γ

DOE/ER/03077-240

Courant Mathematics and Computing Laboratory

U.S. Department of Energy

Shock Waves, Increase of Entropy and Loss of Information

Peter D. Lax

Research and Development Report

Supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy under Contract DE-AC02-76ER03077

Mathematics and Computers
October 1984

New York University

DISCLAIMER

Printed in U.S.A.

Available from

National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

UNCLASSIFIED

DOE/ER/03077-240
UC-32
Mathematics and Computers

Courant Mathematics and Computing Laboratory

New York University

DOE/ER/03077--240

DE85 013645

SHOCK WAVES, INCREASE OF ENTROPY AND LOSS OF INFORMATION

Peter D. Lax

October 1984

Supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U. S. Department of Energy under Contract No. DE-AC02-76ER03077

UNCLASSIFIED

21p

Table of Contents

Shock Waves, Increase of Entropy and Loss of Information

By Peter D. Lax[1]

1. Introduction

We present an informal review of the topics in the title as they pertain to solutions of hyperbolic systems of conservation laws. These are systems of the form

$$u_t^i + f_x^i = 0, \; i = 1, \ldots, n; \tag{1.1}$$

the subscripts t and x denote partial derivatives. Each u^i is a density, f^i the corresponding flux. Each f^i is a function of all u^j's, so f_x^i can be expressed as a linear combination of u_x^j. In matrix notation (1.1) can be written as

$$u_t + Au_x = 0 \tag{1.2}$$

where u denotes the column vector with components u^j, and A the matrix whose i^{th} row is the u-gradient of f^i:

$$a_{ij} = \frac{\partial f^i}{\partial u^j} \tag{1.3}$$

The matrix A is a function of u, unless f^i are linear functions of u; in this talk we deal with systems that are genuinely nonlinear, in a sense to be made precise

[1]Research supported under contract DE-AC02-76ER03077 of the Applied Mathematical Sciences Research Program of the Office of Basic Energy Sciences, DOE.

The system (1.2) is called <u>strictly</u> <u>hyperbolic</u> if the matrix A has real and distinct eigenvalues $a_1, a_2, \ldots, a_n$ for all values of u.

We are interested in solving the initial value problem:

(1.4) $u(x,0) = u_0(x)$,

in particular, we want to study the nature of the dependence of solutions on their initial data. We denote by S(t) the operator relating solutions at time t to their initial values:

$$S(t):u(x,0) \to u(x,t).$$

The main facts of life :

(i) The initial value problem has no proper, i.e., differentiable solution for all time t, no matter how smooth the initial data are.

(ii) The initial value problem can be solved for all time if we admit solutions of (1.1) in the integral sense, i.e., in the sense of distributions.

Solutions in the distribution sense that are piecewise continuous satisfy the Rankine-Hugoniot relation

(1.5) $s[u^i] = [f^i]$, $i = 1, \ldots, n$,

where [ ] denotes the difference between the value on the left side and the right side of the discontinuity; s is the velocity with which the discontinuity propagates.

(iii) Solutions in the distributions sense are not determined uniquely by their initial data.

In these notes we shall describe various criteria that are used to accept or reject distribution solutions. The criteria are suggested by physical facts, and are analyzed mathematically. The analysis shows, or at least leads one to expect,

(iv) The various criteria all pick out the same class of distribution solutions; we shall call these relevant solutions.

(v) Each member of the class of relevant solutions is uniquely determined by its initial data. The initial data can be prescribed arbitrarily within the class of L^∞ functions.

We denote by S(t) the operator linking relevant solutions to their initial data.

The main theme of this talk is the following remarkable property of relevant solutions:

(vi) The set of relevant solution is compact; in particular, the operators S(t), t>0, are compact in appropriate pairs of topologies.

Note that (vi) is a nonlinear property; for linear hyperbolic equations S(t) is invertible, in most cases unitary. In the rest of these notes we shall explore how nonlinearity brings about compactness.

Very little is known about the compactness of S(t) in more than one space dimension.

The organization of these notes is as follows:

In Section 2 we discuss, for the simplified model of a single conservation law, the concepts of genuine nonlinearity, breakdown of classical solutions, solutions in the distribution sense and their nonuniqueness, the viscosity method, finite difference methods, and the shock condition.

In Section 3 we discuss, for the scalar model, the compactness of solutions constructed by the viscosity and difference methods, and derive the entropy inequality for such solutions.

In Section 4 we derive Glimm's estimate for the total variation of solutions of scalar equations that satisfy the shock condition, we show that a discontinuous solution that satisfies the shock condition also satisfies the entropy condition.

In Section 5 we indicate how to extend the notions developed in Sections 2, 3, and 4 for systems.

Section 6 contains scattered remarks about the equations of compressible flow: the increase of entropy, some consequences of Carnot's theorem, and the equipartition of energy in the wake of strong shocks.

Compactness of the family of all solutions places a limit on the amount of information contained in that family. It is natural to ask what this implies about the amount of computing needed to solve an initial value problem to meet a specified degree of resolution; see [14] for some crude notions.

[13] contains a bird's eye view of conservation laws; [16] is a thorough treatise.

2. Single Conservation Laws

We study (1.1) for $n = 1$:

$$u_t + f_x = 0, \tag{2.1}$$

f some function of u; we denote

$$\frac{df}{du} = a(u). \tag{2.2}$$

As in (1.2), Equation (2.1) can be rewritten as

$$u_t + a\, u_x = 0; \tag{2.3}$$

this equation asserts that u is constant along trajectories $x = x(t)$ satisfying

$$\frac{dx}{dt} = a. \tag{2.4}$$

In view of this interpretation, a is called the signal velocity; the trajectories defined by (2.4) are called characteristics. Since according to (2.3), u is constant along characteristics, and since a is a function of u, it follows from (2.4) that the characteristics are straight lines.

When the initial value u_0 of u is specified, we can construct through each point $(y,0)$ on the initial line a characteristic line:

(2.5) $x = y + a(u_0(y))t.$

Suppose y_1 and y_2 are two points, $y_1<y_2$, and the inequality

(2.6) $a(u_1)>a(u_2)$,

holds, where

(2.6)' $u_1 = u_0(y_1)$, $u_2 = u_0(y_2)$.

(2.5) shows that the two characteristics issuing from $(y_1,0)$, respectively, intersect at the time

$$T = \frac{y_2 - y_1}{a(u_1) - a(u_2)}$$

As we saw before, u has the value u_1 along the whole characteristic issuing from $(y_1,0)$, and the value u_2 along the characteristics issuing from $(y_2,0)$. So at the point of intersection u has to be equal to both u_1 and u_2; since (2.6) shows that $u_1 \neq u_2$, this is impossible and shows that no differentiable solution $u(x,t)$ exists beyond the time T.

Note that the crucial inequality (2.6) can hold only if $a(u)$ is a genuine function of u, which makes $f(u)$ a genuinely nonlinear function of u. It is convenient to assume that $\frac{d}{du} a \neq 0$ for all u; in view of (2.2) this implies that

(2.7) $\frac{d^2 f}{du^2} \neq 0,$

i.e., that f is strictly convex or concave.

As noted in the introduction, solutions that cannot be continued beyond some critical time T can, nevertheless, be contrived as solutions of (2.1) in the sense of distributions:

$$(2.8) \quad \iint_{t\geq 0} (u\, w_t + f\, w_x)\, dxdt + \int u_0\, w_0 dx = 0$$

for every test function w in C_0^∞. This is consistent with the point of view of physics that conservation laws are integral relations:

$$(2.8)' \quad \frac{d}{dt}\int_G u dx + \int_{\partial G} f \cdot n dS = 0$$

for every domain G. Relations (2.8) can be deduced from relations (2.8) by a simple process of approximation.

As remarked in Section 1, a piecewise continuous solution is a solution in the distribution sense if the Rankine-Hugoniot relation (1.5) is satisfied across the discontinuity;

$$(2.9) \quad s[u] = [f],$$

where [ ] denotes the difference across the discontinuity, and s is the velocity with which the discontinuity propagates.

A simple example of such a discontinuous solution of Equation (2.1), with

$$(2.10) \quad f(u) = u^2/2$$

is

$$(2.11) \quad u_1(x,t) = \begin{cases} 1 & \text{for } x < t/2 \\ 0 & \text{for } t/2 < x. \end{cases}$$

The discontinuity is across the line

$$(2.12) \quad x = t/2$$

which propagates with velocity $s = 1/2$; across the discontinuity $[u] = 1$ $[f] = 1/2$, so relation (2.9) is satisfied.

The function

$$(2.13)\quad u_2(x,t) = \begin{cases} 0 & \text{for } x < t/2 \\ 1 & \text{for } t/2 < x \end{cases}$$

also is discontinuous across the line (2.12), and $[u] = -1$, $[f] = -1/2$, so that relation (2.9) is satisfied. Note, however, that the function

$$(2.14)\quad u_3(x,t) = \begin{cases} 0 & \text{for } x \leq 0 \\ x/t & \text{for } 0 \leq x \leq t \\ 1 & \text{for } t \leq x \end{cases}$$

satisfies Equation (2.1) with f given by (2.10), so that

$$(2.15)\quad u_t + uu_x = 0,$$

in each of the three regions, and is <u>continuous</u> across the boundaries $x = 0$ and $x = t$ separating the regions. Since $u_2(x,0) \equiv u_3(x,0)$, solutions in the distribution sense are <u>not uniquely determined by their initial data</u>. This shows that an additional criterion is needed, based on physical principles and buttressed by mathematical ones, which rejects certain distribution solutions. The remaining acceptable ones must have the property that every initial value problem has exactly one acceptable solution.

There are several ways of formulating such a criterion of rejection or acceptance; happily, they turn out to be equivalent. We list the most important ones:

(a) The acceptable solutions u of (2.1) are the limits of solutions $u^{(\epsilon)}$ of a family of equations obtained by augmenting the flux f by a small viscous term, and letting the viscosity tend to 0:

$$(2.16)\quad u = \lim_{\epsilon \to 0} u^{(\epsilon)}.$$

The viscous term in the flux is $-\epsilon u_x$, $\epsilon>0$, so that $u^{(\epsilon)}$ satisfies the equation

$$(2.17)\quad u_t^{(\epsilon)} + f_x^{(\epsilon)} = \epsilon u_{xx}^{(\epsilon)}, \qquad f^{(\epsilon)} = f(u^{(\epsilon)}),$$

and has the same initial value as u:

$$(2.18)\quad u^{(\epsilon)}(x,0) = u_0(x).$$

(b) The acceptable solutions are limits of solutions v of a difference approximation as Δt, Δx tend to zero. Denoting the value of v at $x = k\Delta x$, $t = n\Delta t$ by v_k^n, the difference equation is of the form

$$(2.19)\quad v_k^{n+1} = \frac{1}{2}\left(u_{k+1}^n + u_{k-1}^n\right) - \left(f_{k+1}^n - f_{k-1}^n\right)\frac{\Delta t}{2\Delta x};$$

here we use the abbreviation

$$f_j^n = f\left(u_j^n\right).$$

The initial values are

$$v_k^0 = u_0(k\Delta x).$$

The rationale for using limits of (2.19) is the close relation between (2.19) and (2.17): this can be seen by using the Taylor approximations

$$v_k^{n+1} = u + \Delta t\ u_t + \frac{1}{2}(\Delta t)^2\ u_{tt} + 0(\Delta^3)$$

$$\frac{1}{2}\left(u_{k+1}^n + u_{k-1}^n\right) = u + \frac{1}{2}(\Delta x)^2\ u_{xx} + 0(\Delta^3)$$

$$\frac{1}{2}\left(f_{k+1}^n - f_{k-1}^n\right) = \Delta x\ f_x + 0(\Delta^3).$$

Setting these into (2.19) and using (2.1) to calculate the higher derivatives of u we get that v approximates solutions of (2.17), with

$$(2.20) \quad \epsilon = \frac{\Delta t}{2}\left[\left(\frac{\Delta x}{\Delta t}\right)^2 - a^2\right]$$

Since ϵ has to be positive, we must have

$$(2.21) \quad \frac{\Delta x}{\Delta t} \geq |a|;$$

this is the celebrated Courant-Friedricks-Lewy convergence criterion.

We remark that (2.19) is merely one of a variety of difference approximations we may use. Another class of approximations, combining viscosity in space and discretization in time, is Avron Douglis' layering method; it is more flexible than either method (a) or (b), see [3].

An entirely different criterion for accepting or rejecting distribution solutions can be based on the analysis of the mechanism that causes the breakdown of smooth solutions: the intersection of characteristics. Thus, the characteristics issuing from the initial line for the initial data of u_1, given by (2.11), cross in a wedge-shaped region

Fig. 1

The role of the discontinuity in (2.11) along $x = t/2$ is to keep the characteristics from crossing. This is in contrast to the behavior of the characteristics issuing from the initial line for the data of u_2, given by (2.13); these diverge and don't cross at all:

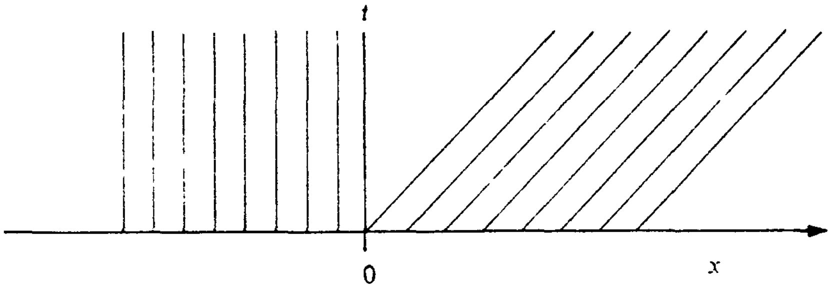


Fig. 2

This leads to the concept of a <u>shock</u>:

A discontinuity of a piecewise continuous distribution solution is called a <u>shock</u> if the characteristics on either side impinge in the forward t direction on the discontinuity. Denoting by u_L and u_R the value of u on the left and right sides of the discontinuity, and by a_L and a_R the corresponding signal velocities:

(2.22) $a_L = a(u_L),\ a_R = a(u_R),$

we can express this condition by the inequality

(2.23) $a_L > s > a_R.$

where s is the velocity of propagation of the discontinuity. We recall from (2.9) that

(2.24) $$s = \frac{f_R - f_L}{u_R - u_L},$$

where

(2.25) $f_L = f(u_L),\ f_R = f(u_R)$.

(c) A distribution solution of (2.1) is acceptable if all its discontinuities are shocks, i.e., satisfy condition (2.23).

3. Viscosity Methods

As a start, let's assume that for arbitrary bounded measurable initial data u_0 of compact support, the solutions $u^{(\epsilon)}(x,t)$ of (2.17), (2.18) exist and converge in the $L^1(dxdt)$ norm over $\mathbb{R} \times (0,T)$, T any value > 0. What can we deduce about the limit u?

Multiplying equation (2.17) by any C_0^2 test function $w(x,t)$ and integrating by parts over all x and $t \geqslant 0$; we get

$$(3.1)\quad -\iint \left[u^{(\epsilon)} w_t + f(u^{\epsilon}) w_x\right] dxdt - \int u_0 w_0 dx = \epsilon \iint u\ w_{xx}\ dxdt.$$

As ϵ tends to zero, $u^{(\epsilon)}$ tends to u and $f(u^{(\epsilon)})$ tends to $f(u)$ in the L^1 norm. Therefore, (3.1) tends to relation (2.8), characterizing distribution solutions; this shows that the L^1 limit of solutions $u^{(\epsilon)}$ of (2.17) is a solution of (2.1) in the distribution sense.

Let $u^{(\epsilon)}$ be a solution of (2.17); since the initial data $u_0(x)$ has compact support, $u^{(\epsilon)}(x,t)$ tends to 0 rapidly as $|x| \to \infty$. So the function

$$(3.2)\quad U^{(\epsilon)}(x,t) = \int_{-\infty}^{x} u^{(\epsilon)}(y,t)dy$$

is well defined and bounded. Clearly,

$$(3.2)'\quad U_x^{(\epsilon)} = u^{(\epsilon)}.$$

Integrating (2.17) with respect to x gives

$$(3.3)\quad U^{(\epsilon)} + f\left[U_x^{(\epsilon)}\right] = \epsilon U_{xx}^{(\epsilon)}.$$

provided that the flux is normalized so that

(3.4) $f(0) = 0.$

Let $v^{(\epsilon)}$ be another solution of (2.17), $V^{(\epsilon)}$ its x-integral, satisfying an analogue of (3.3). Subtracting the two equations from each other, and applying the mean value theorem, we deduce that the difference

(3.5) $D = U^{(\epsilon)} - V^{(\epsilon)}$

satisfies

(3.6) $D_t + \bar{a}D_x = \epsilon D_{xx}$,

where

$$\bar{a} = \frac{f(u^{\epsilon}) - f(v^{\epsilon})}{u^{\epsilon} - v^{\epsilon}}$$

We apply now the <u>maximum</u> <u>principle</u> to the parabolic equation (3.6) and deduce that

(3.7) $|D|_{max}(t) = \underset{x}{Max} \; |D(x,t)|$

is a <u>decreasing</u> function of t.

Clearly,

$$|U^{(\epsilon)}(t)|_{max} \leq \int_{\mathbb{R}} |u^{(\epsilon)}(x,t)| \, dt,$$

so

(3.8) $\int_0^T |U^{(\epsilon)}|_{max}(t) \, dt \leq \int_0^T \int_{\mathbb{R}} |u^{(\epsilon)}(x,t)| \, dxdt.$

By assumption, $u^{(\epsilon)}$ converges to u in the L^1 norm; it follows, therefore, from (3.8) that $U^{(\epsilon)}$ converges to U in the L^1(Max) norm on the left of (3.8). Similarly, $V^{(\epsilon)}$ converges to V; by the triangle inequality, for each t

$$\left| |U^{(\epsilon)}-V^{(\epsilon)}|_{max} - |U-V|_{max} \right| \leq |U^{(\epsilon)}-U|_{max} + |V^{(\epsilon)}-V|_{max}$$

It follows from this that $|U^{(\epsilon)} - V^{\epsilon}|_{max}(t)$ tends to $|U - V|_{max}(t)$ in the $L^1(dt)$ norm. Since by (3.5) and (3.7) the former is a decreasing function of t, and since the L^1 limits of decreasing functions are decreasing, it follows that

$$(3.9) \qquad |U - V|_{max}(t)$$

is a decreasing function of t.

The quantity $|U|_{max}$ is called the $W^{-1,\infty}$ norm u; so property (3.9) can be expressed in terms of the solution operator S(t) relating initial data to data at time t:

<u>Theorem 3.1</u>: <u>The operators</u> S(t) <u>are contractions in the</u> $W^{-1,\infty}$ <u>norm</u>.

Since equation (2.17) is parabolic, another application of the maximum principle shows that

$$\operatorname*{Max}_{x} |u^{(\epsilon)}(x,t)| = |u^{(\epsilon)}|_{max}(t)$$

is a decreasing function of t. We conclude, as before, that $|u^{\epsilon}|_{max}(t)$ tends to $|u|_{max}(t)$ in the L^1 norm, and, therefore, that the latter also is a decreasing function of time. This property can be expressed so:

Theorem 3.2: The operators S(t) map into itself any ball in L^{∞} centered at 0.

Theorems 3.1 and 3.2 derived in [8]; it was observed there that they have the following surprising consequence:

Theorem 3.3: Suppose equation (2.1) is genuinely nonlinear, i.e., the function f is strictly convex or concave. Then the operators S(t), t>0, map any bounded subset of L^{∞} supported on a given interval into a compact subset of L^{1}.

Proof: Let $u_0^{(n)}$ be a uniformly bounded sequence of functions supported on a common interval of the x-axis. Such a set belongs to a compact subset of $W^{-1,\infty}$, i.e., a subsequence of $u_0^{(n)}$ converges to a limit u_0 in the $W^{-1,\infty}$ norm. This limit is also bounded and has compact support. According to Theorem 3.1. for each $t, u^{(n)}(x,t)$ converges in the $W^{-1,\infty}$ norm to $u(x,t)$, where $u^{(n)}$ and u are the solutions of (2.1) constructed by the viscosity method with initial values $u_0^{(n)}$ and u_0, respectively. Both $u^{(n)}$ and u satisfy Equation (2.1) in the distribution sense (2.8): for any C_0^{∞} test function w,

$$(3.10)\quad \iint \left[u^{(n)}w_t + f(u^{(n)})w_x\right] dxdt + \int u_0^{(n)} w_0 dx = 0,$$

and

$$(3.10)'\quad \iint (u\, w_t + f(u)w_x)\, dxdt + \int u_0 w_0 dx = 0$$

Since $u^{(n)}$ tends to u in the $W^{-1,\infty}$ norm, it follows that the first and the third term in (3.10) tends to the first and third term in (3.10)'. Therefore, it follows that so does the second term:

$$(3.11)\quad \iint \left[f(u^{(n)}) - f(u)\right] w_x\, dxdt \to 0$$

Next we make use of the fact that not only for genuine solutions but also for distribution solutions, signals propagate with finite speed.

Since $u_0^{(n)} = u_0 = 0$ outside some finite x-interval, it follows that $u^{(n)}$ and u are zero outside a bounded act in x,t space for $0 \leqslant t \leqslant T$. We choose w so that $w_x \equiv 1$ on this bounded set; since $f(u^n) = f(u)$ outside this set, we can rewrite (3.11) as

$$(3.11)' \quad \iint \left[f(u^{(n)}) - f(u)\right] dxdt \to 0.$$

We have assumed that f is strictly, convex, say $f''(u) \geqslant R > 0$. Then

$$(3.12) \quad f(u^{(n)}) - f(u) \geqslant f'(u)(u^{(n)} - u) + \frac{R}{2}(u^{(n)} - u)^2$$

Integrating this we get

$$(3.12)' \quad \iint \left[f(u^{(n)}) - f(u)\right] dxdt \geqslant \iint \left[f'(u)\,(u^{(n)} - u)\right] dxdt$$
$$+ \frac{R}{2} \iint \left[u^{(n)} - u\right]^2 dxdt$$

Since $u^{(n)} - u$ tends to zero in the $W^{-1,\infty}$ norm, and since $u^{(n)} - u$ is uniformly bounded, it follows readily, by approximating $f'(u)$ in the L^1 sense by smooth functions, that the first term on the right in (3.12)' tends to zero. Since by (3.11)' the left side tends to zero, we deduce from (3.12)' that

$$(3.13) \quad \int_0^T \int (u^{(n)} - u)^2 \, dxdt \to 0.$$

Since $u^{(n)} - u$ is supported on a bounded set, it follows that

$$(3.13)' \quad \int^T \int |u^{(n)} - u| \, dxdt \to 0.$$

This proves $L^1(dxdt)$ convergence of $u^{(n)}$

To prove $L^1(dx)$ convergence, we appeal to the following theorem of Barbara Keyfitz [7]:

If u and v are two distribution solutions of (2.1) that are limits of solutions of (2.17), then

$$\int |v(x,t) - u(x,t)| \, dx$$

is a decreasing function of time. This result applies in particular to $v = u^{(n)}$; from this and (3.13)' we conclude that

$$\int |u^{(n)}(x,t) - u(x,t)| \, dx \to 0$$

for each $t>0$. This completes the proof of Theorem 3.3.

It was already observed in [8] that Theorems 3.1, 3.2 and 3.3 hold also for solutions constructed by the difference scheme (2.19); the same is true when solutions are constructed by more general difference schemes, as long as they are of monotone type.

A basic hypothesis of Theorems 3.1-3.3 is that the solutions $u^{(\epsilon)}$ of the parabolic equation (2.17) converge to a limit u in the L^1 norm for all bounded initial data u_0. How does one prove such a result? In the case (2.10) of a quadratically nonlinear f, E. Hopf, used the fact that in this case Equation (3.3) is changed by the transformation

$$U^{(\epsilon)} = -2\epsilon \log V^{(\epsilon)}$$

into

$$V_t^{(\epsilon)} = \epsilon V_{xx}^{(\epsilon)}, \quad V(x,0) = \exp \{U_0^{(\epsilon)}/2\epsilon\}$$

A solution of the heat equation can be expressed as an integral of its initial values. Using this formula, Hopf was able to show that as $\epsilon \to 0$, $u^{(\epsilon)}$ tends to a limit u; he even obtained a fairly explicit formula for this limit. It was remarked in [9] that a version of Theorem 3.3 can be derived from this formula.

The convergence of solutions of the difference scheme (2.19) can be proved in a similar fashion for the special choice $f(u) = \log (a + be^{-u})$, and of other monotonic schemes for other special choices.

Needless to say, these methods are very special. Recently, see [15], [17] a new method, capable of dealing with more general cases, has been introduced by L. Tartar and F. Murat; we give a brief description of their ideas.

The first idea, going back to L. C. Young, is a precise description of weak convergence. Let $u^{(\epsilon)}(y)$ be a uniformly bounded sequence of functions; then it has a subsequence with the property that for <u>every</u> continuous function g the weak* limits in the sense of L^∞ exist:

$$(3.14) \quad g(u^{(\epsilon)}) \rightharpoonup u_g.$$

Clearly, the weak* limits u_g depend linearly and positively on g. It is not hard to show that u_g can be represented as an integral of g with respect to a family of probability measures $\nu(y)$:

$$(3.15) \quad u_g(y) = \int g(v)\, d\nu(v,y) \equiv \langle g, \nu(y) \rangle.$$

Secondly, it is not hard to show, see (3.12)', that (3.14) is convergence in the L^1 sense if the measure $\nu(y)$ is concentrated at a single point for almost all values of y. In this case

$$(3.16) \quad u_g = g(u)$$

for every g, u being the limit of $u^{(\epsilon)}$.

Tartar takes for $u^{(\epsilon)}$ the solutions of (2.17) with prescribed initial values. By Theorem 3.2 this sequence is uniformly bounded; therefore, it has a subsequence for which (3.14), (3.15) holds. Tartar shows that in this case ν is concentrated at a single point by using the notion of

<u>Compensated</u> <u>Compactness</u> (Tartar):

Let $h^{(\epsilon)}$ and $k^{(\epsilon)}$ be two sequences of vector functions defined in some domain of y-space, satisfying the following conditions:

(i) $h^{(\epsilon)}$ and $k^{(\epsilon)}$ are uniformly bounded in L^2, and converge weakly in L^2:

(3.17) $h^{(\epsilon)} \rightharpoonup h,\ k^{(\epsilon)} \rightharpoonup k.$

(ii) $-\mathrm{div}\ h^{(\epsilon)}$ and curl $k^{(\epsilon)}$ belong to compact subsets of H^{-1}_{loc}.

<u>Conclusion</u>: The scalar product of $h^{(\epsilon)}$ and $k^{(\epsilon)}$ tend in the distribution sense to

(3.18) $h^{(\epsilon)} \cdot k^{(\epsilon)} \rightharpoonup h \cdot k.$

Take y to be t, x, the domain to be $(0,T) \times [x_1,x_2]$ and the vector function $h^{(\epsilon)}$ to be

(3.19) $h^{(\epsilon)} = (u^{(\epsilon)},\ f^{\epsilon}).$

where $f^{(\epsilon)} = f(u^{\epsilon})$. Using Equation (2.17) we see that

(3.20) $\mathrm{div}\ h^{(\epsilon)} = u_t^{(\epsilon)} + f_x^{(\epsilon)} = \epsilon\ u_{xx}^{\epsilon}$

Clearly, the H^{-1}_{loc} norm of div $h^{(\epsilon)}$ is bounded by $\epsilon\ \|u_x^{\epsilon}\|$, where $\|\ \|$ denotes L^2 norm. This quantity can be estimated by multiplying (2.17) by $u^{(\epsilon)}$ and integrating over $\mathbb{R} \times (0,T)$. Since $uf_x = uf'(u)u_x$ is a perfect x derivative, and $u^{(\epsilon)}$ tends to zero as $|x| \to \infty$, we get, after integrating byparts that

(3.21) $$\frac{1}{2} \int |u^{(\epsilon)}|^2 dx \Big|_0^T = -\epsilon \int_0^T |u_x^{(\epsilon)}|^2 dxdt$$

It follows that

(3.21)' $$\epsilon\ \|u_x^{(\epsilon)}\|^2 \equiv \epsilon \int_0^T \int |u_x^{(\epsilon)}|^2 dxdt \leq \text{const} = \frac{1}{2} \int u_0^2 dx.$$

Clearly, $\epsilon \|u_x^{(\epsilon)}\| \leq \epsilon^{1/2}$ const tends to zero as $\epsilon \to 0$; this shows that div $h^{(\epsilon)}$ belongs to a compact set in H^{-1}_{loc}.

To construct $k^{(\epsilon)}$ we take any C^2 function $\eta(v)$, and define $\varphi(v)$ by

$$\varphi(v) = \int_{v_0}^{v} \eta'(u) f'(u) \lambda u.$$

Then

(3.22) $\varphi' = \eta' f'$

Multiply (2.17) by η'; using (3.22) we can write the resulting equation as

(3.23) $\eta_t^{(\epsilon)} + \varphi_x^{(\epsilon)} = \epsilon \eta' u_{xx}^{(\epsilon)} = \epsilon \eta_{xx}^{(\epsilon)} - \epsilon \eta'' u_x^{(\epsilon)^2}$

We take now the vector function $k^{(\epsilon)}$ to be

(3.24) $k^{(\epsilon)} = \left\{\varphi^{(\epsilon)}, -\eta^{(\epsilon)}\right\}$

Using (3.23) we see that

(3.25) curl $k^{(\epsilon)} = \eta_t^{(\epsilon)} + \varphi_x^{(\epsilon)} = \epsilon \eta_{xx}^{(\epsilon)} - \epsilon \eta'' u_x^{(\epsilon)^2}$.

We claim that the right side lies in a compact set in H^{-1}_{loc}. Clearly, the H^{-1}_{loc} norm of the first term $\epsilon \eta_{xx}^{(\epsilon)}$ is bounded by

$$\epsilon \|\eta_x^{(\epsilon)}\| = \epsilon \|\eta' u_x^{(\epsilon)}\|$$

It follows, as before, from (3.21)' that this tends to zero. It follows from (3.21)' that the second term in (3.23), $\epsilon \eta'' u_x^{(\epsilon)^2}$, is bounded in L^1. This does not imply H^{-1}_{loc} compactness; to get that we note that since by Theorem 3.2, $u^{(\epsilon)}$ is uniformly bounded in L^∞, it follows that so are $\eta^{(\epsilon)}$ and $\varphi^{(\epsilon)}$. It follows that curl $k^{(\epsilon)} = \eta_t^{(\epsilon)} + \varphi_x^{(\epsilon)}$ belong to a bounded set in $W^{-1,\infty}$. Now Tartar appeals to

<u>Murat's</u> <u>Lemma</u>:

Suppose the set of functions $\{c\}$ satisfies these conditions:

(a) $\{c\}$ belongs to a bounded set in $W^{-1,\infty}$.

(b) Each c can be decomposed as

$$c = c_1 + c_2$$

where $\{c_1\}$ belongs to a compact set in H^{-1}_{loc}, and $\{c_2\}$ to a bounded set in L^1.

<u>Assertion</u>: The set $\{c\}$ belongs to a compact subset of H^{-1}_{loc}.

Having verified the hypotheses of compensated compactness, we can conclude as in (3.18) that

(3.26) $u^{(\epsilon)}\varphi^{(\epsilon)} - f^{(\epsilon)}\eta^{(\epsilon)} \rightharpoonup u\varphi(u) - f(u)\eta(u)$

We use now formula (3.15) describing the weak limits of functions of $u^{(\epsilon)}$; taking $g(v) = v\varphi(v) - f(v)\eta(v)$ we see that the weak limit of the left side of (3.26) is

$$<v\varphi - f\eta, \nu>$$

The various terms on the right of (3.26) can be similarly described; since the two sides of (3.26) are equal, we conclude that

(3.27) $<v\varphi - f\eta,\nu> = <v,\nu><\varphi,\nu> - <f,\nu><\eta,\nu>$, where $\nu = \nu(x,t)$; (3.27) holds for a.a.(x,t).

We introduce the abbreviations

$$(3.28) \quad u = \langle v, \nu\rangle, \qquad \bar{f} = \langle f(v), \nu\rangle;$$

then (3.27) can be rewritten as

$$(3.29) \quad \langle (v-u)\varphi, \nu\rangle = \langle (f-\bar{f})\eta, \nu\rangle.$$

In the derivation of (3.27) we used second derivatives of η, but (3.27) itself depends continuously on η in the C topology; therefore, (3.37) remains true for η piecewise C^1. We choose

$$(3.30) \quad \eta(v) = |v-u|;$$

then from (3.22)

$$(3.30)' \quad \varphi(v) = \begin{cases} f(v) - f(u) & \text{for } u>v \\ f(u) - f(v) & \text{for } u<v \end{cases}$$

Setting these choices for η and φ into (3.29) gives

$$\langle |v-u|(f-f(u)), \nu\rangle = \langle (f-\bar{f})|v-u, \nu\rangle;$$

we deduce from this that

$$(3.31) \quad (f(u)-\bar{f}) \langle |v-u|, \nu\rangle = 0.$$

For f strictly convex, it follows from (3.28) and Jensen's inequality that the first factor $(f(u)-\bar{f})$ in (3.31) is positive unless ν is concentrated at the single point u; the same, of course, is true for the second factor $\langle |v-u|, \nu\rangle$. Thus, it follows from (3.31) that ν is concentrated at the single point u, and so $u^{(\epsilon)}$ tends to u in the L^1

sense. It follows then, as shown at the beginning of this section, that $u = \lim u^{(\epsilon)}$ satisfies (2.1) in the sense of distributions.

The argument outlined above shows that for every sequence of $\epsilon \to 0$ we can select a subsequence such that the solutions $u^{(\epsilon)}$ of (2.17) with prescribed initial value u_0 tend in the L^1 sense to a distribution solution u of (2.1) with initial value u_0. To prove that $\lim_{\epsilon \to 0} u^{(\epsilon)}$ exist, we have to show that any two subsequences have the same limit. For this we need the following characterization of such limits. see [11]:

<u>Theorem</u> <u>3.4</u>: Let u be the L^1 limit of a subsequence $u^{(\epsilon)}$ of solutions of (2.27). Let η be any <u>convex</u> function, and φ related to η by (3.22). Then

(3.32) $\eta(u)_t + \varphi(u)_x \leqslant 0$

in the sense of distribution.

The proof follows from (3.23); for when η is convex, $\eta'' \geqslant 0$, and so (3.23) implies

$$\eta_t^{(\epsilon)} + \varphi_x^{(\epsilon)} \leqslant \epsilon \eta_{xx}^{(\epsilon)}.$$

(3.32) is the limit in the distribution sense of this relation as $\epsilon \to 0$.

Condition (3.32) is called an <u>entropy</u> <u>condition</u>; this notion will be elaborated in Sections 5 and 6.

It can be shown that distribution solutions of (2.1) that satisfy the entropy conditions (3.32) are <u>uniquely</u> <u>determined</u> by their initial data. This proves that the L' limits of two different subsequences of $u^{(\epsilon)}$ are the same; this completes the proof of convergence.

Tartar observed that the argument above for the L^1 convergence of $u^{(\epsilon)}$ also proves the compact dependence of solutions u on their initial data. For let $u_o^{(n)}$ be a uniformly bounded sequence of data with common support, converging in the weak* topology of L^∞ to u_0; denote by $u^{(n)}$ the corresponding solutions of (2.1) constructed by the viscosity method. A subsequence can be selected so that for every continuous g, $g(u^{(n)})$ converges in the weak* topology of L^∞. Again, we look at the two vector valued functions

$$h^{(n)} = (u^{(n)}, f^{(n)}) \qquad \text{and} \qquad k^{(n)} = (\varphi^{(n)}, -\eta^{(n)})$$

By (2.1),

$$\operatorname{grad} h^{(n)} = 0,$$

and by taking the limit of (3.5)

$$\operatorname{curl} k^{(n)}$$

are measures, uniformly bounded. Using compensated compactness we conclude as before that $u^{(n)}$ converges in the L^1 norm to a limit u, a solution of (2.1) with initial value u_0. Since each $u^{(n)}$ satisfies the entropy inequality (3.32), taking the distribution limit of (3.32) shows that so does the limit u. Then by the uniqueness theorem quoted above we conclude that u is the solution of (2.1) with initial value u_0 obtained by the viscosity method. Thus, the L^1 limits of two different subsequences of $u^{(n)}$ are the same; this completes the proof of the compactness of the mapping $u_0 \to u$ from L^∞ to L^1.

We close this section by remarking that compactness is a property one usually associates with the dependence of solutions of parabolic equations such as (2.17) on their initial data; unlike in the linear case, this property is preserved as ϵ tends to zero. It is instructive to see in some detail how this happens. From (3.21) we get that

$$(3.33) \quad \frac{d}{dt} \int \frac{|u^{(\epsilon)}|^2}{2} dx = -\epsilon \int |u^{(\epsilon)}|^2 dx,$$

i.e., that $\epsilon \int u_x{}^2 dx$ is the <u>rate of energy dissipation</u>. We saw earlier that the limit u of $u^{(\epsilon)}$ has discontinuities; it can be shown that, for ϵ small, $u^{(\epsilon)}$ bridges this discontinuity in a layer $\ell(\epsilon)$ of width ϵ; in this layer it has the shape of a traveling wave:

$$(3.34) \quad u^{(\epsilon)}(x,t) \simeq w\left(\frac{x-st}{\epsilon}, t\right),$$

where $w(\xi)$ is a solution of

$$-s\dot{w} + f'\dot{w} = \ddot{w}, \qquad \dot{} = \frac{d}{d\xi}.$$

Differentiating (3.34) we see that

$$\epsilon \int_{\ell(\epsilon)} |u_x^{(\epsilon)}|^2 dx \simeq \frac{1}{\epsilon} \int \dot{w}\left(\frac{x-st}{\epsilon}\right)^2 dx = \int \dot{w}(\xi)^2 d\xi .$$

This shows that the <u>rate of energy dissipation does not tend to zero as</u> ϵ tends <u>to zero</u>. This is in sharp contrast to the linear case:

$$u_t^{(\epsilon)} = \epsilon u_{xx}^{(\epsilon)},$$

when $u^{(\epsilon)}$ tends to a discontinuous limit. Here the transition layer has width $\epsilon^{1/2}$, and the shape of the wave is

$$u^{(\epsilon)}(x,t) \simeq w\left(\frac{x}{\epsilon^{1/2}}, t\right).$$

So the rate of energy dissipation

$$\epsilon \int_{\ell} |u_x^{(\epsilon)}|^2 \simeq \int \dot{w}\left(\frac{x}{\epsilon^{1/2}}\right)^2 dx = \epsilon^{1/2} \int \dot{w}^2 d\xi$$

tends to zero as ϵ tends to zero.

4. Consequences of the Shock Condition

In this section we study those distribution solutions of (2.1) all of whose discontinuities are shocks, i.e., satisfy condition (2.23). We present a result of James Glimm that shows that such solutions form a compact set.

The shock condition relates the signal velocities a_L and a_R to the left and right of the discontinuity with the velocity s with which the discontinuity propagates:

$$(4.1) \quad a_L > s > a_R.$$

(4.1) expresses the fact that characteristics, i.e., curves propagating with signal speed drawn in the forward direction, intersect the shock. It follows from this that characteristics drawn in the <u>backward</u> direction cannot intersect any shock. It follows from this that in a solution all whose discontinuities are shocks, every point (x,t) can be connected by a characteristic to a point on the initial line.

For simplicity we shall study solutions u(x,t) whose initial data--and, therefore, themselves--are periodic functions of x with period L:

$$u(x + L,t) = u(x,t).$$

We shall estimate the total variations of u(x,t) with respect to x per period. We start by estimating the total variation of $a = a(u(x,t))$; the total variation is the sum of the increasing variation A^+ and the decreasing variation A^-; for a period function the two are the same:

$$(4.2) \quad \text{Total variation of a per period at time } t = 2A^+(t)$$

To estimate the increasing variation of a we note that according to the shock condition (4.1), the discontinuities of a contribute only to the decreasing variation of a. Therefore, we can calculate $A^+(t)$ by dividing an interval of length L at t into subintervals by points

$$x_0(t) < x_1(t) < \cdot < x_N(t) = x_0(t) + L$$

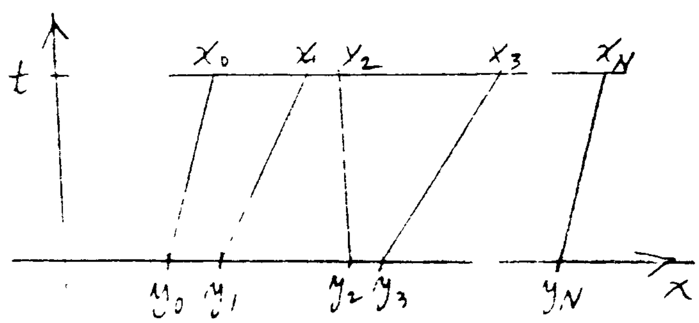
so that a is alternately increasing and decreasing along each interval. Then

(4.3) $\quad A^+(t) = \Sigma\, a_{2n+1} - a_{2n},$

where

(4.4) $\quad a_j = a(u(x_j,t)).$

We connect now the points x_j,t to points y_j on the initial line by characteristics:



By (2.5)

(4.5) $\quad x_j(t) = y_j + a_j t$

Since $x_N = x_0 + L$,

$$\Sigma\, x_{2n+1}(t) - x_{2n}(t) \leq L$$

Using formula (4.5) for x_j we get

$$\Sigma\, y_{2n+1} - y_{2n} + t\, \Sigma\, a_{2n+1} - a_{2n} \leq L$$

Since the first sum is positive, we conclude, using formula (4.3), that

$$A^{+}(t) \leqslant \frac{L}{t}.$$

So by (4.2) we conclude that

$$\text{(4.6)} \quad \text{Total variation of } a(u(x,t)) \text{ per period} \leqslant \frac{2L}{t}$$

We use now the fact that (2.1) is genuinely nonlinear, i.e., that a is nonconstant function of u. Denote R a lower bound for $|a'(u)|$; then Total Var $a(u) \geqslant R$ Total Var u, so we conclude from (4.6):

<u>Theorem</u> <u>4.1</u> <u>(Glimm)</u>: Suppose equation (2.1) is genuinely nonlinear in the sense that

$$\text{(4.7)} \quad |f''(u)| \geqslant R.$$

Let u be a distribution solution of (2.1) all whose discontinuities are shocks, and which is periodic in x with period L. Then the total variation of u(x,t) in x per period is

$$\text{(4.8)} \quad \leqslant \frac{2L}{Rt}.$$

It follows from (4.8) that for any $t>0$, the measures $u_x(x,t)$ are uniformly bounded and, therefore, weakly compact; it is remarkable that nothing need be assumed about the initial data.

We can obtain a compactness result about the u themselves if we can get bound for their integral. This is easy because of the conservation form of the equation: if we integrate (2.1) with respect to x, we deduce that

$$\text{(4.9)} \quad \int_0^L u(x,t)dx$$

is independent of t.

There is an analogue of (4.8) for solutions that are not periodic in x but of compact support. The total variation of $u(x,t)$ in x is less than

$$(4.8)' \quad \frac{C}{Rt^{1/2}},$$

where the constant depends on the length of the support of the initial data but on nothing else, see [12].

We conclude this section by showing that solutions which satisfy the shock condition (4.1) satisfy the entropy condition (3.32) as well, and conversely. Suppose that f is convex, so that a is an increasing function of u; then (4.1) is equivalent with

$$(4.10) \quad u_L > u_R.$$

For a piecewise continuous solution, the left side of (3.32) is zero except along a discontinuity $x = y(t)$, where it has the value

$$(4.11) \quad (\eta_L - \eta_R)s - (\varphi_L - \varphi_R)$$

times $\delta(x-y(t))$; here s denotes the shock velocity

$$s = \frac{dy}{dt}.$$

Since by (4.10), $u_R < u_L$, we can write (4.11) as

$$(4.11)' \quad \int_{u_R}^{u_L} (\eta' s - \varphi')dw.$$

Using relation (3.22) this can be written as

$$(4.11)'' \quad \int_{u_R}^{u_L} \eta'(s - f')du$$

We recall now the Rankine-Hugoniot condition (2.9):

$$s(u_L-u_R) = f_L-f_R;$$

this can be expressed by saying that

$$f(u) - su$$

has the same value at u_L and u_R; denoting this common value by g we can integrate (4.11)" by parts to obtain

$$(4.11)''' \qquad \int_{u_R}^{u_L} \eta'' \,(f - su - g)dn$$

Since η is assumed convex, $\eta'' > 0$; since f is assumed convex, it lies below any secant, i.e.

$$f < su + g$$

for u between u_R and u_L. This shows that (4.11)''' is < 0; thus, so is (4.11), and the entropy condition (3.32) is fulfilled. The proof of the converse is the same.

5. Systems of Conservation Laws

In this section we note very briefly the extension of the ideas in Sections 2, 3 and 4 to systems discussed in Section 1. First of all, we have to assume that the matrix A given by (1.3) and appearing in Equation (1.2),

$$(5.1) \quad u_t + Au_x = 0,$$

has real and distinct eigenvalues $\alpha_1<\alpha_2<\ldots<\alpha_n$; this makes equation (5.1) hyperbolic. The eigenvalues α are functions of u, and so are the eigenvectors r:

(5.2) $Ar = \alpha r$.

Genuine nonlinearity, see [8], is defined to be

(5.3) $\alpha_j' \cdot r_j \neq 0$

for $j = 1, 2, \ldots, n$: here ' denotes the gradient with respect to u.

The curves that satisfy

(5.4) $\dfrac{dx}{dt} = \alpha_j$

are still called characteristics, but they are no longer straight lines in general.

A function w of u is caled a <u>Riemann</u> <u>invariant</u> for the j^{th} field if i for any solution u of (5.1), w(u) is constant along characteristics of the j^{th} field. It follows easily that this is the case if w', the gradient of w, is a left eigenvector of A:

(5.5) $w'A = \alpha_j w'$.

For then multiplying (5.1) by w' gives

(5.5)' $w_t + \alpha_j w_x = 0$.

Since eigenvectors can be rescaled, system (5.5) can always be solved when $n = 2$; for $n>2$ solutions exist only in exceptional cases.

As in the scalar case, solutions in the classical sense cannot be continued beyond a finite time, see [10] and [6], so again we have to turn to distribution solutions. As before, a piecewise continuous solution is a distribution solution if the Rankine-Hugoniot condition (2.9) is satisfied for all n components. Again, the class of distribution solutions is too broad, and has to be narrowed by imposing some criterion of acceptance and rejection. The ones that make physical and mathematical sense are the same ones that were employed for scalar equations:

(a) The acceptable solutions are limits of solutions of a family of equations obtained by augmenting the flux by a small viscous term, and letting the viscosity tend to zero. The simplest way of augmenting the flux leads to the parabolic systems

(5.6) $u_t + f_x = \epsilon u_{xx}$,

or more generally

(5.6)' $u_t + f_x = \epsilon D u_{xx}$,

D some symmetric, positive matrix. We remark that physical viscosity and heat conductivity lead to a matrix D that is merely nonnegative, and is a function of u. For this reason equation (5.6) is said to contain <u>artificial viscosity</u>.

In [11], an entropy condition is formulated for limits of solutions of (5.6). Two functions $\eta(u)$ and $\varphi(u)$ are needed, with these properties:

(i) η and φ satisfy the differential equation

(5.7) $\eta' A = \varphi'$,

where η' and φ' denoted gradients with respect to u.

(ii) $\eta(u)$ is strictly convex

It follows from (i) that if $u^{(\epsilon)}$ satisfies (5.6), $\eta^{(\epsilon)} = \eta(u^{\epsilon})$ and $\varphi^{(\epsilon)} = \varphi(u^{(\epsilon)})$ satisfy the vector analogue of (3.23). From this relation and (ii) we deduce that the L^1 limit u of $u^{(\epsilon)}$ satisfies the entropy inequality (3.32).

For $n = 2$, equation (5.7) has many convex solutions. For $n > 2$, (5.7) is an overdetermined system which only exceptionally has a solution; these exceptional cases happily include most of physical interest, see Section 6.

Theorem 3.1 has no analogy for systems; nor does E. Hopf's trick of linearizing (5.6) work for $n > 1$. Happily Tartar's approach is not tied to $n = 1$, and in fact Ron Di Perna succeeded in extending it to significant cases for $n = 2$, see [1] and [2].

(b) The finite difference approach analogous to (2.19) can be set up, as well as other more general ones. These work very well in practice--some better than others--but it is very hard to prove anything rigorously about their convergence. In [5] Glimm succeeded in proving the a.e. convergence of a scheme that is the mixture of a discrete scheme due to Godunov and a Monte Carlo-type scheme.

(c) The notion of a shock can be extended to piecewise continuous solutions of hyperbolic systems of conservation laws, as follows. In place of condition (4.1) we require that there be an index k such that

$$\begin{aligned} &a_k(u_L) > s > a_{k-1}(u_L), \\ &a_{k+1}(u_R) > s > a_k(u_R) \end{aligned} \tag{5.8}$$

Thus, there are n distinct families of shocks, corresponding to $k = 1, \dots, n$.

In a k-shock, $n - k + 1$ characteristics impinge from the left on the line of discontinuity, and k from the right, altogether $n + 1$. The information carried by these characteristics combined with the Rankine-Hugoniot relations (2.9) serve to determine uniquely the solution on either side of the curve of discontinuity, and determine the curve itself.

Glimm's method for estimating the total variation of solutions sadisfying the shock condition, described in Section 4, has been extended in [5] to solutions with small oscillation of pairs of conservation laws.

It was shown in [11] that a piecewise continuous solution that satisfies the shock condition also satisfies the entropy condition, at least for sufficiently weak shocks.

6. Thermodynamics and Gas Dynamics

There are many thermodynamical variables; some of them, like

Density, ρ

Temperature, T

Pressure, p

Soundspeed, c

are quantities palpable in everyday experience. Others, such as

Internal Energy, e

Entropy, S

Enthalpy, i

per unit mass, are theoretical constructs. In a single-phase system at thermodynamic equilibrium any three are functionally related, i.e., are linked by an equation called equation of state. Any of the variables can be expressed as function of any other two.

The equations governing the flow of compressible gas are the laws of conservation of mass momentum and energy. Let ρ, m and E denote mass, momentum and total energy per unit volume, u the velocity of the flow; set

$$m = \rho u, \quad E = \rho e + \frac{1}{2}\rho u^2. \tag{6.1}$$

The conservation laws for flows depending on a single space variable x are:

$$\begin{aligned} &\rho_t + m_x = 0 \\ &m_t + (um + p)_x = 0 \\ &E_t + (u(E + p))_x = 0. \end{aligned} \tag{6.2}$$

It is not hard to show that this system is hyperbolic; the three signal velocities are

(6.3) $\quad a_1 = u - c,\ a_2 = u,\ a_3 = u + c,\ c = \left(\frac{\partial p}{\partial \rho}\right)^{1/2}.$

a_1 and a_3 are genuinely nonlinear in the sense of (5.3). The signal velocity $a_2 = u$ is linearly degenerate; the corresponding characteristic curves are particle paths.

It follows that if equations (6.2) are satisfied at every point, then

(6.4) $\quad S_t + uS_x = 0.$

In words: in a smooth flow entropy per unit mass is constant along particle paths.

Combining equation (6.4) with $(6.2)_1$ we deduce the conservation law

(6.5) $\quad (\rho S)_t + (u\rho S)_x = 0$

for entropy per unit volume. It is not hard to show that ρS is a <u>concave</u> function of ρ, m and E; since (6.5) is a consequence of (6.1), the pair

$$\eta = \rho S \text{ and } \varphi = u\rho S$$

must satisfy relation (5.7). Thus, it follows that distribution solutions of (6.1) which are the limits of artificially viscous flows (5.6) satisfy the entropy inequality (3.32):

(6.6) $\quad (\rho S)_t + (u\rho S)_x \geqslant 0.$

We have seen in equation (4.11) that for a piecewise continuous solution relation (6.6) means that across each discontinuity

$$[(\rho S)_L - (\rho S)_R]\ s - (u\rho S)_L + (u\rho S)_R \geqslant 0$$

Rearranging turns this into

(6.7) $\quad (s - u_L)\rho_L S_L - (s - u_R)\rho_R S_R \geqslant 0.$

According to the Rankine-Hugoniot relation applied to the conservation of mass equation $(6.1)_1$,

$$s(\rho_L - \rho_R) = (u\rho)_L - (u\rho)_R,$$

which can be rearranged as below:

(6.8) $\quad (s - u_L)\rho_L = (s - u_R)\rho_R.$

We turn now to the shock relations (5.8); for k = 1 we get from (5.8), using (6.3), that

$(6.9)_1 \quad u_L > s, \qquad u_R > s$

while for k = 3 we get that

$(6.9)_3 \quad u_L < s, \qquad u_R < s$

Thus in case k = 1, (6.8) is negative, in case k = 3, (6.8) is positive. Setting this into (6.7) we conclude that for k = 1

$(6.10)_1 \; S_L < S_R,$

while for k = 3,

$(6.10)_3 \; S_L > S_R.$

It follows from $(6.9)_1$ that in case k = 1, particles cross the shock from left to right, while $(6.9)_3$ says that in case k = 3 particles cross the shock from right to left. So both $(6.10)_1$ and $(6.10)_3$ can be summarized in the single statement

Theorem 6.1: When gas crosses a shock, its entropy increases.

We shall now derive an integral form of the increase of entropy. We assume that the gas is in the same state and that it travels at the same velocity at $x = \infty$ as at $x = -\infty$. We normalize entropy so that its value $= 0$ at $x = \pm\infty$. Then integrating (6.6) with respect to x gives

$$(6.11) \quad \frac{d}{dt} \int \rho S \, dx \geq 0;$$

in words.

Theorem 6.2: The total amount of entropy in the flow field is an increasing function of time.

We compare now this result with Glimm's result in Section 4, according to which the total variation of a solution tends to decrease with time. This decrease of total variation can be thought of as a loss of detail, or loss of information. Ever since Maxwell's demon, Boltzmann's H theorem, Smoluchowsky's study of fluctuations, Szilard's thesis and Shannon's work on information and its transmission there have been many attempts to link increase of entropy to decrease of information about the gas. We have shown here that shock waves do both: increase entropy and decrease information, although not on the molecular level but pertaining to organized motion in the whole flow-field.

We turn now to another manifestation of the second law of thermodynamics.

Carnot's Theorem: Consider an engine that extracts heat energy in the amount Q_h from a hot reservoir whose temperature is T_h, and dumps Q_c amount of heat energy into a cold reservoir whose temperature is T_c; the difference is turned into mechanical energy

$$W = Q_h - Q_c.$$

The engine operates cyclically, i.e., everything is in the same state at the end of the cycle as at the beginning.

<u>Assertion</u>: The efficiency of such an engine in converting heat energy into mechanical energy cannot exceed

$$\frac{T_h - T_c}{T_h}$$

<u>Proof</u>: During the absorption of heat energy in the amount of Q_h from a reservoir at temperature T_c, the engine's entropy is increased by the amount

$$(6.12) \quad S = \frac{Q_h}{T_h}$$

Since at the end of the cycle the engine returns to its original state, this amount of entropy must be returned to the outside when Q_c amount of heat is released to a reservoir at temperature T_c. Assuming that no further entropy is generated during the cycle,

$$(6.12)' \quad S = \frac{Q_c}{T_c}.$$

The amount of mechanical energy W extracted is the difference of the heat energy absorbed and released:

$$W = Q_h - Q_c.$$

Using (6.12) and (6.12)' we get the following formula for the efficiency W/Q_h:

$$(6.13) \quad \frac{W}{Q_h} = \frac{Q_h - Q_c}{Q_h} = \frac{ST_h - ST_c}{ST_h} = \frac{T_h - T_c}{T_h}$$

This proves the theorem. Note that if there is entropy production during the cycle, e.g., by shocks produced by the operation of the engine, then an additional amount S' of entropy is produced and has to be gotten rid of. In this case, relation (6.12)' has to be modified to

$$(6.12)'' \quad S + S' = \frac{Q_c}{T_c}.$$

Clearly, this lowers the efficiency to

$$\frac{W}{Q_h} = \frac{T_h - T_c\ (1+S'/S)}{T_h}.$$

Clearly, for sake of efficiency one must keep the entropy production and, thus, the strength of shock waves formed during the operation of the engine to a minimum.

What does Carnot's theorem imply for gas dynamics? Imagine an infinite tube with unit cross section, filled from a cold reservoir, except for a finite section (-L,L) which is filled from a hot reservoir. The initial data are

$$u_0(x) \equiv 0$$

$$(6.14) \quad e_0(x) = \begin{cases} e_h & \text{for } |x|<L \\ e_c & \text{for } |x|>L \end{cases}$$

$$T_0(x) = \begin{cases} T_h & \text{for } |x|<L \\ T_c & \text{for } |x|>L \end{cases}$$

The added heat energy is

$$2L\rho(e_h - e_c).$$

According to Carnot's theorem, the amount that may be converted into mechanical energy does not exceed

$$2L\rho_h e_h \frac{T_h - T_c}{T_h}.$$

Therefore, kinetic energy, which is mechanical energy, can at no future time exceed this amount.

This argument can be extended to variable initial data:

$$u_0(x) = 0 \text{ for } |x|>L$$

$$(6.14)' \quad e_0(x) \begin{cases} \geq e_c & \text{for } |x|< \\ = e_c & \text{for } |x|> \end{cases}$$

$$T_0(x) \begin{cases} \geq T_c & \text{for } |x|<L \\ = T_c & \text{for } |>L. \end{cases}$$

We can think of constructing such data by filling a finite section of the tube with gas taken from a collection of reservoirs and setting it in motion. The amount of heat energy that can be converted to mechanical energy is

$$(6.15) \quad \int \rho_0(x) e_0(x) \frac{T_0(x) - T_c}{T_0(x)} \, dx.$$

Kinetic energy at any time $t>0$ cannot exceed this amount plus the kinetic energy initially imparted to the system:

<u>Theorem</u> <u>6.3</u>: Let u, ρ, e, T denote the velocity, density, internal energy and temperature of a gas moving in a tube, whose initial data satisfy the restrictions (6.14)'. Then for all $t>0$

$$(6.16) \quad \frac{1}{2}\int \rho u^2 dx \leq \frac{1}{2}\int \rho_0 u_0^2 dx + \int \rho_0 e_0 \frac{T_0 - T_c}{T_0} dx.$$

This inequality holds in any number of dimensions; it seems desirable to obtain a proof of it by PDE methods.

We conclude by recounting a curious result about strong shocks that seem to echo a theme in kinetic theory. The Rankine-Hugoniot relations (2.9) for the conservation laws of momentum and energy, see (6.1), are

$$
(6.17)\qquad \begin{aligned} s[m] &= [um+p], \\ s[E] &= [u(E+p)]. \end{aligned}
$$

Assume that a <u>strong</u> <u>shock</u> is impinging on a gas <u>at</u> <u>rest</u>; we denote the quantities behind the shock by capital letters: E, M, P, in the front by lower case letters: e, m, p; note that $u = m = 0$. We rewrite (6.17) as

$$sM = UM+P-p$$

$$s(E-e) = UE+UP$$

We solve the first equation for s and set it into the second:

$$\left(U+\frac{P-p}{M}\right)(E-e) = UE+UP.$$

from which

$$\frac{P-p}{M}(E-e) = U(P+e)$$

Multiplying by $\frac{M}{P}$ we get

$$(6.18)\qquad \left(1-\frac{p}{P}\right)\left(1-\frac{e}{E}\right)E = MU\left(1+\frac{e}{P}\right)$$

For a strong shock, p/P, e/E and e/P are small; so we get

$$(6.18)' \quad E = MU\,(1+\epsilon),$$

ϵ small. Thus, the total energy E behind the shock is approximately twice the kinetic energy $\frac{1}{2}$MU. In words:

<u>Theorem 6.4</u>: When a strong shock impinges on a gas at rest, the energy imparted to the gas is equipartitioned; that is, approximately half of it goes into kinetic, the other half into internal energy.

Note that it follows from (6.18), (6.18)' that the internal energy is always a little greater than the kinetic.

One would like to know what this kind of equipartition of energy has to do with the equipartitioning that is the hallmark of thermodynamic equilibrium.

Bibliography

[1] DiPerna, R. J., Convergence of approximate solutions to conservation laws. Arch. Rat. Mech. Anal., Vol 82, 1983, pp 27-70.

[2] DiPerna, R. J., Convergence of the viscosity method for isentropic gas dynamics, Comm. Math. Phys., 1983, Vol 91, pp 1-30.

[3] Douglis, A., Layering methods for nonlinear partial differential equations of first order, Ann. Inst., Fourier Grenoble, Vol 22, 1972 pp 141-227.

[4] Glimm, J., Solutions in the large for nonlinear hyperbolic systems of equations, Comm. Pure Appl. Math. 18, 1965, pp 697-715.

[5] Glimm, J., Lax, P. D., Decay of solutions of systems of nonlinear hyperbolic conservation laws, Mem. Am. Math. Soc. 101, 1970.

[6] John, F., Formation of singularities in one-dimensional nonlinear wave propagation. Comm. Pure Appl. Math., Vol 27, 1974, pp 377-405.

[7] Keyfitz, B., Solutions with shocks: An example of an L^1-contractive semigroup, Comm. Pure Appl. Math., Vol 24, 1971, pp 125-132.

[8] Lax, P. D., Weak solutions of nonlinear hyperbolic equations and their numerical computation, Comm. Pure Appl. Math. 7, 1954, pp 159-193.

[9] Lax, P. D., Hyperbolic systems of conservation laws, II, Comm. Pure Appl. Math. 10, 1957, pp 537-566.

[10] Lax, P. D., Development of singularities of solutions of nonlinear hyperbolic partial differential equations, J. Math. Phys., Vol 5, 1964, pp 611-613.

[11] Lax, P. D., Shock waves and entropy, in Contributions to Nonlinear Functional Analysis, ed. E. A. Zarantonello, Academic Press, 1971 pp 603-634.

[12] Lax, P. D., The formation and decay of shock waves, Amer. Math. Monthly, Vol 79, 1972, pp 227-241.

[13] Lax, P. D., Hyperbolic systems of conservation laws, and the mathematical theory of shock waves, Regional Conf. Series in Appl. Math., 1973, SIAM, Phila.

[14] Lax, P. D., Accuracy and resolution in the computation of solutions of linear and nonlinear equations, Recent Adv. in Num. Anal., Acad. Press, 1978, pp 107-117.

[15] Murat. F., Compacité par compensation, Ann. Scuola Norm. Sup. Pisa Sci. Fis. Mat. 5, 1978, pp 489-507.

[16] Smoller, J., Shock waves and reaction-diffusion equation, 1982, Springer Verlag.

[17] Tartar, L., Compensated compactness and applications to partial differential equations, in Research Notes in Mathematics 39, Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol 4, ed. R. J. Knops, Pitman Press. 1979.