (In a multicomponent gas, e.g., a stellar atmosphere, the seed electrons may come from trace elements such as Na, Mg, Al, K, and Ca, which have low ionization potentials and are thus easily ionized while the dominant constituents H and He, which have much higher ionization potentials, remain completely neutral.) Thus the seed electrons rapidly produce yet more electrons and generate an *electron avalanche*, which runs away exponentially until ultimately quenched when the ion density becomes large enough that recombinations can equilibrate against the rate of ionization.

Because the amount of energy required to ionize each atom is typically much larger than the average thermal energy of an electron, only the electrons far out on the Maxwellian tail are effective. After they collisionally ionize an atom these electrons end up with much lower velocities, hence the electron gas is cooled. The tail of the distribution can be replenished in a few electron self-collision times, but the factor limiting the replenishment (hence the growth rate of the electron avalanche) is the rate of energy transfer from the shock-heated atoms to the electrons. Initially this energy exchange is very slow because electron-atom collisions transfer

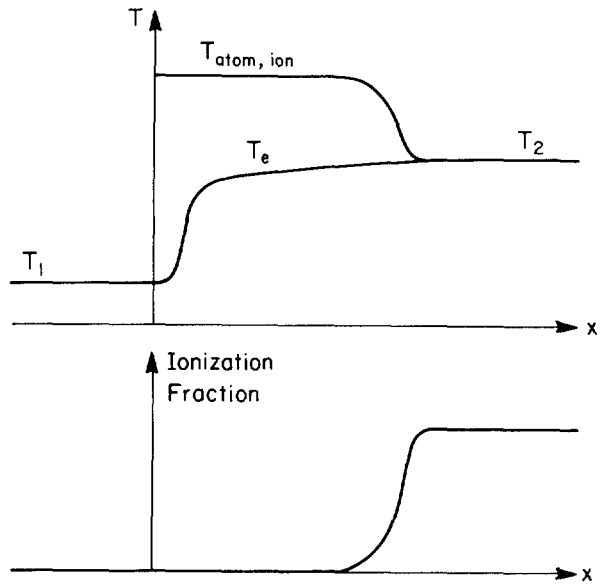

Fig. 57.9 Temperature and ionization-fraction variation as a function of spatial position in a shock in ionizing material.

energy inefficiently. But as the plasma becomes ionized, energy is transferred more rapidly in a two-step process: (1) atoms transfer energy to ions efficiently (because they have equal masses) and (2) the ions transfer energy to the electrons via Coulomb collisions (which are effective because of the long-range potential). Eventually the drain of energy to the electrons cools the atoms and ions, and the temperatures of all three particle species equalize at T_2, as sketched in Figure 57.9. Similarly, the ionization fraction saturates to its equilibrium value.

In most situations of astrophysical interest, radiation plays a more dominant role in determining excitation and ionization within the shock front and relaxation zone than the mechanisms just described. The radiating shock problem is more complex because photon mean free paths usually greatly exceed particle mean free paths. Hence radiation can force a nonlocal coupling of conditions at one point to those at widely separated points and can drive substantial departures from local thermodynamic equilibrium. Examples of such phenomena are described in (**K1**) and (**K2**); we discuss radiating shocks in Chapter 8.

58. Propagation of Weak Shocks

Having considered steady shocks in some detail we turn to the propagation and dissipation of shocks (both single pulses and trains of shocks) in a stratified medium such as a stellar atmosphere. For the present we confine

attention to weak shocks ($m \lesssim 1$ or $\mathscr{M} \lesssim 1$), which can be treated analytically, returning to strong-shock propagation in §60.

A major goal of weak-shock theory is to account for shock-wave dissipation. Dissipation is important because it bleeds energy from shocks and ultimately quenches them. Indeed, from the outset dissipation retards the growth of acoustic disturbances into the nonlinear regime and thus raises the height of shock formation in the atmosphere. Furthermore, once a shock forms, dissipation reduces (or at least retards the growth of) the shock's amplitude, hence extends the range of validity of weak-shock theory. Finally, dissipation provides a basic mechanism for nonradiative heat input into the atmosphere, a matter of great interest in astrophysical calculations.

In constructing the theory we must make several simplifying assumptions. (1) The material is a perfect gas with constant ratio of specific heats γ. We thus neglect ionization effects, which can be an important sink of thermal energy in shock-heated gas. (2) The shocks propagate strictly vertically in an isothermal atmosphere in hydrostatic equilibrium. (This model provides a rough caricature of the temperature-minimum region of the solar atmosphere.) We thus suppress refraction and reflection effects. (3) We ignore the back reaction of the shocks on the ambient medium. This is an important omission because shock heating may significantly alter the thermodynamic state of the atmosphere, and deposition of shock momentum may extend the atmosphere (i.e., increase its scale height). (4) We ignore the gravitational potential energy (buoyancy energy) in, and transported by, the wave. Therefore the theory can be accurate only for waves with frequencies much higher than the acoustic cutoff frequency. (5) Finally, we ignore radiative energy exchange, which is important in astrophysical applications; we return to this aspect of the problem in Chapter 8.

Despite its obvious limitations, weak-shock theory provides useful insight into the physics of shock propagation and has been popular in a wide variety of applications. The theory developed here follows the approach in (**S13**), to which the reader is referred for further details; see also (**B10,** Chaps. 6 and 7), (**B5**), and (**B6**).

PROPAGATION OF N WAVES

Consider the propagation of a small-amplitude, periodic acoustic wave. Because the phase velocity

$$v_p(u) = a_0 + \tfrac{1}{2}(\gamma+1)u \tag{58.1}$$

is largest at wave crests and smallest in the troughs, the wave steepens, the crests eventually overtake the troughs, shock, and produce a propagating **N** wave.

If the velocity profile is initially sinusoidal, $u = \frac{1}{2}u_0 \sin(2\pi z/\Lambda)$, then the crest of the wave overtakes the wave front with a speed $\frac{1}{4}(\gamma+1)u_0$, which is

also the speed with which the wave front overtakes the preceding trough. Therefore the peak and trough coalesce into a vertical front when

$$\tfrac{1}{4}(\gamma+1)\int_0^t u_0(t')\,dt' = \tfrac{1}{4}(\gamma+1)a_0^{-1}\int_0^Z u_0(z)\,dz = \tfrac{1}{4}\Lambda. \tag{58.2}$$

For a uniform medium $u_0(z)\equiv u_0$, hence shocks form after the wave propagates a distance

$$Z = a_0\Lambda/(\gamma+1)u_0. \tag{58.3}$$

In an isothermal atmosphere the velocity amplitude of an acoustic wave scales as $u_0(z)\propto\rho_0^{-1/2}\propto\exp(z/2H)$ where H is the scale height. In this case, (58.2) yields a shock-formation distance

$$Z = 2H\ln\left[\frac{a_0\Lambda}{2(\gamma+1)u_0H}+1\right]. \tag{58.4}$$

Notice that in both cases the distance for shock formation increases with increasing Λ, hence short-period waves steepen into shocks sooner than long-period waves (see Figure 58.1).

After the wave travels a distance Z it has become an N wave with velocity profile

$$u(z) = \tfrac{1}{2}u_0[1-(2z/\Lambda)], \qquad (0\le z\le\Lambda), \tag{58.5}$$

which implies that at a fixed position in the medium the velocity varies as

$$u(t) = \tfrac{1}{2}u_0[1-(2t/\tau)], \qquad (0\le t\le\tau). \tag{58.6}$$

The shock travels with the velocity of the wave crest, namely

$$v_{\text{shock}} = a_0+\tfrac{1}{4}(\gamma+1)u_0 \tag{58.7}$$

hence

$$m = M^2-1\approx\tfrac{1}{2}(\gamma+1)u_0/a_0. \tag{58.8}$$

The energy transported by the wave in one period is

$$E_w = \tau\phi_w = \int_0^\tau (p-p_0)u\,dt \qquad (\text{ergs cm}^{-2}). \tag{58.9}$$

But for a weak shock $p-p_0\approx\gamma p_0u/a_0$ [cf. (48.24a) or (56.44)], hence

$$E_w = (\gamma p_0/a_0)\int_0^\tau u^2\,dt. \tag{58.10}$$

In particular, for an N wave

$$E_w = \tfrac{1}{12}\gamma p_0u_0^2\tau/a_0 = \tfrac{1}{12}\rho_0u_0^2\Lambda = \gamma p_0m^2\Lambda/3(\gamma+1)^2. \tag{58.11}$$

Taking the logarithmic derivative of (58.11) and rearranging, we find an

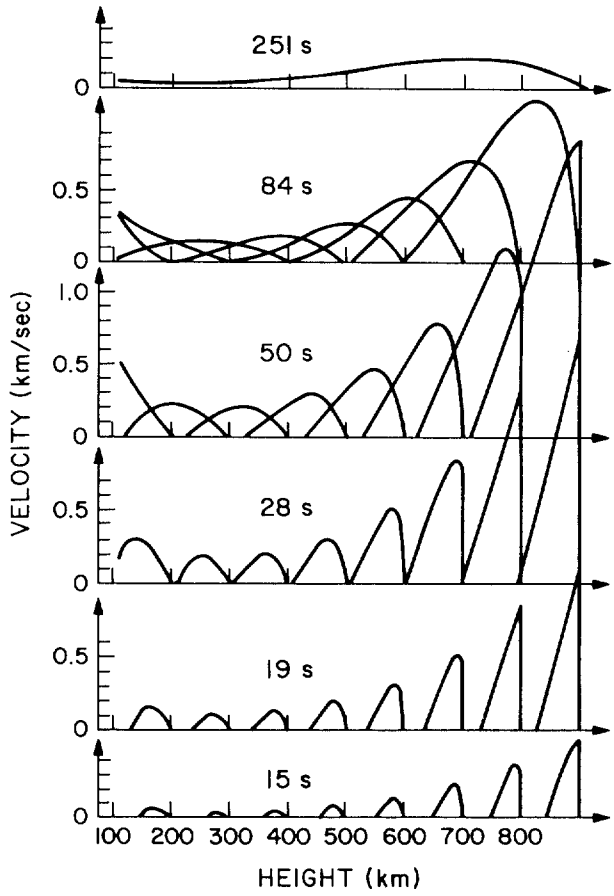

Fig. 58.1 Development of monochromatic acoustic waves into shocks in a model solar atmosphere. From (**U3**), by permission.

equation for the variation of the Mach number with height:

$$\frac{1}{m}\frac{dm}{dz}=\frac{1}{2}\left(\frac{1}{H}+\frac{1}{E_w}\frac{dE_w}{dz}-\frac{1}{\Lambda}\frac{d\Lambda}{dz}\right), \tag{58.12}$$

where H is the pressure scale height of the ambient atmosphere.

We can evaluate (dE_w/dz) by noting, from (56.56), that the amount of energy dissipated, per gram, by the shock is

$$\Delta q = T\,\Delta s = 2\gamma p_0 m^3/3(\gamma+1)^2\rho_0 \qquad (\text{ergs g}^{-1}) \tag{58.13}$$

while the mass flux into the shock front is $\rho_0 v_{\text{shock}}$ (g cm^{-2} s^{-1}), hence the average rate of energy loss by the shock is $(dE_w/dt) = -\rho_0 v_{\text{shock}}\Delta q$ (ergs cm^{-2} s^{-1}). Dividing by the shock velocity we obtain

$$(dE_w/dz) = -2\gamma p_0 m^3/3(\gamma+1)^2. \tag{58.14}$$

Combining (58.11) and (58.14) we see that for an **N** wave

$$(dE_w/dz) = -2mE_w/\Lambda. \tag{58.15}$$

In a weak **N** wave, each pulse peak moves as fast as the one that precedes it, hence the spacing between pulses remains constant, $\Lambda \equiv \Lambda_0$. Using this result and (58.15) in (58.12) we obtain the propagation equation

$$\frac{d}{dz}\left(\frac{1}{m}\right) + \frac{1}{2Hm} = \frac{1}{\Lambda_0}, \tag{58.16}$$

which has the solution

$$\frac{1}{m} = \frac{2H}{\Lambda_0} + \left(\frac{1}{m_0} - \frac{2H}{\Lambda_0}\right) e^{-z/2H}, \tag{58.17}$$

where m_0 measures the Mach number at a convenient reference height $z = 0$. From (58.17) we see that for $z \gg H$

$$m \to \Lambda_0/2H, \tag{58.18}$$

which leads to the remarkable conclusion that in an isothermal atmosphere an **N** wave propagates asymptotically with *constant shape* (i.e., constant wavelength Λ_0 and amplitude m). The tendency for the wave amplitude to grow exponentially as the density decreases is exactly balanced by the increased rate of dissipation, hence damping, resulting from a larger amplitude. This result is fortuitous; we will now show that a single pulse behaves quite differently.

PROPAGATION OF A PULSE

Let us now consider the propagation of a single pulse of total width Λ. Suppose the initial velocity profile is the sinusoid $u = u_0 \sin(\pi z/\Lambda)$. A shock forms when some part of the profile becomes vertical. One can easily show that this condition is first met right at the front of the wave, in a time T given, for a uniform medium, by

$$x = \tfrac{1}{2}(\gamma + 1)u_0 \sin(\pi x/\Lambda)T. \tag{58.19}$$

For $x \ll \Lambda$, (58.19) implies that the shock forms in a distance

$$Z = a_0 T = 2a_0\Lambda/\pi(\gamma + 1)u_0, \tag{58.20}$$

which is a factor of $(2/\pi)$ smaller than for an **N** wave. We can apply the same factor to (58.4) to estimate the distance for shock formation in a stratified atmosphere.

Once the pulse has steepened into a shock the velocity profile becomes

$$u(z) = u_0[1 - (z/\Lambda)], \qquad (0 \le z \le \Lambda), \tag{58.21}$$

so that at a fixed location the velocity varies in time as

$$u(t) = u_0[1 - (t/\tau)], \qquad (0 \le t \le \tau). \tag{58.22}$$

Using (58.22) in (58.10) we find that the energy transported by a pulse is

$$E_w = \tfrac{1}{3}\rho_0 u_0^2 \Lambda = 4\gamma p_0 m^2 \Lambda/3(\gamma+1)^2, \tag{58.23}$$

that is, exactly four times the energy in a single period of an N wave of the same wavelength and total velocity jump at the front. The N wave transports less energy because the downward motion in the tail of the wave partially cancels the effect of the upward motion at the head of the wave.

Unlike an N wave, a pulse changes shape as it propagates because the head of the pulse, traveling with speed Ma_0, always outruns the tail of the pulse, traveling with speed a_0. Thus

$$(d\Lambda/dt) = (M-1)a_0 \approx \tfrac{1}{2}ma_0, \tag{58.24}$$

or

$$(d\Lambda/dz) = \tfrac{1}{2}m. \tag{58.25}$$

Using (58.23) in (58.14) we have

$$(dE_w/dz) = -mE_w/2\Lambda, \tag{58.26}$$

and using (58.26) and (58.25) in (58.12) we have

$$m^{-1}(dm/dz) = (1/2H) - (m/2\Lambda). \tag{58.27}$$

Equations (58.26) and (58.27) completely describe the propagation of the pulse.

Using (58.25) to eliminate z from (58.27) and solving we find

$$m = (\Lambda/2H) - (K/\Lambda), \tag{58.28}$$

where K is a constant. Substituting (58.28) into (58.25) we then have

$$(d\Lambda/dz) = (\Lambda/4H) - (K/2\Lambda) \tag{58.29}$$

which has the solution

$$\Lambda^2 = K'e^{z/2H} + 2HK. \tag{58.30}$$

Evaluating the constants so that $\Lambda = \Lambda_0$ and $m = m_0$ at $z = 0$, we find

$$\Lambda = \Lambda_0[(2Hm_0/\Lambda_0)(e^{z/2H}-1)+1]^{1/2} \tag{58.31}$$

and

$$m = m_0 e^{z/2H}[(2Hm_0/\Lambda_0)(e^{z/2H}-1)+1]^{-1/2}. \tag{58.32}$$

From (58.32) we see that for $(z/H) \gg 1$,

$$m \propto e^{z/4H} \propto \rho_0^{-1/4}, \tag{58.33}$$

in contrast to small-amplitude waves, for which the velocity amplitude increases as $\rho_0^{-1/2}$; this slower growth is attributable to a loss of wave energy via dissipation as the wave propagates.

APPLICATIONS TO THE SOLAR CHROMOSPHERE

Following the early work of Schatzman (**S3**), various forms of weak shock theory have been applied to the propagation and dissipation of shocks in

the solar chromosphere by several authors [see e.g., (**O1**), (**J1**), (**U1**), (**U2**), (**U3**), and the summary in (**B10,** Chaps. 6 and 7)]. Most of this work makes allowance for radiative energy losses in addition to viscous dissipation. Some typical results showing the steepening of monochromatic acoustic waves are displayed in Figure 58.1. The waves are drawn whenever the wavefront reaches a multiple of 100 km. One sees clearly that short-period waves steepen into shocks sooner than long-period waves, as predicted by (58.4). The waves are heavily damped by radiative losses at heights below about 500 km, but develop into shocks quickly thereafter. The calculations cited suggest that the observed radiative energy loss by the chromosphere can be sustained by the dissipation of weak, short-period shocks with $m \leqslant 0.4$ to 0.5, for which weak-shock theory should be valid.

A critical assessment of the accuracy of weak-shock theory can be made by comparing its predictions with the results of full nonlinear calculations, as was done by Stein and Schwartz (**S13**), (**S14**). They find that the theory gives reasonable results so long as the wave period τ is much shorter than the acoustic cutoff period $\tau_a \approx 200$ s; as $\tau \to \tau_a$ the contribution of gravitational terms omitted from the theory described above become increasingly important, and the quality of the results deteriorates rapidly.

Stein and Schwartz also found that weak-shock theory always tends to overestimate the rate of growth of m with height; the effect is small for $\tau \leqslant 25$ s, but is major for $\tau \geqslant 50$ s. A consequence of this too-rapid growth of m is that the dissipation rate predicted by weak-shock theory is too large; for $\tau \sim 100$ s it is in error by almost a factor of 10 (**S14**), and for $\tau \sim 400$ s the error is several orders of magnitude. Indeed, even for the same Mach number, weak-shock theory predicts a larger dissipation rate than the nonlinear theory, by about 10 percent for $m = 0.28$ and about 50 percent for $m = 3$ ($M = 2$); the latter value should not be surprising, however, because weak-shock theory explicitly assumes that $m \ll 1$. The calculations show that almost 90 percent of the shock energy is deposited of heights less than 2000 km, and the damping length for short-period waves is only 500 km. Furthermore, it is found that it is essential to account for ionization effects and radiative losses in calculating the shock-induced temperature rise of the material. The temperature increase calculated by assuming that the gas is adiabatic is a factor of three too large.

CRITIQUE

Despite its frequent application in astrophysics, it is clear that weak-shock theory has only limited validity; an interesting critique of this approach is given in (**C10**). Besides being only linear, the theory contains numerous other approximations, which must be invoked in order to obtain analytical results. Ultimately it is based on Brinkley-Kirkwood theory (**B11**), which makes very simplified assumptions about the thermodynamic path followed by the material through and behind the shock, and further assumes that the postshock flow can be described by a similarity solution (cf. §60).

Most astrophysical calculations using weak-shock theory have been made for one-dimensional, infinite, monochromatic wave trains. This picture is a gross oversimplification for flows (e.g., in the solar chromosphere) containing a field of large-amplitude waves having different periods and directions of propagation, and neglects completely the possibility of wave-wave interactions (e.g., when a shock overruns another shock or a rarefaction). Moreover the theory formulated above omits gravitational energy terms, which are very important for waves with $\tau \approx \tau_{ac}$; allowance for such terms has been made in a theory developed by Saito (**S1**), which is discussed briefly in (**B10,** p. 296 and p. 343). The most serious flaw in the theory discussed thus far is the omission of radiation losses; some 60 to 80 percent of the energy in short-period waves is lost to radiative damping. We remedy this defect in Chapter 8.

The difficulties described above demonstrate the need for a more powerful method. We therefore turn to numerical techniques, which not only can handle the full nonlinear equations, but are also versatile and flexible enough to (1) permit a detailed description of the microphysics of the gas, (2) allow for structural complexities in the ambient medium (e.g., temperature and ionization gradients), (3) allow for wave-wave interactions and the back reaction of the waves on the background atmosphere, (4) be generalized easily to include the transport of energy and momentum by radiation, and (5) account for radiation-induced departures from local thermodynamic equilibrium.

59. Numerical Methods

One of the most effective methods for solving the equations of hydrodynamics is to replace the original differential equations by a set of *finite difference equations* that determine the physical properties of the fluid on discrete space and time meshes. Given suitable initial and boundary conditions we follow the evolution of the fluid by solving this discrete algebraic system at successive timesteps. Two problems to be faced are that (1) we must assure that the finite difference equations are numerically *stable* and (2) an efficient scheme must be found for handling shocks, which can produce discontinuities in the solution at or between mesh points.

In this section we do not attempt to discuss state-of-the-art methods, but confine our attention to one basic technique that has been successful in many applications; this example provides a good introduction to the vast literature on the subject. A fundamental reference on the numerical solution of fluid-flow problems is the classic book by Richtmyer and Morton (**R4**). More powerful modern techniques are discussed in (**W7**).

NUMERICAL SIMULATION OF ACOUSTIC WAVES

To obtain insight we start with a simple physical problem: the propagation of adiabatic acoustic waves in a perfect gas with no external forces. In

planar geometry the Lagrangean dynamical equations are

$$(Dv/Dt) = -(\partial p/\partial m), \tag{59.1}$$

$$(Dx/Dt) = v, \tag{59.2}$$

$$V \equiv 1/\rho = (\partial x/\partial m), \tag{59.3}$$

and

$$(De/Dt) = -p(DV/Dt). \tag{59.4}$$

The system is closed by adjoining an equation of state of the form $p = p(\rho, e)$.

We now discretize the system in both space (i.e., mass) and time. First we divide the medium into a set of *mass cells* by choosing a set of values $\{m_j\}$ giving the Lagrangean coordinate of *cell surfaces*, as sketched in Figure 59.1a. The cell surfaces are located at spatial positions $\{x_j\}$, which

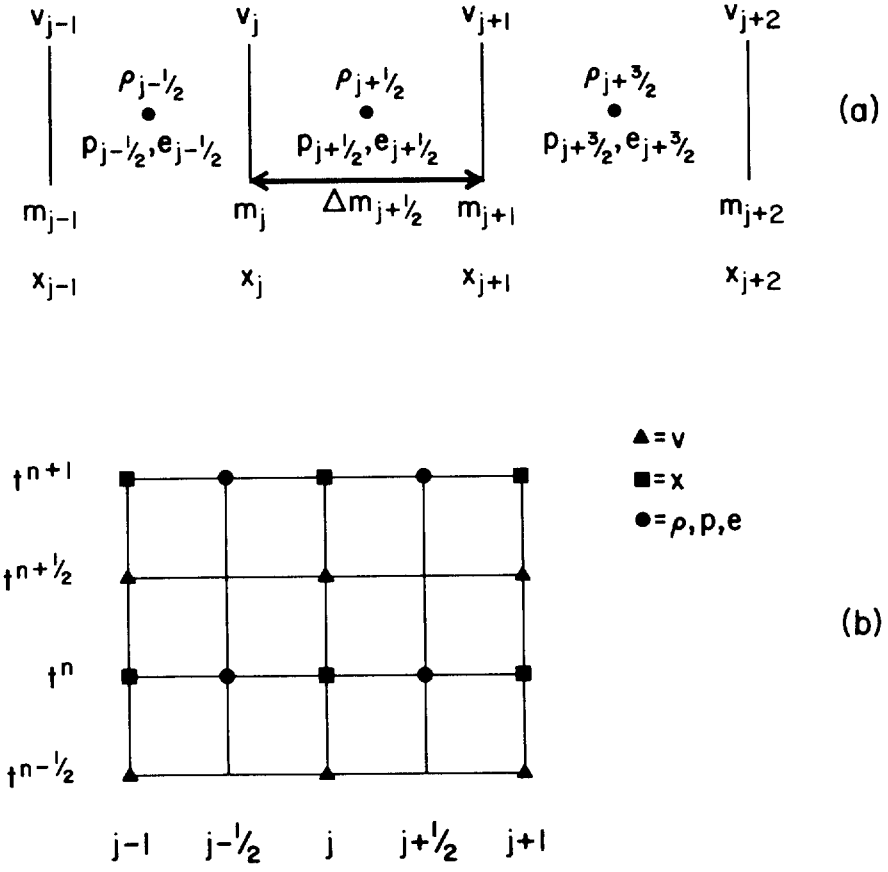

Fig. 59.1 (a) Centering of physical variables on Lagrangean mesh. (b) Spacetime centering of physical variables on Lagrangean mesh.

vary in time as the fluid moves; in contrast the Lagrangean coordinates $\{m_j\}$ remain fixed in time. We assign all thermodynamic properties to *cell centers*, and denote them by half-integer subscripts: $\rho_{j+(1/2)}$, $p_{j+(1/2)}$, $e_{j+(1/2)}$. Similarly, the (fixed) mass contained in a cell is $\Delta m_{j+(1/2)} = m_{j+1} - m_j$. In order to track the motion of the Lagrangean grid $\{m_j\}$ through physical space (x), velocities are assigned to cell surfaces (e.g., v_j).

Next we choose a discrete set of times $\{t^n\}$ at which the thermodynamic properties of the fluid are to be determined. We label variables with superscripts corresponding to their location in time [e.g., $\rho^n_{j+(1/2)}$ for ρ at the center of cell $(j, j+1)$ at time t^n]. Noting that we wish to know interface positions at the same time levels as the thermodynamic variables (because the density of a cell of fixed mass is determined by its volume), it is clear that velocities should be time centered midway between these levels [i.e., at $t^{n+(1/2)} \equiv \frac{1}{2}(t^n + t^{n+1})$]. The location of the variables in spacetime is sketched in Figure 59.1b.

The differential equations (59.1) to (59.4) are then replaced by

$$[v_j^{n+(1/2)} - v_j^{n-(1/2)}]/\Delta t^n = -[p^n_{j+(1/2)} - p^n_{j-(1/2)}]/\Delta m_j, \tag{59.5}$$

$$(x_j^{n+1} - x_j^n)/\Delta t^{n+(1/2)} = v_i^{n+(1/2)}, \tag{59.6}$$

$$V^{n+1}_{j+(1/2)} = 1/\rho^{n+1}_{j+(1/2)} = (x^{n+1}_{j+1} - x_j^{n+1})/\Delta m_{j+(1/2)}, \tag{59.7}$$

and

$$e^{n+1}_{j+(1/2)} - e^n_{j+(1/2)} = -p^n_{j+(1/2)}[V^{n+1}_{j+(1/2)} - V^n_{j+(1/2)}]. \tag{59.8}$$

Here $\Delta t^{n+(1/2)} \equiv t^{n+1} - t^n$, $\Delta t^n \equiv \frac{1}{2}[\Delta t^{n-(1/2)} + \Delta t^{n+(1/2)}]$, and $\Delta m_j \equiv \frac{1}{2}[\Delta m_{j-(1/2)} + \Delta m_{j+(1/2)}]$. Given $\{v_j\}^{n-(1/2)}$, $\{x_i\}^n$, $\{\rho_{j+(1/2)}\}^n$, $\{p_{j+(1/2)}\}^n$ and $\{e_{j+(1/2)}\}^n$, either from initial conditions or from a preceding integration step, we solve (59.5) to (59.8) in the order indicated to obtain $\{v_i\}^{n+(1/2)}$, $\{x_j\}^{n+1}$, $\{\rho_{j+(1/2)}\}^{n+1}$, $\{e_{j+(1/2)}\}^{n+1}$, and finally $\{p_{j+(1/2)}\}^{n+1} = \{p[V^{n+1}_{j+(1/2)}, e^{n+1}_{j+(1/2)}]\}$.

As written, the system is *explicit*, that is, each variable is determined by direct evaluation using information already available. But it should be noted that the accuracy of (59.8) is impaired because p in (59.2) should really be evaluated at the *midpoint* of the integration interval, say as $\frac{1}{2}[p^n_{j+(1/2)} + p^{n+1}_{j+(1/2)}]$, instead of at one end (t^n). If this is done, then *two* unknowns, $p^{n+1}_{j+(1/2)}$ and $e^{n+1}_{j+(1/2)}$, appear in (59.8), which must now be solved simultaneously with the equation of state. In the special case of a perfect gas with constant specific heats, we can use the relation $p = (\gamma - 1)\rho e$ to eliminate either p or e and thus recover an explicit equation for either $p^{n+1}_{j+(1/2)}$ or $e^{n+1}_{j+(1/2)}$. But in general, when the physics of the gas is more complicated (e.g., by ionization), some kind of iteration procedure must be used to solve the coupled equations. As we will see below, in practice it is almost always necessary to make this additional effort anyway to assure numerical stability and energy conservation.

THE PROBLEM OF NUMERICAL STABILITY

If one attempts to solve (59.5) to (59.8) numerically, starting from smooth initial conditions, one finds that for sufficiently small timesteps Δt the solution remains smooth, and provides a good approximation to the analytical solution of (59.1) to (59.4). If, however, Δt is greater than some critical value, then starting from the same initial data the numerical solution quickly develops unphysical oscillations that rapidly grow and eventually swamp the true solution. This behavior is the result of a *numerical instability* of the finite difference equations, which, under certain circumstances, can allow small errors (e.g., roundoff) in the calculation to become progressively amplified until they dominate the computation. Experience has shown that in solving complex physical problems it is essential to perform a *stability analysis* of the difference equations used to model the system.

One of the basic tools of stability theory for initial-value problems is the *von Neumann local stability analysis*, which exploits the fact that a difference equation

$$y_j^n = L(x, t, y), \tag{59.9}$$

where L is a *linear difference operator* on a *uniform* spacetime mesh ($x_j = j\,\Delta x$, $t^n = n\,\Delta t$) can be solved exactly by a Fourier series of the general form

$$y_j^n = \sum_k A_k e^{ikj\,\Delta x} \xi_k^n. \tag{59.10}$$

In (59.10) the coefficients are determined by the imposed initial and boundary conditions. Each harmonic grows or decays independently of the others; ξ_k is the *amplification factor* (or *growth factor*) for the kth Fourier component over the time interval Δt. If the solution is to be stable, we must guarantee that no harmonic can become unbounded. Thus a *necessary* condition for stability is that the modulus $\|\xi_k\| \leq 1$ for all k. If this condition is met, no harmonic will be amplified; if it is violated, some harmonic can grow without limit and the solution becomes unstable.

As an example, consider the equation

$$(\partial\rho/\partial t) + v_0(\partial\rho/\partial x) = 0, \tag{59.11}$$

which describes the advection of material by a constant velocity field; we assume that $v_0 > 0$. The general solution of (59.11) is $\rho = f(x - v_0 t)$ where f is an arbitrary function fixed by initial conditions; this solution is a traveling wave in which the original distribution of the material is displaced to the right without an alteration in form. Suppose we represent (59.11) by the difference equation

$$(\rho_j^{n+1} - \rho_j^n)/\Delta t = -v_0(\rho_{j+1}^n - \rho_{j-1}^n)/2\,\Delta x, \tag{59.12}$$

which is centered in space and explicit in time. Taking a trial solution of

the form

$$\rho_j^n(k) = \xi_k^n e^{ikj\,\Delta x}, \tag{59.13}$$

we find that (59.12) implies that

$$\xi_k = 1 - i\alpha \sin k\Delta x, \tag{59.14}$$

where $\alpha \equiv v_0\,\Delta t/\Delta x$. The modulus of ξ_k is thus

$$\|\xi_k\| = (1+\alpha^2 \sin^2 k\Delta x)^{1/2} \geq 1. \tag{59.15}$$

Hence the difference scheme is *unconditionally unstable*, that is, it is unstable for *any* finite Δt, no matter how small!

Alternatively, suppose we represent (59.11) by

$$(\rho_j^{n+1} - \rho_j^{n-1})/2\,\Delta t = -v_0(\rho_{j+1}^n - \rho_{j-1}^n)/2\,\Delta x, \tag{59.16}$$

which is centered in space and a *leapfrog* in time. Substituting (59.13) we now find

$$\xi_k^2 + 2i\beta\xi_k - 1 = 0, \tag{59.17}$$

where $\beta \equiv \alpha \sin k\Delta x$. Equation (59.17) has the solution

$$\xi_k = -i\beta \pm (1-\beta^2)^{1/2}. \tag{59.18}$$

When $\beta^2 > 1$, $\|\xi_k\| > 1$; when $\beta^2 \leq 1$, $\|\xi_k\| = 1$. Thus (59.16) is *conditionally stable*, that is, it is stable provided that $|(v_0\,\Delta t/\Delta x)\sin k\Delta x| \leq 1$ for all k, which implies that we must choose Δt such that

$$v_0\,\Delta t/\Delta x \leq 1. \tag{59.19}$$

Equation (59.19) is an example of the famous *Courant condition* (**C7**) which, in physical terms, states that the timestep must be sufficiently small that the wave cannot propagate over more than one spatial cell Δx in the interval Δt. We will encounter the Courant condition again in many other contexts.

Another stable difference-equation representation of (59.11) is obtained by using *upstream* (or *upwind*) *differencing*, writing

$$(\rho_j^{n+1} - \rho_j^n)/\Delta t = -v_0(\rho_j^n - \rho_{j-1}^n)/\Delta x, \tag{59.20}$$

which is explicit. Physically, (59.20) recognizes that material flows into cell j from cell $j-1$. Substituting (59.13) and calculating the amplification factor, we find

$$\|\xi_k\| = 1 + 2\alpha(1-\alpha)(\cos k\Delta x - 1). \tag{59.21}$$

For $0 \leq \alpha \leq 1$, $\|\xi_k\|$ reaches its maximum when $\cos k\Delta x = 1$, in which case $\|\xi_k\| = 1$, and the difference equation is stable. For $\alpha > 1$, $\|\xi_k\|$ is maximized when $\cos k\Delta x = -1$, in which case $\|\xi_k\| = (2\alpha - 1)^2 > 1$. Hence (59.20) is stable only if the Courant condition is satisfied. We thus see that a given differential equation may have more than one stable difference-equation representation, as well as unstable representations.

The above analysis applies to a single difference equation. In practice, we are more often interested in the stability of a system of M equations of the form

$$\mathbf{y}^n = \mathbf{L}(x, t, \mathbf{y}), \tag{59.22}$$

where $\mathbf{L}$ is a linear difference operator coupling $\mathbf{y}^n$ and $\mathbf{y}^{n-1}$, and $\mathbf{y}^n(x)$ denotes the numerical approximation, obtained from the difference equations, to the true solution $\mathbf{y}(x, n\Delta t) = [y_1(x, n\Delta t), \ldots, y_M(x, n\Delta t)]$ of the differential equations. Representing the solution by a Fourier series with amplitudes $\mathbf{a}_k$, one can show that (59.22) is equivalent to

$$\mathbf{a}_k^n = \mathbf{G}(\Delta t, k)\mathbf{a}_k^{n-1} \tag{59.23}$$

where $\mathbf{G}(\Delta t, k)$ is the $M \times M$ *amplification matrix* for the kth harmonic. The stability of the system of difference equations clearly depends on the behavior of $[\mathbf{G}(\Delta t, k)]^n$ for $0 \le n\Delta t \le T$, and intuitively it is obvious that we can achieve stability only if there exists some number τ such that for $(0 \le \Delta t \le \tau)$ and $(0 \le n\Delta t \le T)$ the matrices $[\mathbf{G}(\Delta t, k)]^n$ are *uniformly bounded* for all k, in the sense defined below.

We define the *bound* of a matrix $\mathbf{F}$ to be

$$\|\mathbf{F}\| \equiv \max_{|\mathbf{v}|=1} |\mathbf{F}\mathbf{v}| = \max_{|\mathbf{v}|\neq 0} (|\mathbf{F}\mathbf{v}|/|\mathbf{v}|); \tag{59.24}$$

here $|\mathbf{v}|$ is the usual Euclidian norm of $\mathbf{v}$, that is, $|\mathbf{v}| = (v_1^2 + \ldots + v_M^2)^{1/2}$. Thus the bound of $\mathbf{F}$ is the maximum norm of the vectors resulting from the operation of a transformation $\mathbf{F}$ on the M-dimensional vector space ($\mathbf{v}, \mathbf{w}, \ldots$ etc.). If the eigenvalues of $\mathbf{F}$ are $\lambda_1, \ldots, \lambda_M$, we define the *spectral radius* R of $\mathbf{F}$ to be $\max |\lambda_i|$, $(i = 1, \ldots, M)$. It is easy to see that $\|\mathbf{F}\| \ge R$ because the maximum value of $|\mathbf{F}\mathbf{v}|/|\mathbf{v}|$ cannot be less than the value obtained when $\mathbf{v}$ is the eigenvector corresponding to R, and may exceed this value if we can choose $\mathbf{v}$ astutely (e.g., as a linear combination of eigenvectors) such that $|\mathbf{v}|$ becomes very small while $|\mathbf{F}\mathbf{v}|$ remains of order R.

The spectral radius of $\mathbf{F}^n$ is R^n because the eigenvalues of $\mathbf{F}^n$ are $\lambda_1^n, \ldots, \lambda_M^n$. Moreover

$$\|\mathbf{F}^2\| = \max_{|\mathbf{v}|\neq 0} \frac{|\mathbf{F}(\mathbf{F}\mathbf{v})|}{|\mathbf{v}|} = \max_{|\mathbf{v}|\neq 0} \frac{|\mathbf{F}(\mathbf{F}\mathbf{v})|}{|\mathbf{F}\mathbf{v}|} \frac{|\mathbf{F}\mathbf{v}|}{|\mathbf{v}|} \le \max_{\substack{|\mathbf{v}|\neq 0 \\ |\mathbf{w}|\neq 0}} \frac{|\mathbf{F}\mathbf{w}|}{|\mathbf{w}|} \frac{|\mathbf{F}\mathbf{v}|}{|\mathbf{v}|} = \|\mathbf{F}\|^2, \tag{59.25}$$

therefore $R^2 \le \|\mathbf{F}^2\| \le \|\mathbf{F}\|^2$, hence by induction $R^n \le \|\mathbf{F}^n\| \le \|\mathbf{F}\|^n$. In particular if $R(\Delta t, k)$ is the spectral radius of $\mathbf{G}(\Delta t, k)$, it follows from (59.25) that

$$[R(\Delta t, k)]^n \le \|[\mathbf{G}(\Delta t, k)]^n\| \le \|\mathbf{G}(\Delta t, k)\|^n, \tag{59.26}$$

hence a *necessary* condition for stability is that there exist numbers τ and C_1 such that for $(0 \le \tau \le \Delta t)$, $(0 \le n\Delta t \le T)$,

$$[R(\Delta t, k)]^n \le C_1 \tag{59.27}$$

for all k.

Without loss of generality we can assume $C_1 \geq 1$, hence $R(\Delta t, k) \leq C_1^{1/n}$, and in particular

$$R(\Delta t, k) \leq C_1^{\Delta t/T} \tag{59.28}$$

because $n \leq T/\Delta t$. For $0 \leq \Delta t \leq \tau$, the exponential $C_1^{\Delta t/T}$ can always be bounded from above by an expression of the form $(1 + C_2 \Delta t)$. Hence, we find the *von Neumann necessary condition for stability* is that if $\lambda_1, \ldots, \lambda_M$ are the eigenvalues of the amplification matrix $\mathbf{G}(\Delta t, x)$, then

$$|\lambda_i| \leq 1 + O(\Delta t) \tag{59.29}$$

for all $i = 1, \ldots, M$, for $0 \leq \Delta t \leq \tau$, and for all k.

Equation (59.29) merits two comments. (1) In our earlier analysis we required that the modulus of the amplification factors $\|\xi_k\|$ all be less than unity. The present result is less restrictive, and allows for the possibility of a legitimate exponential growth of the solution. (2) Although (59.29) is a *necessary* condition for stability, it is not, in general, *sufficient*. Furthermore, it is a purely *local* criterion and does not account for boundary conditions. In practice (59.29) is sometimes found to be both necessary *and* sufficient, but in general the derivation of sufficient stability criteria for a given system requires a much more extensive (and difficult) analysis; see (**R4**).

To illustrate the von Neumann analysis, consider the system (59.1) to (59.4). In order to simplify the algebra, rewrite (59.1) as

$$(Dv/Dt) = -(\partial p/\partial \rho)_s (\partial \rho/\partial m) = (a^2/V^2)(\partial V/\partial m), \tag{59.30}$$

where a is the adiabatic sound speed. Difference (59.30) as

$$v_j^{n+(1/2)} - v_j^{n-(1/2)} = (\Delta t/\Delta m)(a^2/V^2)_j^n [V_{j+(1/2)}^n - V_{j-(1/2)}^n], \tag{59.31}$$

and couple this equation to the time difference of (59.7), namely

$$V_{j+(1/2)}^{n+1} - V_{j+(1/2)}^n = (\Delta t/\Delta m)[v_{j+1}^{n+(1/2)} - v_j^{n+(1/2)}]. \tag{59.32}$$

Now examine the growth of the kth Fourier component, taking trial solutions

$$v_j^{n+(1/2)}(k) = (A_k)^{n+(1/2)} e^{ikj\,\Delta m} \tag{59.33}$$

and

$$V_{j+(1/2)}^n(k) = (B_k)^n e^{ik[j+(1/2)]\Delta m}. \tag{59.34}$$

One finds that the amplification matrix is

$$\mathbf{G}(\Delta t, k) = \begin{pmatrix} 1 & i\beta^2/\alpha \\ i\alpha & 1 - \beta^2 \end{pmatrix}, \tag{59.35}$$

where $\alpha \equiv 2(\Delta t/\Delta m) \sin(\frac{1}{2} k \Delta m)$ and $\beta \equiv (a/V)\alpha$. The eigenvalues of $\mathbf{G}$ are

$$\lambda = \tfrac{1}{2}\{(2 - \beta^2) \pm [(2 - \beta^2)^2 - 4]^{1/2}\}. \tag{59.36}$$

If $\beta^2 \leq 4$, λ lies on the unit circle; if $\beta^2 > 4$, $\|\lambda\| > 1$. Thus for stability we must

have $-1 \leq (\rho a\, \Delta t/\Delta m) \sin(\frac{1}{2}k\Delta m) \leq 1$, which is guaranteed if

$$\Delta t \leq (\Delta m/\rho)/a = \Delta x/a, \tag{59.37}$$

that is, if the Courant condition is satisfied. In practice we choose the smallest value of Δt found from (59.37) for the entire mesh.

IMPLICATIONS OF SHOCK DEVELOPMENT

As we saw in §55, nonlinear effects in wave propagation inevitably lead to shock formation. We further saw in §57 that the thickness of a shock is of the order of a few particle mean free paths, and indeed that in an ideal fluid the shock is a mathematical discontinuity. Once shocks form, the differential equations governing the flow must be supplemented by jump conditions, in effect internal boundary conditions within the flow, in order to obtain a unique solution.

If we were to attempt to simulate a flow containing shocks using the Lagrangean difference equations written above, we would immediately encounter severe difficulties. First, these equations contain no dissipative terms, and therefore cannot account for the entropy increase produced by a shock; consequently they will not yield even approximately the right answers behind a shock. Second, even if we were to include dissipative terms using realistic values of the molecular viscosity and thermal conductivity of the gas, the shock thickness would generally be several orders of magnitude smaller than the spacing of grid points on which the difference equations are solved. A brute-force attempt to reduce the grid size to a mean free path is doomed from the outset because if we reduce $\Delta x \rightarrow \Delta x/k$, it follows from the Courant condition that we must also reduce the timestep $\Delta t \rightarrow \Delta t/k$, hence the total computing effort needed to follow the flow for a definite time rises as k^2, which rapidly becomes prohibitive.

One option is to use the *method of characteristics* to follow discontinuities, and then impose the Rankine-Hugoniot relations to do *shock fitting*. Although this method has been highly developed [see e.g., (**H16**)], it is relatively complex and cumbersome because generally the shock is not at a gridpoint, and an iterative solution of a moderately complicated set of nonlinear equations is required to locate it before the jump conditions can be applied [see (**H16**), (**R4,** 304–311)]. While shock fitting is relatively easy to apply when shocks propagate into undisturbed fluid or at least occur regularly in a train (**U5**), (**U6**), it becomes harder to use when, say, shocks overrun one another, or one shock collides with another propagating in the opposite direction. Moreover, because shocks can develop spontaneously anywhere in a flow, one must also develop strategies for *detecting* when shocks have formed.

To overcome these difficulties, von Neumann and Richtmyer (**V4**) devised a scheme that handles shocks *automatically*, wherever and whenever they arise. The essence of the method is to introduce into the difference equations an artificial dissipative process that models the real dissipation

mechanisms in a gas and that gives the correct entropy jump (hence the correct physical properties in the postshock flow), but which smears the shock over a few mesh points in the difference-equation grid, instead of leaving it unresolved on a subgrid level. The shock discontinuity is thus replaced by a transition layer within which the fluid properties vary rapidly but not discontinuously, across which the basic conservation relations are satisfied, and whose thickness can be adjusted to match the grid size, which is chosen according to the physical requirements of the problem. The difference equations then apply everywhere, and the computations proceed completely automatically, without shock fitting.

THE VON NEUMANN–RICHTMYER ARTIFICIAL-VISCOSITY METHOD

We showed in §57 that both viscosity and thermal conduction produce entropy in a shock. We found that viscosity yields a smooth variation of all quantities in shocks of all strengths, whereas thermal conduction yields a smooth transition only for shocks below a certain strength, while stronger inviscid conducting shocks sustain discontinuities in density and pressure. As smoothness of the numerical solution is of paramount importance, we choose viscosity as our dissipation mechanism.

In §§26 and 27 we saw that the momentum and energy equations for one-dimensional planar viscous flow can be written

$$\rho(Dv/Dt) = f - [\partial(P+Q)/\partial z] \tag{59.38}$$

and

$$(De/Dt) + (p+Q)[D(1/\rho)/Dt] = \dot{q}, \tag{59.39}$$

where $\dot{q}$ is the energy input per unit mass from "external" sources (e.g., radiation or thermonuclear reactions), and Q is an equivalent viscous pressure

$$Q = -\tfrac{4}{3}\mu'(\partial v/\partial z). \tag{59.40}$$

Here $\mu' = \mu + \frac{3}{4}\zeta$ is the effective viscosity. To handle shocks in the difference equations we propose to use an *artificial viscosity* (or *pseudoviscosity*), choosing a suitable form for Q. One option would be to use (59.40) with a large constant value for μ', chosen such that the implied molecular mean free path would be of the same order as the grid spacing Δz. This approach is not satisfactory, however, for as we saw in §57 the thickness of a shock, for a given λ, is inversely proportional to its strength, hence we would obtain sharp strong shocks, but weak shocks would be spread over many gridpoints. Moreover such a large viscosity would spuriously reduce the Reynolds number of the flow in regions devoid of shocks, and would therefore seriously degrade the accuracy of the overall solution.

Von Neumann and Richtmyer realized that these problems could be overcome by using a *nonlinear* artificial viscosity that is large in shocks but very small elsewhere. In particular, they suggested using a Q that is

quadratic in the rate of shear, and adopted

$$Q = \begin{cases} \frac{4}{3}\rho l^2(\partial v/\partial z)^2 & \text{for} \quad (\partial v/\partial z) < 0, \quad (59.41a) \\ 0 & \text{for} \quad (\partial v/\partial z) \geq 0. \quad (59.41b) \end{cases}$$

Recalling that μ has dimensions (dyne s cm^{-2}) = (g cm^{-3})(cm^2 s^{-1}) one sees that l has the dimensions of length. Typically l is chosen to be some small multiple of the grid spacing Δz.

We can also view (59.41) as a viscous pressure that is linear in the rate of shear

$$Q = -\tfrac{4}{3}\mu_Q(\partial v/\partial z), \quad (59.42)$$

with an artificial viscosity coefficient that is proportional to the rate of compression of the fluid:

$$\mu_Q = \begin{cases} l^2(D\rho/Dt) & \text{for} \quad (D\rho/Dt) > 0, \quad (59.43a) \\ 0 & \text{for} \quad (D\rho/Dt) \leq 0. \quad (59.43b) \end{cases}$$

The relevance of this interpretation will emerge below.

The artificial viscosity given by (59.41) to (59.43): (1) comes into action only when the gas is compressed (both a prerequisite and a characteristic property of shock formation) and (2) is very small or zero in regions away from shocks. Both of these properties are highly desirable.

From the analysis of §57 it follows that (59.38) to (59.40) will lead to the Rankine-Hugoniot relations regardless of the precise analytical form of Q, provided only that $Q \to 0$ where velocity gradients vanish in the upstream and downstream flow far from a shock front. As the von Neumann-Richtmyer pseudoviscosity manifestly meets this requirement, we are assured that it will produce the correct jump in entropy (and all other variables) across the shock, as well as the correct shock propagation speed. [These attributes can also be verified by direct analysis (**V4**), (**R4,** 314–316).] A large body of computational work has demonstrated that the von Neumann-Richtmyer method gives good results as long as the resulting shock thickness, typically 3 to 4 Δz, is not too coarse to permit an accurate representation of other physical processes of interest (e.g., radiation transport).

THE IMPLICATIONS OF DIFFUSIVE ENERGY TRANSPORT

We have thus far assumed that the gas is essentially adiabatic and have ignored energy transport by diffusion processes such as thermal conduction (or radiation diffusion). However, we know that these processes do occur in a real fluid, and we should inquire into their consequences for the numerical stability of finite-difference representations of the energy equation.

To gain insight, consider a simple linear heat-conduction problem as posed by the parabolic equation

$$(\partial T/\partial t) = \sigma(\partial^2 T/\partial x^2) \quad (59.44)$$

along with suitable initial and boundary conditions. Suppose we represent (59.44) by the *explicit formula*

$$(T_j^{n+1} - T_j^n)/\Delta t = \sigma(T_{j+1}^n - 2T_j^n + T_{j-1}^n)/\Delta x^2. \tag{59.45}$$

Using (59.13) as a trial solution for the kth Fourier component we find

$$\xi_k = 1 + \alpha(\cos k\Delta x - 1) \tag{59.46}$$

where $\alpha \equiv 2\sigma\,\Delta t/\Delta x^2$. Clearly ξ_k is always ≤ 1. To bound ξ_k away from -1 we must have $\alpha \leq 1$; hence the stability criterion is

$$\Delta t = \Delta x^2/2\sigma, \tag{59.47}$$

which is quite restrictive because if we refine the spatial grid by a factor $1/k$, we must decrease Δt by a factor $1/k^2$, so the computing effort to span a given time interval increases as k^3. The timestep set by (59.47) may be much smaller than the natural hydrodynamic timestep, particularly in regions of high conductivity and/or low-heat capacity (i.e., large σ). Thus use of an explicit formula for diffusion processes may seriously impair one's ability to follow a flow numerically. We must therefore find a more stable difference scheme.

The difficulty just described is overcome by using an *implicit* difference equation, that is, one that contains information about T^{n+1} in both the time and space operators. In particular, consider formulae of the general form

$$T_j^{n+1} - T_j^n = \tfrac{1}{2}\alpha[\theta(\delta^2 T)_j^{n+1} + (1-\theta)(\delta^2 T)_j^n], \tag{59.48}$$

with $0 \leq \theta \leq 1$. Here $(\delta^2 T)_j$ denotes the centered second difference $(T_{j+1} - 2T_j + T_{j-1})$ at the time levels indicated. The von Neumann local stability analysis now leads to

$$\xi_k = [1 - (1-\theta)\alpha(1 - \cos k\Delta x)]/[1 + \theta\alpha(1 - \cos k\Delta x)]. \tag{59.49}$$

As before, ξ_k is always ≤ 1. Furthermore, ξ_k is a monotone decreasing function of $\gamma \equiv \alpha(1 - \cos k\Delta x)$, hence for a given α and θ, ξ_k is minimized when $\cos k\Delta x = -1$. To bound ξ_k away from -1 we therefore demand that

$$-1 \leq [1 - 2(1-\theta)\alpha]/(1 + 2\alpha), \tag{59.50}$$

which implies that

$$(1 - 2\theta)\alpha \leq 1. \tag{59.51}$$

For $0 \leq \theta \leq \frac{1}{2}$, Δt is restricted by

$$\Delta t \leq \Delta x^2/[2\sigma(1 - 2\theta)], \tag{59.52}$$

which reduces to (59.47) when $\theta = 0$. For $\frac{1}{2} \leq \theta \leq 1$, (59.51) is satisfied for *all* α; the difference equation (59.48) is then *unconditionally stable*, and can be solved using arbitrarily large values of Δt, thus surmounting any incompatibility with the hydrodynamically determined timestep. For $\theta = \frac{1}{2}$

we have the *Crank-Nicholson scheme*, which is time centered and has a truncation error of $O(\Delta t^2)$. For $\theta = 1$ we have the *backward Euler* (or *fully implicit*) *scheme*, which has a worse truncation error, $0(\Delta t)$, but is very stable.

The advantages of an implicit scheme are manifest; but to enjoy them we must pay a price. Unlike the explicit scheme, where T_j^{n+1} is calculated directly from preexisting information, in an implicit scheme, T_j^{n+1} is coupled to T^{n+1} at adjacent space points, hence we must solve a linear system of the form

$$-a_j T_{j-1}^{n+1} + b_j T_j^{n+1} - c_j T_{j+1}^{n+1} = r_j, \qquad (j = 1, \ldots, J). \tag{59.53}$$

Boundary conditions guarantee that $a_1 = c_J \equiv 0$. Equation (59.53) is solved by *Gaussian elimination*; we first perform a forward elimination to compute

$$d_j \equiv c_j/(b_j - a_j d_{j-1}) \tag{59.54}$$

and

$$e_j \equiv (r_j + a_j e_{j-1})/(b_j - a_j d_{j-1}) \tag{59.55}$$

for $j = 1, \ldots, J$, and then calculate T_j^{n+1} by back substitution

$$T_j^{n+1} = d_j T_{j+1}^{n+1} + e_j \tag{59.56}$$

for $j = J, J-1, \ldots, 1$. The computational effort required to solve (59.53) scales only linearly with the number of mesh points [as does an update of T^{n+1} via the explicit scheme (59.45)]. In any realistic problem, solving either (59.45) or (59.53) is typically only a small fraction of the total effort required to advance the dynamics a timestep, and the additional cost of using an implicit system is usually greatly outweighed by the ability to take large timesteps.

In this book we are not concerned with solving the heat-conduction equation per se, but rather with solving the equations of hydrodynamics when the energy equation contains diffusive terms. The main lesson we learn from the above analysis is that even if we can use an explicit scheme for the continuity and momentum equations, we must generally use an implicit energy equation in order to avoid unacceptable timestep restrictions; therefore in what follows we will always write the energy equation in implicit form.

ILLUSTRATIVE DIFFERENCE EQUATIONS

Let us now examine some illustrative *examples* of difference equations for solving one-dimensional Lagrangean flow problems in a gravitationally stratified medium.

(*a*) *Explicit Hydrodynamics*; *Planar Geometry* Consider first explicit hydrodynamics in planar geometry. Choose the independent variable to be the column mass measured inward into the medium from an upper boundary

(as would be appropriate for a stellar atmosphere viewed from the outside). Then

$$m(z) \equiv \int_z^\infty \rho(z')\, dz'; \tag{59.57}$$

note that now $dm = -\rho\, dz$, which is opposite to the sign convention used in (59.1) to (59.8). The equations to be solved are

$$(Dv/Dt) = [\partial(p+Q)/\partial m] - g, \tag{59.58}$$

$$(Dz/Dt) = v, \tag{59.59}$$

$$V = 1/\rho = -(\partial z/\partial m), \tag{59.60}$$

and

$$(De/Dt) + (p+Q)[D(1/\rho)/Dt] = \dot{q}. \tag{59.61}$$

Choose a mass grid $\{m_d\}$, $d = 1, \ldots, D+1$ marking the surfaces of D discrete slabs. At the upper boundary, m_1 will be nonzero if we assume there is material lying above that surface. The momentum equation is represented by the explicit formula

$$[v_d^{n+(1/2)} - v_d^{n-(1/2)}]/\Delta t^n = -g + [p_{d+(1/2)}^{n+\lambda} - p_{d-(1/2)}^{n+\lambda} + Q_{d+(1/2)}^{n-(1/2)} - Q_{d-(1/2)}^{n-(1/2)}]/\Delta m_d. \tag{59.62}$$

Here Δm_d, $\Delta m_{d+(1/2)}$, Δt^n, and $\Delta t^{n+(1/2)}$ are defined as in (59.5) to (59.8). In general, the timesteps $\Delta t^{n-(1/2)}$ and $\Delta t^{n+(1/2)}$ are unequal; to improve the accuracy of (59.62) we can center the pressure gradient by defining

$$t^{n+\lambda} \equiv \tfrac{1}{2}\{[t^n - \tfrac{1}{2}\Delta t^{n-(1/2)}] + [t^n + \tfrac{1}{2}\Delta t^{n+(1/2)}]\} = t^n + \tfrac{1}{4}[\Delta t^{n+(1/2)} - \Delta t^{n-(1/2)}], \tag{59.63}$$

and using the approximate extrapolation

$$p_{d+(1/2)}^{n+\lambda} \equiv p_{d+(1/2)}^n + \tfrac{1}{4}[\Delta t^{n+(1/2)} - \Delta t^{n-(1/2)}][p_{d+(1/2)}^n - p_{d+(1/2)}^{n-1}]/\Delta t^{n-(1/2)}. \tag{59.64}$$

The artificial viscosity is computed from a difference representation of (59.41), that is,

$$Q_{d+(1/2)}^{n-(1/2)} = k_Q^2 \tfrac{1}{2}[\rho_{d+(1/2)}^{n-1} + \rho_{d+(1/2)}^n][v_{d+1}^{n-(1/2)} - v_d^{n-(1/2)}]^2 \tag{59.65}$$

if $\rho_{d+(1/2)}^n > \rho_{d+(1/2)}^{n-1}$, and $Q_{d+(1/2)}^{n-(1/2)} = 0$ otherwise. Here $k_Q \equiv l/\Delta z$ is a pure number, typically 1.5 to 2. In astrophysical applications, where zone thicknesses may range over several orders of magnitude in a single flow, it is often more satisfactory to use a fixed length l in (59.41) than to choose a constant k_Q (**W7**). Notice that the pseudoviscous pressure is *lagged* at $t^{n-(1/2)}$ in (59.62). In general this lack of centering does not produce large errors; it is possible to improve the centering, but the result is inconvenient and may even be unstable (**R4,** 319).

To find $v_1^{n+(1/2)}$ and $v_{D+1}^{n+(1/2)}$ we apply boundary conditions. At the lower

boundary we assume that v is a known function of time:

$$v_{D+1}^{n+(1/2)} = f[t^{n+(1/2)}]. \tag{59.66}$$

Typical choices are $v_{D+1} \equiv 0$, or v_{D+1} equals the velocity of a driving "piston." At the upper boundary there are several commonly used choices. For a *free boundary* (no net force across the first cell) we have

$$v_1^{n+(1/2)} \equiv v_2^{n+(1/2)}. \tag{59.67}$$

For a *transmitting boundary* we demand

$$(\partial v/\partial t) = -a(\partial v/\partial z) = a\rho(\partial v/\partial m), \tag{59.68}$$

where a is the local sound speed, which implies that an incident wave is propagated through the boundary without alteration [cf. (59.11)]. In finite form,

$$v_1^{n+(1/2)} - v_1^{n-(1/2)} = \Delta t^n a_{3/2}^n \rho_{3/2}^n [v_2^{n-(1/2)} - v_1^{n-(1/2)}]/\Delta m_{3/2}. \tag{59.69}$$

For *zero surface pressure*, that is $(p+Q)_1 \equiv 0$, we have

$$v_1^{n+(1/2)} - v_1^{n-(1/2)} = -g\,\Delta t^n + [(\Delta t^n/\Delta m_1)(p_{3/2}^{n+\lambda} + Q_{3/2}^{n-(1/2)})], \tag{59.70}$$

where

$$\Delta m_1 \equiv \tfrac{1}{2}\Delta m_{3/2} + m_1 \tag{59.71}$$

includes any atmospheric mass (m_1) assumed to lie above the first cell.

Having updated velocities we calculate

$$z_d^{n+1} = z_d^n + v_d^{n+(1/2)}\,\Delta t^{n+(1/2)} \tag{59.72}$$

and

$$V_{d+(1/2)}^{n+1} = 1/\rho_{d+(1/2)}^{n+1} = (z_d^{n+1} - z_{d+1}^{n+1})/\Delta m_{d+(1/2)}, \tag{59.73}$$

and can then calculate the artificial viscosity $Q_{d+(1/2)}^{n+(1/2)}$ at $t^{n+(1/2)}$.

Finally we solve the implicit energy equation

$$\begin{aligned} e_{d+(1/2)}^{n+1} - e_{d+(1/2)}^n + \{\tfrac{1}{2}[p_{d+(1/2)}^n + p_{d+(1/2)}^{n+1}] + Q_{d+(1/2)}^{n+(1/2)}\}[V_{d+(1/2)}^{n+1} - V_{d+(1/2)}^n] \\ = \Delta t^{n+(1/2)}[(1-\theta)\dot{q}_{d+(1/2)}^n + \theta\dot{q}_{d+(1/2)}^{n+1}] \equiv \Delta t^{n+(1/2)}\langle \dot{q}_{d+(1/2)}\rangle^{n+(1/2)} \end{aligned} \tag{59.74}$$

given constitutive equations $e = e(\rho, T)$ and $p = p(\rho, T)$ or $e = e(\rho, p)$. Notice that in (59.74) the pseudoviscous pressure is time centered so that the correct total work is computed. In calculating $\langle \dot{q} \rangle$ one would choose $\theta = \frac{1}{2}$ for accuracy, and $\theta = 1$ to enhance stability or for quasistatic energy transport (e.g., by radiation—see introduction to §6.5 and also §98).

To solve (59.74) we start from an estimate of T^{n+1}, say T^*, and then *linearize*:

$$p_{d+(1/2)}^{n+1} \approx p[\rho_{d+(1/2)}^{n+1}, T_{d+(1/2)}^*] + [(\partial p/\partial T)_\rho]_{d+(1/2)}\,\delta T_{d+(1/2)} \tag{59.75}$$

and

$$e_{d+(1/2)}^{n+1} \approx e[\rho_{d+(1/2)}^{n+1}, T_{d+(1/2)}^*] + [(\partial e/\partial T)_\rho]_{d+(1/2)}\,\delta T_{d+(1/2)}. \tag{59.76}$$

We likewise linearize the external source $\dot{q}$. If $\dot{q}$ depends only on local values of (ρ, T) we can solve (59.74) pointwise, iterating to consistency (i.e., $\delta T \to 0$). If $\dot{q}$ contains diffusive terms, the linearization process yields a banded system like (59.53), which is solved by Gaussian elimination and iterated to convergence.

A stability analysis of the full system (59.62) plus (59.72) to (59.74) is complicated, so we will only quote results (**R4,** 12.12). Outside a shock, where the pseudoviscosity vanishes, the usual Courant condition must be satisfied. In a very strong shock, the analysis implies that the timestep must be restricted further by a factor $f_Q = \gamma^{1/2}/2k_Q$, which is about $\frac{1}{3}$ in typical problems. Trial calculations show that this theoretical restriction is too strict, and that choosing $f_Q \approx \frac{1}{2}$ is usually sufficient to assure stability. One must also impose timestep restrictions to assure *accuracy* of the solution as well as stability; thus one may restrict the fractional change in any variable to be less than some prechosen value, whose size is set from experience.

The efficacy of the artificial viscosity method is illustrated in Figure 59.2, which shows two test calculations for a propagating shock with a pressure ratio $p_2/p_1 = 5$ in a gas with $\gamma = 2$. The results in part (a) were obtained using $k_Q = 2$, and in part (b) using $k_Q = 0$. Without pseudoviscosity, there are large oscillations in the postshock fluid, and the shock speed is 10 per cent too low. These oscillations do not grow in time, and are not numerical instabilities (Δt was chosen to be 0.22 times the Courant limit). They are motions of the mass cells reminiscent of random motions of molecules in the shock-heated gas; the effect of artificial viscosity is to damp these motions and to convert their kinetic energy into internal energy of the gas.

The differential equations (59.58) to (59.61) can be combined into a total energy equation

$$\frac{D}{Dt}(e+\tfrac{1}{2}v^2) - \frac{\partial}{\partial m}[(p+Q)v] = -gv + \dot{q}, \tag{59.77}$$

which has the integral

$$\frac{D}{Dt}\int (e+\tfrac{1}{2}v^2+gz)\,dm = \int \dot{q}\,dm + v_{D+1}(p+Q)_{D+1} - v_1(p+Q)_1. \tag{59.78}$$

Physically, (59.78) states that the change in the total (internal plus kinetic plus potential) energy of the fluid over some time interval equals the time-integrated energy input from external sources minus the work done by the fluid against its boundaries. A discretized energy conservation relation is obtained by replacing the integrals over mass and time by sums over mass cells and timesteps, starting from some initial time t^0 when the

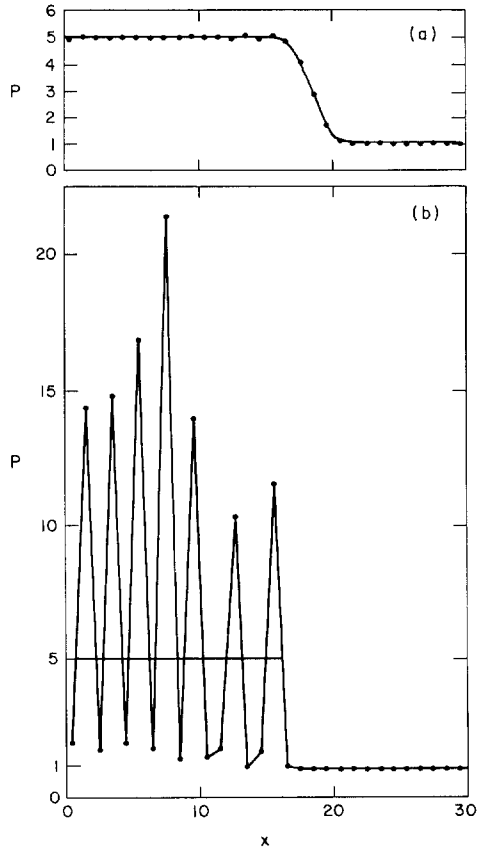

Fig. 59.2 Shock computed (a) with and (b) without artificial viscosity. From (**R4**) by permission.

state of the flow is known. Thus at t^{n+1}

$$\begin{aligned}\mathscr{E}^{n+1} &= \sum_d \{e^{n+1}_{d+(1/2)}\,\Delta m_{d+(1/2)} + [g(z^{n+1}_d - z^0_d) + \tfrac{1}{2}(v^{n+1}_d)^2]\,\Delta m_d\} \\ &\quad + \mathscr{W}^{n+1}_1 - \mathscr{W}^{n+1}_{D+1} + \mathscr{q}^{n+1} \\ &= \sum_d [e^0_{d+(1/2)}\,\Delta m_{d+(1/2)} + \tfrac{1}{2}(v^0_d)^2\,\Delta m_d] = \text{Constant},\end{aligned} \tag{59.79}$$

where

$$\mathscr{W}^{n+1}_i \equiv \sum_{k=0}^{n} [\tfrac{1}{2}(p^k_i + p^{k+1}_i) + Q^{k+(1/2)}_i](z^{k+1}_i - z^k_i) \tag{59.80}$$

and

$$\mathscr{q}^{n+1} \equiv \sum_{k=0}^{n} \Delta t^{k+(1/2)} \sum_d \langle \dot{q}_{d+(1/2)} \rangle^{k+(1/2)}\,\Delta m_{d+(1/2)}. \tag{59.81}$$

In (59.79), v^{n+1} must be estimated by interpolation between $v^{n+(1/2)}$ and $v^{n+(3/2)}$. For problems using the boundary conditions $(p+Q)_1 \equiv 0$ and $v_{D+1} \equiv 0$ (hence $z_{D+1} \equiv$ constant), both boundary work terms $\mathcal{W}^{n+1}$ vanish identically.

While $\mathcal{E}$ should, ideally, be constant, this property is not guaranteed exactly by the explicit difference equations. Rather, $\mathcal{E}$ is calculated after each integration step and is monitored as a check on the quality of the solution, cf. (**C8**), (**F1**). If satisfactory energy conservation is not obtained, the timestep is decreased and the integration step is done over.

(*b*) *Explicit Hydrodynamics; Spherical Geometry* For one-dimensional Lagrangean flow calculations in spherical geometry, the independent variable is M_r, the mass interior to radius r, increasing outward in the medium. The choice of a pseudoviscosity in spherical flows requires some care; the customary approach [see e.g., (**C5**), (**C8**), (**F1**)] is to use a scalar pseudoviscous pressure as in the planar case. The equations to be solved are then

$$(Dv/Dt) = -(GM_r/r^2) - 4\pi r^2[\partial(p+Q)/\partial M_r], \tag{59.82}$$

$$(Dr/Dt) = v, \tag{59.83}$$

and

$$V = 1/\rho = \tfrac{4}{3}\pi(\partial r^3/\partial M_r), \tag{59.84}$$

with

$$Q \equiv \tfrac{4}{3}\rho l^2(\partial v/\partial r)^2 \tag{59.85}$$

if $(\partial v/\partial r) < 0$, and $Q = 0$ otherwise. The energy equation is again (59.61).

These equations are discretized on a mass grid $\{M_i\}$, $i = 1, \ldots, I+1$, defining the surfaces of I spherical shells; M_i is the mass interior to the ith surface, and the mass of the ith shell is $\Delta M_{i+(1/2)} \equiv M_{i+1} - M_i$. An explicit difference representation of the momentum equation is

$$\begin{aligned}[v_i^{n+(1/2)} - v_i^{n-(1/2)}]/\Delta t^n = &-[GM_i/(r_i^{n+\lambda})^2] \\ &- 4\pi(r_i^{n+\lambda})^2[p_{i+(1/2)}^{n+\lambda} - p_{i-(1/2)}^{n+\lambda} + Q_{i+(1/2)}^{n-(1/2)} - Q_{i-(1/2)}^{n-(1/2)}]/\Delta M_i,\end{aligned} \tag{59.86}$$

where $Q_{i+(1/2)}^{n-(1/2)}$ is obtained from a discretized version of (59.85), $p^{n+\lambda}$ is defined by (59.64),

$$r_i^{n+\lambda} \equiv r_i^n + \tfrac{1}{4}[\Delta t^{n+(1/2)} - \Delta t^{n-(1/2)}]v_i^{n-(1/2)} \tag{59.87}$$

and

$$\Delta M_i \equiv \tfrac{1}{2}[\Delta M_{i-(1/2)} + \Delta M_{i+(1/2)}]. \tag{59.88}$$

At the inner boundary we impose (59.66). At the outer boundary we impose (59.67) to (59.71), modified for spherical geometry.

After updating velocities we compute

$$r_i^{n+1} = r_i^n + v_i^{n+(1/2)}\,\Delta t^{n+(1/2)} \tag{59.89}$$

and

$$V_{i+(1/2)}^{n+1} = 1/\rho_{i+(1/2)}^{n+1} = \tfrac{4}{3}\pi[(r_{i+1}^{n+1})^3 - (r_i^{n+1})^3]/\Delta M_{i+(1/2)}, \tag{59.90}$$

and solve the same energy equation as in the planar case.

Although (59.86) to (59.90) have been applied widely, using (59.85) for the pseudoviscosity can cause serious numerical difficulties in certain problems, for example, accretion flows in star formation (**A1**), (**T5**). In particular, because radii converge as $r \to 0$, inflowing material ($v < 0$) can experience compression even when $(\partial v/\partial r) > 0$; this material should be subject to a pseudoviscous pressure, yet (59.85) predicts $Q = 0$ in this case. Furthermore, in some problems (59.85) produces a spurious diffusion of radial momentum.

These difficulties are overcome by use of a *tensor artificial viscosity* (**T5**). Writing $\mathsf{T} = -p\mathsf{I} + \mathsf{Q}$, we demand that Q have the same analytical form as the molecular viscosity $\boldsymbol{\sigma}$, but permit a different viscosity coefficients. Thus we write

$$Q_{ij} = \mu_Q(v_{i;j} + v_{j;i} - \tfrac{2}{3}v^k_{;k}\,\delta_{ij}) \tag{59.91}$$

where, by analogy with the planar case, we set

$$\mu_Q = -\rho l^2 v^k_{;k} = l^2(D\rho/Dt) \tag{59.92}$$

for $v^k_{;k} < 0$ or $(D\rho/Dt) > 0$, and $\mu_Q = 0$ otherwise. Equations (59.91) and (59.92) have the following desirable properties: (1) μ_Q is positive for compression and zero for expansion regardless of the direction of the flow. (2) Trace $\mathsf{Q} = 0$, hence the pseudoviscosity is zero for homologous contraction ($\mathbf{v} = -k\mathbf{r}$), as is also true for molecular viscosity (the no-slip condition). (3) They reduce to the previous formulae in the planar limit, while discriminating correctly between the velocity divergence (scalar) and velocity gradient (tensor), which are fortuitously identical in one-dimensional planar flows. In practice (59.91) and (59.92) give very satisfactory results.

For one-dimensional spherical flow (59.91) is

$$\mathsf{Q} = 2\mu_Q \begin{pmatrix} (\partial v/\partial r) - \frac{1}{3}\boldsymbol{\nabla}\cdot\mathbf{v} & 0 & 0 \\ 0 & (v/r) - \frac{1}{3}\boldsymbol{\nabla}\cdot\mathbf{v} & 0 \\ 0 & 0 & (v/r) - \frac{1}{3}\boldsymbol{\nabla}\cdot\mathbf{v} \end{pmatrix}. \tag{59.93}$$

From (A3.91) and the fact that Q is traceless we find the radial component of the pseudoviscous force to be

$$(\boldsymbol{\nabla}\cdot\mathsf{Q})_r = r^{-3}[\partial(r^3 Q_{rr})/\partial r] \tag{59.94}$$

where

$$Q_{rr} = 2\mu_Q[(\partial v/\partial r) + \tfrac{1}{3}(D\ln\rho/Dt)] \equiv -Q. \tag{59.95}$$

The momentum equation is

$$(Dv/Dt) = -(GM_r/r^2) - 4\pi r^2(\partial p/\partial M_r) - (4\pi/r)[\partial(r^3 Q)/\partial M_r], \tag{59.96}$$

and the energy equation is

$$(De/Dt) + p[D(1/\rho)/Dt] = \dot{q} + \Phi, \tag{59.97}$$

where, from (27.31), the dissipation function is

$$\Phi = 3\mu_Q[(\partial v/\partial r) + \tfrac{1}{3}(D \ln \rho/Dt)]^2. \tag{59.98}$$

The momentum equation has the discrete representation

$$\frac{v_i^{n+(1/2)} - v_i^{n-(1/2)}}{\Delta t^n} = \frac{-GM_i}{(r_i^{n+\lambda})^2} - \frac{4\pi}{\Delta M_i}\left((r_i^{n+\lambda})^2[p_{i+(1/2)}^{n+\lambda} - p_{i-(1/2)}^{n+\lambda}] + \frac{1}{r_i^n}\{[r_{i+(1/2)}^n]^3 Q_{i+(1/2)}^{n-(1/2)} - [r_{i-(1/2)}^n]^3 Q_{i-(1/2)}^{n-(1/2)}\}\right) \tag{59.99}$$

where the cell-center radius $r_{i+(1/2)}$ is chosen so as to contain half the volume (or mass) of the shell (r_i, r_{i+1}),

$$r_{i+(1/2)}^3 \equiv \tfrac{1}{2}(r_i^3 + r_{i+1}^3), \tag{59.100}$$

and where $r^{n+\lambda}$ and $p^{n+\lambda}$ are defined by (59.87) and (59.64). Equations (59.89) and (59.90) remain unchanged, while the energy equation becomes

$$e_{i+(1/2)}^{n+1} - e_{i+(1/2)}^n + \tfrac{1}{2}[p_{i+(1/2)}^n + p_{i+(1/2)}^{n+1}][V_{i+(1/2)}^{n+1} - V_{i+(1/2)}^n] = \Delta t^n[\langle \dot{q}_{i+(1/2)} \rangle^{n+(1/2)} + \Phi_{i+(1/2)}^{n-(1/2)}]. \tag{59.101}$$

In (59.99) and (59.101)

$$Q_{i+(1/2)}^{n-(1/2)} = -2(\mu_Q)_{i+(1/2)}^{n-(1/2)}\left[\frac{v_{i+1}^{n-(1/2)} - v_i^{n-(1/2)}}{r_{i+1}^{n-(1/2)} - r_i^{n-(1/2)}} + \frac{1}{3}\frac{\ln \rho_{i+(1/2)}^n - \ln \rho_{i+(1/2)}^{n-1}}{\Delta t^{n-(1/2)}}\right] \tag{59.102}$$

and

$$\Phi_{i+(1/2)}^{n+(1/2)} = 3(\mu_Q)_{i+(1/2)}^{n+(1/2)}\left[\frac{v_{i+1}^{n+(1/2)} - v_i^{n+(1/2)}}{r_{i+1}^{n+(1/2)} - r_i^{n+(1/2)}} + \frac{1}{3}\frac{\ln \rho_{i+(1/2)}^{n+1} - \ln \rho_{i+(1/2)}^n}{\Delta t^{n+(1/2)}}\right]^2, \tag{59.103}$$

where

$$(\mu_Q)_{i+(1/2)}^{n-(1/2)} = l^2[\rho_{i+(1/2)}^n - \rho_{i+(1/2)}^{n-1}]/\Delta t^{n-(1/2)} \tag{59.104}$$

if $\rho_{i+(1/2)}^n > \rho_{i+(1/2)}^{n-1}$, and is zero otherwise. In (59.104), either $l = k_Q \Delta r$ where k_Q is a number of order unity, or l is a prechosen, fixed length. The radii $r^{n\pm(1/2)}$ in (59.102) and (59.103) must be estimated by interpolation in time.

It follows from (27.4) [cf. also (27.33)] that for one-dimensional, spherically symmetric viscous flows we have a total energy equation of the form

$$\frac{D}{Dt}(e + \tfrac{1}{2}v^2) + \frac{\partial}{\partial M_r}[4\pi r^2 v(p + Q)] = \frac{-GM_r v}{r^2} + \dot{q}, \tag{59.105}$$

where Q is a scalar pressure as in (59.85) or the radial component of a

tensor as in (59.95). Equation (59.105) yields a conservation relation like (59.78). As was true in planar geometry, this conservation relation follows from the *differential* equations governing the flow, but is not guaranteed by the explicit *difference* equations written above; rather, it is again monitored as a check on the quality of the solution. However, we will see below that it is possible to write a set of implicit difference equations that do yield an exact total energy conservation relation provided that the pseudoviscosity enters as a scalar pressure in both the energy and momentum equations. This is not the case when one uses the tensor formulation described above. A compromise is to use a pseudoviscosity that is mathematically equivalent to a scalar, but which is tailored to mimic the basic physical properties of the tensor pseudoviscosity. In particular, if we replace v/r in (59.93) by $(\partial v/\partial r)$ we obtain the isotropic tensor

$$\mathbf{Q} = 2\mu_Q[(\partial v/\partial r) - \tfrac{1}{3}\nabla \cdot \mathbf{v}]\mathbf{I} \equiv -Q\mathbf{I}, \tag{59.106}$$

where μ_Q is given by (59.92). This choice, while heuristic, has the following desirable properties. (1) Q is nonzero for compression and zero for expansion; (2) Q is zero for homologous contraction; (3) $\mathbf{Q}$ is isotropic, hence we may use Q as a scalar pressure. Note, however, that (59.106) does not yield trace $\mathbf{Q} = 0$, as did (59.93).

(*c*) *Implicit Hydrodynamics; Spherical Geometry* The schemes described above use explicit hydrodynamics. A rationale for that approach is that if we model wave phenomena and/or shocks, the physical system changes significantly in a flow time $t_f \sim l/a$, which is generally of the same order as the Courant time $\Delta x/a$. Because we want to follow these very changes there is no reason to take longer timesteps. A counterexample to this argument is provided by stellar evolution calculations, where we are interested in changes on a nuclear burning—rather than hydrodynamic—timescale, and we need to use very large timesteps to be able to follow the evolution of a star through a long sequence of near-equilibrium states. Furthermore, it is necessary to use an implicit scheme in order to avoid unnecessarily restrictive timestep limitations from thin zones and/or regions of high sound speed (e.g., in a stellar interior).

An implicit scheme (**K4**) suitable for calculations of both quasistatic stellar evolution and hydrodynamic events such as nova explosions is

$$(r_i^{n+1} - r_i^n)/\Delta t^{n+(1/2)} = \langle v \rangle^{n+(1/2)}, \tag{59.107}$$

$$(v_i^{n+1} - v_i^n)/\Delta t_{n+(1/2)} = \langle a_i \rangle^{n+(1/2)}, \tag{59.108}$$

$$V_{i+(1/2)}^{n+1} = 1/\rho_{i+(1/2)}^{n+1} = \tfrac{4}{3}\pi[(r_{i+1}^{n+1})^3 - (r_i^{n+1})^3]/\Delta M_{i+(1/2)}, \tag{59.109}$$

and

$$e_{i+(1/2)}^{n+1} - e_{i+(1/2)}^n + \langle (p+Q)_{i+(1/2)} \rangle^{n+(1/2)} [V_{i+(1/2)}^{n+1} - V_{i+(1/2)}^n] = \langle \dot{q}_{i+(1/2)} \rangle^{n+(1/2)} \Delta t^{n+(1/2)}. \tag{59.110}$$

Here the time averages $\langle\ \rangle^{n+(1/2)}$ are defined as

$$\langle x\rangle^{n+(1/2)} \equiv (1-\theta)x^n + \theta x^{n+1}, \tag{59.111}$$

and in (59.108) the acceleration is

$$a_i^m \equiv -GM_i/(r_i^m)^2 - 4\pi(r_i^m)^2[(p+Q)_{i+(1/2)}^m - (p+Q)_{i-(1/2)}^m]/\Delta M_i. \tag{59.112}$$

For Q we use a discrete representation of the scalar pressure defined by (59.106). Notice that there is no difficulty in time-centering variables in an implicit scheme, and special interpolation or extrapolation procedures such as are used in explicit schemes become unnecessary. In stellar evolution calculations, the equations are often rewritten in terms of the logarithms of physical variables, such as ρ and p, that run over several orders of magnitude (**K4**).

The system written above is unconditionally stable for $\frac{1}{2} \le \theta \le 1$. Nevertheless, (59.107) and (59.108) may be unsatisfactory for collapse problems on a Kelvin-Helmholtz timescale because a centered formula ($\theta = \frac{1}{2}$) may lead to growing oscillations. Equation (59.107) should then be made fully implicit ($\theta = 1$). But then (59.108) with $\theta = 1$ may artificially damp real oscillations; in such cases special formulae (e.g., containing three time levels) may be needed (**B7**).

Given constitutive relations for p and e, and a specification of $\dot{q}$, (59.107) to (59.112) are linearized around a trial solution at t^{n+1} and iterated to consistency. The linearized system is typically block tridiagonal, and is solved by Gaussian elimination (see §§83, 97, and 98).

As Fraley (**F2**) pointed out, the momentum equation can also be written

$$\begin{aligned}(v_i^{n+1} - v_i^n)/\Delta t^{n+(1/2)} = &-GM_i\langle 1/r^2\rangle_i \\ &- 4\pi\langle r^2\rangle_i[\langle (p+Q)_{i+(1/2)}\rangle^{n+(1/2)} - \langle (p+Q)_{i-(1/2)}\rangle^{n+(1/2)}]/\Delta M_i\end{aligned} \tag{59.113}$$

where the angular brackets denote suitable time averages. In particular, if we choose the special averages

$$\langle r^2\rangle_i \equiv \tfrac{1}{3}[(r_i^n)^2 + r_i^n r_i^{n+1} + (r_i^{n+1})^2] \tag{59.114}$$

and

$$\langle 1/r^2\rangle_i \equiv 1/(r_i^n r_i^{n+1}), \tag{59.115}$$

then on multiplying (59.113) through by $v_i^{n+(1/2)} \equiv \frac{1}{2}(v_i^n + v_i^{n+1})$ and using the fact that $r_i^{n+1} - r_i^n = v_i^{n+(1/2)}\,\Delta t^{n+(1/2)}$ we obtain the *exact* conservation relation

$$\begin{aligned}&\sum_i \{[\tfrac{1}{2}(v_i^{n+1})^2 - (GM_i/r_i^{n+1})]\,\Delta M_i - \langle (p+Q)_{i+(1/2)}\rangle^{n+(1/2)} V_{i+(1/2)}^{n+1}\,\Delta M_{i+(1/2)}\} \\ &= \sum_i \{[\tfrac{1}{2}(v_i^n)^2 - (GM_i/r_i^n)]\,\Delta M_i - \langle (p+Q)_{i+(1/2)}\rangle^{n+(1/2)} V_{i+(1/2)}^n\,\Delta M_{i+(1/2)}\}.\end{aligned} \tag{59.116}$$

Here we assumed, for simplicity, zero contribution from the work terms at the boundary surfaces. Provided that we use precisely the same average for $\langle p+Q\rangle$ in both (59.113) and (59.110) we then obtain an exact total energy conservation relation

$$\sum_i \{e_{i+(1/2)}^{n+1}\,\Delta M_{i+(1/2)} + [\tfrac{1}{2}(v_i^{n+1})^2 - (GM_i/r_i^{n+1})]\,\Delta M_i\} - \sum_{k=0}^{n} \Delta t^{k+(1/2)} \sum_i \langle \dot{q}_{i+(1/2)}\rangle^{k+1}\,\Delta M_{i+(1/2)} = \text{constant}. \tag{59.117}$$

Fraley's form of the momentum equation with $\theta = \frac{1}{2}$ in $\langle p+Q\rangle$ is often used in stellar pulsation calculations (**C1**), (**W9**) where it is important to obtain precise total energy conservation in order to avoid artificial damping of self-excited oscillations.

CRITIQUE

In numerical simulations of fluid flow, the choice of the best computational method requires good judgment because one's desire for accuracy and stability must be balanced against limitations in computer speed and capacity. In addition, one must sometimes face (perhaps unresolved) questions about the ability of the equations used to model faithfully the real physics of the flow. A short but illuminating discussion of these points is given in (**H1,** 86–90).

The worst problem inherent in one-dimensional Lagrangean schemes is their limited ability to resolve features with very steep gradients of material properties as they move through the fluid. Important examples are propagating shocks, and the cyclic motion of the hydrogen ionization zone in pulsating stars. In the case of shocks, artificial viscosity smears the front over a few zones. Although the Rankine-Hugoniot relations are still satisfied in the upstream and downstream flows, and the effect of the shock on the large-scale structure of the ambient medium is given correctly, each zone contains much more mass, and therefore is much more opaque, than the actual shock front itself; hence a calculation of radiative transport through the shock can be falsified badly. One approach to overcoming this difficulty is to use upstream and downstream conditions determined from coarse-zone calculations to do after-the-fact shock fitting with an extremely fine-zone model that resolves the shock structure and permits an accurate transfer calculation (**H7**).

The hydrogen-helium ionization zone in pulsating stars is even more troublesome because it contains the thermodynamic "engine" that drives the pulsation, and accurate modeling of its structure is therefore essential. As the star pulsates, a steep temperature and ionization front having a characteristic thickness of about a thousandth of a scale height sweeps back and forth over several pressure scale heights, hence through several Lagrangean mass zones. To avoid prohibitively small timesteps, relatively

coarse zones are used; but coarse zoning produces unphysical bumps in the luminosity light curve and in the surface velocity. The only satisfactory solution is to use *rezoning* or *adaptive-mesh* schemes, in which the computational mesh is neither fixed in space (Eulerian), nor attached to definite material elements (Lagrangean). Instead, the mesh moves both in space and through the fluid in such a way as to track physically significant phenomena, such as shocks and ionization fronts, in the flow. Adaptive-mesh algorithms are discussed in (**C1**), (**T5**), and (**W7**).

60. *Propagating Strong Shocks*

LARGE-AMPLITUDE WAVES IN THE SOLAR CHROMOSPHERE

The propagation of strong shocks through a stratified atmosphere can be modeled using the methods discussed in §59. Instructive examples are provided by the work of Stein and Schwartz (**S13**), (**S14**), who solved (59.58) to (59.61) for periodic strong shocks in an isothermal atmosphere ($\gamma \equiv 1$). They chose $T = 5700$ K and $g \approx 3 \times 10^4$ cm s^{-2} (appropriate for the Sun), which imply a sound speed $a \approx 7$ km s^{-1} and an acoustic-cutoff period

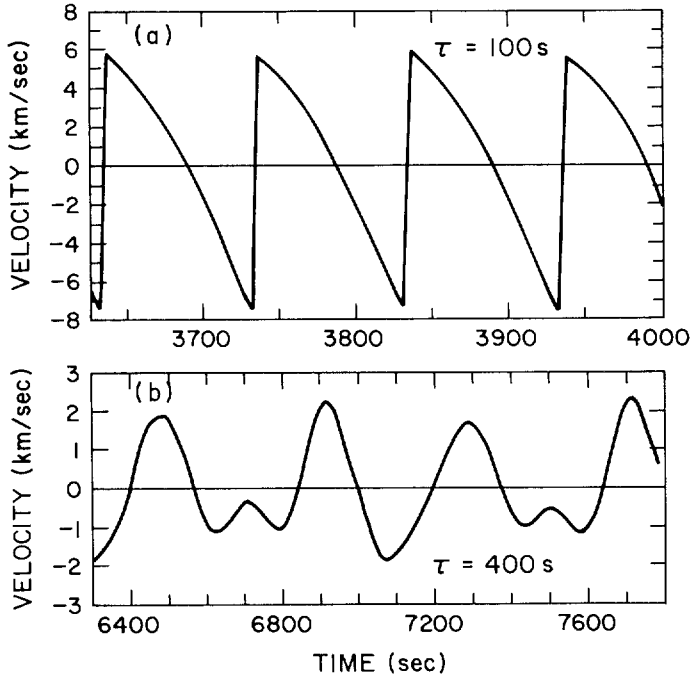

Fig. 60.1 Periodic wave trains in an isothermal stratified atmosphere. (a) Shock train generated by disturbance with period of 100 s. (b) Oscillation generated by disturbance with period of 400 s. From (**S14**) by permission.

$\tau_{ac} \approx 200$ s. The waves were excited by imposing a periodic velocity $v = 0.32 \sin(t/\tau)$ km s^{-1} at the lower boundary.

Their results reveal an important qualitative difference between waves with $\tau < \tau_{ac}$ and those with $\tau > \tau_{ac}$. The short-period waves steepen into shocks and form N waves, as expected from the discussion in §§58 and 59. The steady-state velocity variation of a mass element whose initial height in the atmosphere was 1000 km is shown in Figure 60.1a for a wave with period $\tau = 100$ s.

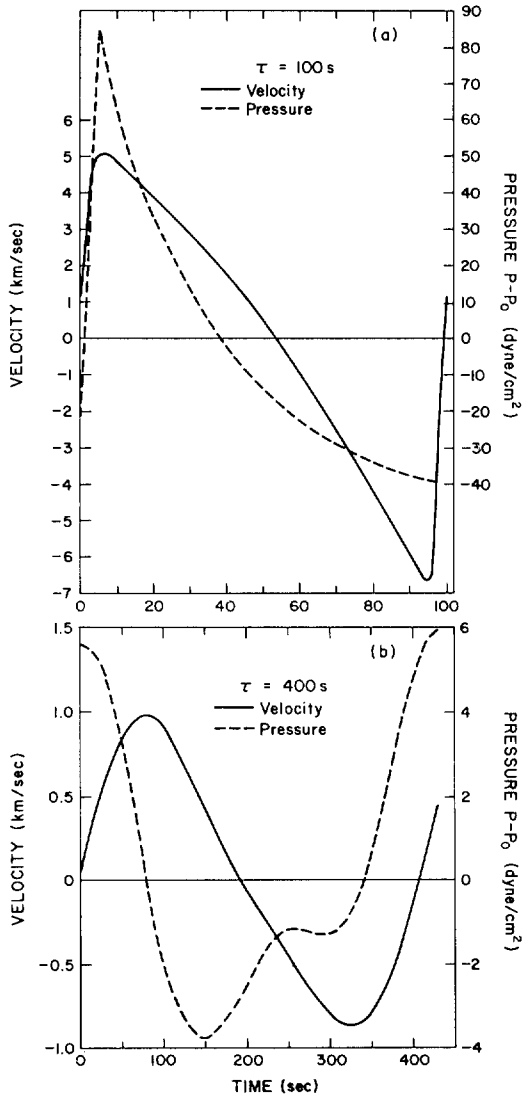

Fig. 60.2 Velocity and pressure perturbations for a single cycle of waves shown in Figure 60.1. From (**S14**) by permission.

In contrast, waves with $\tau > \tau_{ac}$ tend to lift the medium quasi-rigidly. The velocity variation of a mass element whose initial height was 1000 km is shown in Figure 60.1b for a wave with period $\tau = 400$ s. Note that the velocity varies nearly sinusoidally, shows no indication of shocks, and has only about one-third the amplitude of the short-period wave even though both are excited with the same driving term. The 400 s oscillation also shows beats with the 200 s natural period of the atmosphere, which indicates that a steady state has not yet been achieved in the calculation even after 19 full wave periods.

As shown in Figure 60.2, the velocity and pressure perturbations are nearly in phase for the 100 s wave, and approximately 90° out of phase for the 400 s wave. The 100 s wave transports a large energy flux, whereas the 400 s wave is almost a standing wave with the velocity nearly in phase at all heights, and transports very little energy. Even though the amplitude of the 400 s wave increases substantially with height, its small energy transport retards shock formation and inhibits wave-energy dissipation.

The dissipation per period in the 100 s wave as a function of height is shown in Figure 60.3a. Above 1000 km the dissipation rate is nearly constant, and the estimate given by weak shock theory is almost a factor of 10 too low, which is not surprising because the shock is no longer weak

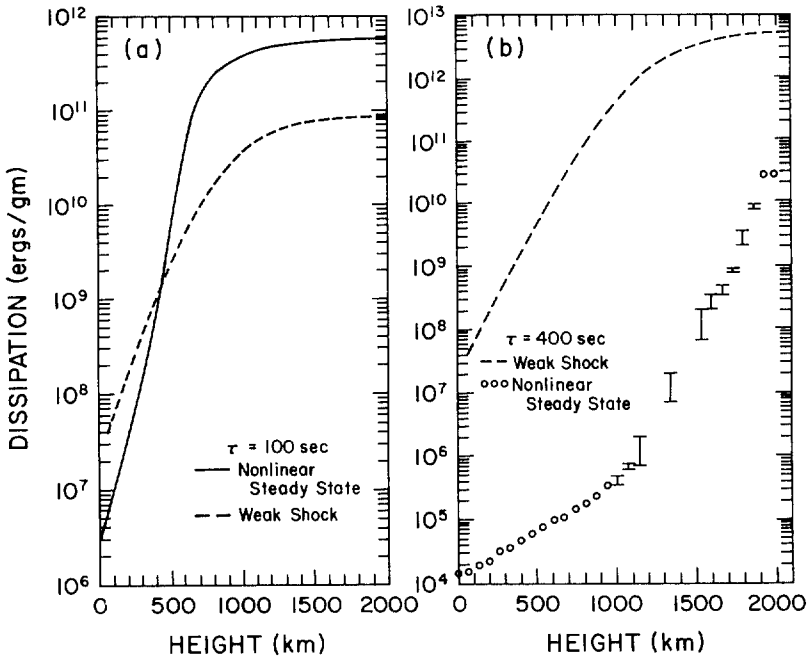

Fig. 60.3 Dissipation per period as a function of height for the waves shown in Figure 60.1. From (**S14**) by permission.

($M \approx 2$) at these heights. Below 500 km, weak shock theory predicts too large a dissipation rate because it assumes the existence of a shock before one has actually developed. As shown in Figure 60.3b the failure of weak shock theory for the 400-s wave is much more dramatic. The actual dissipation in this (nearly) standing wave is about 10^4 times smaller than predicted by weak shock theory. These results clearly show that weak shock theory must be used with great caution, and that the full nonlinear equations must be solved before meaningful estimates of chromospheric heating by shock dissipation can be made.

The results discussed above are only illustrative because radiative losses are omitted; more realistic calculations are discussed in §105.

SIMILARITY SOLUTIONS

An alternative to numerical modeling of shocks is to develop analytical solutions for idealized problems that are reasonable representations of situations of interest; this approach offers physical insight and provides benchmarks against which numerical calculations can be compared. An effective method of constructing such solutions is to carry out *similarity* (or *dimensional*) *analyses* of *self-similar flows*. In these flows the spatial *profiles* (i.e., *distributions*) of the physical variables are time-independent functions of an appropriate *similarity variable*; the time evolution of the flow is described fully by the time variations of the *scaling* of these profiles and of the similarity variable. Similarity methods have been highly developed by Sedov and his co-workers, and have been applied to a wide variety of problems (**S9**); here we consider only two examples of astrophysical interest.

Consider first the *blast wave* driven by a *point explosion*. Here we imagine the essentially instantaneous release of a large amount of energy $\mathscr{E}$ into a very small volume, which drives a spherical shock into a homogeneous medium of density ρ_1. We assume that the material is a perfect gas with a constant ratio of specific heats γ. We seek to describe the motion of the shock at a time when the mass of the material set in motion by the blast is large compared to that in which the initial energy release occurred, but when the shock strength is still so large that we can neglect the exterior gas pressure (*backpressure*) relative to the postshock pressure. We thus neglect the internal energy of the ambient gas compared to the explosion energy.

From the conservation laws (56.6) and (56.7) we have

$$\rho_1 v_s = \rho_2 (v_s - v_2) \tag{60.1}$$

and

$$\rho_1 v_s^2 = \rho_2 (v_s - v_2)^2 + p_2, \tag{60.2}$$

where v_s is the lab-frame shock velocity, and subscripts 1 and 2 denote pre- and post-shock quantities, respectively. Neglecting p_1 is equivalent to assuming an infinite shock strength, hence

$$\rho_2/\rho_1 = (\gamma+1)/(\gamma-1). \tag{60.3}$$

From (60.1) and (60.2) we obtain

$$p_2 = \rho_1 v_s v_2, \tag{60.4}$$

and from (60.1) and (60.3) we obtain

$$v_2 = 2v_s/(\gamma+1), \tag{60.5}$$

whence we find

$$p_2 = 2\rho_1 v_s^2/(\gamma+1). \tag{60.6}$$

For a self-similar expansion of the shock front, the pressure, density, and velocity distribution in the flow can be written

$$p(r, t) = p_2(t)\hat{p}(\xi), \tag{60.7a}$$

$$\rho(r, t) = \rho_2(t)\hat{\rho}(\xi), \tag{60.7b}$$

and

$$v(r, t) = v_2(t)\hat{v}(\xi). \tag{60.7c}$$

The profiles $\hat{p}$, $\hat{\rho}$, and $\hat{v}$ depend on the dimensionless variable ξ, which is related to physical distance by a transformation of the form

$$\xi = r/R(t). \tag{60.8}$$

We can determine $R(t)$ from dimensional arguments. The nature of the flow depends only on the two quantities $\mathscr{E}$ and ρ_1, having dimensions $[\mathscr{E}] = \text{ergs} = \text{g cm}^2\,\text{s}^{-2}$ and $[\rho_1] = \text{g cm}^{-3}$. The only combination of $\mathscr{E}$ and ρ_1 that contains only length and time is the ratio $\mathscr{E}/\rho_1$, which has dimensions $\text{cm}^5\,\text{s}^{-2}$. Hence self-similar motion of the flow can depend on length and time only through the dimensionless combination $\mathscr{E}t^2/\rho_1 r^5$, which implies that

$$\xi = (\rho_1/\mathscr{E})^{1/5}(r/t^{2/5}) \tag{60.9}$$

is an appropriate similarity variable.

For a given value of γ the shock front is characterized by a fixed value of ξ, say ξ_s; hence the position of the shock at time t is given by

$$r_s = \xi_s(\mathscr{E}/\rho_1)^{1/5}t^{2/5}, \tag{60.10}$$

and the shock speed is

$$v_s = (dr_s/dt) = \tfrac{2}{5}\xi_s(\mathscr{E}/\rho_1)^{1/5}t^{-3/5} = \tfrac{2}{5}\xi_s^{5/2}(\mathscr{E}/\rho_1)^{1/2}r_s^{-3/2}. \tag{60.11}$$

The speed of the gas behind the front, v_2, then follows from (60.5). One can understand the scaling in (60.11) intuitively by noting that the kinetic energy per unit volume in the blast wave as measured by either $\rho_2 v_s^2$ or $\rho_2 v_2^2$ must scale as $\mathscr{E}/r_s^3$.

According to (60.6) and (60.11) the pressure behind the shock is

$$p_2 = K\xi_s^2(\rho_1^3\mathscr{E}^2/t^6)^{1/5} = K\xi_s^5\mathscr{E}/r_s^3, \tag{60.12}$$

where $K \equiv 8/25(\gamma+1)$. Thus (for a given γ and ρ_1) the postshock pressure in

explosions of different strengths reaches the same value at times and distances that are proportional to $\mathscr{E}^{1/3}$. We can understand this scaling intuitively by recalling that the pressure in a gas is proportional to the average energy per unit volume, hence to $\mathscr{E}/r_s^3$.

The density behind the shock remains fixed at its limiting value (60.3) as long as the postshock pressure p_2 is much greater than the back pressure p_1.

To complete the solution, one must determine the dimensionless shock-position parameter ξ_s and the profile functions $\hat{p}$, $\hat{\rho}$, and $\hat{v}$. The latter follow from solving three first-order ordinary differential equations obtained by transforming the equations of gasdynamics into dimensionless variables and converting derivatives with respect to r and t into derivatives with respect to ξ. The solution must satisfy the constraints $\hat{p} = \hat{\rho} = \hat{v} = 1$ at $\xi = \xi_s$. The value of ξ_s is obtained by demanding total energy conservation, which implies

$$\int_0^{r_s} (e + \tfrac{1}{2}v^2) 4\pi r^2 \, dr = \mathscr{E}, \tag{60.13}$$

where e is the internal energy of the gas.

An exact solution of the point-explosion problem was obtained by Sedov (**S9**). It shows that the density drops very rapidly behind the shock front. In fact, nearly all the gas contained within r_s is concentrated into a very thin shell ($\Delta r/r_s \lesssim 0.08$) near the front because the shock slows as it sweeps up more and more material. Therefore, gas closer to the origin has a higher expansion speed than gas farther out in the flow, and tends to overrun the front. The pressure drops by a factor of 2 to 3 a short distance behind the shock and then remains roughly constant everywhere inside.

A rigorous derivation of the solution is moderately complicated. But by making the simplifying assumption that all the mass inside the blast wave is concentrated into the thin shell behind the shock, one can derive (**C4**), (**Z1,** 97–99) the approximate formula

$$\xi_s \approx [75(\gamma - 1)(\gamma + 1)^2/16\pi(3\gamma - 1)]^{1/5}, \tag{60.14}$$

which is found to be in good agreement with the exact results.

Another astrophysically interesting problem that can be treated by similarity methods is the propagation of a strong shock in an exponentially stratified medium. Examples are a global shock resulting from a stellar pulsation or supernova explosion passing outward through a stellar envelope, or perhaps a shock emanating from a point source such as a man-made explosion in the Earth's atmosphere or an impulsive flare in the Sun's atmosphere. A shock from a point explosion may initially weaken as it sweeps up material (cf. discussion above), but eventually density-gradient effects dominate and the upward-propagating part of the shock strengthens monotonically as it passes outward into regions of ever-decreasing density. A global shock, of course, responds only to the density drop and

strengthens continuously. Ultimately the shock's speed becomes so large that it escapes to infinity in a finite time; a blast wave can thus *break out* (or *vent*) from an atmosphere, perhaps even transporting material from the heart of the explosion into space if the explosion is sufficiently strong and occurs at a sufficiently high altitude.

Self-similar upward-propagating shocks are discussed in (**G1**), (**H5**), (**H6**), and (**R1**) [see also (**Z1,** Chap. 12)]. Consider a planar atmosphere of an ideal gas with constant γ, having a density stratification $\rho(z)=\rho_0 e^{-z}$, where z is the distance above some convenient reference level, in units of the scale height H. Then similarity analysis shows (**G1**) that the density, shock speed, postshock material speed, and pressure have the forms

$$\rho = A\psi(\zeta)e^{-Z}, \tag{60.15a}$$

$$v_s = v_{s0}e^{\alpha Z}, \tag{60.15b}$$

$$v_2 = v_s\theta(\zeta), \tag{60.15c}$$

and

$$p = Av_{s0}^2\phi(\zeta)e^{-\beta Z}, \tag{60.15d}$$

where $\zeta \equiv z - Z$ is the distance behind the shock (in scale heights). Note that $\zeta \leq 0$, and that velocities are measured in (scale heights s^{-1}). The parameters α and β depend only on the ratio of specific heats γ, and are tabulated in (**G1**) for a wide range of γ; for example, $(\alpha, \beta) = (0.176, 0.646)$ for $\gamma = \frac{4}{3}$ and $(0.204, 0.591)$ for $\gamma = \frac{5}{3}$. Near the shock front ($\zeta = 0$) one has

$$\rho(\zeta) \approx \rho(0)e^{-\zeta}, \tag{60.16a}$$

$$p(\zeta) \approx p(0)e^{-\beta\zeta}, \tag{60.16b}$$

$$v(\zeta) \approx v_s e^{\alpha\zeta}, \tag{60.16c}$$

and

$$\mu(\zeta) \approx e^{(1-\beta)\zeta} \tag{60.16d}$$

where $\mu \equiv \gamma p/\rho v_s^2 = a^2/v_s^2 = 1/M^2$. Equations (60.15) and (60.16) are found to be in excellent agreement with numerical computations.

By integrating (60.15b), we find that the shock position as a function of time is

$$Z = \frac{1}{\alpha}\ln\left(\frac{1}{1-\alpha v_{s0}t}\right) = \frac{1}{\alpha}\ln\left(\frac{t_\infty}{t_\infty - t}\right), \tag{60.17}$$

which shows that the shock approaches infinity in the finite time

$$t_\infty = 1/\alpha v_{s0}. \tag{60.18}$$

Because the shock accelerates rapidly, it can outrun even large perturbations in the flow if they are located beyond some (small) critical distance behind the front; such shocks are sometimes called *self-propagating*. As a result, the asymptotic behavior of the shock is very insensitive to details of the original explosion; in particular it is essentially identical for any

energy-deposition time Δt on the range $0 \leq \Delta t \leq t_\infty$, and is modified significantly only if the deposition rate is strongly singular near $t = t_\infty$.

Moreover, for $t \geq t_\infty$ the fluid variables at points near the origin are independent of the shock position. Thus the ratios p_∞/p_1 and v_∞/v_1 are the same for all fluid elements in the flow, where v_1 and p_1 are an element's initial speed and pressure, and v_∞ and p_∞ are its speed and pressure for $t \geq t_\infty$. Similarly, the distance l_∞ through which a fluid element moves between the time the shock passes over it and t_∞ is the same for all elements. Numerical calculations (**G1**) yield $(l_\infty, v_\infty/v_1, p_\infty/p_1) = (4.60, 1.72, 0.099)$ for $\gamma = \frac{5}{3}$ and $(6.23, 1.75, 0.075)$ for $\gamma = \frac{4}{3}$.

5.4 Thermally Driven Winds

The Sun is surrounded by a tenuous, extremely hot envelope ($n_e \sim 4 \times 10^8\ \mathrm{cm}^{-3}$, $T \sim 1.5 \times 10^6\ \mathrm{K}$) called the corona; other late-type stars have similar structures. As first recognized by Chapman (**C3**), at such high temperatures thermal conduction by electrons becomes an efficient energy transport mechanism, and an unavoidable consequence of this fact is that the corona must extend far out into interplanetary space, even enveloping the Earth in a low-density, high-temperature plasma. Subsequently, Parker showed (**P1**) that if the corona were hydrostatic, the gas pressure at infinite distance from the Sun would exceed the total pressure in the surrounding interstellar medium by orders of magnitude; therefore a static corona is actually impossible. Instead, the corona undergoes a continuous *dynamical expansion* and produces a transsonic flow known as the *solar wind.* The solar wind is driven, ultimately, by conversion of the high specific enthalpy of coronal gas into kinetic energy of fluid motion; such winds are known as *thermal winds* to distinguish them from winds accelerated by direct momentum input to the fluid by intense radiation fields (cf. §107).

Several excellent reviews of thermal winds are available [e.g., (**B9**), (**H13**), (**H14**), (**H15**), (**H18**), and (**P2**)]. We discuss here only the most elementary aspects of the theory.

61. Basic Model

To obtain insight into the dynamics of thermal winds we consider the highly idealized problem of a steady, spherically symmetric flow of a fully ionized pure hydrogen plasma in a gravitational field. We first briefly recapitulate Chapman's and Parker's arguments about the inevitability of coronal expansion.

CORONAL EXPANSION

In a fully ionized hydrogen plasma, the electron heat flux is $\mathbf{q} = -K \nabla T$, where the thermal conductivity $K = K_0 T^{5/2}$ with $K_0 \approx 8 \times 10^{-7}$ ergs $\mathrm{cm}^{-1}\ \mathrm{s}^{-1}\ \mathrm{K}^{-7/2}$ (cf. §33). At coronal temperatures the conductivity of the

energy-deposition time Δt on the range $0 \leq \Delta t \leq t_\infty$, and is modified significantly only if the deposition rate is strongly singular near $t = t_\infty$.

Moreover, for $t \geq t_\infty$ the fluid variables at points near the origin are independent of the shock position. Thus the ratios p_∞/p_1 and v_∞/v_1 are the same for all fluid elements in the flow, where v_1 and p_1 are an element's initial speed and pressure, and v_∞ and p_∞ are its speed and pressure for $t \geq t_\infty$. Similarly, the distance l_∞ through which a fluid element moves between the time the shock passes over it and t_∞ is the same for all elements. Numerical calculations (**G1**) yield $(l_\infty, v_\infty/v_1, p_\infty/p_1) = (4.60, 1.72, 0.099)$ for $\gamma = \frac{5}{3}$ and $(6.23, 1.75, 0.075)$ for $\gamma = \frac{4}{3}$.

5.4 Thermally Driven Winds

The Sun is surrounded by a tenuous, extremely hot envelope ($n_e \sim 4 \times 10^8\ \mathrm{cm}^{-3}$, $T \sim 1.5 \times 10^6\ \mathrm{K}$) called the corona; other late-type stars have similar structures. As first recognized by Chapman (**C3**), at such high temperatures thermal conduction by electrons becomes an efficient energy transport mechanism, and an unavoidable consequence of this fact is that the corona must extend far out into interplanetary space, even enveloping the Earth in a low-density, high-temperature plasma. Subsequently, Parker showed (**P1**) that if the corona were hydrostatic, the gas pressure at infinite distance from the Sun would exceed the total pressure in the surrounding interstellar medium by orders of magnitude; therefore a static corona is actually impossible. Instead, the corona undergoes a continuous *dynamical expansion* and produces a transsonic flow known as the *solar wind.* The solar wind is driven, ultimately, by conversion of the high specific enthalpy of coronal gas into kinetic energy of fluid motion; such winds are known as *thermal winds* to distinguish them from winds accelerated by direct momentum input to the fluid by intense radiation fields (cf. §107).

Several excellent reviews of thermal winds are available [e.g., (**B9**), (**H13**), (**H14**), (**H15**), (**H18**), and (**P2**)]. We discuss here only the most elementary aspects of the theory.

61. Basic Model

To obtain insight into the dynamics of thermal winds we consider the highly idealized problem of a steady, spherically symmetric flow of a fully ionized pure hydrogen plasma in a gravitational field. We first briefly recapitulate Chapman's and Parker's arguments about the inevitability of coronal expansion.

CORONAL EXPANSION

In a fully ionized hydrogen plasma, the electron heat flux is $\mathbf{q} = -K \nabla T$, where the thermal conductivity $K = K_0 T^{5/2}$ with $K_0 \approx 8 \times 10^{-7}$ ergs $\mathrm{cm}^{-1}\ \mathrm{s}^{-1}\ \mathrm{K}^{-7/2}$ (cf. §33). At coronal temperatures the conductivity of the

plasma exceeds that of ordinary metallic laboratory conductors. Following Chapman, suppose that thermal conduction is the only energy transport mechanism operative; then $\nabla \cdot \mathbf{q} = 0$, which implies that

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2 K_0 T^{5/2}\frac{dT}{dr}\right) = 0. \tag{61.1}$$

Integrating (61.1) and demanding that $T \to 0$ as $r \to \infty$ we find

$$T(r) = T_0(r_0/r)^{2/7}, \tag{61.2}$$

where r_0 is a suitable reference level (say $r_0 = R_\odot \approx 7\times 10^{10}$ cm). The falloff predicted by (61.2) is very slow; for $T_0 \sim 1.5\times 10^6$ K, $T \sim 3\times 10^5$ K at the Earth's orbit ($r_\oplus \approx 1.5\times 10^{13}$ cm).

Suppose further that the corona is in hydrostatic equilibrium so that

$$(dp/dr) = -G\mathcal{M}_\odot \rho/r^2. \tag{61.3}$$

For fully ionized hydrogen, $n_e = n_p \equiv n$, $p = 2nkT$, and $\rho = nm_H$. If we assume that $T(r)$ is given by (61.2), then (61.3) becomes

$$\frac{d}{dx}(nx^{-2/7}) = -\left(\frac{r_0}{H}\right)\frac{n}{x^2}. \tag{61.4}$$

Here $x \equiv r/r_0$, and

$$H \equiv 2kT_0 r_0^2/G\mathcal{M}_\odot m_H \tag{61.5}$$

is a scale height; for the solar corona $H \sim 10^5$ km. Integrating (61.4) we find

$$n(r) = n_0 x^{2/7}\exp\left[-\tfrac{7}{5}(r_0/H)(1-x^{-5/7})\right], \tag{61.6}$$

which implies that $n \sim 10^5$ near the Earth's orbit if $n_0 \sim 4\times 10^8$ in the corona. We thus arrive at Chapman's conclusion that the Earth must be enveloped in hot, dense (compared to the interstellar medium) plasma extending from the solar corona.

Combining (61.2) and (61.6) we see that the pressure distribution in a static corona,

$$p(r) = p_0\exp\left[-\tfrac{7}{5}(r_0/H)(1-x^{-5/7})\right], \tag{61.7}$$

implies that the coronal pressure does not vanish as $x \to \infty$, but instead approaches a finite value. For $p_0 \sim 0.2$ dynes cm^{-2} we find that $p_\infty \sim 10^{-5}$ dynes cm^{-2}, which is six to seven orders of magnitude larger than the total pressure in the interstellar medium. We thus arrive at Parker's conclusion that the corona *must* expand.

STEADY ONE-DIMENSIONAL FLOW

Consider now a steady, one-dimensional wind flow in spherical geometry. We must solve the continuity equation

$$r^{-2}[d(r^2\rho v)/dr] = 0, \tag{61.8}$$

the momentum equation

$$\rho v(dv/dr) = -(dp/dr) - G\mathcal{M}_\odot \rho / r^2, \tag{61.9}$$

and the energy equation

$$\frac{1}{r^2}\frac{d}{dr}[r^2 \rho v(\tfrac{1}{2}v^2 + h)] = -\rho v\left(\frac{G\mathcal{M}_\odot}{r^2}\right) + \frac{1}{r^2}\frac{d}{dr}\left(r^2 K \frac{dT}{dr}\right), \tag{61.10}$$

where $h = e + (p/\rho) = 5kT/m_H$ is the specific enthalpy of the plasma.

The continuity equation has the integral

$$4\pi r^2 n m_H v = \dot{\mathcal{M}} \equiv m_H \mathcal{F} \tag{61.11}$$

where $\mathcal{F}$ is the particle flux. The energy equation has the integral

$$\dot{\mathcal{M}}[\tfrac{1}{2}v^2 + h - (G\mathcal{M}_\odot/r)] - 4\pi r^2 K(dT/dr) = \mathcal{E} \tag{61.12}$$

where $\mathcal{E}$ is the total energy flux. The two first-order differential equations (61.9) and (61.12) determine the structure of a thermal wind. When this system is integrated, two more integration constants appear; hence a total of four conditions (boundary conditions, specifications of the behavior of the solution, or choices of free parameters) must be imposed in order to determine a unique solution.

ISOTHERMAL WINDS

Before discussing the general problem posed above, it is instructive to consider the case of an isothermal wind. In physical terms we, in effect, invoke some hypothetical heating mechanism to maintain a constant temperature in the face of the tendency of the gas to cool adiabatically as it expands. In practical terms we can then dispense with (61.12) and solve only (61.9) subject to (61.11).

Setting $T \equiv T_0$, and using (61.11) to eliminate n we can rewrite (61.9) as

$$\tfrac{1}{2}[1 - (2kT_0/m_H v^2)](dv^2/dr) = -(G\mathcal{M}_\odot/r^2)[1 - (4kT_0 r/G\mathcal{M}_\odot m_H)]. \tag{61.13}$$

Equation (61.13) admits several types of solution. Notice first that in the solar corona $(4kT_0 r_0/G\mathcal{M}_\odot m_H) \sim 0.3$, hence the right-hand side of (61.13) passes from negative to positive as r increases from r_0 to ∞, and vanishes at the *critical radius*

$$r_c \equiv G\mathcal{M}_\odot m_H / 4kT_0. \tag{61.14}$$

At $r = r_c$ the left-hand side of (61.13) must therefore vanish. There are two options: either

$$(dv/dr)_{r_c} = 0, \tag{61.15}$$

or (dv/dr) is nonzero but v equals the *critical velocity*

$$v_c = (2kT_0/m_H)^{1/2}, \tag{61.16}$$

which is also the isothermal sound speed at $r = r_0$.

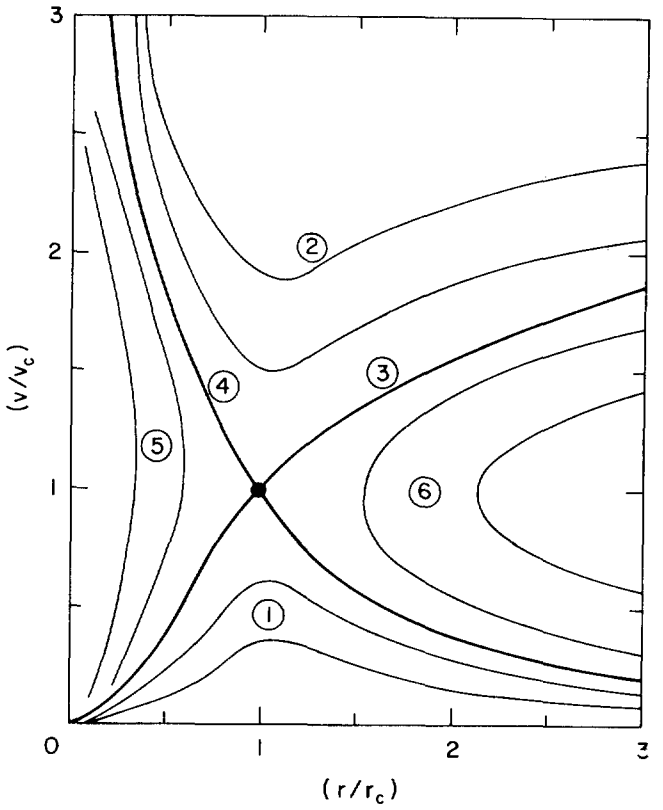

Fig. 61.1 Topology of thermal wind solutions.

If we demand that both v and (dv/dr) be single valued and continuous, we find four types of solutions as sketched in Figure 61.1. First, if we suppose $(dv/dr)_{r_c} = 0$ we can construct solutions in which $[1-(2kT_0/m_H v^2)]$ has the same sign for all r. If we choose $v(r_c) < v_c$, then $v(r)$ has an absolute maximum at r_c, and v is everywhere subsonic (type 1). If $v(r_c) > v_c$, then $v(r)$ has an absolute minimum at r_c, and v is everywhere supersonic (type 2). Alternatively, if we assume that $v(r_c) = v_c$, then we obtain two unique *critical solutions* that pass through (r_c, v_c) with finite slope. Both solutions are transonic, either with v increasing monotonically from subsonic ($v < v_c$) for $r < r_c$ to supersonic for $r > r_c$ (type 3), or with v decreasing monotonically from supersonic to subsonic (type 4). If we drop the requirement that v be single valued we find two additional families of solutions, types 5 and 6 in Figure 61.1; their significance will emerge below.

To choose a model for the real solar wind, one appeals to observation. It is known that flow velocities at the base of the corona are much smaller than $v_c \approx 170 \text{ km s}^{-1}$, hence we can immediately exclude solutions of types 2 and 4. To choose between solutions of types 1 and 3 we integrate (61.13)

to obtain

$$(v/v_c)^2 - \ln (v/v_c)^2 = 4[\ln (r/r_c) + (r/r_c)] + C. \tag{61.17}$$

For solutions of type 1, $v < v_c$, and v decreases as $r \to \infty$; hence for $r \gg r_c$ the dominant term on the left-hand side is $\ln (v/v_c)^2$ and on the right-hand side it is $4 \ln (r/r_c)$. Thus for type-1 solutions $v \propto r^{-2}$ as $r \to \infty$, which implies [cf. (61.11)] that n, hence p, remains finite. In fact these solutions yield values of p_∞ that greatly exceed interstellar pressures, and can therefore be rejected on the same grounds as the hydrostatic solution was.

In contrast, for the critical solution (type 3), $v > v_c$ and increases with increasing r, hence $v(r) \sim 2v_c\ [\ln (r/r_c)]^{1/2}$ for large r, which implies that $n(r) \propto (1/r^2 v) \to 0$ as $r \to \infty$. The critical solution can thus match a zero-pressure boundary condition at infinity. This fact led Parker to conclude (**P1**) that the solar wind is an accelerating transonic flow, a conclusion verified by observations from space vehicles.

The unbounded velocity of the critical solution as $r \to \infty$ is unphysical. It is an artifact of insisting the flow be isothermal, for, as noted earlier, this assumption implies a continuous deposition of thermal energy into the gas, leading to an infinite reservoir of energy which can accelerate the flow without limit. Parker overcame this difficulty by demonstrating that one can obtain satisfactory wind models, in which $n \to 0$ and v approaches a finite value v_∞ as $r \to \infty$, by assuming that the corona either (1) is isothermal for $R_\odot \leq r \leq r_*$ and then expands adiabatically ($\gamma = \frac{5}{3}$) for $r \geq r_*$, or (2) is everywhere polytropic with $\gamma \leq \frac{3}{2}$. Although it is obviously highly oversimplified, model (1) provides at least a plausible caricature of a thermal wind.

HEAT-CONDUCTING WINDS

While instructive, the isothermal winds just discussed are too idealized; to model realistic thermal winds we must retain the energy equation (61.12) in order to determine $T(r)$ consistently with $n(r)$ and $v(r)$. For a systematic survey of solutions, it is convenient to transform to dimensionless variables (**C2**), writing

$$\tau \equiv T/T_0, \tag{61.18a}$$

$$\psi \equiv v^2 \mu m_H / kT_0, \tag{61.18b}$$

and

$$\lambda \equiv G\mathcal{M}_\odot \mu m_H / kT_0 r, \tag{61.18c}$$

where μ is the number of atomic mass units per particle ($\frac{1}{2}$ for ionized hydrogen). In these variables the dynamical equations become

$$n\psi^{1/2}/\lambda^2 = (kT_0/\mu m_H)^{3/2} \mathscr{F}/(4\pi G^2 \mathcal{M}_\odot) \equiv \mathscr{C}, \tag{61.19}$$

$$\tfrac{1}{2}[1 - (\tau/\psi)](d\psi/d\lambda) = 1 - 2(\tau/\lambda) - (d\tau/d\lambda), \tag{61.20}$$

and

$$\mathscr{A}\tau^{5/2}(d\tau/d\lambda) = \varepsilon_\infty - \tfrac{1}{2}\psi + \lambda - \tfrac{5}{2}\tau \tag{61.21}$$

where

$$\varepsilon_\infty \equiv \mu \mathscr{E}/kT_0\mathscr{F} \tag{61.22}$$

is the residual energy, per particle, at infinity, in units of kT_0, and

$$\mathscr{A} \equiv 4\pi K_0 G\mu^2 m_H \mathscr{M}_\odot T_0^{3/2}/k^2\mathscr{F}. \tag{61.23}$$

To integrate these equations we must specify ε_∞ and $\mathscr{A}$, and impose the requirements that $\tau \to 0$ as $\lambda \to 0$ and that the solution pass through the critical point if appropriate (see below).

The solutions of (61.19) to (61.23) fall into two basic classes: (1) transonic critical *winds*, resembling the isothermal solution of type 3 and (2) subsonic *breezes*, similar to the isothermal solutions of type 1. The winds all have $\mathscr{E} > 0$, whereas the breezes all have $\mathscr{E} = 0$. Breeze solutions played an important role in the development of stellar wind theory (**C2**), (**R7**), but will not be discussed further here.

The wind solutions display three distinct asymptotic behaviors of $T(r)$ as $r \to \infty$ [(**D1**), (**D2**), (**D3**), (**H18,** 47), (**P3**), (**R7**), (**W4**)]. The behavior of any particular solution is determined by the dominant heat-transport mechanism as $r \to \infty$. Suppose first that the heat-conduction flux at infinity, $\mathscr{E}_c(\infty)$, remains finite while the enthalpy flux goes to zero. Then $[r^2T^{5/2}(dT/dr)]_\infty =$ constant, which implies $T(r) \propto r^{-2/7}$ asymptotically; this is the kind of solution discussed by Chapman and by Parker (**P3**). Next suppose that both the conduction and enthalpy fluxes go to zero in such a way that their ratio $(nvr^2 5kT)/[r^2K_0T^{5/2}(dT/dr)]$ remains finite as $r \to \infty$. This condition can be achieved if $T(r) \propto r^{-2/5}$ asymptotically, the solution first discovered by Whang and Chang (**W4**). Finally, suppose that the conduction flux vanishes more rapidly than the enthalpy flux as $r \to \infty$; in this case there is no energy exchange within the gas and it expands adiabatically. Thus as $r \to \infty$, $T \propto \rho^{(\gamma-1)} \propto r^{-2(\gamma-1)}$ (because $\rho \propto r^{-2}$); hence $T(r) \propto r^{-4/3}$ for $\gamma = \frac{5}{3}$, the solution first discovered by Durney (**D1**).

In fact, a continuous sequence of solutions exhibiting these different asymptotic behaviors can be obtained by fixing the coronal base temperature T_0 and choosing different values for the base density n_0. For small values of n_0, the critical solutions have large values of $e(\infty) \equiv \mathscr{E}/\mathscr{F}$, the energy flux at infinity per particle, and $T(r) \propto r^{-2/7}$. As n_0 is increased, an ever-larger fraction of $e_c(\infty)$, the conduction flux at infinity, is consumed in driving a more and more massive flow. Eventually at some critical value of n_0, say n_0^*, $e_c(\infty) \to 0$, and $T \propto r^{-2/5}$ for this particular solution. If we increase n_0 still further, $e_c(\infty)$ remains zero and at large r the gas expands adiabatically with $T \propto r^{-4/3}$. As more and more material is added to the flow, more and more thermal energy is consumed in the adiabatic expansion, and $e(\infty)$ decreases. Finally, for a sufficiently large n_0, $e(\infty)$ vanishes, and the solution abruptly changes from a transonic wind to a subsonic breeze.

For wind solutions (only) the further transformation $\tau_* \equiv \tau/e_\infty$, $\psi_* \equiv \psi/e_\infty$,

and $\lambda_* \equiv \lambda/e_\infty$ reduces the number of arbitrary constants to one; equation (61.20) has the same form in the new variables, while (61.21) becomes

$$\mathcal{K}\tau_*^{5/2}(d\tau_*/d\lambda_*) = 1 - \tfrac{1}{2}\psi_* + \lambda_* - \tfrac{5}{2}\tau_* \tag{61.24}$$

where $\mathcal{K} \equiv e_\infty^{3/2}\mathcal{A}$. A large number of solutions for a wide range of $\mathcal{K}$ are given in (**D2**); each of these can be redimensionalized into an infinite number of physical solutions for differing choices of, say, n_0 and T_0. For a typical solar wind model, one finds that near the Earth's orbit the wind speed is $\sim 300\ \mathrm{km\ s^{-1}}$ and the particle density is $\sim 10\ \mathrm{cm^{-3}}$. The rate of mass loss in the wind is $\sim 10^{-14}\,\mathcal{M}_\odot$/year, which is negligible compared to the rate of mass loss via thermonuclear energy release.

TRANSITION TO THE INTERSTELLAR MEDIUM

The above discussion tacitly assumes that the flow expands into a vacuum. In reality, the wind ultimately stagnates against the ambient interstellar medium, forming a stationary shock across which the flow velocity drops suddenly from highly supersonic to subsonic, while both the temperature and density rise sharply. The wind solution thus jumps discontinuously from the critical solution to one of the solutions of type 6 in Figure 61.1; the correct subsonic solution is chosen by matching conditions in the interstellar medium as $r \to \infty$.

By imposing the Rankine-Hugoniot relations across the jump, one can determine the radius r_s at which the shock front is located. To obtain a rough estimate of r_s, we can equate the *impact pressure* $nm_H v_\infty^2$ of the flow to the interstellar gas pressure p_i. Noting that $r^2 n(r) = r_\oplus^2 n(r_\oplus)$ because the flow speed is already essentially v_∞ at $r = r_\oplus$, we find

$$r_s/r_\oplus \approx (n_\oplus m_H v_\infty^2/p_i)^{1/2}. \tag{61.25}$$

For typical values of v_∞, p_i, and $n_\oplus$ one obtains $r_s \gtrsim 30 r_\oplus$.

Accretion flows, in which material falls in from infinity onto a star, are also possible (**H15**); in this case the solution runs inward along the critical solution of type 4 and jumps discontinuously to a solution of type 5 through an *accretion shock* near the surface of the star.

62. Physical Complications

While the model described in §61 gives basic insight into the nature of thermal winds, it is a terribly oversimplified description of the real solar wind. It is therefore worthwhile to mention some of the physical complications that occur in the real solar wind as an introduction to more sophisticated treatments that include phenomena which change the picture substantially, sometimes even qualitatively. Good general reviews are given in (**H13**) and (**H14**).

FLUID PROPERTIES

In §16 we assumed that a thermal wind can be considered to be a steady flow of an equilibrium, inviscid (although heat-conducting) fluid. Each of these assumptions requires scrutiny. Dimensional arguments (**P2**) show that because the physical scale lengths in stellar winds are so large, the Knudsen number is usually (but not always) small enough that the material can be treated as a continuum rather than as individual particles. On the other hand, interparticle collision frequencies are so low (because of low densities) that collisional equilibrium between electrons and protons in the plasma cannot be maintained (**S15**), (**H2**). Instead, we must use a two-fluid model, allowing the electrons and protons to have different temperatures. Indeed, space observations show that near $r_{\oplus}$, $T_e \approx 1.5 \times 10^5$ K while $T_p \approx 4 \times 10^4$ K; the higher electron temperature is maintained by the larger heat flux transported by the electrons. The analytical properties of two-fluid polytropic flows are discussed in (**S17**) and a comprehensive review of two-fluid, solar-wind models is given in (**H11**). A problem with many two-fluid models of the solar wind is that the computed difference between T_e and T_p is much larger than is observed; a number of additional "noncollisional" energy-exchange mechanisms, including a variety of plasma instabilities, have been hypothesized to remedy this problem.

The inclusion of viscous terms has a large impact on the mathematical structure of the stellar wind problem because they eliminate the singularity at the critical point (**W5**), raise the order of the system of differential equations, and admit a richer variety of boundary conditions. Early calculations including viscous terms [e.g., (**S2**)] predicted major changes in the flow; these conclusions are now known to be incorrect because the integration scheme yielded spurious solutions of the Navier-Stokes equations (**S16**). More recent calculations [e.g., (**W8**)] show that viscous terms have relatively little effect on the dynamics of the flow but do contribute to heating the proton component of the plasma.

Thermal conduction by electrons is a major energy-transport mechanism in thermal winds. But the density in the flow is so low that the correctness of the standard Spitzer-Härm conduction coefficient, valid in a collision-dominated plasma, becomes questionable. In the outer parts of the flow one finds that the ratio of the collisional mean free path λ to the scale length l of the temperature gradient exceeds unity. In this regime, the thermal flux predicted by the standard formula $\mathbf{q} \propto -\lambda \nabla T$ may exceed the physical upper bound set by assuming the entire electron thermal energy $\frac{3}{2} n_e k T$ is transported at the mean thermal speed v_{th} of the electrons. We must therefore impose *flux limiting* on the thermal conduction (a problem discussed in the context of radiative energy diffusion at the end of §97). A variety of schemes [e.g., (**P4**), (**H10**), (**W8**)] that inhibit the heat flux when $\lambda \gtrsim l$ have been suggested. These modifications typically improve agreement between theory and observation; nevertheless they are ad hoc, and the difficult problem of calculating accurate transport coefficients in a

collisionless plasma remains to be solved, although progress is being made (**S11**).

MAGNETIC FIELDS

The assumption of spherically symmetric steady flow is also inadequate for the real solar wind. It has been known for some time that the properties of the solar wind measured near the Earth vary over a wide range on time scales of hours to days. Prominent features include strong flare-induced shocks, and high-speed *plasma streams* often observed to recur with a solar rotation period. In the customary picture of the solar wind, these phenomena were viewed as distinct events superposed on a "quiet solar wind" presumably described by models like those discussed in §61. This picture had to be abandoned in the face of X-ray observations, which show the corona to be extremely inhomogeneous and highly structured by magnetic fields (**Z2**). These structures strongly affect the character of the wind emanating from any particular volume element. On one hand, little, if any, wind originates from the hot, dense, magnetically confined *coronal loops*; these structures are cooled mainly by thermal conduction downward into the chromosphere and by radiative losses. On the other hand, high-speed wind streams originate in the *coronal holes*, which are rarefied, relatively cool regions with rapidly diverging magnetic fields that open into interplanetary space; here the wind itself is an important cooling mechanism. Thus the solar wind is not a smooth, steady flow perturbed by occasional "atypical" events. Instead, the observed strongly fluctuating, complex flow is *representative* of the wind, indeed *is* the wind.

Magnetic fields may also strongly affect energy and momentum input into the wind. It is now believed that the dominant coronal heating mechanism is direct dissipation of magnetic energy into the plasma (**V1**), (**L4**). The earlier concept of shock heating seems inconsistent with current observations. Recent work has shown that the dissipation of Alfven waves can deposit substantial energy and momentum in a wind flow; models based on this mechanism seem to provide a satisfactory representation of the winds observed in many late-type stars (**H3**), (**M1**).

More realistic wind models attempt to allow for rapidly diverging flow geometries and nonthermal momentum and energy sources. These phenomena have dramatic effects on the flow. For example, the solutions may exhibit multiple critical points. A sketch of possible topologies for three-critical-point solutions (**H12**) is shown in Figure 62.1; one sees that the flow can become rather complicated. Momentum and energy inputs to the flow can also strongly affect the mass flux and terminal flow speed of a wind. Thus energy input to the supersonic part of the flow increases the terminal flow speed but has no effect on the mass flux (as expected because the mass flux is already fixed in the subsonic part of the flow because information about changes in the supersonic part of the flow cannot propagate upstream). In contrast, energy addition to the subsonic part of

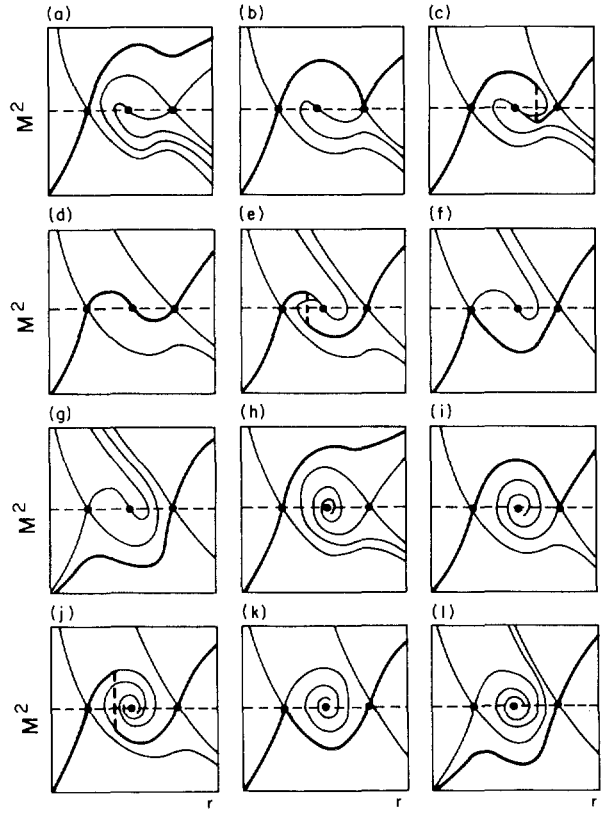

Fig. 62.1 Topology of thermal wind solutions with multiple critical points. From (**H12**) by permission.

the flow increases the mass flux but has little effect on the terminal wind speed. Momentum input in the subsonic part of the flow may actually reduce the terminal wind speed (**L3**).

Magnetic fields can also induce an azimuthal component in the flow in a stellar wind from a rotating star, thereby allowing the wind to exert a torque on the star and to act as a sink of stellar angular momentum (**W1**), (**B9,** §3.7). Whereas the rate of mass loss into the solar wind is negligible, the rate of angular-momentum loss is substantial and is responsible for the Sun's present low rotation rate.

Present research in this area attempts to address fully three-dimensional magnetohydrodynamic flow from a structured corona, a topic that lies far outside the scope of this book.

References

(A1) Appenzeller, I. (1970) *Astron. and Astrophys.*, **5,** 355.

(A2) Athay, R. G. (1976) *The Solar Chromosphere and Corona: Quiet Sun.* Dordrecht: Reidel.

(B1) Becker, E. (1968) *Gas Dynamics.* New York: Academic Press.
(B2) Becker, R. (1922) *Z. fur Physik,* **8,** 321.
(B3) Biermann, L. (1946) *Naturwiss.,* **33,** 118.
(B4) Biermann, L. (1948) *Z. fur Astrophys.,* **25,** 161.
(B5) Bird, G. A. (1964) *Astrophys. J.,* **139,** 684.
(B6) Bird, G. A. (1964) *Astrophys. J.,* **140,** 288.
(B7) Bodenheimer, P. (1968) *Astrophys. J.,* **153,** 483.
(B8) Bond, J. W., Watson, K. M., and Welch, J. A. (1965) *Atomic Theory of Gas Dynamics.* Reading: Addison-Wesley.
(B9) Brandt, J. C. (1970) *Introduction to the Solar Wind.* San Francisco: Freeman.
(B10) Bray, R. J. and Loughhead, R. E. (1974) *The Solar Chromosphere.* London: Chapman and Hall.
(B11) Brinkley, S. R. and Kirkwood, J. G. (1947) *Phys. Rev.,* **71,** 606.
(C1) Castor, J. I., Davis, C. G., and Davison, D. K. (1977) *Los Alamos Scientific Laboratory Report No. LA-6664.* Los Alamos: University of California.
(C2) Chamberlain, J. (1961) *Astrophys. J.,* **133,** 675.
(C3) Chapman, S. (1959) *Proc. Roy. Soc. (London),* **A253,** 450.
(C4) Chernyi, G. G. (1957) *Doklady Akad. Nauk SSR,* **112,** 213.
(C5) Christy, R. F. (1964) *Rev. Mod. Phys.,* **36,** 555.
(C6) Courant, R. and Friedrichs, K. O. (1948) *Supersonic Flow and Shock Waves.* New York: Interscience.
(C7) Courant, R., Friedrichs, K. O., and Lewy, H. (1928) *Math. Ann.,* **100,** 32.
(C8) Cox, A. N., Brownlee, R. R., and Eilers, D. D. (1966) *Astrophys. J.,* **144,** 1024.
(C9) Cox, J. P. and Giuli, R. T. (1968) *Principles of Stellar Structure.* New York: Gordon and Breach.
(C10) Cram, L. E. (1977) *Astron. Astrophys.,* **59,** 151.
(D1) Durney, B. R. (1971) *Astrophys. J.,* **166,** 669.
(D2) Durney, B. R. and Roberts, P. H. (1971) *Astrophys. J.,* **170,** 319.
(D3) Durney, B. R. and Werner, N. E. (1972) *Astrophys. J.,* **171,** 609.
(E1) Eckart, C. (1960) *Hydrodynamics of Oceans and Atmospheres.* Oxford: Pergamon.
(F1) Falk, S. W. and Arnett, W. D. (1977) *Astrophys. J. Suppl.,* **33,** 515.
(F2) Fraley, G. S. (1968) *Astrophys. Space Sci.,* **2,** 96.
(G1) Grover, R. and Hardy, J. W. (1966) *Astrophys. J.,* **143,** 48.
(H1) Harlow, F. H. and Amsden, A. A. (1971) *Los Alamos Scientific Laboratory Report No. LA-4700.* Los Alamos: University of California.
(H2) Hartle, R. E. and Sturrock, P. A. (1968) *Astrophys. J.,* **151,** 1155.
(H3) Hartmann, L. and MacGregor, K. B. (1980) *Astrophys. J.,* **242,** 260.
(H4) Hayes, W. D. (1960) *Gasdynamic Discontinuities.* Princeton: Princeton University Press.
(H5) Hayes, W. D. (1968) *J. Fluid Mech.,* **32,** 305.
(H6) Hayes, W. D. (1968) *J. Fluid Mech.,* **32,** 317.
(H7) Hill, S. J. (1972) *Astrophys. J.,* **178,** 793.
(H8) Hines, C. O. (1960) *Can. J. Phys.,* **38,** 1441.
(H9) Hines, C. O. (1968) Syllabus on *Internal Gravity and Acoustic Waves in Planetary and Solar Atmospheres.* Boulder: University of Colorado.
(H10) Hollweg, J. V. (1976) *J. Geophys. Res.,* **81,** 1649.
(H11) Hollweg, J. V. (1978) *Rev. Geophys. Space Phys.,* **16,** 689.
(H12) Holzer, T. E. (1977) *J. Geophys. Res.,* **82,** 23.

(H13) Holzer, T. E. (1979) in *Solar System Plasma Physics*. ed. C. F. Kennel, L. J. Lanzerotti, and E. N. Parker. p. 101. Amsterdam: North-Holland.
(H14) Holzer, T. E. (1980) in *Cool Stars, Stellar Systems, and the Sun*. SAO Special Report No. 389. ed. A. K. Dupree. p. 153. Cambridge: Smithsonian Astrophysical Observatory.
(H15) Holzer, T. E. and Axford, W. I. (1970) *Ann. Rev. Astron. Astrophys.*, **8,** 31.
(H16) Hoskin, N. E. (1964) in *Methods in Computational Physics*. ed. B. Alder, S. Fernbach, and M. Rotenberg. Vol. **3,** p. 265. New York: Academic Press.
(H17) Hugoniot, A. (1889) *J. de l'Ecole Polytech.*, **58,** 1.
(H18) Hundhausen, A. J. (1972) *Coronal Expansion and Solar Wind*. New York: Springer.
(I1) Imshinnik, V. S. (1962) *Soviet Phys.-JETP*, **15,** 167.
(I2) Israel, W. (1960) *Proc. Roy. Soc. London*, **A259,** 129.
(J1) Jordan, S. D. (1970) *Astrophys. J.*, **161,** 1189.
(K1) Klein, R. I., Stein, R. F., and Kalkofen, W. (1976) *Astrophys. J.*, **205,** 499.
(K2) Klein, R. I., Stein, R. F., and Kalkofen, W. (1978) *Astrophys. J.*, **220,** 1024.
(K3) Kuperus, M. (1969) *Space Sci. Rev.*, **9,** 713.
(K4) Kutter, G. S. and Sparks, W. M. (1972) *Astrophys. J.*, **175,** 407.
(L1) Lamb, H. (1945) *Hydrodynamics*. New York: Dover.
(L2) Landau, L. D. and Lifshitz, E. M. (1959) *Fluid Mechanics*. Reading: Addison-Wesley.
(L3) Leer, E. and Holzer, T. E. (1980) *J. Geophys. Res.*, **85,** 4681.
(L4) Leibacher, J. W. and Stein, R. F. (1981) in *Second Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*. SAO Special Report No. 392. ed. M. S. Gimpapa and L. Golub, p. 23. Cambridge: Smithsonian Astrophysical Observatory.
(L5) Leibacher, J. W. and Stein, R. F. (1981) in *The Sun as a Star*. ed. S. D. Jordan, p. 263. Washington: National Aeronautics and Space Administration.
(L6) Leighton, R. B., Noyes, R. W., and Simon, G. W. (1962) *Astrophys. J.*, **135,** 474.
(L7) Liang, E. P. (1977) *Astrophys. J.*, **211,** 361.
(L8) Lichnerowicz, A. (1967) *Relativistic Hydrodynamics and Magnetohydrodynamics*. New York: Benjamin.
(L9) Lighthill, J. (1978) *Waves in Fluids*. Cambridge: Cambridge University Press.
(M1) MacGregor, K. B. (1981) in *Second Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*. SAO Special Report No. 392. ed. M. S. Giampapa and L. Golub, p. 83. Cambridge: Smithsonian Astrophysical Observatory.
(M2) McKee, C. R. and Colgate, S. A. (1973) *Astrophys. J.*, **181,** 903.
(M3) Mihalas, B. R. W. (1979) *Ph.D. Thesis*, University of Colorado, Boulder.
(M4) Mihalas, B. W. and Toomre, J. (1981) *Astrophys. J.*, **249,** 349.
(M5) Mihalas, B. W. and Toomre, J. (1982) *Astrophys. J.*, **263,** 386.
(N1) Noyes, R. W. and Leighton, R. B. (1963) *Astrophys. J.*, **138,** 631.
(O1) Osterbrock, D. E. (1961) *Astrophys. J.*, **134,** 347.
(O2) Owczarek, J. A. (1964) *Fundamentals of Gas Dynamics*. Scranton: International Textbook Company.
(P1) Parker, E. N. (1958) *Astrophys. J.*, **128,** 664.

(P2) Parker, E. N. (1963) *Interplanetary Dynamical Processes.* New York: Interscience.
(P3) Parker, E. N. (1965) *Astrophys. J.*, **141,** 1463.
(P4) Perkins, F. W. (1973) *Astrophys. J.*, **179,** 637.
(R1) Raizer, Yu. P. (1964) *Zh. Prikl. Math. Tech. Fiz.*, **4,** 49.
(R2) Rankine, W. J. M. (1870) *Phil. Trans. Roy. Soc.*, **160,** 277.
(R3) Rayleigh, J. W. S. Lord (1910) *Proc. Roy. Soc. (London)*, **A84,** 247.
(R4) Richtmyer, R. D. and Morton, K. W. (1967) *Difference Methods for Initial-Value Problems.* (2nd ed.) New York: Interscience.
(R5) Riemann, B. (1953) *Collected Works of Bernhard Riemann.* (2nd ed.) New York: Dover.
(R6) Roberts, P. H. (1971) *Astrophys. Letters,* **9,** 79.
(R7) Roberts, P. H. and Soward, A. M. (1972) *Proc. Roy. Soc. (London)*, **A328,** 185.
(S1) Saito, M. (1964) *Pub. Astron. Soc. Japan,* **16,** 179.
(S2) Scarf, F. L. and Noble, L. M. (1965) *Astrophys. J.*, **141,** 1479.
(S3) Schatzman, E. (1949) *Ann. d. Astrophys.*, **12,** 203.
(S4) Schatzman, E. and Souffrin, P. (1967) *Ann. Rev. Astron. Astrophys.*, **5,** 67.
(S5) Schmitz, F. and Ulmschneider, P. (1980) *Astron. Astrophys.*, **84,** 93.
(S6) Schmitz, F. and Ulmschneider, P. (1980) *Astron. Astrophys.*, **84,** 191.
(S7) Schmitz, F. and Ulmschneider, P. (1981) *Astron. Astrophys.*, **93,** 178.
(S8) Schwarzschild, M. (1948) *Astrophys. J.*, **107,** 1.
(S9) Sedov, L. I. (1959) *Similarity and Dimensional Methods in Mechanics.* New York: Academic Press.
(S10) Shafranov, V. D. (1957) *Soviet Phys.-JETP,* **5,** 1183.
(S11) Shoub, E. C. (1977) *Astrophys. J. Supp.*, **34,** 259.
(S12) Stein, R. F. and Leibacher, J. (1974) *Ann. Rev. Astron. Astrophys,* **12,** 407.
(S13) Stein, R. F. and Schwartz, R. A. (1972) *Astrophys. J.*, **177,** 807.
(S14) Stein, R. F. and Schwartz, R. A. (1973) *Astrophys. J.*, **186,** 1083.
(S15) Sturrock, P. A. and Hartle, R. E. (1966) *Phys. Rev. Letters,* **16,** 628.
(S16) Summers, D. (1980) *Astrophys. J.*, **241,** 468.
(S17) Summers, D. (1982) *Astrophys. J.*, **257,** 881.
(T1) Taub, A. (1948) *Phys. Rev.*, **74,** 328.
(T2) Taub, A. (1978) *Ann. Rev. Fluid Mech.*, **10,** 301.
(T3) Thomas, J. H., Clark, P. A., and Clark, A. (1971) *Solar Phys.*, **16,** 51.
(T4) Thorne, K. S. (1973) *Astrophys. J.*, **179,** 897.
(T5) Tscharnuter, W. M. and Winkler, K.-H. (1979) *Comp. Phys. Comm.*, **18,** 171.
(U1) Ulmschneider, P. (1970) *Solar Phys.*, **12,** 403.
(U2) Ulmschneider, P. (1971) *Astron. Astrophys.*, **12,** 297.
(U3) Ulmschneider, P. (1971) *Astron. Astrophys.*, **14,** 275.
(U4) Ulmschneider, P. and Kalkofen, W. (1977) *Astron. Astrophys.*, **57,** 199.
(U5) Ulmschneider, P., Kalkofen, W., Nowak, T., and Bohn, H. U. (1977) *Astron. Astrophys.* **54,** 61.
(U6) Ulmschneider, P., Schmitz, F., Kalkofen, W., and Bohn, H. U. (1978) *Astron. Astrophys.*, **70,** 487.
(V1) Vaiana, G. S. (1980) in *Cool Stars, Stellar Systems, and the Sun.* SAO Special Report No. 389. ed. A. K. Dupree, p. 195. Cambridge: Smithsonian Astrophysical Observatory.

(V2) Vernazza, J. E., Avrett, E. H., and Loeser, R. (1973) *Astrophys. J.*, **184,** 605.
(V3) Vernazza, J. E., Avrett, E. H., and Loeser, R. (1976) *Astrophys. J. Supp.* **30,** 1.
(V4) Von Neumann, J. and Richtmyer, R. D. (1950) *J. Appl. Phys.*, **21,** 232.
(W1) Weber, E. J. and Davis, L. (1967) *Astrophys. J.*, **148,** 217.
(W2) Weinberg, S. (1971) *Astrophys. J.*, **168,** 175.
(W3) Weymann, R. (1960) *Astrophys. J.*, **132,** 452.
(W4) Whang, Y. C. and Chang, C. C. (1965) *J. Geophys. Res.*, **70,** 4175.
(W5) Whang, Y. C., Liu, C. K., and Chang, C. C. (1965) *Astrophys. J.*, **145,** 255.
(W6) Whitam, G. B. (1974) *Linear and Nonlinear Waves.* New York: Wiley.
(W7) Winkler, K.-H. (1984) in *Astrophysical Radiation Hydrodynamics.* ed. K.-H. Winkler and M. Norman. Dordrecht: Reidel.
(W8) Wolff, C. L., Brandt, J. C., and Southwick, R. G. (1971) *Astrophys. J.*, **165,** 181.
(W9) Wood, P. R. (1974) *Astrophys. J.*, **190,** 605.
(Y1) Yuan, S. W. (1967) *Foundations of Fluid Mechanics.* Englewood Cliffs: Prentice-Hall.
(Z1) Zel'dovich, Ya. B. and Raizer, Yu. P. (1966) *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena.* New York: Academic Press.
(Z2) Zirker, J. B. (1977) *Coronal Holes and High Speed Wind Streams.* Boulder: Colorado Associated University Press.