Preface

These notes were written beginning in 1989 and arose from a graduate course in the numerical solution of scalar hyperbolic conservation laws that I taught a number of times at Purdue University. They're rather focused, and travel in a straight line to the results that I want to get to without much discussion of related topics.

I thought at one time that I might make them into a book, but I would need to learn what Bernardo Cockburn and his collaborators called "a posteriori" error bounds for numerical methods for conservation laws in order to "finish" them. I also never got around to adding the Besov space regularity results by me and Ron DeVore to the notes.

In 1999 Bernardo and I talked, and I found out that he was thinking of writing a book on conservation laws, and our material was almost complementary, so we thought we might put our material together (and I wouldn't have to learn about "a posteriori" analysis) to make a book. And, indeed, Bernardo did a lot of work to merge the two sets of notes.

Then the book *Front tracking for hyperbolic conservation laws* by Helge Holden and Nils Henrik Risebro came out, and the first half of their book covers basically the same material as these notes (but with a somewhat different emphasis and point of view) and Bernardo and I abandoned the thought of a joint book.

As it is, there has been great progress in the area since the time these notes were composed, especially by Bressan and his collaborators.

But these notes still exist, and I decided to write up this preface and distribute them privately before I'm able to clean them up for more general distribution.

So, if you got these notes from me, please don't distribute them further.

And, if you got them from someone else, you shouldn't have them; please e-mail me and I'll send you the latest version.

There are no references in the notes, so we gather them here, together with the later papers on Besov space regularity on conservation laws.

Bradley Lucier, June 2009

REFERENCES

[1] A. J. Chorin, *Random choice solution of hyperbolic systems*, J. Comp. Phys., 22 (1976), pp. 517–533.

[2] B. Engquist and S. Osher, *Stable and entropy satisfying approximations for transonic flow calculations*, Math. Comp., 34 (1980), pp. 45–75.

[3] M. G. Crandall and A. Majda, *Monotone difference approximations for scalar conservation laws*, Math. Comp., 34 (1980), pp. 1–21.

[4] C. M. Dafermos, *Polygonal approximations of solutions of the initial value problem for a conservation law*, J. Math. Anal. Appl., 38 (1972), pp. 33–41.

[5] R. A. DeVore and B. J. Lucier, *High order regularity for conservation laws*, Indiana Univ. Math. J., 39 (1990), pp. 413–430.

[6] ———, *On the size and smoothness of solutions to nonlinear hyperbolic conservation laws*, SIAM J. Math. Anal., 27 (1996), pp. 684–707.

[7] J. Glimm, *Solutions in the large for nonlinear hyperbolic systems of equations*, Comm. Pure App. Math., 18 (1965), pp. 697–715.

[8] S. K. Godunov, *A finite difference method for the numerical computation of discontinuous solutions of the equations of fluid dynamics*, Mat. Sb., 47 (1959), pp. 271–290.

[9] A. Harten and P. D. Lax, *A random choice finite difference scheme for hyperbolic conservation laws*, SIAM J. Numer. Anal., 18 (1981), pp. 289–315.

[10] G. W. Hedstrom and G. H. Rodrigue, *Adaptive-grid methods for time-dependent partial differential equations*, in Multigrid Methods, W. Hackbusch and U. Trottenberg, eds., Springer Verlag, 1982, pp. 474–484.

[11] N. N. Kuznetsov, *Accuracy of some approximate methods for computing the weak solutions of a first-order quasi-linear equation*, USSR Comp. Math. and Math. Phys., 16 (1976), pp. 105–119.

[12] P.D. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock waves*, Regional Conference Series in Applied Mathematics 11, SIAM, Philadelphia, 1973.

[13] R. J. LeVeque, *Large time step shock-capturing techniques for scalar conservation laws*, SIAM J. Numer. Anal., 19 (1982), pp. 1091–1109.

[14] B. J. Lucier, *A moving mesh numerical method for hyperbolic conservation laws*, Math. Comp., 46 (1986), pp. 59–69.

[15] ———, *Error bounds for the methods of Glimm, Godunov and LeVeque*, SIAM J. Numer. Anal., 22 (1985), pp. 1074–1081.

[16] ———, *Regularity through approximation for scalar conservation laws*, SIAM J. Math. Anal., 19 (1988), pp. 763–773.

[17] R. Sanders, *On convergence of monotone finite difference schemes with variable spatial differencing*, Math. Comp., 40 (1983), pp. 91–106.

Chapter 1

Introduction

We shall be concerned with the mathematical properties of hyperbolic conservation laws, which are differential equations that arise, typically, as laws of conservation in physics. Our motivating example of such laws will be the Euler equations that describe the conservation of mass, momentum, and energy in an inviscid, perfect gas. Given, for each point $x \in \mathbb{R}^3$, the initial values of the density ρ, the momentum in each of the coordinate directions m_1, m_2, and m_3, and the total energy E, we attempt to find the values of the vector $U = (\rho, m_1, m_2, m_3, E)$ for all points x and all positive time, t. The laws of conservation of mass, momentum, and energy state that there is a matrix function $F = F(U)$ such that the flux of U across any surface S at a point x on S in the direction of the unit normal vector ν is equal to $F(U) \cdot \nu$. Thus, if Ω is a region in $\mathbb{R}^3$ with boundary S and unit *outward* normal ν, we have

$$\frac{\partial}{\partial t} \iiint_\Omega U \, dx = - \iint_S F(U) \cdot \nu \, d\sigma.$$

Since this is true for all Ω, we have by the divergence theorem if U is smooth

$$\frac{\partial U}{\partial t} + \nabla \cdot F(U) = 0, \qquad x \in \mathbb{R}^3, \ t > 0. \tag{0.1}$$

We shall be concerned with the mathematical properties and numerical approximation of solutions of (0.1). One point of mathematical interest is the fact that, regardless of the use of derivatives in formulating the problem, solutions of (0.1) generally do not remain differentiable or even continuous as time progresses. This leads one to consider the existence and properties of discontinuous solutions of (0.1) through the addition of side conditions called *entropy* conditions. Numerically, the computation of the solution of the Euler equations and other special cases of (0.1) are important in many areas of computational modeling.

The mathematical theory of (0.1) is far from complete. If U is a scalar variable, then existence, uniqueness and continuous dependence of the solution on the initial data is known in any number of spatial dimensions. As for regularity, if U has bounded variation initially, then the variation does not increase with time. In one space dimension, more is known—roughly speaking, the solution can be approximated by moving-grid finite-elements for positive time with the same accuracy as at the initial time, no matter how many discontinuities arise in the solution.

Much less is known for systems. In one dimension, Glimm has proved under various conditions that a solution exists for all time. This theory has been extended, but as of yet there is no general uniqueness theorem, even less a proof of continuous dependence. In several space dimensions there is not even a mathematical proof of the existence of solutions.

These equations are so important, however, that the mere fact that mathematicians cannot prove that a solution is unique, or even exists, should not dissuade people from trying to compute solutions numerically! Many different numerical schemes have been proposed for (0.1); in the first part of the course we shall emphasize methods for which proofs of convergence are available. This will, of course, restrict us to the scalar case, and for high order methods, to one dimensional problems.

Historically, numerical methods have played an important role in the mathematical theory of (0.1). We shall take a numerical approach to almost all the properties of solutions of (0.1). Specifically, existence of solutions for scalar equations will be proved using monotone numerical methods, while uniqueness and continuous dependence of solutions will be proved using an approximation theorem of Kuznetsov. Later we shall consider Glimm's scheme and various moving grid numerical schemes for the scalar equation in one dimension.

REMARKS. The book *Shock Waves and Reaction-Diffusion Equations* by JOEL SMOLLER deals with earlier approaches to the scalar problem and with systems in one space dimension. LAX'S monograph *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves* is of interest in one space dimension. The book *Introduction to Partial Differential Equations with Applications* by ZACHMANOGLOU AND THOE contains a good explanation of the C^1 theory of nonlinear, first-order, partial differential equations.

§1. Motivation of Mathematical Properties

In this section we derive in a formal and nonrigorous way several properties that we expect will hold for solutions of scalar conservation laws. It will be necessary in later chapters to prove that solutions, as we define them, will indeed satisfy these properties.

Consider the scalar equation in one space dimension

$$\begin{aligned} &u_t + f(u)_x = 0, \qquad x \in \mathbb{R},\ t > 0, \\ &u(x,0) = u_0(x), \qquad x \in \mathbb{R}. \end{aligned} \tag{1.1}$$

We can rewrite the differential equation as

$$\nabla_{x,t} u \cdot (f'(u), 1) = 0,$$

so that in the x-t plane u is constant in the direction $(f'(u), 1)$. Because this direction depends only on u, u is constant along the *characteristic line* $x = x_0 + f'(u)t$. Because of the initial condition, in fact we have that

$u = u_0(x_0)$ for $x = x_0 + f'(u)t$. Eliminating x_0 from these equations gives the implicit formula

$$u = u_0(x - f'(u)t) \tag{1.2}$$

for the solution $u(x,t)$ of (1.1).

EXAMPLE 1.1. Consider the C^1 initial data

$$u_0(x) = \begin{cases} \cos^2(x), & \text{for } x \in [-\pi/2, \pi/2], \\ 0, & \text{otherwise,} \end{cases}$$

and the flux $f(u) = u^2$. Then for $x_0 = 0$, $u = u_0(x_0) = 1$ along the line $x = x_0 + f'(u_0(x_0))t = 2t$, while for $x_0 = \pi/2$, $u = u_0(x_0) = 0$ along the line $x = x_0 + f'(u_0(x_0))t = \pi/2$. Clearly we shall have problems when $t = \pi/4$ and $x = \pi/2$ — u cannot take on two distinct values there! (See Figure 1.)

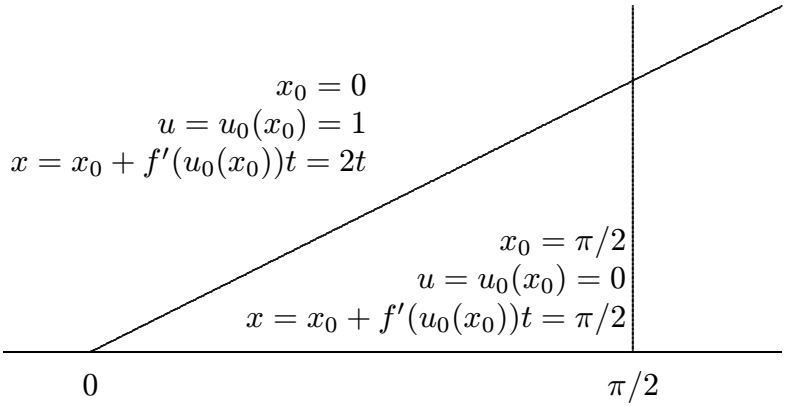

FIGURE 1. Characteristic lines in x-t space can cross.

So, discontinuities occur in u. We can see this in a different way by differentiating (1.2) with respect to x:

$$u_x = u_0'(x - f'(u)t) \times (1 - tf''(u)u_x),$$

or

$$u_x = \frac{u_0'(x_0)}{1 + u_0'(x_0)f''(u_0(x_0))t}$$

where $x_0 = x - f'(u)t$. Thus, one can see that u_x is well-defined as long as $1 + u_0'(x_0)f''(u_0(x_0))t$ is not zero, and conversely, the first value of t for which $1 + u_0'(x_0)f''(u_0(x_0))t = 0$ for some x_0 is the time at which C^1 solutions no longer exist. See Figure 2.

The above examples show that we must consider the existence and properties of discontinuous solutions of (1.1). We shall attempt to extend to discontinuous solutions selected qualitative properties of C^1 solutions, which we now describe.

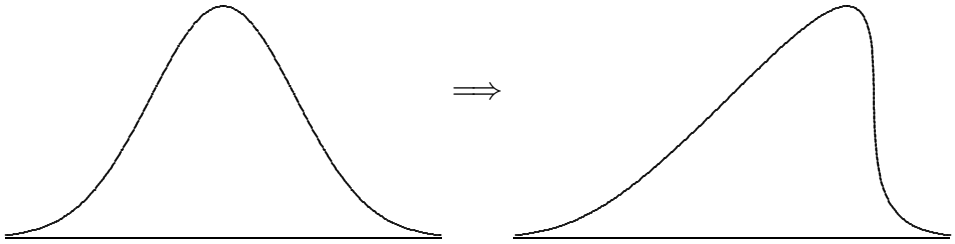

FIGURE 2. The solution of $u_t + (u^2)_x = 0$ with $u(x,0) = \exp(-x^2/2)/\sqrt{2\pi}$ at time 2, just before the shock time.

First, we shall assume that if

$$\|u_0\|_{L^1(\mathbb{R})} := \int_{\mathbb{R}} |u_0(x)|\,dx < \infty$$

then the integral of $u(\,\cdot\,,t)$ does not change with time, that is

$$\int_{\mathbb{R}} u(x,t)\,dx = \int_{\mathbb{R}} u_0(x)\,dx.$$

This is clear formally for smooth solutions in $L^1(\mathbb{R})$ with u_x bounded because

$$\begin{aligned}\frac{\partial}{\partial t}\int_{\mathbb{R}} u(x,t)\,dx &= \int_{\mathbb{R}} u_t(x,t)\,dx \\ &= -\int_{\mathbb{R}} f(u(x,t))_x\,dx \\ &= -\lim_{R\to\infty}\left(f(u(R,t)-f(u(-R,t))\right) \\ &= 0,\end{aligned}$$

because $u(R,t) \to 0$ as $R \to \pm\infty$.

Next, we assume that the mapping $u_0 \to u(\,\cdot\,,t)$ is a contraction in $L^1(\mathbb{R})$, that is, for any two solutions u and v with initial data u_0 and v_0, respectively,

$$\|u(\,\cdot\,,t) - v(\,\cdot\,,t)\|_{L^1(\mathbb{R})} \le \|u_0 - v_0\|_{L^1(\mathbb{R})}. \tag{1.3}$$

For two smooth solutions u and v one can show formally that equality holds in (1.3). For each t, assume we can find a partition $\{x_i\}$ of $\mathbb{R}$ with no limit points such that $u(x,t) > v(x,t)$ on $I_i := (x_{2i}, x_{2i+1})$ and $u(x,t) < v(x,t)$ on $J_i := (x_{2i-1}, x_{2i})$; see Figure 3

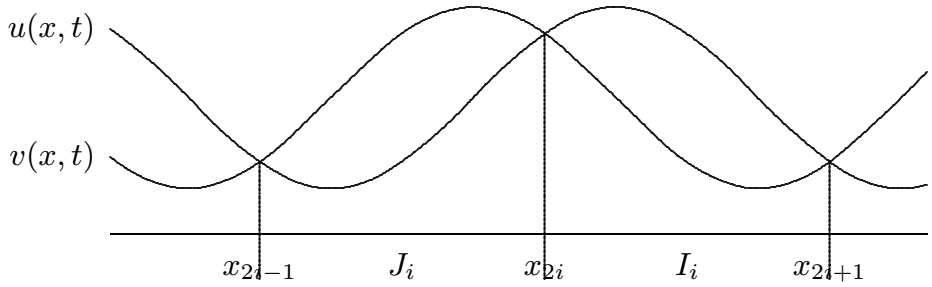

FIGURE 3. We assume that $\mathbb{R}$ can be partitioned into intervals where $u(x,t) > v(x,t)$ and $u(x,t) < v(x,t)$.

We can write formally

$$\begin{aligned}
&\frac{\partial}{\partial t}\|u(\,\cdot\,,t) - v(\,\cdot\,,t)\|_{L^1(\mathbb{R})} \\
&\qquad = \frac{\partial}{\partial t}\sum_i \int_{I_i} (u-v)\,dx - \frac{\partial}{\partial t}\sum_i \int_{J_i} (u-v)\,dx \\
&\qquad = \sum_i \int_{I_i} (u_t - v_t)\,dx - \sum_i \int_{J_i} (u_t - v_t)\,dx \\
&\qquad = -\sum_i \int_{I_i} (f(u)_x - f(v)_x)\,dx + \sum_i \int_{J_i} (f(u)_x - f(v)_x)\,dx \\
&\qquad = -\sum_i (f(u) - f(v))\big|_{I_i} + \sum_i (f(u) - f(v))\big|_{J_i} = 0
\end{aligned}$$

because $f(u(x_i,t)) = f(v(x_i,t))$. Physically, we expect discontinuous solutions of (1.1) to be limits as $\epsilon \to 0$ of solutions of the viscous equation

$$\begin{aligned}
&u_t + f(u)_x = \epsilon\, u_{xx}, && x \in \mathbb{R},\ t > 0,\ \epsilon > 0, \\
&u(x,0) = u_0(x), && x \in \mathbb{R}.
\end{aligned} \tag{1.4}$$

The above argument applied to (1.4) shows that

$$\begin{aligned}
&\frac{\partial}{\partial t}\|u(\,\cdot\,,t) - v(\,\cdot\,,t)\|_{L^1(\mathbb{R})} \\
&\qquad = -\sum_i \int_{I_i} (f(u)_x - \epsilon u_{xx} - f(v)_x + \epsilon v_{xx})\,dx \\
&\qquad\qquad + \sum_i \int_{J_i} (f(u)_x - \epsilon u_{xx} - f(v)_x + \epsilon v_{xx})\,dx \\
&\qquad = -\sum_i (-\epsilon u_x + \epsilon v_x)\big|_{I_i} + \sum_i (-\epsilon u_x + \epsilon v_x)\big|_{J_i}
\end{aligned}$$

Because $u > v$ on I_i and $u < v$ on J_i, we have that

$$u_x(x_{2i},t) \geq v_x(x_{2i},t) \qquad \text{and} \qquad u_x(x_{2i+1},t) \leq v_x(x_{2i+1},t);$$

substituting this into the previous equality shows that

$$\frac{\partial}{\partial t}\|u(\,\cdot\,,t)-v(\,\cdot\,,t)\|_{L^1(\mathbb{R})}\le 0.$$

Thus, if we expect the solution of the inviscid problem (with $\epsilon=0$) to be the limit of the viscous solutions as $\epsilon\to 0$, then we expect (1.3) to hold for solutions of (1.1).

A third property that is not obvious even for smooth solutions is that if $u_0(x)\ge v_0(x)$ for all x then $u(x,t)\ge v(x,t)$ for all x and t. This may be surprising because the characteristics coming into a point (x,t) will generally start at two different points $(x_0,0)$ for u and $(x_1,0)$ for v, so $u_0(x_0)$ and $v_0(x_1)$ are not directly comparable. Nevertheless, this property follows from the following useful lemma.

Lemma 1.1. (Crandall and Tartar)**.** *Assume that $(\Omega,d\mu)$ is a measure space (e.g., $\mathbb{R}^n$ with the usual Lebesgue measure dx) and that the possibly nonlinear mapping $T: L^1(\Omega)\to L^1(\Omega)$ satisfies for all $u\in L^1(\Omega)$*

$$\int_\Omega Tu\,d\mu=\int_\Omega u\,d\mu. \tag{1.5}$$

Then the following two properties are equivalent:

(1) *For all $u,v\in L^1(\Omega)$, $\|Tu-Tv\|_{L^1(\Omega)}\le\|u-v\|_{L^1(\Omega)}$.*
(2) *For all $u,v\in L^1(\Omega)$, $u\ge v$ a.e. $(d\mu)$ implies $Tu\ge Tv$ a.e. $(d\mu)$.*

Remark 1.1. For any fixed $t>0$ we can apply this lemma to the mapping T: $u_0\to u(\,\cdot\,,t)$ to substantiate the claimed property.

Proof of Lemma 1.1. Assume that (1) holds and let $u\ge v$ *a.e.* $(d\mu)$. Then

$$\begin{aligned}\int_\Omega |Tu-Tv|\,d\mu=\|Tu-Tv\|_{L^1(\Omega)}&\le\|u-v\|_{L^1(\Omega)}\\ &=\int_\Omega u-v\,d\mu\\ &=\int_\Omega Tu-Tv\,d\mu\qquad\text{by (1.5).}\end{aligned}$$

Therefore, $Tu-Tv\ge 0$ *a.e.* $(d\mu)$.

Conversely, assume that (2) holds and consider $u\vee v=\max(u,v)$ and $u\wedge v=\min(u,v)$. Then by (2), $T(u\vee v)\ge T(u)$ and $T(u\vee v)\ge T(v)$, so

$T(u \vee v) \geq T(u) \vee T(v)$. Also, $|u - v| = u \vee v - u \wedge v$. So

$$\begin{aligned}\|Tu - Tv\|_{L^1(\Omega)} = \int_\Omega |Tu - Tv|\, d\mu &= \int_\Omega Tu \vee Tv - Tu \wedge Tv\, d\mu \\ &\leq \int_\Omega T(u \vee v) - T(u \wedge v)\, d\mu \\ &= \int_\Omega u \vee v - u \wedge v\, d\mu \qquad \text{by (1.5)} \\ &= \|u - v\|_{L^1(\Omega)}. \qquad \square\end{aligned}$$

Finally, we shall assume that u satisfies a maximum principle: for all $t > 0$ and $u(\,\cdot\,, t) \in L^1(\mathbb{R})$

$$\begin{aligned}\operatorname*{ess\,sup}_{x \in \mathbb{R}} u(x,t) &\leq \operatorname*{ess\,sup}_{x \in \mathbb{R}} u_0(x), \quad \text{and} \\ \operatorname*{ess\,inf}_{x \in \mathbb{R}} u(x,t) &\geq \operatorname*{ess\,inf}_{x \in \mathbb{R}} u_0(x).\end{aligned}$$

Because of (1.2), this is clear for smooth solutions of (1.1).

§2. Weak Solutions and the Entropy Condition

The example in the previous section shows that continuous solutions of (1.1) generally do not exist. In this section we shall give examples to show that so-called *weak* solutions of (1.1) are not unique. We shall then go on to motivate the *entropy condition*, which we shall prove in later chapters specifies a unique weak solution of (1.1).

If $u(x,t)$ is a smooth solution of (1.1) then for any C^1 function $\phi(x,t)$ with bounded support and any value of $T > 0$

$$0 = \int_0^T \int_{\mathbb{R}} (u_t + f(u)_x)\phi\, dx\, dt$$

or, after integrating by parts in x and t,

$$\begin{aligned}-\int_0^T \int_{\mathbb{R}} u\phi_t + f(u)\phi_x\, dx\, dt \\ + \int_{\mathbb{R}} u(x,T)\phi(x,T)\, dx - \int_{\mathbb{R}} u_0(x)\phi(x,0)\, dx = 0,\end{aligned} \tag{2.1}$$

where, of course, we have used the fact that $u(x,0) = u_0(x)$.

Definition. If u is bounded and measurable and satisfies (2.1) for all $\phi \in C^1$ with bounded support, then we say that u is a *weak solution* of (1.1) in $\mathbb{R} \times [0,T]$.

We can readily give a necessary and sufficient condition that a piecewise smooth function be a weak solution of (1.1). Assume that $u(x,t)$ is a weak solution of (1.1) in a rectangle $\Omega := [x_0, x_1] \times [t_0, t_1]$, that u is a pointwise solution of (1.1) in each of $\Omega_1 := \{(x,t) \in \Omega \mid x < s(t)\}$ and $\Omega_2 := \{(x,t) \in$

$\Omega \mid x > s(t)\}$, and that u has well-defined, continuous, limits from the left and right as x approaches the curve $S := \{(x,t) \in \Omega \mid x = s(t)\}$; see Figure 4.

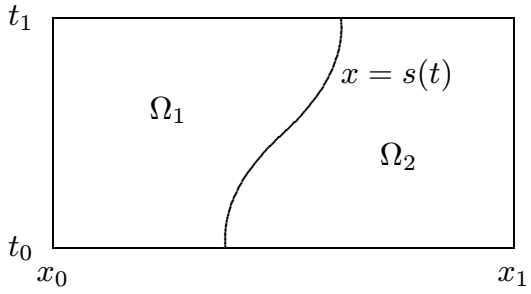

FIGURE 4. The weak solution $u(x,t)$ is assumed to be smooth in Ω_1 and Ω_2, with a discontinuity along $S = \{x = s(t)\}$.

By the definition of weak solutions, we have for any $\phi \in C_0^1(\Omega)$

$$\begin{aligned} 0 &= \int_\Omega u\phi_t + f(u)\phi_x \, dx \, dt \\ &= \int_{\Omega_1} u\phi_t + f(u)\phi_x \, dx \, dt + \int_{\Omega_2} u\phi_t + f(u)\phi_x \, dx \, dt. \end{aligned} \tag{2.2}$$

Because u is a pointwise solution of (1.1) in Ω_1, we can write

$$\begin{aligned} &\int_{\Omega_1} u\phi_t + f(u)\phi_x \, dx \, dt \\ &\qquad = \int_{\Omega_1} (u\phi)_t + (f(u)\phi)_x \, dx \, dt - \int_{\Omega_1} u_t\phi + f(u)_x\phi \, dx \, dt \\ &\qquad = \int_{\Omega_1} (u\phi)_t + (f(u)\phi)_x \, dx \, dt \\ &\qquad = \int_{\partial\Omega_1} (f(u)\phi, u\phi) \cdot \nu \, d\sigma \end{aligned} \tag{2.3}$$

by the divergence theorem; here ν is the unit outward normal of Ω_1. By assumption, ϕ is zero on the boundary of Ω_1, except possibly on S. Because of this, calculus shows that (2.3) is equal to

$$\int_{t_0}^{t_1} \phi(s(t),t)[f(u(s(t)^-,t)) - s'(t)u(s(t)^-,t)] \, dt,$$

where $u(s(t)^-,t)$ means the limit of $u(x,t)$ as you approach $(s(t),t)$ from the left, i.e., from Ω_1. Substituting into (2.2) this value and a similar one for the integral over Ω_2 shows that

$$\int_{t_0}^{t_1} \phi(s(t),t)\{[f(u(s(t)^-,t)) - f(u(s(t)^+,t))] - s'(t)[u(s(t)^-,t) - u(s(t)^+,t)]\} dt$$

is zero for all $\phi \in C_0^1(\Omega)$. Because ϕ is arbitrary, this is true if and only if the quantity in braces is identically zero, i.e., for all t,

$$s'(t) = \frac{f(u(s(t)^+,t)) - f(u(s(t)^-,t))}{u(s(t)^+,t) - u(s(t)^-,t)} =: \frac{[f(u)]}{[u]}, \tag{2.4}$$

where we have introduced the notation $[g(u)]$ to mean the jump in a quantity $g(u)$ at a point. The relation (2.4) is called the *Rankine-Hugoniot condition.* Thus, if $u(x,t)$ is a piecewise smooth function that satisfies (1.1) pointwise where it is C^1, and whose jumps satisfy (2.4), then it is a weak solution of (1.1).

EXAMPLE 2.1. Let $f(u) = u^2$ and let

$$u_0(x) = \begin{cases} 1, & x \in [0,1], \\ 0, & \text{otherwise.} \end{cases}$$

It is left to the reader to verify that both

$$u(x,t) = \begin{cases} 1, & x \in [t, 1+t], \\ 0, & \text{otherwise,} \end{cases}$$

and

$$u(x,t) = \begin{cases} x/2t, & x \in [0,2t], \\ 1, & x \in [2t, 1+t], \\ 0, & \text{otherwise,} \end{cases}$$

are weak solutions of (1.1) on the strip $\mathbb{R} \times [0,1]$; see Figure 5.

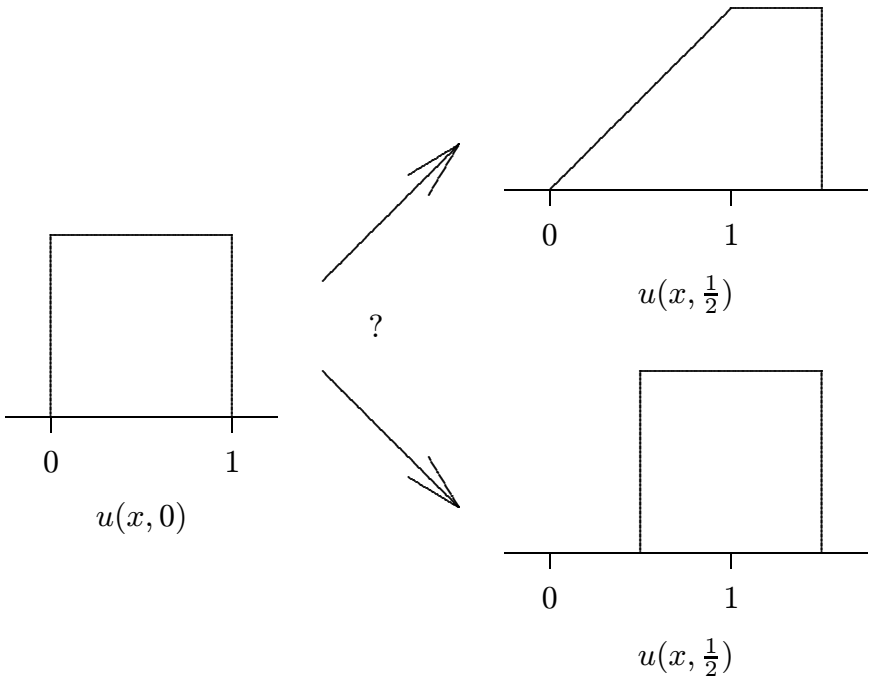

FIGURE 5. There are at least two weak solutions at time 1/2 with $u_0 = \chi_{[0,1]}$ and $f(u) = u^2$.

EXAMPLE 2.2. A rather more striking example of lack of uniqueness is the following. Let $f(u) = u^2$ and $u_0(x) = 0$ for all x. Then of course $u(x,t) \equiv 0$ is a weak solution of (1.1). Nonetheless, the function

$$u(x,t) = \begin{cases} -1, & -1 \le x/t \le 0, \\ 1, & 0 < x/t \le 1, \\ 0, & \text{otherwise}, \end{cases}$$

is also a weak solution of (1.1) with zero initial data.

So, classical solutions do not exist for all time, and weak solutions are not unique. However, we shall prove later that weak solutions that satisfy the properties listed in the previous section *are* unique. First, we use an idea of Crandall and Majda to show that the properties in the previous section, together with the assumption that the correct weak solution for the data $u_0(x) \equiv c$ is $u(x,t) \equiv c$, imply that u satisfies the so-called *entropy condition*: for all $c \in \mathbb{R}$

$$|u-c|_t + (\operatorname{sgn}(u-c)(f(u)-f(c)))_x \le 0.$$

This inequality is to be understood in the weak sense, i.e., for all C^1 $\phi \ge 0$ with compact support and all $T > 0$,

$$\begin{aligned} &-\int_0^T \int_{\mathbb{R}} |u-c|\phi_t + \operatorname{sgn}(u-c)(f(u)-f(c))\phi_x \, dx \, dt \\ &+ \int_{\mathbb{R}} |u(x,T)-c|\phi(x,T)\, dx - \int_{\mathbb{R}} |u_0(x)-c|\phi(x,0)\, dx \le 0. \end{aligned} \tag{2.5}$$

Inequality (2.5) can be derived for piecewise smooth solutions as follows. Assume that for any continuous, piecewise C^1 initial data u_0 one can solve (1.1) for a short time and that the solution will satisfy the properties of the previous section. For any $T > 0$ consider the solution to the problem

$$\begin{aligned} &v_t + f(v)_x = 0, && x \in \mathbb{R},\ t > T, \\ &v(x,T) = u(x,T) \vee c, && x \in \mathbb{R}. \end{aligned}$$

Because we hope that the solution operator of (1.1) is order-preserving, for $t > T$ we know that $v(x,t) \ge u(x,t)$ and $v(x,t) \ge c$ (the function $u(x,t) = c$ for all x and t is, by assumption, a solution of (1.1)), so $v(x,t) \ge u(x,t) \vee c$. Therefore

$$\frac{v(x.t) - v(x,T)}{t-T} = \frac{v(x.t) - u(x,T) \vee c}{t-T} \ge \frac{u(x.t) \vee c - u(x,T) \vee c}{t-T}.$$

Let $t \to T$ to see that

$$-f(u \vee c)_x\big|_{t=T} = -f(v)_x\big|_{t=T} = v_t\big|_{t=T} \ge (u \vee c)_t\big|_{t=T}.$$

Therefore

$$(u \vee c)_t + f(u \vee c)_x \le 0.$$

Similarly,

$$(u \wedge c)_t + f(u \wedge c)_x \geq 0.$$

Because $|u-c| = u \vee c - u \wedge c$ and $\operatorname{sgn}(u-c)(f(u)-f(c)) = f(u \vee c) - f(u \wedge c)$, (2.5) follows. Any u that satisfies (2.5) will be called an *entropy weak solution* of (1.1).

If u is a bounded entropy weak solution of (1.1), then it is also satisfies (2.1). For if $c \leq \inf u$, then $\operatorname{sgn}(u-c) = 1$, $|u-c| = u-c$, and (2.5) implies that

$$\begin{aligned} 0 \geq & -\int_0^T \int_{\mathbb{R}} (u-c)\phi_t + (f(u)-f(c))\phi_x \, dx \, dt \\ & + \int_{\mathbb{R}} (u(x,T)-c)\phi(x,T)\, dx - \int_{\mathbb{R}} (u_0(x)-c)\phi(x,0)\, dx \\ = & -\int_0^T \int_{\mathbb{R}} u\phi_t + f(u)\phi_x \, dx \, dt \\ & + \int_{\mathbb{R}} u(x,T)\phi(x,T)\, dx - \int_{\mathbb{R}} u_0(x)\phi(x,0)\, dx. \end{aligned} \tag{2.6}$$

(The terms involving the constant c integrate to zero.) On the other hand, if $c \geq \sup u$, then $\operatorname{sgn}(u-c) = -1$, $|u-c| = c-u$, and we have

$$\begin{aligned} 0 \geq & +\int_0^T \int_{\mathbb{R}} (u-c)\phi_t + (f(u)-f(c))\phi_x \, dx \, dt \\ & - \int_{\mathbb{R}} (u(x,T)-c)\phi(x,T)\, dx + \int_{\mathbb{R}} (u_0(x)-c)\phi(x,0)\, dx \\ = & +\int_0^T \int_{\mathbb{R}} u\phi_t + f(u)\phi_x \, dx \, dt \\ & - \int_{\mathbb{R}} u(x,T)\phi(x,T)\, dx + \int_{\mathbb{R}} u_0(x)\phi(x,0)\, dx. \end{aligned} \tag{2.7}$$

Inequalities (2.6) and (2.7) together imply (2.1).

In the same way that (2.1) implies (2.4) for piecewise smooth solutions of (1.1), the entropy inequality (2.5) will imply an inequality, which we now derive, for the speed of valid, or *entropy satisfying,* shocks. We assume that $u(x,t)$ is an entropy weak solution of (1.1) in a rectangle $\Omega := [x_0, x_1] \times [t_0, t_1]$, that u is a pointwise solution of (1.1) in each of $\Omega_1 := \{(x,t) \in \Omega \mid x < s(t)\}$ and $\Omega_2 := \{(x,t) \in \Omega \mid x > s(t)\}$, and that u has well-defined, continuous, limits from the left and right as x approaches the curve $S := \{(x,t) \in \Omega \mid x = s(t)\}$; see Figure 4. For any $c \in \mathbb{R}$ and any nonnegative $\phi \in C_0^1(\Omega)$, we now have

$$0 \leq \int_\Omega |u-c|\phi_t + \operatorname{sgn}(u-c)(f(u)-f(c))\phi_x \, dx \, dt. \tag{2.8}$$

Assume for a particular $t \in (t_0, t_1)$ that no two of $u(s(t)^+, t)$, c, and $u(s(t)^-, t)$ are the same. Then, because u is continuous in both Ω_1 and

Ω_2, there is a ball B around $(s(t),t)$ such that c, $u(x,t)$ for $(x,t) \in B \cap \Omega_1$, and $u(x,t)$ for $(x,t) \in B \cap \Omega_2$ are in the same order as c, $u(s(t)^-,t)$, and $u(s(t)^+,t)$; for example, $u(s(t)^-,t) < c < u(s(t)^+,t)$. In $B \cap \Omega_1$ we obviously have

$$|u-c|_t + [\operatorname{sgn}(u-c)(f(u)-f(c))]_x = 0 \quad \text{for } (x,t) \in \Omega_1 \tag{2.9}$$

since $\operatorname{sgn}(u-c)$ has a fixed sign in this ball. Therefore, when we consider the weak equation we can restrict ϕ to have support in this ball, and all the following arguments are valid. Whenever $c = u(s(t)^-,t)$ or $c = u(s(t)^+,t)$ then a different argument is needed.

Proceeding now in the same way as we did to derive (2.2) and (2.3), we see that

$$\int_{\Omega_1} |u-c|\phi_t + \operatorname{sgn}(u-c)(f(u)-f(c))\phi_x \, dx \, dt$$

is equal to

$$\int_{t_0}^{t_1} \phi[\operatorname{sgn}(u_L-c)(f(u_L)-f(c)) - s'(t)|u_L-c|] \, dt,$$

where $u_L := u(s(t)^-,t)$ and $\phi = \phi(s(t),t)$. Similarly,

$$\int_{\Omega_2} |u-c|\phi_t + \operatorname{sgn}(u-c)(f(u)-f(c))\phi_x \, dx \, dt$$

is equal to

$$-\int_{t_0}^{t_1} \phi[\operatorname{sgn}(u_R-c)(f(u_R)-f(c)) - s'(t)|u_R-c|] \, dt,$$

where $u_R := u(s(t)^+,t)$. Since (2.8) holds for all $\phi \geq 0$, adding the integrals over Ω_1 and Ω_2 implies that

$$\begin{aligned} \operatorname{sgn}(u_L-c)(f(u_L)-f(c)) - \operatorname{sgn}(u_R-c)(f(u_R)-f(c)) \\ - s'(t)[|u_L-c| - |u_R-c|] \geq 0 \quad \text{for all } c \in \mathbb{R}. \end{aligned} \tag{2.10}$$

When c is either less than both u_L and u_R or greater than u_L and u_R, then $\operatorname{sgn}(u_L-c) = \operatorname{sgn}(u_R-c)$. When $c \leq \min(u_L,u_R)$, then (2.10) implies that

$$f(u_L) - f(u_R) - s'(t)[u_L-u_R] \geq 0.$$

Similarly, when $c \geq \max(u_L,u_R)$, then

$$f(u_L) - f(u_R) - s'(t)[u_L-u_R] \leq 0,$$

so (2.10) implies the Rankine-Hugoniot condition

$$f(u_L) - f(u_R) - s'(t)[u_L-u_R] = 0. \tag{2.11}$$

(This is good, because (2.5) implies (2.1), from which the Rankine-Hugoniot condition was derived!)

Let us restrict our attention for the moment to the case $u_L > u_R$. For all $u_L > c > u_R$, (2.10) implies that

$$f(u_L) - f(c) + f(u_R) - f(c) - s'(t)[u_L - c - c + u_R] \geq 0. \tag{2.12}$$

Adding (2.11) to (2.12) shows that

$$2f(u_L) - 2f(c) - s'(t)[2u_L - 2c] \geq 0.$$

So, we must have for all $u_L > c > u_R$,

$$\frac{f(u_L) - f(u_R)}{u_L - u_R} = s'(t) \leq \frac{f(u_L) - f(c)}{u_L - c}. \tag{2.13}$$

Geometrically, this says that the slope of the line joining the points $(u_L, f(u_L))$ and $(u_R, f(u_R))$ must be less than the slope of the line joining the points $(u_L, f(u_L))$ and $(c, f(c))$ for all c between u_L and u_R. This is equivalent to saying that the line joining $(u_L, f(u_L))$ and $(u_R, f(u_R))$ must be above the graph of the function $f(u)$ for u between u_L and u_R. See Figure 6

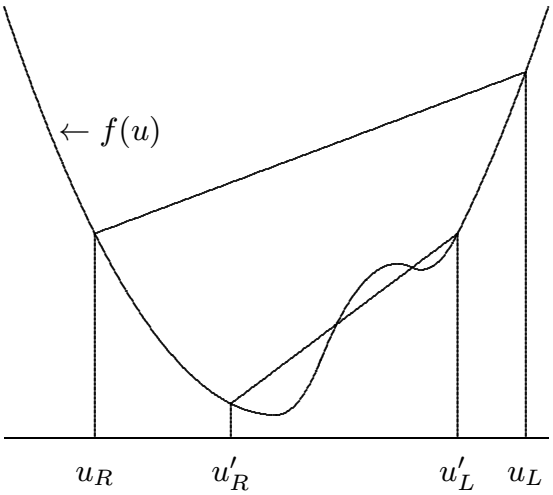

FIGURE 6. The constant states u_L and u_R can be joined by an entropy satisfying shock; the states u'_L and u'_R cannot be so joined.

We can derive an inequality that is equivalent to (2.13), but which involves the value at the right endpoint u_R. By subtracting (2.11) from (2.12) we obtain

$$2f(u_R) - 2f(c) - s'(t)[2u_R - 2c] \geq 0,$$

which gives for all $u_L > c > u_R$, because $u_R - c < 0$,

$$\frac{f(u_L) - f(u_R)}{u_L - u_R} = s'(t) \geq \frac{f(u_R) - f(c)}{u_R - c}. \tag{2.14}$$

Now, by letting $c \to u_L$ in (2.13) and $c \to u_R$ in (2.14), we derive the following inequality that is a necessary condition for a shock to be entropy-satisfying:

$$f'(u_L) \geq s'(t) = \frac{f(u_L) - f(u_R)}{u_L - u_R} \geq f'(u_R).$$

Geometrically, this says that the characteristic speed on the left of the shock must be greater than the shock speed, which in turn must be greater than the characteristic speed on the right. In other words, the characteristics must point into the shock.

We leave it to the reader to show that whenever $u_L < c < u_R$ the same inequality (2.13) results. However, (2.13) has a different graphical interpretation in this case: it requires that the line joining $(u_L, f(u_L))$ and $(u_R, f(u_R))$ be *below* the graph of $f(u)$ for u between u_L and u_R. Inequality (2.13) (or one algebraically equivalent to it, using (2.11)) is called the *Oleinik entropy condition* for piecewise smooth solutions of (1.1).

If the flux $f(u)$ is convex ($f''(u) \geq 0$), then the line joining $(u_L, f(u_L))$ and $(u_R, f(u_R))$ is *always* above the graph of $f(u)$ for u between u_L and u_R. Therefore, the entropy condition requires for convex $f(u)$ that any entropy-satisfying shock with left state u_L and right state u_R have $u_L > u_R$. Similarly, if $f(u)$ is concave then the line joining $(u_L, f(u_L))$ and $(u_R, f(u_R))$ is always below the graph of $f(u)$ for u between u_L and u_R, so any entropy-satisfying shock must have $u_L < u_R$.

We can use (2.13) to see which (if any) of our several weak solutions given in Examples 2.1 and 2.2 are entropy weak solutions. In both examples, $f(u) = u^2$, so the entropy condition requires that all discontinuities have $u_L > u_R$. By this criterion, it is easily seen that neither

$$u(x,t) = \begin{cases} 1, & x \in [t, 1+t], \\ 0, & \text{otherwise}, \end{cases}$$

nor

$$u(x,t) = \begin{cases} -1, & -1 \leq x/t \leq 0, \\ 1, & 0 < x/t \leq 1, \\ 0, & \text{otherwise}, \end{cases}$$

satisfies the entropy condition, whereas both $u(x,t) \equiv 0$ and

$$u(x,t) = \begin{cases} x/2t, & x \in [0, 2t], \\ 1, & x \in [2t, 1+t], \\ 0, & \text{otherwise}, \end{cases} \quad \text{for } 0 \leq t \leq 1,$$

are entropy weak solutions of (1.1) with $f(u) = u^2$. This leaves open the questions of whether there are *other* entropy weak solutions of (1.1) with the same initial data (in our case $u_0(x) = 0$ and $u_0(x) = \chi_{[0,1]}(x)$), or if entropy weak solutions of (1.1) exist for other initial data. It will be a corollary of Kuznetsov's approximation theorem in Chapter 2 that entropy

weak solutions of (1.1), as defined by (2.5), are unique, and we shall show in Chapter 3 that, under fairly general conditions, entropy weak solutions of (1.1) exist.

§3. Norms and Spaces

$\mathbb{R}^n$ will denote the set of n-tuples of real numbers $x := (x_1, \dots, x_n)$, and $\mathbb{Z}^n$ will denote the set of n-tuples of integers $\nu := (\nu_1, \dots, \nu_n)$. In either case e_i will denote the n-vector with the ith component equal to 1 and all other components 0. The norm of x will always be taken to be the L^∞ norm, $|x| := \max_{1 \le j \le n} |x_j|$.

The space of real integrable functions will be given by

$$L^1(\mathbb{R}^n) := \{f \colon \mathbb{R}^n \to \mathbb{R} \mid \|f\|_{L^1(\mathbb{R}^n)} := \int_{\mathbb{R}^n} |f(x)|\, dx < \infty\}.$$

Similarly, the space of "integrable" functions on $\mathbb{Z}^n$ will be given by

$$L^1(\mathbb{Z}^n) := \{f \colon \mathbb{Z}^n \to \mathbb{R} \mid \|f\|_{L^1(\mathbb{Z}^n)} := \sum_{\nu \in \mathbb{Z}^n} |f_\nu| < \infty\}.$$

The variation of a function on $\mathbb{R}^n$ is defined to be

$$\|f\|_{\mathrm{BV}(\mathbb{R}^n)} := \sum_{j=1}^{n} \sup_{\tau \in \mathbb{R}} \frac{1}{|\tau|} \int_{\mathbb{R}^n} |f(x + \tau e_j) - f(x)|\, dx.$$

(BV stands for "bounded variation.") This is not really a norm, because if there is a constant c such that for all $x \in \mathbb{R}^n$, $f(x) = c$, then $\|f\|_{\mathrm{BV}(\mathbb{R}^n)} = 0$ but $f \neq 0$. The variation of a function on $\mathbb{Z}^n$ is defined similarly:

$$\|f\|_{\mathrm{BV}(\mathbb{Z}^n)} := \sum_{j=1}^{n} \sum_{\nu \in \mathbb{Z}^n} |f_{\nu + e_j} - f_\nu|.$$

If X is a normed linear space (such as $\mathbb{R}^n$, $L^1(\mathbb{R}^n)$, and $L^1(\mathbb{Z}^n)$) then $C([0,T], X)$ is the space of continuous functions $f : [0,t] \to X$. This means that for all $t \in [0,T]$, $\lim_{t' \to t} \|f(t) - f(t')\|_X = 0$. If X is a Banach space (a complete, normed, linear space, again such as $\mathbb{R}^n$, $L^1(\mathbb{R}^n)$, and $L^1(\mathbb{Z}^n)$), then so is $C([0,T], X)$, with norm given by

$$\|f\|_{C([0,T],X)} := \sup_{t \in [0,T]} \|f(t)\|_X.$$

Chapter 2

Kuznetsov's Approximation Theorem

In this chapter we consider entropy weak solutions of the scalar conservation law in several space dimensions

$$\begin{aligned} &u_t + \nabla_x \cdot f(u) = 0, \qquad && x \in \mathbb{R}^n,\ t > 0, \\ &u(x,0) = u_0(x), && x \in \mathbb{R}^n. \end{aligned} \tag{0.1}$$

More precisely, we first consider *approximate* entropy solutions of (0.1), and we bound the difference of such approximate solutions in $L^1(\mathbb{R}^n)$ at time $T > 0$ in terms of their difference at time zero, their smoothness, and how well they satisfy the entropy condition in a certain technical sense. Next, assuming that entropy solutions *do* exist (which we prove in the next chapter), we show that they are unique, and depend continuously on the initial data; i.e., that problem (0.1) is *well-posed* in the sense of Hadamard. Finally, we state the approximation theorem originally presented by N. N. KUZNETSOV that compares an entropy solution of (0.1) to an approximate entropy solution.

§1. Comparing Approximate Entropy Weak Solutions

Definition. The bounded measurable function u is an entropy weak solution of (0.1) if for all positive $\phi \in C^1(\mathbb{R}^{n+1})$ with compact support, all $c \in \mathbb{R}$, and all positive T

$$\begin{aligned} &\Lambda(u,c,T,\phi) := \\ &\quad -\int_0^T \int_{\mathbb{R}^n} |u-c|\,\phi_t + \operatorname{sgn}(u-c)(f(u)-f(c)) \cdot \nabla_x \phi \, dx\, dt \\ &\quad + \int_{\mathbb{R}^n} |u(x,T)-c|\,\phi(x,T)\,dx - \int_{\mathbb{R}^n} |u_0(x)-c|\,\phi(x,0)\,dx \le 0. \end{aligned} \tag{1.1}$$

We introduce a smooth, nonnegative, function $\eta(s)$, $s \in \mathbb{R}$, with support in $[-1,1]$, integral 1, decreasing for positive s, and satisfying $\eta(-s) = \eta(s)$. For each positive ϵ, define $\eta_\epsilon(s) = \frac{1}{\epsilon}\eta(\frac{s}{\epsilon})$. Assume that for each parameter pair (x',t') there is a value $v(x',t')$ and set

$$c = v(x',t')$$

$$\phi(x,t) = \omega(x-x',t-t') := \eta_{\epsilon_0}(t-t') \prod_{i=1}^n \eta_\epsilon(x_i - x_i');$$

ϵ and ϵ_0 are to be chosen later. Next define

$$\Lambda_\epsilon^{\epsilon_0}(u,v,T) := \int_0^T \int_{\mathbb{R}^n} \Lambda(u, v(x',t'), T, \omega(\,\cdot - x', \cdot - t'))\, dx'\, dt'.$$

If u is an entropy weak solution of (0.1), then $\Lambda_\epsilon^{\epsilon_0}(u,v,T) \le 0$ for all v and $T > 0$. In addition, if $\Lambda_\epsilon^{\epsilon_0}(u,v,T) \le C$ for some small positive constant C, then we say that u is an approximate entropy solution of (0.1). (Of course, any function u is an approximate entropy solution of (0.1) for some value of C, but you get the idea.) In fact, if u and v are two approximate solutions of (0.1) we can bound $\|u(\,\cdot\,,T) - v(\,\cdot\,,T)\|_{L^1(\mathbb{R}^n)}$ in terms of the difference in the initial data $\|u(\,\cdot\,,0) - v(\,\cdot\,,0)\|_{L^1(\mathbb{R}^n)}$, and the *average weak truncation errors* $\Lambda_\epsilon^{\epsilon_0}(u,v,T)$, and $\Lambda_\epsilon^{\epsilon_0}(v,u,T)$, together with the following measures of smoothness: For any $w \in L^\infty(\mathbb{R}^n)$, we define the L^1 modulus of smoothness in space

$$\omega_1(w,\epsilon) := \sup_{|\xi|<\epsilon} \int_{\mathbb{R}^n} |w(x+\xi) - w(x)|\, dx,$$

and for any $u\colon [0,T] \longrightarrow L^\infty(\mathbb{R}^n)$, we define the modulus of smoothness in time

$$\nu(u,t,\epsilon) := \sup_{\max(t-\epsilon,0)<t'<\min(T,t+\epsilon)} \|u(\,\cdot\,,t') - u(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)}.$$

Our most general result is the following theorem.

Theorem 1.1. *If $\epsilon_0 \le T$ and $u, v\colon [0,T] \longrightarrow L^\infty(\mathbb{R}^n)$ then*

$$\begin{aligned}
\|u(T) - v(T)\|_{L^1(\mathbb{R}^n)} &\le \|u(0) - v(0)\|_{L^1(\mathbb{R}^n)} \\
&\quad + \frac{1}{2}\left\{\omega_1(u(T),\epsilon) + \omega_1(v(T),\epsilon) + \omega_1(u(0),\epsilon) + \omega_1(v(0),\epsilon)\right\} \\
&\quad + \frac{1}{2}\left\{\nu(u,0,\epsilon_0) + \nu(v,0,\epsilon_0) + \nu(u,T,\epsilon_0) + \nu(v,T,\epsilon_0)\right\} \\
&\quad + \Lambda_\epsilon^{\epsilon_0}(u,v,T) + \Lambda_\epsilon^{\epsilon_0}(v,u,T),
\end{aligned} \tag{1.2}$$

whenever the quantities on the right hand side of (1.2) *exist and are finite.*

Proof. It follows from the symmetries of $\omega(x-x',t-t')$ that

$$\nabla_{x'}\omega(x-x',t-t') = -\nabla_x\omega(x-x',t-t'),$$

and

$$\frac{\partial}{\partial t'}\omega(x-x',t-t') = -\frac{\partial}{\partial t}\omega(x-x',t-t').$$

In addition, we obviously have that

$$|u(x,t) - v(x',t')| = |v(x',t') - u(x,t)|$$

and

$$\operatorname{sgn}(u(x,t)-v(x',t'))(f(u(x,t))-f(v(x',t'))) = \\ \operatorname{sgn}(v(x',t')-u(x,t))(f(v(x',t'))-f(u(x,t))).$$

Therefore, if we fully expand the definitions of $\Lambda_\epsilon^{\epsilon_0}(u,v,T)$ and $\Lambda_\epsilon^{\epsilon_0}(v,u,T)$ into integrals, the quadruple integrals will cancel and we are left with

$$\begin{aligned}
&\Lambda_\epsilon^{\epsilon_0}(u,v,T)+\Lambda_\epsilon^{\epsilon_0}(v,u,T)\\
&\quad=\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|u(x,T)-v(x',t')|\,\omega(x-x',T-t')\,dx\,dx'\,dt'\\
&\quad-\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|u(x,0)-v(x',t')|\,\omega(x-x',0-t')\,dx\,dx'\,dt'\\
&\quad+\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|v(x,T)-u(x',t')|\,\omega(x-x',T-t')\,dx\,dx'\,dt'\\
&\quad-\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|v(x,0)-u(x',t')|\,\omega(x-x',0-t')\,dx\,dx'\,dt'.
\end{aligned}\tag{1.3}$$

This is the fundamental identity on which Kuznetsov's theorem is based; no approximations are involved, and various inequalities can be derived based on how we write the right side of (1.3).

All the terms in the right hand side of (1.3) have the same form; we shall analyze the first. We can write by the triangle inequality

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|u(x,T)-v(x',t')|\,\omega(x-x',T-t')\,dx\,dx'\,dt'\\
&\quad\ge\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|u(x',T)-v(x',T)|\,\omega(x-x',T-t')\,dx\,dx'\,dt'\\
&\quad-\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|u(x,T)-u(x',T)|\,\omega(x-x',T-t')\,dx\,dx'\,dt'\\
&\quad-\int_0^T\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}|v(x',T)-v(x',t')|\,\omega(x-x',T-t')\,dx\,dx'\,dt'.
\end{aligned}\tag{1.4}$$

When the first term on the right of (1.4) is integrated over x and t' we get

$$\frac{1}{2}\|u(T)-v(T)\|_{L^1(\mathbb{R}^n)}.$$

Because $\omega(\xi,t)$ is zero if $|\xi|\ge\epsilon$, the second term on the right of (1.4) can

be bounded as

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x,T) - u(x',T)|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\quad = \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x'+\xi,T) - u(x',T)|\, \omega(\xi, T-t')\, d\xi\, dx'\, dt' \\
&\quad \le \int_0^T \int_{\mathbb{R}^n} \sup_{|\xi|<\epsilon} \left[\int_{\mathbb{R}^n} |u(x'+\xi,T) - u(x',T)|\, dx' \right] \omega(\xi, T-t')\, d\xi\, dt' \\
&\quad = \frac{1}{2}\omega_1(u(T),\epsilon).
\end{aligned}$$

Finally, because $\omega(s,t)$ is zero if $|t| \ge \epsilon_0$, the third term on the right of (1.4) is bounded by

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x',T) - v(x',t')|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\quad \le \int_0^T \sup_{T-\epsilon_0<t'<T} \left[\int_{\mathbb{R}^n} |v(x',T) - v(x',t')|\, dx' \right] \eta_{\epsilon_0}(T-t')\, dt' \\
&\quad = \frac{1}{2}\nu(v,T,\epsilon_0).
\end{aligned}$$

So the first term on the right hand side of (1.3) satisfies

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x,T) - v(x',t')|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\qquad\qquad \ge \frac{1}{2}\|u(T) - v(T)\|_{L^1(\mathbb{R}^n)} - \frac{1}{2}\omega_1(u(T),\epsilon) - \frac{1}{2}\nu(v,T,\epsilon_0).
\end{aligned}$$

Similarly, we find for the second term on the right hand side of (1.3),

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x,0) - v(x',t')|\, \omega(x-x', 0-t')\, dx\, dx'\, dt' \\
&\qquad\qquad \le \frac{1}{2}\|u(0) - v(0)\|_{L^1(\mathbb{R}^n)} + \frac{1}{2}\omega_1(u(0),\epsilon) + \frac{1}{2}\nu(v,0,\epsilon_0).
\end{aligned}$$

The other two terms in (1.3) can be bounded as

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x,T) - u(x',t')|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\qquad\qquad \ge \frac{1}{2}\|v(T) - u(T)\|_{L^1(\mathbb{R}^n)} - \frac{1}{2}\omega_1(v(T),\epsilon) - \frac{1}{2}\nu(u,T,\epsilon_0),
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x,0) - u(x',t')|\, \omega(x-x', 0-t')\, dx\, dx'\, dt' \\
&\qquad\qquad \le \frac{1}{2}\|v(0) - u(0)\|_{L^1(\mathbb{R}^n)} + \frac{1}{2}\omega_1(v(0),\epsilon) + \frac{1}{2}\nu(u,0,\epsilon_0),
\end{aligned}$$

which proves the theorem. □

REMARK 1.1. It is left as an exercise to analyze the case $T < \epsilon_0$.

§2. Consequences of the Approximation Theorem

At this point we shall move a little ahead of ourselves to discover what properties of entropy weak solutions of (0.1) are implied by Theorem 1.1. In the next chapter we shall prove that if $u(0) \in L^\infty(\mathbb{R}^n)$ satisfies $\omega_1(u(0), \epsilon) \to 0$ as $\epsilon \to 0$, and if f is Lipschitz continuous, then an entropy weak solution of (0.1) exists. Furthermore, we shall show that $u(x,t)$ is a bounded, measurable function of $x \in \mathbb{R}^n$ and $t \in [0,T]$ and that $u - u(0)$ is a continuous mapping of $[0,T]$ into $L^1(\mathbb{R}^n)$. Of course, an entropy weak solution u also satisfies $\Lambda_\epsilon^{\epsilon_0}(u,w,T) \le 0$ for all $\epsilon > 0$, $\epsilon_0 > 0$, $T > 0$, and bounded, measurable w. In this section we shall consider only entropy weak solutions of (0.1) with the preceding properties.

Corollary 2.1 (Continuous dependence)**.** *If u and v are entropy weak solutions of* (0.1)*, then for all $T > 0$,*

$$\|u(T) - v(T)\|_{L^1(\mathbb{R}^n)} \le \|u(0) - v(0)\|_{L^1(\mathbb{R}^n)}. \tag{2.1}$$

PROOF. We shall show that the terms in the right hand side of (1.2) that depend on ϵ and ϵ_0 are either zero or negative in the limit as ϵ and ϵ_0 tend to zero. We remark explicitly about the terms involving u; the same arguments hold for v.

By assumption, $\Lambda_\epsilon^{\epsilon_0}(u,v,T) \le 0$ for all ϵ and ϵ_0. Also by assumption, $\omega_1(u(0),\epsilon)$ vanishes as ϵ approaches zero.

Because $u - u(0)$ is assumed to be in $C([0,T], L^1(\mathbb{R}^n))$, we know for any $t \in [0,T]$ that $\|u(t') - u(t)\|_{L^1(\mathbb{R}^n)} \to 0$ as $t' \to t$. Therefore $\nu(u,0,\epsilon_0)$ and $\nu(u,T,\epsilon_0)$ approach zero as $\epsilon_0 \to 0$.

Finally, we note that

$$\omega_1(u(T),\epsilon) \le \omega_1(u(T) - u(0),\epsilon) + \omega_1(u(0),\epsilon).$$

By assumption, the second term on the right vanishes as $\epsilon \to 0$; since $u(T) - u(0) \in L^1(\mathbb{R}^n)$, so does the first term. Therefore, $\omega_1(u(T),\epsilon)$ tends to zero as $\epsilon \to 0$. The theorem follows by letting ϵ and ϵ_0 approach zero. □

Corollary 2.2 (Uniqueness)**.** *If u and v are entropy weak solutions of* (0.1) *and $u(0) = v(0)$, then for all $T > 0$ we have $u(T) = v(T)$.*

PROOF. The result obviously follows from Corollary 2.1. □

Corollary 2.2 implies that if we let $v(x,0) := u(x + \xi, 0)$, then the entropy weak solutions u and v of (0.1) satisfy for all $T > 0$,

$$v(x,T) = u(x + \xi, t),$$

since our assumptions are invariant under translation of the initial data.

Corollary 2.3. (Spatial smoothness)**.** *If u is an entropy weak solution of* (0.1) *then for all $T > 0$ and $\epsilon > 0$*

$$\omega_1(u(T), \epsilon) \le \omega_1(u(0), \epsilon). \tag{2.2}$$

PROOF. For each $\xi \in \mathbb{R}^n$ with $|\xi| < \epsilon$ and for all $T > 0$, (2.1) implies:

$$\int_{\mathbb{R}^n} |u(x+\xi, T) - u(x, T)|\, dx \le \int_{\mathbb{R}^n} |u(x+\xi, 0) - u(x, 0)|\, dx.$$

Therefore, (2.2) follows from the definition of ω_1. □

Corollaries 2.2 and 2.3 imply that the operator $S_t\colon u(0) \longrightarrow u(t)$ is a semigroup: $S_t S_s(u(0)) = S_{s+t}(u(0)) = u(s+t)$. Explicitly, if $w(x, 0) := u(x, s)$ for all x then for all $T > 0$,

$$w(x, T) = u(x, T + s).$$

Corollary 2.4. (Temporal smoothness)**.** *If u is an entropy weak solution of* (0.1) *then for all $T > 0$ and $\epsilon > 0$*

$$\nu(u, T, \epsilon) \le \nu(u, 0, \epsilon). \tag{2.3}$$

PROOF. For each $0 < s < \epsilon$ and for all $T > 0$, we have by (2.1):

$$\|u(T+s) - u(T)\|_{L^1(\mathbb{R}^n)} \le \|u(s) - u(0)\|_{L^1(\mathbb{R}^n)}.$$

Therefore, (2.3) follows from the definition of ν. □

§3. Bounding the Error of Approximations

Being symmetric in u and v, the right side of (1.2) is suitable for comparing two approximate solutions of (0.1). We can use the *a priori* information derived in the previous section to give a new, asymmetric, bound for $\|u(T) - v(T)\|_{L^1(\mathbb{R}^n)}$ that is useful when comparing an entropy weak solution v to an approximate solution u. In addition to the properties proved in the previous section, we shall prove in Chapter 3 that

$$\nu(v, 0, \epsilon_0) \le L\, \omega_1(v(0), \epsilon_0),$$

where $L = \sup_{\xi \in \mathbb{R}} \sum_{j=1}^n |f_j'(\xi)|$.

Specifically, we bound (1.4) in the following way:

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x,T) - v(x',t')|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\quad \geq \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x,T) - v(x,T)|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\qquad - \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x,t') - v(x',t')|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\qquad - \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x,T) - v(x,t')|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\quad \geq \frac{1}{2}\|u(T) - v(T)\|_{L^1(\mathbb{R}^n)} - \frac{1}{2}\omega_1(v(0),\epsilon) - \frac{1}{2} L\, \omega_1(v(0),\epsilon_0).
\end{aligned}$$

Similarly, the second term of (1.3) satisfies

$$\begin{aligned}
&- \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x,0) - v(x',t')|\, \omega(x-x', 0-t')\, dx\, dx'\, dt' \\
&\quad \geq - \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u(x,0) - v(x,0)|\, \omega(x-x', 0-t')\, dx\, dx'\, dt' \\
&\qquad - \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x,t') - v(x',t')|\, \omega(x-x', 0-t')\, dx\, dx'\, dt' \\
&\qquad - \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x,0) - v(x,t')|\, \omega(x-x', 0-t')\, dx\, dx'\, dt' \\
&\quad \geq -\frac{1}{2}\|u(0) - v(0)\|_{L^1(\mathbb{R}^n)} - \frac{1}{2}\omega_1(v(0),\epsilon) - \frac{1}{2} L\, \omega_1(v(0),\epsilon_0).
\end{aligned}$$

The final two terms of (1.3) can be simplified to

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x,T) - u(x',t')|\, \omega(x-x', T-t')\, dx\, dx'\, dt' \\
&\qquad\qquad \geq \frac{1}{2}\|v(T) - u(T)\|_{L^1(\mathbb{R}^n)} - \frac{1}{2}\omega_1(v(0),\epsilon) - \frac{1}{2}\nu(u,T,\epsilon_0),
\end{aligned}$$

and

$$\begin{aligned}
&- \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |v(x,0) - u(x',t')|\, \omega(x-x', 0-t')\, dx\, dx'\, dt' \\
&\qquad\qquad \geq -\frac{1}{2}\|v(0) - u(0)\|_{L^1(\mathbb{R}^n)} - \frac{1}{2}\omega_1(v(0),\epsilon) - \frac{1}{2}\nu(u,0,\epsilon_0).
\end{aligned}$$

Anticipating Chapter 3 somewhat, we can therefore state the following theorem, which is more or less how Kuznetsov stated it in the first place.

Theorem 3.1. *Assume that v_0 is a bounded, measurable function,*

with $\omega_1(v_0,\epsilon)\to 0$ *as* $\epsilon\to 0$*, that* f *is Lipschitz continuous, with*

$$L := \sup_{\xi\in\mathbb{R}} \sum_{j=1}^{n} |f_j'(\xi)| < \infty,$$

and that $v(x,t)$ *is the entropy weak solution of*

$$\begin{aligned} v_t + \nabla\cdot f(v) &= 0, && x\in\mathbb{R}^n,\ t>0,\\ v(x,0) &= v_0(x), && x\in\mathbb{R}^n. \end{aligned}$$

Then for any bounded, measurable u *defined on* $\mathbb{R}^n\times[0,T]$ *we have*

$$\begin{aligned} \|u(T)-v(T)\|_{L^1(\mathbb{R}^n)} \le \|u(0)-v(0)\|_{L^1(\mathbb{R}^n)} &+ 2\omega_1(v_0,\epsilon) + L\,\omega_1(v_0,\epsilon_0)\\ &+ \frac{1}{2}\nu(u,0,\epsilon_0) + \frac{1}{2}\nu(u,T,\epsilon_0) + \Lambda_\epsilon^{\epsilon_0}(u,v,T). \end{aligned}$$

Chapter 3

Monotone Finite Difference Methods

In this chapter we study monotone finite difference methods for the approximation of $u(x,t)$, the solution of the scalar equation

$$\begin{aligned} &u_t + \nabla_x \cdot f(u) = 0, && x \in \mathbb{R}^n,\ t > 0, \\ &u(x,0) = u_0(x), && x \in \mathbb{R}^n. \end{aligned} \tag{0.1}$$

The solutions of the finite difference equations will be used as approximate solutions of (0.1) and analyzed using KUZNETSOV'S theorem in Chapter 2. We study in depth only the ENGQUIST-OSHER scheme; this will allow us to prove existence of solutions of (0.1). The analysis of other monotone schemes is so similar that it does not warrant repeating. The references for this chapter are the papers by SANDERS and CRANDALL AND MAJDA.

§1. The Engquist-Osher Method

Let us first consider the equation in one space dimension,

$$\begin{aligned} &u_t + f(u)_x = 0, && x \in \mathbb{R},\ t > 0, \\ &u(x,0) = u_0(x), && x \in \mathbb{R}. \end{aligned} \tag{1.1}$$

To derive a numerical method for (1.1) it is natural to substitute finite difference approximations to the derivatives in (1.1). To that end, let h and Δt be fixed positive numbers and introduce the quantities

$$\begin{aligned} &U_i^k \approx u(ih, k\Delta t), && i \in \mathbb{Z},\ k \geq 0, \\ &\Delta_t^+ (U^k)_i := \frac{U_i^{k+1} - U_i^k}{\Delta t} \approx \frac{\partial u(ih, k\Delta t)}{\partial t}, && i \in \mathbb{Z},\ k \geq 0, \\ &\Delta_x (f(U^k))_i := \frac{f(U_{i+1}^k) - f(U_{i-1}^k)}{2h} \approx \frac{\partial f(u(ih, k\Delta t))}{\partial x}, && i \in \mathbb{Z},\ k \geq 0. \end{aligned}$$

The quantity $\Delta_x (f(U^k))_i$ is a second order, centered difference approximation to $f(u(ih, k\Delta t))_x$. Then a *possible* finite difference scheme is

$$\begin{aligned} &\Delta_t^+ (U^k)_i + \Delta_x (f(U^k))_i = 0, && i \in \mathbb{Z},\ k \geq 0, \\ &U_i^0 = \frac{1}{h} \int_{ih}^{(i+1)h} u_0(x)\, dx, && i \in \mathbb{Z}. \end{aligned} \tag{1.2}$$

In this equation we can solve explicitly for U_i^{k+1} if we know U_i^k for all i:

$$U_i^{k+1} = U_i^k - \frac{\Delta t}{2h}(f(U_{i+1}^k) - f(U_{i-1}^k)). \tag{1.3}$$

We would like the finite-difference scheme to have the same qualitative properties that we expect the entropy solution of (1.1) to have. In particular, we would like it to be the case that if $U_i^k \geq V_i^k$ for all $i \in \mathbb{Z}$, then $U_i^{k+1} \geq V_i^{k+1}$ for all $i \in \mathbb{Z}$. But this will not be so for (1.3)—if f is increasing, for example, then increasing U_{i+1}^k will *decrease* U_i^{k+1}. Thus, there is no chance that the finite difference formula (1.3) will be order preserving.

One can attempt to rectify this problem by introducing a one-sided divided difference in x,

$$\Delta_x^-(f(U^k))_i := \frac{f(U_i^k) - f(U_{i-1}^k)}{h} \approx \frac{\partial f(u(ih, k\Delta t))}{\partial x}, \qquad i \in \mathbb{Z},\ k \geq 0,$$

and solving the problem

$$\begin{aligned} &\Delta_t^+(U^k)_i + \Delta_x^-(f(U^k))_i = 0, && i \in \mathbb{Z},\ k \geq 0, \\ &U_i^0 = \frac{1}{h}\int_{ih}^{(i+1)h} u_0(x)\,dx, && i \in \mathbb{Z}. \end{aligned} \tag{1.4}$$

For this equation, one has the explicit formula

$$U_i^{k+1} = U_i^k - \frac{\Delta t}{h}(f(U_i^k) - f(U_{i-1}^k)). \tag{1.5}$$

This formula will be order preserving if and only if the function

$$G(r, s) := r - \frac{\Delta t}{h}(f(r) - f(s))$$

is increasing in r and s, i.e., $G_r \geq 0$ and $G_s \geq 0$. We can calculate explicitly

$$G_r(r, s) = 1 - \frac{\Delta t}{h} f'(r) \quad \text{and} \quad G_s(r, s) = \frac{\Delta t}{h} f'(s).$$

Thus, (1.4) is order preserving if and only if

$$\begin{aligned} &f'(\xi) \geq 0 \quad \text{and} \\ &\frac{\Delta t}{h} f'(\xi) \leq 1 \end{aligned}$$

for all $\xi \in \mathbb{R}$. Condition (1.5) means that the scheme is an *upwind* scheme: the characteristic lines have positive slope, and to calculate the value of U_i^{k+1} we are using spatial differences taken to the left of U_i^k, "into the wind." Condition (1.5) is called a CFL condition and was introduced first by Courant, Friedrichs, and Lewy.

So, we have an order preserving scheme if Δt is small enough and f' is positive and bounded. If f' is negative and bounded then we can set

$$\Delta_x^+(f(U^k))_i := \frac{f(U_{i+1}^k) - f(U_i^k)}{h} \approx \frac{\partial f(u(ih, k\Delta t))}{\partial x}, \qquad i \in \mathbb{Z},\ k \geq 0$$

and solve the finite difference equations

$$\begin{aligned} &\Delta_t^+(U^k)_i + \Delta_x^+(f(U^k))_i = 0, && i \in \mathbb{Z},\ k \geq 0, \\ &U_i^0 = \frac{1}{h}\int_{ih}^{(i+1)h} u_0(x)\,dx, && i \in \mathbb{Z}. \end{aligned} \tag{1.6}$$

In the same way as we analyzed (1.4), we see that (1.6) will be order preserving if

$$-\frac{\Delta t}{h} f'(\xi) \leq 1, \quad \xi \in \mathbb{R}.$$

For a general function $f(\xi)$ that increases for some ξ and decreases for others, neither (1.4) or (1.6) is satisfactory. Engquist and Osher introduced the following method that *is* suitable for any f.

Without loss of generality, we let $f(0) = 0$. We shall assume that f is Lipschitz continuous, i.e., there exists a constant L such that for all $\zeta, \xi \in \mathbb{R}$, $|f(\zeta) - f(\xi)| \leq L|\zeta - \xi|$. Then, by the Radon-Nikodym theorem there exists *a.e.* a derivative $f'(\xi)$ with $|f'(\xi)| \leq L$ and for all $\xi, \zeta \in \mathbb{R}$,

$$f(\zeta) - f(\xi) = \int_\xi^\zeta f'(s)\,ds.$$

Using this f', we can decompose f into its increasing and decreasing parts:

$$f^+(\xi) = \int_0^\xi f'(\zeta) \vee 0\,d\zeta \quad \text{and} \quad f^-(\xi) = \int_0^\xi f'(\zeta) \wedge 0\,d\zeta. \tag{1.7}$$

Because $f'(\zeta) \vee 0 \geq 0$, $f'(\zeta) \wedge 0 \leq 0$, and $f'(\zeta) = f'(\zeta) \vee 0 + f'(\zeta) \wedge 0$, we know that f^+ is increasing, f^- is decreasing, and for all $\zeta \in \mathbb{R}$ $f(\zeta) = f^+(\zeta) + f^-(\zeta)$. For later purposes we also define $f^t = f^+ - f^-$; note that $(f^t)' = (f^+)' - (f^-)' = f' \vee 0 - f' \wedge 0 = |f'|$.

Now that we have decomposed f, we can difference the increasing part to the left and the decreasing part to the right to get the Engquist-Osher scheme:

$$\begin{aligned} &\Delta_t^+(U^k)_i + \Delta_x^-(f^+(U^k))_i + \Delta_x^+(f^-(U^k))_i = 0, && i \in \mathbb{Z},\ k \geq 0, \\ &U_i^0 = \frac{1}{h}\int_{ih}^{(i+1)h} u_0(x)\,dx, && i \in \mathbb{Z}. \end{aligned} \tag{1.8}$$

We next consider approximating (0.1) in several space dimensions. In (0.1), $f(u)$ is a vector $(f_1(u), \ldots, f_n(u))$, so one extends the one-dimensional Engquist-Osher method by splitting each of the components $f_j(u)$ into its increasing and decreasing parts, $f_j(u) = f_j^+(u) + f_j^-(u)$. The index $i \in \mathbb{Z}$ will no longer suffice; now points in $\mathbb{Z}^n$ are indexed by the multi-index

$\nu = (\nu_1, \ldots, \nu_n)$. Recall that the unit vector e_j has a single 1 in the jth component. Now the approximations for each $\nu \in \mathbb{Z}^n$ and $k \geq 0$ will be

$$
\begin{aligned}
&U_\nu^k \approx u(\nu h, k\Delta t),\\
&\Delta_t^+(U^k)_\nu := \frac{U_\nu^{k+1} - U_\nu^k}{\Delta t} \approx \frac{\partial u(\nu h, k\Delta t)}{\partial t},\\
&\Delta_j^-(f_j^+(U^k))_\nu := \frac{f_j^+(U_\nu^k) - f_j^+(U_{\nu-e_j}^k)}{h} \approx \frac{\partial f_j^+(u(\nu h, k\Delta t))}{\partial x_j},\\
&\Delta_j^+(f_j^-(U^k))_\nu := \frac{f_j^-(U_{\nu+e_j}^k) - f_j^-(U_\nu^k)}{h} \approx \frac{\partial f_j^-(u(\nu h, k\Delta t))}{\partial x_j}.
\end{aligned}
$$

Using this (increasingly Byzantine) notation, the Engquist-Osher method in several space dimensions can be written as

$$
\begin{aligned}
&\Delta_t^+(U^k)_\nu + \sum_{j=1}^n \left[\Delta_j^-(f_j^+(U^k))_\nu + \Delta_j^+(f_j^-(U^k))_\nu\right] = 0,\\
&\qquad\qquad \nu \in \mathbb{Z}^n,\ k \geq 0,\\
&U_\nu^0 = \frac{1}{h^n}\int_{\nu_1 h}^{(\nu_1+1)h} \cdots \int_{\nu_n h}^{(\nu_n+1)h} u_0(x)\,dx, \qquad \nu \in \mathbb{Z}^n.
\end{aligned}
\tag{1.9}
$$

Remark. A general explicit finite difference scheme of the form

$$
\Delta_t^+(U^k)_i + \frac{F(U_{i-N}^k, \ldots U_{i+N+1}^k) - F(U_{i-N-1}^k, \ldots U_{i+N}^k)}{h} = 0, \quad i \in \mathbb{Z},
$$

for the problem

$$
u_t + f(u)_x = 0, \qquad x \in \mathbb{R},\ t > 0,
$$

is said to be *conservative,* since

$$
\sum_{i\in\mathbb{Z}} U_i^{k+1} = \sum_{i\in\mathbb{Z}} U_i^k
$$

whenever $U^k \in L^1(\mathbb{Z})$. The Engquist-Osher scheme is conservative, with $N = 0$ and

$$
F(U_i^k, U_{i+1}^k) = f^+(U_i^k) + f^-(U_{i+1}^k).
$$

A general method is *consistent* if for all $c \in \mathbb{R}$,

$$
F(c, \ldots, c) = f(c);
$$

the Engquist-Osher method is consistent, since, by (1.7),

$$
F(c, c) = f^+(c) + f^-(c) = f(c).
$$

Finally, a method is *monotone* if, under some conditions on Δt, h, and f, we have that

$$
U_i^k \geq V_i^k \quad \text{for all } i \in \mathbb{Z} \Longrightarrow U_i^{k+1} \geq V_i^{k+1} \quad \text{for all } i \in \mathbb{Z}.
$$

To show this, it is sufficient to show that

$$\frac{\partial U_i^{k+1}}{\partial U_j^k} \geq 0, \qquad j \in \mathbb{Z}.$$

For the Engquist-Osher scheme, we have the explicit formula

$$U_i^{k+1} = U_i^k - \frac{\Delta t}{h}\{f^+(U_i^k) - f^+(U_{i-1}^k) + f^-(U_{i+1}^k) - f^-(U_i^k)\}. \tag{1.10}$$

If $|j-i| > 1$, then U_i^{k+1} does not depend on U_j^k, so $\dfrac{\partial U_i^{k+1}}{\partial U_j^k} = 0$. We calculate from (1.10) that

$$\frac{\partial U_i^{k+1}}{\partial U_{i-1}^k} = \frac{\Delta t}{h}(f^+)'(U_{i-1}^k) \geq 0 \quad \text{and} \quad \frac{\partial U_i^{k+1}}{\partial U_{i+1}^k} = -\frac{\Delta t}{h}(f^-)'(U_{i+1}^k) \geq 0.$$

Finally, we have that

$$\begin{aligned}\frac{\partial U_i^{k+1}}{\partial U_i^k} &= 1 - \frac{\Delta t}{h}\{(f^+)'(U_i^k) - (f^-)'(U_i^k)\} \\ &= 1 - \frac{\Delta t}{h}|f'(U_i^k)|,\end{aligned}$$

which will be nonnegative if the CFL condition

$$\frac{\Delta t}{h}|f'(\xi)| \leq 1, \qquad \xi \in \mathbb{R},$$

holds.

Other examples of consistent, conservative, monotone, finite difference schemes are Godunov's scheme (described later) and the Lax-Friedrich scheme

$$\Delta_t^+(U^k)_i + \frac{f(U_{i+1}^k) - f(U_{i-1}^k)}{2h} - \mu\frac{U_{i-1}^k - 2U_i^k + U_{i+1}^k}{h^2} = 0, \quad i \in \mathbb{Z},\ k \geq 0,$$

when $\mu \geq \frac{h}{2}\max|f'(\xi)|$.

§2. Properties of the Engquist-Osher Method

Theorem 2.1. *Assume that there exists a number L such that*

$$\sup_\xi \sum_i |f_i'(\xi)| \leq L \quad \textit{and} \quad \frac{\Delta t}{h}L \leq 1. \tag{2.1}$$

Assume U^k and V^k are given, and that U^{k+1} and V^{k+1} are defined by (1.9).

Then

(1) *If* $U_\nu^k \ge V_\nu^k$ *for all* $|\nu| \le l+1$, *then* $U_\nu^{k+1} \ge V_\nu^{k+1}$ *for all* $|\nu| \le l$.

(2) $$\sum_{|\nu|\le l} |U_\nu^{k+1} - V_\nu^{k+1}| \le \sum_{|\nu|\le l+1} |U_\nu^k - V_\nu^k|.$$

(3) $$\sup_{|\nu|\le l} U_\nu^{k+1} \le \sup_{|\nu|\le l+1} U_\nu^k \quad \textit{and} \quad \inf_{|\nu|\le l} U_\nu^{k+1} \ge \inf_{|\nu|\le l+1} U_\nu^k.$$

(4) $$\sum_{|\nu|\le l} |U_\nu^{k+1} - U_\nu^k| \le \frac{L\Delta t}{h} \sum_{|\nu|\le l+1} \sum_{j=1}^n |U_\nu^k - U_{\nu-e_j}^k|.$$

(5) $$\sum_{|\nu|\le l} \sum_{j=1}^n |U_\nu^{k+1} - U_{\nu-e_j}^{k+1}| \le \sum_{|\nu|\le l+1} \sum_{j=1}^n |U_\nu^k - U_{\nu-e_j}^k|.$$

Remark 2.1. The following inequalities, which we shall now take for granted, can be derived by letting l tend to infinity in (1) to (5):

(1) If $U_\nu^k \ge V_\nu^k$ for all $\nu \in \mathbb{Z}^n$, then $U_\nu^{k+1} \ge V_\nu^{k+1}$ for all $\nu \in \mathbb{Z}^n$.

(2) $\|U^{k+1} - V^{k+1}\|_{L^1(\mathbb{Z}^n)} \le \|U^k - V^k\|_{L^1(\mathbb{Z}^n)}$.

(3) $$\sup_{\nu\in\mathbb{Z}^n} U_\nu^{k+1} \le \sup_{\nu\in\mathbb{Z}^n} U_\nu^k \quad \text{and} \quad \inf_{\nu\in\mathbb{Z}^n} U_\nu^{k+1} \ge \inf_{\nu\in\mathbb{Z}^n} U_\nu^k.$$

(4) $\|U^{k+1} - U^k\|_{L^1(\mathbb{Z}^n)} \le \dfrac{L\Delta t}{h}\|U^k\|_{\mathrm{BV}(\mathbb{Z}^n)} \le \dfrac{L\Delta t}{h}\|U^0\|_{\mathrm{BV}(\mathbb{Z}^n)}$ (see (5)).

(5) $\|U^{k+1}\|_{\mathrm{BV}(\mathbb{Z}^n)} \le \|U^k\|_{\mathrm{BV}(\mathbb{Z}^n)} \le \|U^0\|_{\mathrm{BV}(\mathbb{Z}^n)}$, by induction.

Proof of Theorem 2.1. We have the explicit formulae for U_ν^{k+1}:

$$\begin{aligned}
U_\nu^{k+1} &= G(U_\nu^k, U_{\nu-e_1}^k, U_{\nu+e_1}^k, \dots, U_{\nu-e_n}^k, U_{\nu+e_n}^k) \\
&:= U_\nu^k - \Delta t \left[\sum_{j=1}^n \left[\Delta_j^-(f_j^+(U^k))_\nu + \Delta_j^+(f_j^-(U^k))_\nu \right] \right]. \\
&= U_\nu^k - \frac{\Delta t}{h} \sum_{j=1}^n (f_j^+(U_\nu^k) - f_j^-(U_\nu^k)) \\
&\qquad - \frac{\Delta t}{h} \sum_{j=1}^n (f_j^-(U_{\nu+e_j}^k) - f_j^+(U_{\nu-e_j}^k)) \\
&= U_\nu^k - \frac{\Delta t}{h} \sum_{j=1}^n (f_j^-(U_{\nu+e_j}^k) + f_j^+(U_\nu^k)) \\
&\qquad + \frac{\Delta t}{h} \sum_{j=1}^n (f_j^-(U_\nu^k) + f_j^+(U_{\nu-e_j}^k))
\end{aligned}$$

It's clear from these formulae that U_ν^{k+1} for $|\nu| \le l$ depend only on U_ν^k for $|\nu| \le l+1$, and that if $U_\nu^k = C$ for $|\nu| \le l+1$, then $U_\nu^{k+1} = C$ for $|\nu| \le l$.

(1) It suffices to show that

$$\frac{\partial G}{\partial U_\nu^k} \ge 0 \quad \text{and} \quad \frac{\partial G}{\partial U_{\nu \pm e_j}^k} \ge 0, \quad j = 1, \dots, n.$$

We calculate explicitly

$$\begin{aligned}
\frac{\partial G}{\partial U_\nu^k} &= 1 - \frac{\Delta t}{h} \sum_{j=1}^n ((f_j^+)'(U_\nu^k) - (f_j^-)'(U_\nu^k)) \\
&= 1 - \frac{\Delta t}{h} \sum_{j=1}^n (f_j^t)'(U_\nu^k) \\
&\ge 1 - \frac{\Delta t L}{h} \\
&\ge 0
\end{aligned}$$

by assumption (2.1). In addition,

$$\frac{\partial G}{\partial U_{\nu+e_j}^k} = -\frac{\Delta t}{h}(f_j^-)'(U_{\nu+e_j}^k) \ge 0$$

and

$$\frac{\partial G}{\partial U_{\nu-e_j}^k} = \frac{\Delta t}{h}(f_j^+)'(U_{\nu-e_j}^k) \ge 0.$$

So U^{k+1} is a monotone function of U^k.

(2) We have from (2.1) that

$$\begin{aligned}
U_\nu^{k+1} - V_\nu^{k+1} = \\
[U_\nu^k - V_\nu^k - \frac{\Delta t}{h} \sum_{j=1}^n (f_j^+(U_\nu^k) - f_j^+(V_\nu^k)) + \frac{\Delta t}{h} \sum_{j=1}^n (f_j^-(U_\nu^k) - f_j^-(V_\nu^k))] \\
+ \frac{\Delta t}{h} \sum_{j=1}^n (f_j^+(U_{\nu-e_j}^k) - f_j^+(V_{\nu-e_j}^k)) \\
- \frac{\Delta t}{h} \sum_{j=1}^n (f_j^-(U_{\nu+e_j}^k) - f_j^-(V_{\nu+e_j}^k)).
\end{aligned}$$

Because f_j^+ increases, f_j^- decreases, and (2.1) holds, the quantity in square brackets above, $(U_\nu^k - V_\nu^k)$, $(f_j^+(U_\nu^k) - f_j^+(V_\nu^k))$, and $-(f_j^-(U_\nu^k) - f_j^-(V_\nu^k))$

all have the same sign. Therefore, taking absolute values we have

$$\begin{aligned}
&|U_\nu^{k+1} - V_\nu^{k+1}| \\
&\le |U_\nu^k - V_\nu^k - \frac{\Delta t}{h}\sum_{j=1}^n (f_j^+(U_\nu^k) - f_j^+(V_\nu^k)) + \frac{\Delta t}{h}\sum_{j=1}^n (f_j^-(U_\nu^k) - f_j^-(V_\nu^k))| \\
&\qquad + \frac{\Delta t}{h}\sum_{j=1}^n |f_j^+(U_{\nu-e_j}^k) - f_j^+(V_{\nu-e_j}^k)| \\
&\qquad + \frac{\Delta t}{h}\sum_{j=1}^n |f_j^-(U_{\nu+e_j}^k) - f_j^-(V_{\nu+e_j}^k)| \\
&= |U_\nu^k - V_\nu^k| - \frac{\Delta t}{h}\sum_{j=1}^n |f_j^+(U_\nu^k) - f_j^+(V_\nu^k)| - \frac{\Delta t}{h}\sum_{j=1}^n |f_j^-(U_\nu^k) - f_j^-(V_\nu^k)| \\
&\qquad + \frac{\Delta t}{h}\sum_{j=1}^n |f_j^+(U_{\nu-e_j}^k) - f_j^+(V_{\nu-e_j}^k)| \\
&\qquad + \frac{\Delta t}{h}\sum_{j=1}^n |f_j^-(U_{\nu+e_j}^k) - f_j^-(V_{\nu+e_j}^k)|.
\end{aligned}$$

Summing this inequality over all $|\nu| \le l$ yields

$$\begin{aligned}
\sum_{|\nu|\le l} |U_\nu^{k+1} - V_\nu^{k+1}| \le \sum_{|\nu|\le l} |U_\nu^k - V_\nu^k| \\
&- \frac{\Delta t}{h}\sum_{j=1}^n \sum_{\substack{\nu_j=0 \\ |\nu|\le l}} |f_j^-(U_{\nu-le_j}^k) - f_j^-(V_{\nu-le_j}^k)| \\
&- \frac{\Delta t}{h}\sum_{j=1}^n \sum_{\substack{\nu_j=0 \\ |\nu|\le l}} |f_j^+(U_{\nu+le_j}^k) - f_j^+(V_{\nu+le_j}^k)|. \\
&+ \frac{\Delta t}{h}\sum_{j=1}^n \sum_{\substack{\nu_j=0 \\ |\nu|\le l}} |f_j^-(U_{\nu+(l+1)e_j}^k) - f_j^-(V_{\nu+(l+1)e_j}^k)| \\
&+ \frac{\Delta t}{h}\sum_{j=1}^n \sum_{\substack{\nu_j=0 \\ |\nu|\le l}} |f_j^+(U_{\nu-(l+1)e_j}^k) - f_j^+(V_{\nu-(l+1)e_j}^k)|.
\end{aligned}$$

The sums here are over the $(n-1)$–dimensional hypercubes $|\nu| \le l$ with $\nu_j = 0$, each of which when added to le_j gives the jth "side" of the n-dimensional hypercube $|\nu| \le l$. Because of (2.1), we only increase the right side of the previous inequality by substituting $|U_{\nu+(l+1)e_j}^k - V_{\nu+(l+1)e_j}^k|$ for

$\frac{\Delta t}{h}|f_j^-(U^k_{\nu+(l+1)e_j}) - f_j^-(V^k_{\nu+(l+1)e_j})|$, etc. Thus we arrive at

$$\sum_{|\nu|\le l} |U_\nu^{k+1} - V_\nu^{k+1}| \le \sum_{|\nu|\le l+1} |U_\nu^k - V_\nu^k|,$$

which proves (2).

(3) Let $V_\nu^k = \sup_{|\mu|\le l+1} U_\mu^k$ for $|\nu| \le l+1$. Property (1) implies that $V_\nu^{k+1} \ge U_\nu^{k+1}$ for all $|\nu| \le l$, which implies the first part of (3). The second part follows in the same way.

(4) Equation (2.1) and condition (2.1) imply that

$$\begin{aligned}
&\sum_{|\nu|\le l} |U_\nu^{k+1} - U_\nu^k| \\
&\qquad \le \frac{\Delta t}{h} \sum_{|\nu|\le l} \sum_{j=1}^n \{|f_j^+(U_\nu^k) - f_j^+(U^k_{\nu-e_j})| + |f_j^-(U^k_{\nu+e_j}) - f_j^-(U_\nu^k)|\} \\
&\qquad \le \frac{\Delta t}{h} \sum_{|\nu|\le l+1} \sum_{j=1}^n |f_j^t(U_\nu^k) - f_j^t(U^k_{\nu-e_j})| \\
&\qquad \le \frac{\Delta t L}{h} \sum_{|\nu|\le l+1} \sum_{j=1}^n |U_\nu^k - U^k_{\nu-e_j}|
\end{aligned}$$

In fact, our argument shows that in the last inequality we can replace $L = \sup_\xi \sum_{j=1}^n |f_j'(\xi)|$ by $\max_j \sup_\xi |f_j'(\xi)|$.

(5) This property follows directly from Property 2. For each j let $V_\nu^k := U^k_{\nu-e_j}$. Then Property 2 implies that

$$\sum_{|\nu|\le l} |U_\nu^{k+1} - U^{k+1}_{\nu-e_j}| \le \sum_{|\nu|\le l+1} |U_\nu^k - U^k_{\nu-e_j}|.$$

Summing this inequality over $j = 1, \dots, n$ gives (5). □

Remark 2.2. Because solutions of (1.9) satisfy the maximum and minimum principles (3), we can replace the definition of L in (2.1) by

$$L = \sup_{\inf_\nu U_\nu^k \le \xi \le \sup_\nu U_\nu^k} \sum_{j=1}^n |f_j'(\xi)|.$$

If we are working with a family of bounded U^k, with $|U_\nu^k| \le M$, say, for all $\nu \in \mathbb{Z}^n$ and $k \ge 0$, then we can take

$$L = \sup_{|\xi|\le M} \sum_{j=1}^n |f_j'(\xi)|.$$

This allows us to consider flux functions like $f_j(u) = u^2$, which otherwise would not satisfy (2.1).

§3. Discrete and Continuous Norms of Approximations

Before using Kuznetsov's Theorem to bound in $L^1(\mathbb{R}^n)$ the difference of two numerical approximations to (0.1), we must first interpret the function U^k_ν, defined for the discrete values of $\nu \in \mathbb{Z}^n$ and $k \geq 0$, as a function $u^h(x)$ defined for all $x \in \mathbb{R}^n$.

As in the previous section, we fix positive numbers h and Δt, and we assume that (2.1) is satisfied. Let $\chi(x) := \chi_{[0,1)^n}(x)$ be the characteristic function of the n-dimensional unit hypercube $I := [0,1)^n \subset \mathbb{R}^n$, i.e.,

$$\chi(x) = \begin{cases} 1, & 0 \leq x_j < 1,\ j = 1, \dots, n, \\ 0, & \text{otherwise.} \end{cases}$$

For each $\nu \in \mathbb{Z}^n$ define $\chi^h_\nu(x) := \chi(\frac{x}{h} - \nu)$, so that χ^h_ν is the characteristic function of the cube $I_{h\nu} := \prod_{j=1}^n [\nu_j h, (\nu_j + 1)h)$. For each $k \geq 0$ let

$$u^h(x, k\Delta t) := \sum_{\nu \in \mathbb{Z}^n} U^k_\nu \chi^h_\nu(x),$$

and for $k\Delta t < t < (k+1)\Delta t$, so that $t = (k+\alpha)\Delta t$ for some $0 < \alpha < 1$, define

$$u^h(x,t) := \alpha u^h(x, (k+1)\Delta t) + (1-\alpha) u^h(x, k\Delta t). \tag{3.1}$$

From the definition, it is clear that $u^h(x,t)$ is piecewise constant in x and piecewise linear in t. The following Lemma allows us to relate continuous norms of $u^h(\,\cdot\,, t)$ to discrete norms of U^k:

Lemma 3.1. *If for all $k \geq 0$, $U^k \in L^1(\mathbb{Z}^n)$ then*
(1) $\|u^h(\,\cdot\,, k\Delta t)\|_{L^1(\mathbb{R}^n)} = h^n \|U^k\|_{L^1(\mathbb{Z}^n)}$, $k \geq 0$.
(2) $\|u^h(\,\cdot\,, k\Delta t)\|_{\mathrm{BV}(\mathbb{R}^n)} = h^{n-1} \|U^k\|_{\mathrm{BV}(\mathbb{Z}^n)}$, $k \geq 0$.
(3) *For $k\Delta t < t < (k+1)\Delta t$,*

$$\|u^h_t(\,\cdot\,, t)\|_{L^1(\mathbb{R}^n)} = \frac{h^n}{\Delta t} \|U^{k+1} - U^k\|_{L^1(\mathbb{Z}^n)}, \qquad k \geq 0.$$

Proof. (1) This is clear from

$$\begin{aligned} \|u^h(\,\cdot\,, k\Delta t)\|_{L^1(\mathbb{R}^n)} &= \sum_{\nu \in \mathbb{Z}^n} |U^k_\nu| \int_{\mathbb{R}^n} \chi^h_\nu \, dx \\ &= \sum_{\nu \in \mathbb{Z}^n} |U^k_\nu| h^n \\ &= h^n \|U^k\|_{L^1(\mathbb{Z}^n)}. \end{aligned}$$

(2) We note that for any $\epsilon > 0$,

$$\|v\|_{\mathrm{BV}(\mathbb{R}^n)} = \sum_{j=1}^n \sup_{|\tau| \leq \epsilon} \frac{1}{|\tau|} \int_{\mathbb{R}^n} |v(x + \tau e_j) - v(x)| \, dx,$$

That is, we can take the supremum over small τ rather than over all τ and get the same result. This is simply because if $|\tau| > \epsilon$, and $\sigma = \tau/d$ satisfies $|\sigma| \le \epsilon$, then

$$\begin{aligned}
&\frac{1}{|\tau|}\int_{\mathbb{R}^n} |v(x+\tau e_j) - v(x)|\,dx \\
&\qquad \le \frac{1}{|\tau|}\int_{\mathbb{R}^n} \sum_{l=1}^{d} |v(x+l\sigma e_j) - v(x+(l-1)\sigma e_j)|\,dx \\
&\qquad = \frac{d}{|\tau|}\int_{\mathbb{R}^n} |v(x+\sigma e_j) - v(x)|\,dx \\
&\qquad = \frac{1}{|\sigma|}\int_{\mathbb{R}^n} |v(x+\sigma e_j) - v(x)|\,dx.
\end{aligned}$$

Therefore the supremum over all τ is the same as the supremum over all σ with $|\sigma| \le \epsilon$ for any positive ϵ.

So we can assume that $|\tau| \le h/2$ when calculating $\|u^h(\,\cdot\,,t)\|_{\mathrm{BV}(\mathbb{R}^n)}$, with $t = k\Delta t$. We calculate

$$\int_{\mathbb{R}^n} |u^h(x+\tau e_j,t) - u^h(x,t)|\,dx = h^{n-1}|\tau| \sum_{\nu\in\mathbb{Z}^n} |U^k_{\nu+e_j} - U^k_\nu|.$$

(See the (missing) picture for what happens in two space dimensions.) So

$$\begin{aligned}
\|u^h(\,\cdot\,,t)\|_{\mathrm{BV}(\mathbb{R}^n)} &= \sum_{j=1}^{n} \sup_{|\tau|\le h/2} \frac{1}{|\tau|}\int_{\mathbb{R}^n} |u^h(x+\tau e_j,t) - v^h(x,t)|\,dx \\
&= \sum_{j=1}^{n} h^{n-1} \sum_{\nu\in\mathbb{Z}^n} |U^k_{\nu+e_j} - U^k_\nu| \\
&= h^{n-1}\|U^k\|_{\mathrm{BV}(\mathbb{Z}^n)}.
\end{aligned}$$

(3) Note that $\alpha = (t - k\Delta t)/\Delta t$, so $d\alpha/dt = 1/\Delta t$. From (3.1), we have that

$$\begin{aligned}
\|u^h_t(x,t)\|_{L^1(\mathbb{R}^n)} &= \left\| \frac{u^h(x,(k+1)\Delta t) - u^h(x,k\Delta t)}{\Delta t} \right\|_{L^1(\mathbb{R}^n)} \\
&= \frac{h^n}{\Delta t}\|U^{k+1} - U^k\|_{L^1(\mathbb{Z}^n)}
\end{aligned}$$

by Property (1). □

The initial data for the numerical method is calculated from the formula

$$U^0_\nu = \frac{1}{h^n}\int_{I_{h\nu}} u_0(x)\,dx, \quad \nu \in \mathbb{Z}^n. \tag{3.2}$$

We now calculate norms of U^0 in terms of norms of u_0.

Lemma 3.2. *Let U^0 be calculated by the formula* (3.2)*. Then*
(1) $h^n\|U^0\|_{L^1(\mathbb{Z}^n)} \le \|u_0\|_{L^1(\mathbb{R}^n)}$,
(2) $h^{n-1}\|U^0\|_{\mathrm{BV}(\mathbb{Z}^n)} \le \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}$, *and*
(3) $\|u_0 - u^h(\,\cdot\,,0)\|_{L^1(\mathbb{R}^n)} \le h\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}$.

PROOF. (1) We have for each $\nu \in \mathbb{Z}^n$,

$$h^n|U^0_\nu| \le \int_{I_{h\nu}} |u_0(x)|\,dx.$$

Summing this over all $\nu \in \mathbb{Z}^n$ gives (1).

(2) For the variation bound, we calculate

$$\begin{aligned}
h^n\|U^0\|_{\mathrm{BV}(\mathbb{Z}^n)} &= h^n \sum_{j=1}^n \sum_{\nu\in\mathbb{Z}^n} |U^0_\nu - U^0_{\nu-e_j}| \\
&= \sum_{j=1}^n \sum_{\nu\in\mathbb{Z}^n} \left| \int_{I_{h\nu}} (u_0(x) - u_0(x - he_j))\,dx \right| \\
&\le h \sum_{j=1}^n \sum_{\nu\in\mathbb{Z}^n} \frac{1}{h} \int_{I_{h\nu}} |u_0(x) - u_0(x - he_j)|\,dx \\
&\le h\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}.
\end{aligned}$$

(3) We apply the triangle inequality to calculate,

$$\begin{aligned}
&\|u_0 - u^h(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)} \\
&= \sum_{\nu\in\mathbb{Z}^n} \int_{I_{h\nu}} \left| u_0(y) - \frac{1}{h^n} \int_{I_{h\nu}} u_0(x)\,dx \right| dy \\
&= \sum_{\nu\in\mathbb{Z}^n} \int_{I_{h\nu}} \left| \frac{1}{h^n} \int_{I_{h\nu}} (u_0(y) - u_0(x))\,dx \right| dy \\
&\le \frac{1}{h^n} \sum_{\nu\in\mathbb{Z}^n} \int_{I_{h\nu}} \int_{I_{h\nu}} |u_0(y) - u_0(x)|\,dx\,dy \\
&\le \frac{1}{h^n} \sum_{\nu\in\mathbb{Z}^n} \int_{I_{h\nu}} \int_{I_{h\nu}} \sum_{j=1}^n |u_0(z(x,y,j)) - u_0(z(x,y,j-1))|\,dx\,dy,
\end{aligned} \tag{3.3}$$

where $z(x,y,j) := (x_1,\dots,x_j,y_{j+1},\dots,y_n)$. We shall examine the jth term in the sum on j, and assume for simplicity that $\nu = 0$, i.e., that the cube is situated at the origin. The jth term does not depend on $y_1,\dots,y_{j-1}$ or $x_{j+1},\dots,x_n$; so, integrating with respect to these variables gives a factor of h^{n-1}. If we set $s = y_j - x_j$ and rename the variables y_{j+1} to y_n to be

x_{j+1} to x_n, we can write

$$\begin{aligned}\int_{I_{h\nu}} \int_{I_{h\nu}} &|u_0(z(x,y,j)) - u_0(z(x,y,j-1))|\,dx\,dy \\ &= h^{n-1} \int_{I_{h\nu}} \int_{-x_j}^{h-x_j} |u_0(x+se_j) - u_0(x)|\,ds\,dx \\ &\le h^{n-1} \int_{-h}^{h} |s| \left\{ \frac{1}{|s|} \int_{I_{h\nu}} |u_0(x+se_j) - u_0(x)|\,dx \right\} ds.\end{aligned}$$

We substitute this into (3.3) to give

$$\begin{aligned}&\|u_0 - u^h(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)} \\ &\quad \le \frac{1}{h^n} \sum_{j=1}^{n} h^{n-1} \int_{-h}^{h} |s| \left\{ \frac{1}{|s|} \sum_{\nu\in\mathbb{Z}^n} \int_{I_{h\nu}} |u_0(x+se_j) - u_0(x)|\,dx \right\} ds \\ &\quad \le \frac{1}{h} \int_{-h}^{h} |s| \sum_{j=1}^{n} \sup_{s\in\mathbb{R}} \left\{ \frac{1}{|s|} \int_{\mathbb{R}^n} |u_0(x+se_j) - u_0(x)|\,dx \right\} ds \\ &\quad = h\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}. \qquad \square\end{aligned}$$

§4. Convergence of the Engquist-Osher Method

In this section we shall prove that the solutions of the Engquist-Osher numerical method are Cauchy in $C([0,T],L^1(\mathbb{R}^n))$, and hence converge to a function $u \in C([0,T],L^1(\mathbb{R}^n))$ as h and Δt tend to zero. This function u satisfies $\Lambda_\epsilon^{\epsilon_0}(u,v,T) \le 0$ for all bounded measurable v and all $T>0$; it requires another simple argument (which we omit) to show that in fact u is an entropy weak solution as defined in Chapter 2. By the previous chapter, this entropy weak solution is unique. Finally, we give error bounds for the Engquist-Osher scheme and give examples to show that these bounds are sharp.

To simplify things somewhat, we shall consider a sequence of approximations $u^h(x,t)$ with $h = 2^{-M}$, $M = 1,2,\ldots$, with Δt chosen as a constant multiple of h such that the assumptions of Theorem 2.1 hold. The initial data will be chosen by the formula (3.2).

Consider now two solutions u_1, with parameters h_1 and Δt_1, and u_2 with parameters $h_2 < h_1$ and $\Delta t_2 < \Delta t_1$. First, we note that it is sufficient to show that $\|u_1(\,\cdot\,,t) - u_2(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)}$ is small when t=$k\Delta t_1$, for if $k\Delta t_1 \le t < (k+\frac{1}{2})\Delta t_1$, then

$$\begin{aligned}\|u_1(\,\cdot\,,t) - u_2(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)} \le{}& \|u_1(\,\cdot\,,k\Delta t_1) - u_2(\,\cdot\,,k\Delta t_1)\|_{L^1(\mathbb{R}^n)} \\ &+ \|u_1(\,\cdot\,,t) - u_1(\,\cdot\,,k\Delta t_1)\|_{L^1(\mathbb{R}^n)} \\ &+ \|u_2(\,\cdot\,,t) - u_2(\,\cdot\,,k\Delta t_1)\|_{L^1(\mathbb{R}^n)}\end{aligned}$$

and

$$\begin{aligned}
&\|u_1(\,\cdot\,,t) - u_1(\,\cdot\,,k\Delta t_1)\|_{L^1(\mathbb{R}^n)} \\
&\le (t - k\Delta t_1) \sup \|\partial_t u_1(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)} && \text{sup over } k\Delta t_1 \le t < (k+1)\Delta t_1 \\
&\le \frac{\Delta t_1}{2} \frac{h_1^n}{\Delta t_1} \|U^{k+1} - U^k\|_{L^1(\mathbb{Z}^n)} && \text{by (3) of Lemma 3.1} \\
&\le \frac{h_1^n}{2} \frac{L\Delta t_1}{h_1} \|U^k\|_{\mathrm{BV}(\mathbb{Z}^n)} && \text{by (4) of Theorem 2.1} \\
&\le \frac{h_1^{n-1}}{2} L\Delta t_1 \|U^0\|_{\mathrm{BV}(\mathbb{Z}^n)} && \text{by (5) of Theorem 2.1 and induction} \\
&\le L\frac{\Delta t_1}{2} \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} && \text{by (3) of Lemma 3.2.}
\end{aligned}$$

If $(k+\frac{1}{2})\Delta t_1 \le t < (k+1)\Delta t_1$, then the same bound holds by comparison with $u_1(x,(k+1)\Delta t_1)$. Because $k\Delta t_1$ is a multiple of Δt_2, a similar argument shows that $\|u_2(\,\cdot\,,t) - u_2(\,\cdot\,,k\Delta t_1)\|_{L^1(\mathbb{R}^n)}$ is bounded in the same way. Thus,

$$\begin{aligned}
(4.1)\qquad &\sup_{0\le t\le T} \|u_1(\,\cdot\,,t) - u_2(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)} \le \\
&\qquad \sup_{0\le k\Delta t_1\le T} \|u_1(\,\cdot\,,k\Delta t_1) - u_2(\,\cdot\,,k\Delta t_1)\|_{L^1(\mathbb{R}^n)} + L\Delta t_1 \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)},
\end{aligned}$$

and we need show only that the right hand side tends to zero as h_1 (and hence h_2) tends to zero.

For the rest of the section, we redefine u_1 by

$$u_1(x,t) := \sum_{\nu\in\mathbb{Z}^n} U_\nu^k \chi_\nu^{h_1}(x) \qquad \text{for } k\Delta t_1 \le t < (k+1)\Delta t_1.$$

We similarly redefine u_2 to be constant in time for $k\Delta t_2 \le t < (k+1)\Delta t_2$. This new way of "filling the gaps" between time steps does not change the values of $u_1(x,k\Delta t_1)$ or $u_2(x,k\Delta t_2)$, so bounding the right side of (4.1) for these new functions will give us the result we desire. We propose to bound the relevant quantities in Theorem 1.1 of Chapter 2 with $u = u_1$, $v = u_2$, and $T = K\Delta t_1$.

We bound the difference of the initial values by

$$\begin{aligned}
&\|u_1(\,\cdot\,,0) - u_2(\,\cdot\,,0)\|_{L^1(\mathbb{R}^n)} \\
&\qquad \le \|u_1(\,\cdot\,,0) - u_0\|_{L^1(\mathbb{R}^n)} + \|u_2(\,\cdot\,,0) - u_0\|_{L^1(\mathbb{R}^n)} \\
&\qquad \le (h_1 + h_2)\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)},
\end{aligned}$$

by (3) of Lemma 3.2. We always have the bound

$$\omega_1(w,\epsilon) \le \epsilon\,\|w\|_{\mathrm{BV}(\mathbb{R}^n)},$$

so

$$(4.2)\qquad \omega_1(u_1(\,\cdot\,,0),\epsilon) \le \epsilon\,\|u_1(\,\cdot\,,0)\|_{\mathrm{BV}(\mathbb{R}^n)} \le \epsilon\,\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}$$

by (2) from Lemmas 3.1 and 3.2. One combines (5) from Theorem 2.1, (2) from Lemma 3.1, and an argument by induction to show that

$$\|u_1(\,\cdot\,,T)\|_{\mathrm{BV}(\mathbb{R}^n)} \le \|u_1(\,\cdot\,,0)\|_{\mathrm{BV}(\mathbb{R}^n)}.$$

Therefore, we can bound $\omega(u_1(\,\cdot\,,T),\epsilon)$ by $\epsilon\,\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}$. The same arguments show that both $\omega(u_2(\,\cdot\,,0),\epsilon)$ and $\omega(u_2(\,\cdot\,,T),\epsilon)$ are bounded by $\epsilon\,\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}$.

We next set out to bound

$$\sup_{0<t<\epsilon_0} \|u_1(\,\cdot\,,0) - u_1(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)}.$$

Note that u_1 is piecewise constant in time, with jumps only at the points $k\Delta t_1$, and that it is right continuous. There are $\lfloor \epsilon_0/\Delta t_1 \rfloor$ such jumps for $0 \le t < \epsilon_0$, where $\lfloor s \rfloor$ denotes the greatest integer less than or equal to s. Therefore, Inequalities (4) and (5) of Theorem 2.1, together with our typical generous application of Lemmas 3.1 and 3.2, show that

$$\begin{aligned}\sup_{0<t<\epsilon_0} \|u_1(\,\cdot\,,0) - u_1(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)} &\le \left\lfloor \frac{\epsilon_0}{\Delta t_1} \right\rfloor L\Delta t_1 \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} \\ &\le \epsilon_0 L \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}.\end{aligned}$$

It is a little more delicate to bound

$$\sup_{T-\epsilon_0<t<T} \|u_1(\,\cdot\,,T) - u_1(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)}.$$

The number of jumps in u_1 between $T-\epsilon_0$ and T is $\lceil \epsilon_0/\Delta t_1 \rceil$, where $\lceil s \rceil$ is the smallest integer greater than or equal to s. This is because there is a jump right at $t = T$; u_1 is continuous from above, not from below, in time. Therefore

$$\begin{aligned}\sup_{T-\epsilon_0<t<T} \|u_1(\,\cdot\,,T) - u_1(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)} &\le \left\lceil \frac{\epsilon_0}{\Delta t_1} \right\rceil L\Delta t_1 \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} \\ &\le (\epsilon_0 + \Delta t_1) L \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}.\end{aligned}$$

The above arguments and Theorem 1.1 of Chapter 2 imply that

$$\begin{aligned}&\|u_1(\,\cdot\,,T) - u_2(\,\cdot\,,T)\|_{L^1(\mathbb{R}^n)} \\ &\qquad \le [h_1 + h_2 + 2\epsilon + 2\epsilon_0 L + \frac{1}{2}(\Delta t_1 + \Delta t_2)L]\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} \\ &\qquad\quad + \Lambda_\epsilon^{\epsilon_0}(u_1,u_2,T) + \Lambda_\epsilon^{\epsilon_0}(u_2,u_1,T).\end{aligned} \tag{4.3}$$

The only thing left to do is show how $\Lambda_\epsilon^{\epsilon_0}(u_1,u_2,T)$ is bounded in terms of h_1, Δt_1, ϵ, and ϵ_0.

In order to bound $\Lambda_\epsilon^{\epsilon_0}(u_1,u_2,T)$, we shall show that u_1 satisfies a numerical entropy condition, the form of which was first derived by Crandall and Majda. We used their argument in "deriving" the entropy condition from our observations about smooth solutions of (0.1) in Chapter 1; we repeat it here.

Let us denote by U_ν^k the discrete solution associated with u_1 at time $k\Delta t_1$. Let $c \in \mathbb{R}$, and define $V_\nu^k := U_\nu^k \vee c$ for all $\nu \in \mathbb{Z}^n$. We calculate V^{k+1} as the solution of the Engquist-Osher scheme with initial data V^k:

$$\Delta_t^+(V^k)_\nu + \sum_{j=1}^n \left[\Delta_j^-(f_j^+(V^k))_\nu + \Delta_j^+(f_j^-(V^k))_\nu\right] = 0, \quad \nu \in \mathbb{Z}^n. \tag{4.4}$$

We note that $V_\nu^k = U_\nu^k \vee c$ is greater than both U_ν^k and c, so by (1) from Theorem 2.1, we have that $V_\nu^{k+1} \geq U_\nu^{k+1} \vee c$ for all $\nu \in \mathbb{Z}^n$. We substitute these expressions into (4.4) to see that for all $\nu \in \mathbb{Z}^n$,

$$\Delta_t^+(U^k \vee c)_\nu + \sum_{j=1}^n \left[\Delta_j^-(f_j^+(U^k \vee c))_\nu + \Delta_j^+(f_j^-(U^k \vee c))_\nu\right] \leq 0. \tag{4.5}$$

Similarly, by setting $V_\nu^k = U_\nu^k \wedge c$, we find that for all $\nu \in \mathbb{Z}^n$,

$$\Delta_t^+(U^k \wedge c)_\nu + \sum_{j=1}^n \left[\Delta_j^-(f_j^+(U^k \wedge c))_\nu + \Delta_j^+(f_j^-(U^k \wedge c))_\nu\right] \geq 0. \tag{4.6}$$

Note that $U_\nu^k \vee c - U_\nu^k \wedge c = |U_\nu^k - c|$. To make our notation a little more concise, we define

$$\bar{F}_j(U_\nu^k, c) := f_j(U_\nu^k \vee c) - f_j(U_\nu^k \wedge c) = \operatorname{sgn}(U_\nu^k - c)(f_j(U_\nu^k) - f_j(c))$$

and

$$\begin{aligned}\tilde{F}_j(U^k, c)_\nu &:= f_j^-(U_{\nu+e_j}^k \vee c) + f_j^+(U_\nu^k \vee c) - f_j^-(U_{\nu+e_j}^k \wedge c) - f_j^+(U_\nu^k \wedge c) \\ &= \operatorname{sgn}(U_{\nu+e_j}^k - c)(f_j^-(U_{\nu+e_j}^k) - f_j^-(c)) \\ &\qquad + \operatorname{sgn}(U_\nu^k - c)(f_j^+(U_\nu^k) - f_j^+(c))\end{aligned}$$

The difference of (4.5) and (4.6) can now be expressed as

$$\Delta_t^+|U^k - c|_\nu + \sum_{j=1}^n \Delta_j^- \tilde{F}_j(U^k, c)_\nu \leq 0 \tag{4.7}$$

for all $\nu \in \mathbb{Z}^n$ and $k \geq 0$. This is our numerical entropy condition.

If we let $u_1 \equiv u_1(x,t)$, $u_2 \equiv u_2(x',t')$, $h \equiv h_1$, and $\Delta t \equiv \Delta t_1$, then

$$\begin{aligned}\Lambda_\epsilon^{\epsilon_0}(u_1, u_2, t) = &- \int_0^T \int_{\mathbb{R}^n} \int_0^T \int_{\mathbb{R}^n} |u_1 - u_2|\, \omega_t(x - x', t - t') \\ &\qquad + \sum_{j=1}^n \bar{F}_j(u_1, u_2)\, \omega_{x_j}(x - x', t - t')\, dx\, dt\, dx' dt' \\ &+ \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u_1(x,T) - u_2(x',t')|\, \omega(x - x', T - t')\, dx\, dx' dt' \\ &- \int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u_1(x,0) - u_2(x',t')|\, \omega(x - x', 0 - t')\, dx\, dx' dt'.\end{aligned}$$

Because $u_1 - u_2$ is piecewise constant in x and t, the time derivative can be integrated to yield

$$
\begin{aligned}
&- \int_0^T \int_{\mathbb{R}^n} \int_0^T \int_{\mathbb{R}^n} |u_1 - u_2| \, \omega_t(x - x', t - t') \, dx \, dt \, dx' dt' \\
&= - \sum_{k=0}^{K-1} \int_0^T \int_{\mathbb{R}^n} \Delta t \sum_{\nu \in \mathbb{Z}^n} \int_{I_\nu} |U_\nu^k - u_2| \, \Delta_t^+ \omega(x - x', t^k - t') \, dx \, dx' dt' \\
&= \sum_{k=0}^{K-1} \int_0^T \int_{\mathbb{R}^n} \Delta t \sum_{\nu \in \mathbb{Z}^n} \int_{I_\nu} \Delta_t^+ |U_\nu^k - u_2| \, \omega(x - x', t^{k+1} - t') \, dx \, dx' dt' \\
&\quad - \int_0^T \int_{\mathbb{R}^n} \sum_{\nu \in \mathbb{Z}^n} \int_{I_\nu} |U_\nu^K - u_2| \, \omega(x - x', T - t') \, dx \, dx' dt' \\
&\quad + \int_0^T \int_{\mathbb{R}^n} \sum_{\nu \in \mathbb{Z}^n} \int_{I_\nu} |U_\nu^0 - u_2| \, \omega(x - x', 0 - t') \, dx \, dx' dt'
\end{aligned}
$$

where $t^k := k\Delta t$ and I_ν is defined to be the hypercube $[\nu, \nu + 1]h$. Now,

$$
\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u_1(x, T) - u_2(x', t')| \, \omega(x - x', T - t') \, dx \, dx' dt' \\
&\qquad = \int_0^T \int_{\mathbb{R}^n} \sum_{\nu \in \mathbb{Z}^n} \int_{I_\nu} |U_\nu^K - u_2| \, \omega(x - x', T - t') \, dx \, dx' dt'
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |u_1(x, 0) - u_2(x', t')| \, \omega(x - x', 0 - t') \, dx \, dx' dt' \\
&\qquad = \int_0^T \int_{\mathbb{R}^n} \sum_{\nu \in \mathbb{Z}^n} \int_{I_\nu} |U_\nu^0 - u_2| \, \omega(x - x', 0 - t') \, dx \, dx' dt',
\end{aligned}
$$

so that $\Lambda_\epsilon^{\epsilon_0}(u_1, u_2, T)$ equals

$$
\begin{aligned}
&\int_0^T \int_{\mathbb{R}^n} \sum_{k=0}^{K-1} \Delta t \sum_{\nu \in \mathbb{Z}^n} \int_{I_\nu} \Delta_t^+ |U_\nu^k - u_2| \, \omega(x - x', t^{k+1} - t') \, dx \, dx' dt' \\
&\qquad - \int_0^T \int_{\mathbb{R}^n} \int_0^T \int_{\mathbb{R}^n} \sum_{j=1}^n \bar{F}_j(u_1, u_2) \, \omega_{x_j}(x - x', t - t') \, dx \, dt \, dx' dt'.
\end{aligned}
$$

We now hope to transform the spatial derivatives into something manageable. Let's fix j and analyze one term alone:

$$
- \int_0^T \int_{\mathbb{R}^n} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{\nu \in \mathbb{Z}^n} \int_{I_\nu} \bar{F}_j(U_\nu^k, u_2) \, \omega_{x_j}(x - x', t - t') \, dx \, dt \, dx' dt'.
$$

We'll first substitute $\tilde{F}_j(U^k, u_2)_\nu$ for $\bar{F}_j(U^k_\nu, u_2)$; on each interval I_ν we incur an error of

$$(4.8)\quad f_j(U^k_\nu \vee u_2) - f_j(U^k_\nu \wedge u_2) \\ - [f_j^-(U^k_{\nu+e_j} \vee u_2) + f_j^+(U^k_\nu \vee u_2) - f_j^-(U^k_{\nu+e_j} \wedge u_2) - f_j^+(U^k_\nu \wedge u_2)].$$

We've defined f_j^+ and f_j^- so that $f_j = f_j^+ + f_j^-$, so we can rewrite (4.8) as

$$f_j^-(U^k_\nu \vee u_2) - f_j^-(U^k_{\nu+e_j} \vee u_2) - f_j^-(U^k_\nu \wedge u_2) + f_j^-(U^k_{\nu+e_j} \wedge u_2).$$

The above expression is bounded by

$$|f_j^-(U^k_{\nu+e_j}) - f_j^-(U^k_\nu)|,$$

as can be seen by examining the cases $u_2 \le \min(U^k_{\nu+e_j}, U^k_\nu)$, $u_2 \ge \max(U^k_{\nu+e_j}, U^k_\nu)$, and $\min(U^k_{\nu+e_j}, U^k_\nu) \le u_2 \le \max(U^k_{\nu+e_j}, U^k_\nu)$. In total, we have incurred an error by this substitution of at most

$$\int_0^T \int_{\mathbb{R}^n} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{\nu\in\mathbb{Z}^n} \int_{I_\nu} |f_j^-(U^k_{\nu+e_j}) - f_j^-(U^k_\nu)|\, |\omega_{x_j}(x-x', t-t')|\, dx\, dt\, dx' dt'.$$

The only factor in the integrand that now depends on x' and t' is $|\omega_{x_j}(x - x', t - t')| = |\omega_{x'_j}(x - x', t - t')|$. When we integrate with respect to t' we get a factor of

$$\int_0^T \eta_{\epsilon_0}(t - t')\, dt' \le 1;$$

the integral equals 1 except possibly when $t < \epsilon_0$ or $t > T - \epsilon_0$. Integrating with respect to each of the variables x'_i, $i \ne j$, yields additional factors of 1; however, integrating with respect to x'_j gives a factor of

$$\frac{1}{\epsilon}\|\eta'\|_{L^1(\mathbb{R})},$$

because of the scaling $\eta_\epsilon(x'_j) := \eta(x'_j/\epsilon)/\epsilon$. The terms involving x and t are constant for $x \in I_\nu$ and $t^k \le t < t^{k+1}$, so integrating with respect to x and t on each space time interval $I_\nu \times [t^k, t^{k+1}]$ gives a factor of $h^n \Delta t$. So, we can bound the sum of the errors incurred in each coordinate direction by

$$\frac{h^n \Delta t}{\epsilon}\|\eta'\|_{L^1(\mathbb{R})} \sum_{k=0}^{K-1} \sum_{j=1}^{n} \sum_{\nu\in\mathbb{Z}^n} |f_j^-(U^k_{\nu+e_j}) - f_j^-(U^k_\nu)|$$

Because $|f_j^-(\xi) - f_j^-(\zeta)| \le M|\xi - \zeta|$, where $M \le L$ and because of (5) from Theorem 2.1 and (2) of Lemma 3.2 and Lemma 3.1, the the error incurred by this substitution is bounded by

$$\frac{h^n \Delta t}{\epsilon}\|\eta'\|_{L^1(\mathbb{R})} KM \|U^0\|_{\mathrm{BV}(\mathbb{Z}^n)} \le \frac{hTM}{\epsilon}\|\eta'\|_{L^1(\mathbb{R})}\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}.$$

Note that there is no error if all the functions f_j are nondecreasing.

ATTENTION: BIG NOTATIONAL SIMPLICATION AHEAD!

The notation in n dimensions for what I want to do next is too arcane to be useful, so for this next estimate (which is the main estimate), I shall restrict attention to one space dimension. In n dimensions one performs the same calculation in each coordinate direction separately.

In one dimension we shall work with

$$-\int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \int_{I_i} \tilde{F}(U_i^k, u_2)\, \omega_x(x-x', t-t')\, dx\, dt\, dx' dt'.$$

On each interval I_i, $\tilde{F}(U_i^k, u_2)$ is constant in x, so we can integrate ω_x to get

$$\begin{aligned}
&-\int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \int_{I_i} \tilde{F}(U_i^k, u_2)\, \omega_x(x-x', t-t')\, dx\, dt\, dx' dt' \\
&= -\int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} h \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \tilde{F}(U_i^k, u_2)\, \Delta_x^+ \omega(x_i - x', t-t')\, dt\, dx' dt' \\
&= \int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} h \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \Delta_x^+ \tilde{F}(U_i^k, u_2)\, \omega(x_{i+1} - x', t-t')\, dt\, dx' dt' \\
&= \int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} h \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \Delta_x^- \tilde{F}(U_i^k, u_2)\, \omega(x_i - x', t-t')\, dt\, dx' dt'
\end{aligned}$$

(Summation by parts does not have a boundary term because $\omega(x-x', t-t')$ has bounded support in x for each fixed x', t, and t'.) So, with an error of at most $\frac{hTM}{\epsilon}\|\eta'\|_{L^1(\mathbb{R})}\|u_0\|_{\mathrm{BV}(\mathbb{R})}$, $\Lambda_\epsilon^{\epsilon_0}(u_1, u_2, T)$ is equal to

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} \Delta t \sum_{i\in\mathbb{Z}} \int_{I_i} \Delta_t^+ |U_i^k - u_2|\, \omega(x-x', t^{k+1} - t')\, dx\, dx' dt' \\
&\qquad + \int_0^T \int_{\mathbb{R}} h \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \Delta_x^- \tilde{F}(U_i^k, u_2)\, \omega(x_i - x', t-t')\, dt\, dx' dt' \\
&= \int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \int_{I_i} \Delta_t^+ |U_i^k - u_2|\, \omega(x-x', t^{k+1} - t')\, dx\, dt\, dx' dt' \\
&\qquad + \int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \int_{I_i} \Delta_x^- \tilde{F}(U_i^k, u_2)\, \omega(x_i - x', t-t')\, dx\, dt\, dx' dt'.
\end{aligned}$$

If ω were evaluated at the same arguments in each term above, then $\Lambda_\epsilon^{\epsilon_0}(u_1, u_2, T)$ would be no greater than zero and all would be well; for (4.7) declares that the terms involving U_i^k are nonpositive for all values of i, k,

and $u_2(x',t')$. That is,

$$\int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \int_{I_i} [\Delta_t^+ |U_i^k - u_2| + \Delta_x^- \tilde{F}(U_i^k, u_2)]\, \omega(x-x', t-t')\, dx\, dt\, dx' dt' \le 0.$$

Now $|\Delta_t^+ |U_i^k - u_2|\,| \le |\Delta_t^+ U_i^k|$ so the error incurred by substituting $\omega(x - x', t - t')$ for $\omega(x - x', t^{k+1} - t')$ is bounded by

$$\int_0^T \int_{\mathbb{R}} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \int_{I_i} |\Delta_t^+ U_i^k|\, |\omega(x-x', t-t') - \omega(x-x', t^{k+1}-t')|\, dx\, dt\, dx' dt'.$$

Only the terms involving ω, which can be written as

$$\omega(x-x', t-t') - \omega(x-x', t^{k+1}-t') = \eta_\epsilon(x-x')[\eta_{\epsilon_0}(t-t') - \eta_{\epsilon_0}(t^{k+1}-t')],$$

depend in any way on x' and t'. When we integrate with respect to x' we get a factor of 1; integrating with respect to t' yields

$$\begin{aligned}\int_{\mathbb{R}} |\eta_{\epsilon_0}(t-t') - \eta_{\epsilon_0}(t^{k+1}-t')|\, dt' &\le \int_{\mathbb{R}} \int_{t-t'}^{t^{k+1}-t'} |\eta'_{\epsilon_0}(\xi)|\, d\xi\, dt' \\ &= \frac{(t^{k+1}-t)}{\epsilon_0} \|\eta'\|_{L^1(\mathbb{R})}.\end{aligned}$$

Therefore, the error in the time term is bounded by

$$\begin{aligned}&\frac{1}{\epsilon_0}\|\eta'\|_{L^1(\mathbb{R})} \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{i\in\mathbb{Z}} \int_{I_i} |\Delta_t^+ U_i^k|\, (t^{k+1}-t)\, dx\, dt \\ &\quad = \frac{\Delta t^2 h}{2\epsilon_0}\|\eta'\|_{L^1(\mathbb{R})} \sum_{k=0}^{K-1} \sum_{i\in\mathbb{Z}} |\Delta_t^+ U_i^k| \\ &\quad \le \frac{LK\Delta t^2}{2\epsilon_0}\|\eta'\|_{L^1(\mathbb{R})}\|U^0\|_{\mathrm{BV}(\mathbb{Z})} \\ &\quad \le \frac{LT\Delta t}{2\epsilon_0}\|\eta'\|_{L^1(\mathbb{R})}\|u_0\|_{\mathrm{BV}(\mathbb{R})}.\end{aligned}$$

For the spatial term, we have that $|\Delta_x^- \tilde{F}(U_i^k, u_2)| \le |\Delta_x^- f^t(U_i^k)|$. Arguing in the same way as for the time term, we can see that substituting $\omega(x - x', t - t')$ for $\omega(x_i - x', t - t')$ introduces an error of at most

$$\begin{aligned}\frac{\Delta t h^2}{2\epsilon}\|\eta'\|_{L^1(\mathbb{R})} \sum_{k=0}^{K-1} \sum_{i\in\mathbb{Z}} |\Delta_x^- f^t(U_i^k)| &\le \frac{LK\Delta t h}{2\epsilon}\|\eta'\|_{L^1(\mathbb{R})}\|U^0\|_{\mathrm{BV}(\mathbb{Z})} \\ &\le \frac{LTh}{2\epsilon}\|\eta'\|_{L^1(\mathbb{R})}\|u_0\|_{\mathrm{BV}(\mathbb{R})}.\end{aligned}$$

In several space dimensions exactly the same arguments are used, only now they are repeated in each of the n spatial dimensions. The same bound

results:

$$\Lambda_\epsilon^{\epsilon_0}(u_1,u_2,T) \le \left[\frac{LT\Delta t_1}{2\epsilon_0} + \frac{LTh_1}{2\epsilon} + \frac{MTh_1}{\epsilon}\right] \|\eta'\|_{L^1(\mathbb{R})}\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}.$$

This bound does not depend in any way on the fact that u_2 is another approximation generated by the Engquist-Osher scheme. In other words, for all measurable, locally integrable $v(x',t')$ we have

$$\Lambda_\epsilon^{\epsilon_0}(u_1,v,T) \le \left[\frac{LT\Delta t_1}{2\epsilon_0} + \frac{LTh_1}{2\epsilon} + \frac{MTh_1}{\epsilon}\right] \|\eta'\|_{L^1(\mathbb{R})}\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}. \tag{4.9}$$

We can determine η as we choose, as long as it satisfies our hypotheses. By letting $\eta \to \frac{1}{2}\chi_{[-1,1]}$, we have $\|\eta'\|_{L^1(\mathbb{R})} \to 1$. Therefore, we can use (4.1), (4.3), and (4.9) for u_1 and u_2 to see that for any $0 \le t \le T$

$$\begin{aligned}
&\|u_1(\,\cdot\,,t) - u_2(\,\cdot\,,t)\|_{L^1(\mathbb{R}^n)} \\
&\qquad \le [h_1 + h_2 + 2\epsilon + 2\epsilon_0 L + (\frac{3}{2}\Delta t_1 + \frac{1}{2}\Delta t_2)L]\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} \\
&\qquad\quad + \left[\frac{LT\Delta t_1}{2\epsilon_0} + \frac{LTh_1}{2\epsilon} + \frac{MTh_1}{\epsilon}\right] \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} \\
&\qquad\quad + \left[\frac{LT\Delta t_2}{2\epsilon_0} + \frac{LTh_2}{2\epsilon} + \frac{MTh_2}{\epsilon}\right] \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}
\end{aligned} \tag{4.10}$$

At this point our purpose is to show that the sequence of numerical approximations is Cauchy in $C([0,T],L^1(\mathbb{R}^n))$. We just note that $h_2 \le h_1$, $\Delta t_2 \le \Delta t_1$, and if we set $\epsilon = \epsilon_0 = (h_1 T)^{1/2}$ we have that

$$\|u_1(\,\cdot\,,T) - u_2(\,\cdot\,,T)\|_{L^1(\mathbb{R}^n)} \le C(h_1 + \Delta t_1 + (h_1 T)^{1/2})\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}.$$

Since $h_1 \in \{2^{-M}\}$, this shows that, indeed, the approximations generated by the Engquist-Osher scheme are Cauchy in $C([0,T],L^1(\mathbb{R}^n))$ for any fixed $T > 0$. Therefore, they converge in $C([0,T],L^1(\mathbb{R}^n))$ to a function $u(x,t)$. By letting h_1 and Δt_1 tend to zero in (4.9), we can conclude by the Lebesgue Dominated Convergence theorem that for fixed ϵ and ϵ_0 and for all bounded v we have

$$\Lambda_\epsilon^{\epsilon_0}(u,v,T) \le 0. \tag{4.11}$$

By Kuznetsov's theorem, there is at most one function in $C([0,T],L^1(\mathbb{R}^n))$ that satisfies (4.11), and it is the limit of our numerical approximations.

We shall now show that u satisfies the entropy condition. From our numerical entropy condition (4.7), we know that for all $c \in \mathbb{R}$, all $T = K\Delta t > 0$, and all $\phi \in C_0^1(\mathbb{R}^{n+1})$ with $\phi \ge 0$, we have

$$0 \ge \sum_{k=0}^{K-1} \int_{t^k}^{t^{k+1}} \sum_{\nu\in\mathbb{Z}^n} \int_{I_\nu} [\Delta_t^+|U_\nu^k - c| + \sum_{j=1}^n \Delta_j^- \tilde{F}_j(U_\nu^k, c)]\,\phi(x,t)\,dx\,dt$$

$$= -\sum_{k=0}^{K-1}\int_{t^k}^{t^{k+1}}\sum_{\nu\in\mathbb{Z}^n}\int_{I_\nu}|U_\nu^{k+1}-c|\Delta_t^+\phi(x,t)+\sum_{j=1}^{n}\tilde{F}_j(U_{\nu-e_j}^k,c)\Delta_j^-\phi(x,t)\,dx\,dt$$
$$+\sum_{\nu\in\mathbb{Z}^n}\int_{I_\nu}|U_\nu^K-c|\phi(x,T)\,dx-\sum_{\nu\in\mathbb{Z}^n}\int_{I_\nu}|U_\nu^0-c|\phi(x,0)\,dx,$$

by summation by parts. As h and Δt tend to zero, $u^h\to u$, $|u^h-c|\to|u-c|$, and $\tilde{F}_j(u^h,c)\to\bar{F}_j(u,c)$ in $C([0,t],L^1(\mathbb{R}))$, while $\Delta_t^+\phi(x,t)\to\phi_t(x,t)$ and $\Delta_j^-\phi(x,t)\to\phi_{x_j}(x,t)$ boundedly in $\mathbb{R}^{n+1}$. As u is uniformly bounded, by the Lebesgue Dominated Convergence theorem we conclude that

$$-\int_0^T\int_{\mathbb{R}^n}|u(x,t)-c|\phi_t(x,t)+\sum_{j=1}^{n}\bar{F}_j(u(x,t),c)\phi_{x_j}(x,t)\,dx\,dt$$
$$+\int_{\mathbb{R}^n}|u(x,T)-c|\phi(x,T)\,dx-\int_{\mathbb{R}^n}|u(x,0)-c|\phi(x,0)\,dx\le 0,$$

or $\Lambda(u,c,\phi,T)\le 0$. Thus, we can conclude that u is the unique entropy solution of (0.1).

§5. Properties of the Entropy Solution

In this section we derive the properties of entropy solutions of (0.1) that we expected to find on the basis of our heuristic calculations with smooth solutions and the viscosity approximation in Chapter 1.

Theorem 5.1. *Assume $F\colon\mathbb{R}\to\mathbb{R}^n$ is Lipschitz continuous, that there exists a constant L such that $\sup_\xi\sum_{j=1}^n|F_j'(\xi)|\le L$, and that that u_0 and v_0 are bounded and in $\mathrm{BV}(\mathbb{R}^n)$. Then there exist unique functions $u(x,t)$ and $v(x,t)$ in $C([0,T],L^1(\mathbb{R}^n))$ that are weak entropy solutions of (0.1) with initial data u_0 and v_0, respectively. Furthermore, we have:*

(1) *If $u_0(x)\ge v_0(x)$ a.e. then $u(x,t)\ge v(x,t)$ a.e.*

(2) *For all $t>0$, $\|u(x,t)-v(x,t)\|_{L^1(\mathbb{R}^n)}\le\|u_0(x)-v_0(x)\|_{L^1(\mathbb{R}^n)}$.*

(3) *For all $t>0$, $\operatorname{ess\,sup}_{x\in\mathbb{R}^n}u(x,t)\le\operatorname{ess\,sup}_{x\in\mathbb{R}^n}u_0(x)$ and $\operatorname{ess\,inf}_{x\in\mathbb{R}^n}u(x,t)\ge\operatorname{ess\,inf}_{x\in\mathbb{R}^n}u_0(x)$.*

(4) *For all $t,t'>0$, $\|u(\,\cdot\,,t)-u(\,\cdot\,,t')\|_{L^1(\mathbb{R}^n)}\le L|t-t'|\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}$.*

(5) *For all $t>0$, $\|u(\,\cdot\,,t)\|_{\mathrm{BV}(\mathbb{R}^n)}\le\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}$.*

Finally, for any $t>0$, $u(x,t)$ for $|x|\le R$ depends only on $u_0(x)$ for $|x|\le R+Lt$.

Proof. We have shown that the limits in $C([0,T],L^1(\mathbb{R}^n))$ of u^h and v^h are entropy weak solutions of (0.1). That u and v are unique follows from Theorem 1.1 of Chapter 2. Our numerical approximations u^h satisfy (1), (2), (3), and (4), so u and v, the limit in $C([0,T],L^1(\mathbb{R}^n))$ and pointwise almost everywhere of u^h and v^h as $h\to 0$, also satisfy these properties. Property (5) follows immediately from Property (2) with $v_0(x):=u_0(x+he_j)$, $j=1,\dots,n$. Finally, if $R=Nh$, then $u^h(x,k\Delta t)$ for $|x|\le R$ depends

on $u^h(x,(k-1)\Delta t)$ for $|x| \le R + h \le R + L\Delta t$. The final remark follows by induction. □

§6. Error Bounds for Monotone Finite Difference Methods

In this section we shall derive rather sharp error bounds for the discrete ($\Delta t > 0$) and the semidiscrete (the limit as $\Delta t \to 0$) Engquist-Osher schemes. We also remark about error bounds for general conservative, consistent, monotone finite difference methods for (0.1). We give an example where we estimate (rather than bound) the error in the Engquist-Osher scheme, and show that our bounds are close to the estimate.

To bound the norm of the error $u - u^h$ we can take the limit of (4.10) as h_2 and Δt_2 tend to zero. This shows that

(6.1)
$$\|u^h(\,\cdot\,,T) - u(\,\cdot\,,T)\|_{L^1(\mathbb{R}^n)} \le [h + 2\epsilon + 2\epsilon_0 L + \frac{3}{2}\Delta t L]\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} + \left[\frac{L\Delta t T}{2\epsilon_0} + \frac{LhT}{2\epsilon} + \frac{MhT}{\epsilon}\right]\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}.$$

We wish to choose ϵ and ϵ_0 to minimize this bound on the error. Any expression of the form $a\epsilon + b/\epsilon$ is minimized when $\epsilon = \sqrt{b/a}$; the minimum value is $2\sqrt{ab}$. Therefore, the minimum value of the right side of (6.1) is

$$\left[h + \frac{3}{2}\Delta t + 2\sqrt{L^2\Delta t T} + 2\sqrt{(L+2M)hT}\right]\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}. \tag{6.2}$$

This is our final error bound for the discrete Engquist-Osher scheme.

Note that as Δt increases our error bound (6.2) also increases. (Remember, though, that our CFL condition requires that $L\Delta t \le h$.) We can imagine a so-called semi-discrete scheme, where the values at each point $\nu \in \mathbb{Z}^n$ are not discrete values U_ν^k, where $k = 0, 1, \ldots$, but are defined for all $t > 0$, $U_\nu(t)$. This corresponds to letting $\Delta t \to 0$ in the definition of the scheme:

$$\frac{dU_\nu(t)}{dt} + \sum_{j=1}^{n}[\Delta_j^- f_j^+(U_\nu(t)) + \Delta_j^+ f_j^-(U_\nu(t))] = 0, \quad \nu \in \mathbb{Z}^n.$$

We can set $\Delta t = 0$ in (6.2) to find an error bound for the semi-discrete method:

$$\|u^h(\,\cdot\,,T) - u(\,\cdot\,,T)\|_{L^1(\mathbb{R}^n)} \le \left[h + 2\sqrt{(L+2M)ht}\right]\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}. \tag{6.3}$$

We claimed at the beginning of this chapter that the analysis for other conservative, consistent, monotone schemes will be so similar as to not warrant repeating. Here we make good on this claim. If our scheme is of the form

$$\Delta_t^+(U^k)_i + \frac{F(U_{i-N}^k, \ldots U_{i+N+1}^k) - F(U_{i-N-1}^k, \ldots U_{i+N}^k)}{h} = 0, \quad i \in \mathbb{Z},$$

then the analysis for the Engquist-Osher scheme goes through if we can find an L such that if $L\Delta t/h \le 1$ then the scheme is monotone and (4) of Theorem 2.1 holds. The constant M is taken so that the bound on $\tilde{F}(U^k_\nu, c) - \bar{F}(U^k_\nu, c)$ holds. In general, we can take

$$L = \sum_{i=-N}^{N+1} \left| \frac{\partial F(U^k_{-N}, \ldots, U^k_{N+1})}{\partial U^k_i} \right|$$

and

$$M = \sum_{i=-N}^{N+1} |i| \left| \frac{\partial F(U^k_{-N}, \ldots, U^k_{N+1})}{\partial U^k_i} \right|.$$

Note that this gives us the correct bounds for the Engquist-Osher scheme.

It is natural to ask if our $O((hT)^{1/2} + (\Delta t T)^{1/2})$ error bound is sharp, i.e., if the error in the numerical scheme is actually of this order. For the simple linear problem

$$\begin{aligned} &u_t + u_x = 0, && x \in \mathbb{R},\ t > 0, \\ &u(x,0) = \begin{cases} 0, & x \le 0, \\ 1, & x > 0, \end{cases} && x \in \mathbb{R}, \end{aligned} \tag{6.4}$$

we can calculate an asymptotic expression for the error in the Engquist-Osher scheme. Of course, the solution of (6.4) is simply $u(x,t) = u_0(x-t)$. For this problem the numerical method is given by

$$\begin{aligned} U^0_i &= \begin{cases} 0, & i < 0, \\ 1, & i \ge 0, \end{cases} && i \in \mathbb{Z}, \text{ and} \\ U^{k+1}_i &= U^k_i - \frac{\Delta t}{h}(U^k_i - U^k_{i-1}) \\ &= \frac{\Delta t}{h} U^k_{i-1} + \big(1 - \frac{\Delta t}{h}\big) U^k_i, && i \in \mathbb{Z},\ k \ge 0. \end{aligned}$$

We claim that U^k_i can be given the following probabilistic interpretation. Consider an off-balance coin with a probability $p := \Delta t/h$ of coming up heads and $q := 1 - p = 1 - \Delta t/h$ of coming up tails, and let X^k be the random variable that is the number of heads that come up in k tosses of our coin. Then if we define $V^k_i := \text{Prob}\{X^k \le i\}$ of our funny coin, we have

$$V^0_i = \begin{cases} 0, & i < 0, \\ 1, & i \ge 0, \end{cases} \quad i \in \mathbb{Z},$$

while

$$\begin{aligned} V_i^{k+1} &= \text{Prob}\{X^{k+1} \le i\} \\ &= \text{Prob}\{X^k \le i-1\} + \text{Prob}\{X^k = i\} \times \text{Prob}\{\text{a tail}\} \\ &= V_{i-1}^k + (V_i^k - V_{i-1}^k)\big(1 - \frac{\Delta t}{h}\big) \\ &= \frac{\Delta t}{h} V_{i-1}^k + \big(1 - \frac{\Delta t}{h}\big) V_i^k, \end{aligned}$$

for $k \ge 0$ and $i \in \mathbb{Z}$. Thus U_i^k and V_i^k satisfy the same initial condition and the same recurrence relation, so $U_i^k = V_i^k$ for all $k \ge 0$ and $i \in \mathbb{Z}$. We let $T = K\Delta t$. By the binomial theorem,

$$U_i^K = \sum_{j=0}^{\min(i,K)} \binom{K}{j} p^j (1-p)^{K-j},$$

and X^k has mean

$$\mu = pK = \frac{\Delta t}{h} K = \frac{T}{h}$$

and variance

$$\sigma^2 = Kp(1-p) = \frac{T}{h}\big(1 - \frac{\Delta t}{h}\big).$$

The Central Limit Theorem says that

$$\frac{X^k - \mu}{\sigma} \longrightarrow N(0,1),$$

where $N(0,1)$ denotes the normally distributed random variable with distribution

$$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

If, as before, we set $u^h(x,T) := \sum_{i\in\mathbb{Z}} U_i^K \chi_{I_i}(x)$, then the error in the numerical method is

$$\begin{aligned} &\frac{1}{\sqrt{h}} \|u(\,\cdot\,,T) - u^h(\,\cdot\,,T)\|_{L^1(\mathbb{R})} \\ &\qquad = \frac{1}{\sqrt{h}} \int_{-\infty}^{\infty} |u(x,T) - u^h(x,T)|\, dx \\ &\qquad = \frac{1}{\sqrt{h}} \int_{-\infty}^{T} u^h(x,T)\, dx - \frac{1}{\sqrt{h}} \int_{T}^{\infty} (1 - u^h(x,T))\, dx \\ &\qquad \longrightarrow \frac{2\sigma\sqrt{h}}{\sqrt{2\pi}} \int_{-\infty}^{0} \int_{-\infty}^{x} e^{-\xi^2/2}\, d\xi\, dx \end{aligned}$$

as $h \to 0$. The last integral has value 1 (change the order of integration), so we have that asymptotically

$$\begin{aligned}\|u(\,\cdot\,,T) - u^h(\,\cdot\,,T)\|_{L^1(\mathbb{R})} &\sim \frac{2h\sigma}{\sqrt{2\pi}} \\ &= \big(\frac{2}{\pi}\big)^{1/2} h \big(\frac{T}{h}\big)^{1/2} \big(1 - \frac{\Delta t}{h}\big)^{1/2} \\ &= \big(\frac{2}{\pi}\big)^{1/2} (hT)^{1/2} \big(1 - \frac{\Delta t}{h}\big)^{1/2}\end{aligned}$$

as h tends to zero.

Thus, the error in our approximation is $\geq C(ht)^{1/2}$, so our error bound is the right order. The error is zero when $\Delta t = h$; this is because the numerical and real discontinuities move exactly one space interval in each time step.

For the Engquist-Osher method applied to (6.4), we have that $L = 1$, $M = 0$ ($f^- = 0$), and $\|u_0\|_{\mathrm{BV}(\mathbb{R})} = 1$, so if we ignore terms of order Δt and h, our bound is

$$\begin{aligned}\|u(\,\cdot\,,T) - u^h(\,\cdot\,,T)\|_{L^1(\mathbb{R})} &\leq 2\sqrt{\Delta t T} + 2\sqrt{hT} \\ &= 2\sqrt{ht}\left(1 + \big(\frac{\Delta t}{h}\big)^{1/2}\right).\end{aligned}$$

So if Δt is much less than $2h$, our bound is no more than 3 times as large as the real error. This is fairly sharp, as error bounds go!

§7. Existence and Error Bounds without Bounded Variation

Existence and uniqueness of entropy weak solutions of (0.1) for initial data u_0 that does not have bounded variation, together with error bounds for numerical methods applied to such solutions, can easily be derived. We shall do the first here with a limiting argument; the second result can be derived by assuming that $\omega_1(u_0, \epsilon) \leq C\epsilon^\alpha$ for some $\alpha \leq 1$ and deriving the other bounds where we have previously used the bounded variation of u_0.

Given any initial data $u_0 \in L^1(\mathbb{R}^n)$ and an $\epsilon > 0$ we can calculate the smoothed initial data

$$u_0^\epsilon = u_0 * \eta_\epsilon,$$

where η_ϵ is our standard molifier introduced in Chapter 2. Then one can derive

$$\|u_0^\epsilon\|_{\mathrm{BV}(\mathbb{R}^n)} \leq \frac{C}{\epsilon}\|u_0\|_{L^1(\mathbb{R}^n)},$$

and

$$\|u_0 - u_0^\epsilon\|_{L^1(\mathbb{R}^n)} \leq C\omega_1(u_0, \epsilon) \longrightarrow 0$$

as $\epsilon \to 0$. If we denote by u^{ϵ_1} and u^{ϵ_2} the entropy weak solutions of (0.1) with initial data $u_0^{\epsilon_1}$ and $u_0^{\epsilon_2}$, respectively, then for all $T > 0$ we have

$$\|u^{\epsilon_1}(\,\cdot\,,T) - u^{\epsilon_2}(\,\cdot\,,T)\|_{L^1(\mathbb{R}^n)} \leq \|u_0^{\epsilon_1} - u_0^{\epsilon_2}\|_{L^1(\mathbb{R}^n)},$$

so the set $\{u^\epsilon\}$ is Cauchy in $C([0,\infty), L^1(\mathbb{R}^n))$. Therefore, the sequence converges to a function $u \in C([0,\infty), L^1(\mathbb{R}^n))$; it is this u that we consider to be the entropy weak solution of (0.1) with initial data u_0. Because each of the functions u^ϵ satisfies the weak entropy condition, u also satisfies the weak entropy condition; u also satisfies

$$\omega_1(u,\epsilon) \le \omega_1(u_0,\epsilon)$$

for all $\epsilon > 0$. Thus, Kuznetsov's theorem shows that u is unique.

I'll have to get around to doing the error estimates for numerical methods for non-BV data later.

TO BE CONTINUED

Chapter 4

Godunov's Method and the Random Choice Method

In this chapter we discuss Godunov's method and Glimm's random choice method, two numerical methods that are of great historical importance both computationally and theoretically, mainly for their applications to hyperbolic systems of nonlinear conservation laws. Both schemes are based on the solution of the Riemann problem, which we discuss in §1 for the scalar problem in one space dimension. In §2 we present a corollary of Theorem 1.2 in Chapter 2, proved originally by Kuznetsov. We use this corollary to provide, in §3, error bounds for the random choice method applied to scalar problems, and then, in §4, error bounds for Godunov's method.

§1. The Riemann Problem

In this section we consider the solution of the so-called Riemann problem

$$
\begin{aligned}
&u_t + f(u)_x = 0, && x \in \mathbb{R},\ t > 0,\\
&u(x,0) = \begin{cases} u_L, & x \le 0,\\ u_R, & x > 0, \end{cases} && x \in \mathbb{R}.
\end{aligned}
\tag{1.1}
$$

In one-dimensional gas dynamics this problem arises when a gas is contained in a tube with an impermeable membrane at the point $x = 0$; the gas has constant density and pressure on each side of the membrane. At time $t = 0$, the membrane is removed and we study the evolution of the system. Our problem is much simpler, as we consider only the scalar equation.

Necessary and sufficient conditions were given in Chapter 1 for a piecewise smooth function u to be a solution to the scalar conservation law; we shall construct the solution of (1.1) from this class of functions. Let us assume first that $u_L > u_R$ and that $f'(u)$ is continuous and in $\mathrm{BV}(\mathbb{R})$. We consider the set S, defined to be the convex hull of the set

$$\{(u,v) \mid v \le f(u),\ u_R \le u \le u_L\},$$

and the function $\tilde{f}(u) := \sup_{(u,v)\in S} v$. Note that as u declines from u_L to u_R, $\tilde{f}'(u)$ increases monotonically, with $\tilde{f}'$ constant on intervals where $\tilde{f}$ is linear. Under our assumptions on f, $\tilde{f}'$ is onto the interval $[\tilde{f}'(u_L), \tilde{f}'(u_R)]$.

We claim that if $u(x,t)$ solves (1.1), then $u(x,t) := u(x/t,1) := (\tilde{f}')^{-1}(x/t)$. (Wherever $\tilde{f}'$ is constant, u has a shock, and we consider

u to be multivalued.) In other words, $x/t = \tilde{f}'(u)$. In an interval where u is smooth, so that $\tilde{f}''$ is strictly negative, then $\tilde{f}(u) = f(u)$ and

$$u_t + f(u)_x = \frac{1}{f''((f')^{-1}(\frac{x}{t}))}(-\frac{x}{t^2}) + f'(u)\frac{1}{f''((f')^{-1}(\frac{x}{t}))}(\frac{1}{t}) = 0.$$

Wherever $\tilde{f}(u) \geq f(u)$ is linear on a maximal interval (u_1, u_2), then there is a discontinuity in the solution $u(x,t)$ between $u = u_1$ and $u = u_2$ at the location $x/t = \tilde{f}'(u) = (f(u_1) - f(u_2))/(u_1 - u_2)$; one sees that both the Rankine-Hugoniot condition and the entropy condition are satisfied for this shock, since the line joining the points $(u_1, f(u_1))$ and $(u_2, f(u_2))$, is, by the definition of $\tilde{f}$, above the graph of $f(u)$.

When $u_L < u_R$, we define S to be the convex hull of the set

$$\{(u,v) \mid v \geq f(u),\ u_L \leq u \leq u_R\},$$

and the function $\tilde{f}(u) := \inf_{(u,v)\in S} v$. The same construction as before now works to define $u(x,t)$ by $x/t = \tilde{f}'(u)$.

We shall also want to solve the Riemann problem in the case when $f(u)$ is a continuous, piecewise linear function in u. By the above construction, $\tilde{f}$ is always piecewise linear, but it is not C^1. Let us consider when $\tilde{f}$ has discontinuities in its first derivative at the points $u_R = u_0 < u_1 < \cdots < u_N = u_L$. Then, we say that $u(x,t) = v$ for $x/t \in [\tilde{f}'(v^-), \tilde{f}'(v^+)]$. We find that u is now piecewise constant, with constant states $u_0, \ldots, u_N$, separated by discontinuities at the points $x/t = \tilde{f}'(u_k^-)$, $k = 1, \ldots, N$. Where u is constant, it obviously solves (1.1), and one can see trivially that the discontinuities satisfy the Rankine-Hugoniot condition and the entropy condition. One can obviously make the same construction when $u_L < u_R$. Therefore, this is the solution of the Riemann problem for any piecewise linear flux f.

§2. A Corollary of Kuznetsov's Theorem

There are many important numerical methods based on the solution of the Riemann problem. In this section we discuss the random choice method for the scalar hyperbolic conservation law

$$\begin{aligned} &u_t + f(u)_x = 0, && x \in \mathbb{R},\ t > 0, \\ &u(x,0) = u_0(x), && x \in \mathbb{R}. \end{aligned} \tag{2.1}$$

Glimm introduced and used this scheme to prove the existence of global weak solutions to hyperbolic systems of conservation laws. In the next section we discuss Godunov's method, one of the first successful numerical methods for hyperbolic conservation laws.

To analyze versions of the random choice method and Godunov's method, we prove the following lemma, due to Kuznetsov.

Lemma 2.1. *If u_0 is in $\mathrm{BV}(\mathbb{R})$, f is Lipschitz continuous, and u is the entropy weak solution of (2.1) on $[0,T]$, and v is a function that is right continuous in t, uniformly bounded in $\mathrm{BV}(\mathbb{R})$ for $t>0$, and an entropy weak solution of $v_t+f(v)_x=0$ in each strip $\mathbb{R}\times[t^k,t^{k+1})$ for $0=t^0<t^1<\cdots<t^K=T$, then*

$$\|v(\,\cdot\,,T-0)-u(\,\cdot\,,T)\|_{L^1(\mathbb{R})}\le\|u_0-v(\,\cdot\,,0)\|_{L^1(\mathbb{R})}+2\epsilon\|u_0\|_{\mathrm{BV}(\mathbb{R})}+\sum_{k=1}^{K-1}[\rho_\epsilon(v(t^k),u(t^k))-\rho_\epsilon(v(t^k-0),u(t^k))],\tag{2.2}$$

where $\rho_\epsilon(w,z)=\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{1}{\epsilon}\eta(\frac{x-y}{\epsilon})|w(x)-z(y)|\,dx\,dy$, and η is any nonnegative smooth function with support in $[-1,1]$, integral one, and $\eta(-x)=\eta(x)$.

Proof. We shall use Theorem 1.2 of Chapter 2 and a bound on $\Lambda_\epsilon^{\epsilon_0}(v,u,T-0)$ to bound the difference $\|v(\,\cdot\,,T-0)-u(\,\cdot\,,T)\|_{L^1(\mathbb{R})}$.

Because v is an entropy weak solution of $v_t+f(v)_x=0$ on each strip $\mathbb{R}\times[t^k,t^{k+1})$, we know that with $\omega:=\omega(x-x',t-t'):=\eta_\epsilon(x-x')\eta_{\epsilon_0}(t-t')$, $u:=u(x',t')$, and $v:=v(x,t)$, we have

$$\begin{aligned}&-\int_0^T\int_{\mathbb{R}}\int_{t^k}^{t^{k+1}}\int_{\mathbb{R}}|v-u|\,\omega_t+\operatorname{sgn}(v-u)(f(v)-f(u))\,\omega_x\,dx\,dt\,dx'dt'\\&+\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}|v(x,t^{k+1}-0)-u(x',t')|\,\omega(x-x',t^{k+1}-t')\,dx\,dx'dt'\\&-\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}|v(x,t^k)-u(x',t')|\,\omega(x-x',t^k-t')\,dx\,dx'dt'\le 0.\end{aligned}$$

(Here we use $v(x,t^{k+1}-0)$ because we assume that v is right continuous, so v may have a jump in t at $t=t^{k+1}$.) We add each of these bounds for $t^k\le t<t^{k+1}$ to see that

$$\begin{aligned}&\Lambda_\epsilon^{\epsilon_0}(v,u,T-0)=\\&\quad-\int_0^T\int_{\mathbb{R}}\int_0^T\int_{\mathbb{R}}|v-u|\,\omega_t+\operatorname{sgn}(v-u)(f(v)-f(u))\,\omega_x\,dx\,dt\,dx'dt'\\&\quad+\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}|v(x,T-0)-u(x',t')|\,\omega(x-x',T-t')\,dx\,dx'dt'\\&\quad-\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}|v(x,0)-u(x',t')|\,\omega(x-x',0-t')\,dx\,dx'dt'\\&\le\sum_{k=1}^{K-1}\Big[\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}|v(x,t^k)-u(x',t')|\,\omega(x-x',t^k-t')\,dx\,dx'dt'\\&\quad-\int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}}|v(x,t^k-0)-u(x',t')|\,\omega(x-x',t^k-t')\,dx\,dx'dt'\Big].\end{aligned}$$

We shall let ϵ_0 tend to zero and see what our bound for $\Lambda_\epsilon^{\epsilon_0}(v,u,T-0)$

tends to. We see that if $\epsilon_0 \le \min_k |t^k - t^{k+1}|$ and $0 < k < K$ then

$$\begin{aligned}\int_0^T \int_{\mathbb{R}} \int_{\mathbb{R}} |v(x,t^k) - u(x',t')|\, \omega(x-x', t^k - t')\, dx\, dx' dt' \\ &= \int_0^T \int_{\mathbb{R}} \int_{\mathbb{R}} |v(x,t^k) - u(x',t^k)|\, \omega(x-x', t^k - t')\, dx\, dx' dt' + E \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} |v(x,t^k) - u(x',t^k)|\, \eta_\epsilon(x-x')\, dx\, dx' + E\end{aligned}$$

where E is an error that is no greater than

$$\begin{aligned}\int_0^T \int_{\mathbb{R}} \int_{\mathbb{R}} |u(x',t') - u(x',t^k)|\, \omega(x-x', t^k - t')\, dx\, dx' dt' \\ &= \int_0^T \int_{\mathbb{R}} |u(x',t') - u(x',t^k)|\, \eta_{\epsilon_0}(t^k - t')\, dx' dt' \\ &\le \sup_{|t'-t^k| \le \epsilon_0} \|u(\,\cdot\,,t') - u(\,\cdot\,,t^k)\|_{L^1(\mathbb{R})} \\ &\le L\epsilon_0 \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}.\end{aligned}$$

Therefore, for any fixed ϵ_0, we have by Theorem 1.2 that

$$\begin{aligned}\|v(\,\cdot\,,T-0) - u(\,\cdot\,,T)\|_{L^1(\mathbb{R})} \\ &\le \|u_0 - v(\,\cdot\,,0)\|_{L^1(\mathbb{R})} + 2\epsilon \|u_0\|_{\mathrm{BV}(\mathbb{R})} + \epsilon_0 L \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} \\ &\quad + \sum_{k=1}^{K-1} [\rho_\epsilon(v(t^k), u(t^k)) - \rho_\epsilon(v(t^k - 0), u(t^k))] \\ &\quad + (K-1)\epsilon_0 L \|u_0\|_{\mathrm{BV}(\mathbb{R}^n)} + \epsilon_0 L \sup_{0<t<T} \|v(\,\cdot\,,t)\|_{\mathrm{BV}(\mathbb{R}^n)}.\end{aligned}$$

We now let $\epsilon_0 \to 0$ to prove the lemma. □

§3. The Random Choice Method

The random choice method, introduced by Glimm, is a probabilistic method for proving existence of solutions for the hyperbolic system of conservation laws (2.1). Glimm showed that if the total variation of u_0 is sufficiently small, and if the equation is hyperbolic and genuinely nonlinear in the sense of Lax, then the approximate solution generated by his scheme converges almost surely to a weak solution of (2.1); later Harten and Lax showed that any Glimm weak solution satisfies the entropy condition. Glimm's scheme has also been applied with some success as a numerical method. Here we bound the expectation of the L^1 error of the approximate solution in the special case where (2.1) is a scalar equation. In particular, we prove the following theorem.

Theorem 3.1. *If $u^n(x,t)$ is the solution of Glimm's method for $t^n \le t < t^{n+1}$, $u(x,t)$ is the entropy solution of (2.1), and $T = (N+1)\Delta t$, then*

$$E(\|u(\,\cdot\,,T) - u^N(\,\cdot\,,T)\|_{L^1(\mathbb{R})}) \le \left(h + \frac{2}{\sqrt{3}}\left(\frac{h}{\Delta t}\right)^{1/2}(hT)^{1/2}\right)\|u_0\|_{\mathrm{BV}(\mathbb{R})},$$

where h is the mesh spacing, Δt is the time step, and $t^n = n\Delta t$.

We note first that although Glimm's scheme is usually defined on alternating meshes (the approximate solution is piecewise constant on the intervals $[ih,(i+1)h)$ at time t^n if n is even, and piecewise constant on $[(i-1/2)h,(i+1/2)h)$ when n is odd) this in no way affects the error estimates given below. Consequently, a fixed mesh is used for all time as a notational convenience.

We prove Theorem 3.1 for the following formulation of Glimm's scheme. We assume that u_0 has bounded variation, and that $\|f'\|_{L^\infty(\mathbb{R})}$ is finite. Choose a positive mesh size h. For each integer i, let I_i be $[ih,(i+1)h)$, and let χ_{I_i} be the characteristic function of I_i. We assume that the time step, Δt, satisfies $0 < \Delta t \le h/(2\|f'\|_{L^\infty(\mathbb{R})})$, and we define $t^n = n\Delta t$. For each nonnegative integer n, we define a function $U^n\colon \mathbb{Z} \to \mathbb{R}$ of bounded variation in the following way. Let

$$U_i^0 = \frac{1}{h}\int_{I_i} u_0(x)\,dx \tag{3.1}$$

for each integer i. If U_i^n has been defined for all i, solve the initial value problem

$$\begin{aligned} &u_t^n + f(u^n)_x = 0, && x \in \mathbb{R},\ t^n < t < t^{n+1},\\ &u^n(x,t^n) = \sum_{i\in\mathbb{Z}} U_i^n \chi_{I_i}(x), && x \in \mathbb{R}. \end{aligned} \tag{3.2}$$

The function $u^n(x,t)$ is found by piecing together the solutions of the Riemann problems at the points ih, $i \in \mathbb{Z}$. The nontrivial parts of these solutions do not overlap because of our bound on Δt.

We now choose a random variable X^{n+1}, uniformly distributed on $[0,h)$, so that the set of random variables $\{X^1,\dots,X^{n+1}\}$ are independent; the values of U_i^{n+1} are then given by

$$U_i^{n+1} = u^n(ih + X^{n+1}, t^{n+1}) \tag{3.3}$$

for every i. Chorin seems to have been the first to use exactly one random choice for all intervals I_i.

As can be seen from the definition, U^n is itself a random variable that depends on the sequence of random variables $X^1,\dots,X^n$; we propose to bound the expected value of the error at time t^{N+1}, $E(\|u(\,\cdot\,,t^{N+1}) - u^N(\,\cdot\,,t^{N+1})\|_{L^1(\mathbb{R})})$.

Because, for any values of X^1 through X^N, the approximate solution satisfies the differential equation exactly for $(x,t) \in \mathbb{R} \times (t^n, t^{n+1})$, Lemma 2.1 applies to bound the error. From this lemma it follows that

$$\begin{aligned} E(\|u(\,\cdot\,,t^{N+1}) - u^N(\,\cdot\,,t^{N+1})\|_{L^1(\mathbb{R})}) \\ &\le \|u_0(\,\cdot\,) - u^0(\,\cdot\,,0)\|_{L^1(\mathbb{R})} + 2\epsilon\|u_0\|_{\mathrm{BV}(\mathbb{R})} \\ &\quad + \sum_{n=1}^{N} E(\rho_\epsilon(u^n(t^n), u(t^n)) - \rho_\epsilon(u^{n-1}(t^n), u(t^n))). \end{aligned}$$

If we let $E^n(f)$ denote the conditional expectation of f given $X^1, \dots, X^{n-1}$ and $X^{n+1}, \dots, X^N$, then (writing t for t^n, X for X^n, and $\eta_\epsilon(x)$ for $\frac{1}{\epsilon}\eta(\frac{x}{\epsilon})$),

$$\begin{aligned} (3.4)\quad & E^n(\rho_\epsilon(u^n(t), u(t)) - \rho_\epsilon(u^{n-1}(t), u(t))) \\ &= \int_{[0,h)} \int_{\mathbb{R}} \sum_{i\in\mathbb{Z}} \int_{I_i} \eta_\epsilon(x-y) \\ &\quad \times \frac{1}{h}\left\{|u^{n-1}(ih+X,t) - u(y,t)| - |u^{n-1}(x,t) - u(y,t)|\right\} dx\, dy\, dX \\ &= \int_{\mathbb{R}} \sum_{i\in\mathbb{Z}} \int_{I_i} \int_{I_i} \eta_\epsilon(x-y) \\ &\qquad \times \frac{1}{h}\left\{|u^{n-1}(z,t) - u(y,t)| - |u^{n-1}(x,t) - u(y,t)|\right\} dz\, dx\, dy \\ &= \frac{1}{2}\int_{\mathbb{R}} \sum_{i\in\mathbb{Z}} \int_{I_i} \int_{I_i} (\eta_\epsilon(x-y) - \eta_\epsilon(z-y)) \\ &\qquad \times \frac{1}{h}\left\{|u^{n-1}(z,t) - u(y,t)| - |u^{n-1}(x,t) - u(y,t)|\right\} dz\, dx\, dy \\ &\le \frac{1}{2}\int_{\mathbb{R}} \sum_{i\in\mathbb{Z}} \int_{I_i} \int_{I_i} |\eta_\epsilon(x-y) - \eta_\epsilon(z-y)| \\ &\qquad\qquad \times \frac{1}{h}|u^{n-1}(z,t) - u^{n-1}(x,t)|\, dz\, dx\, dy. \end{aligned}$$

If we now integrate over y, we find that

$$\int_{\mathbb{R}} |\eta_\epsilon(x-y) - \eta_\epsilon(z-y)| dy \le |z-x| \frac{\|\eta'\|_{L^1(\mathbb{R})}}{\epsilon}.$$

Trivially, $|u^{n-1}(z,t) - u^{n-1}(x,t)| \le \|u^{n-1}(t)\|_{BV(I_i)}$. So it follows that

(3.4) is bounded by

$$\begin{aligned}
&\frac{1}{2}\sum_{i\in\mathbb{Z}}\int_{I_i}\int_{I_i}\frac{|z-x|}{h}dz\,dx\frac{\|\eta'\|_{L^1(\mathbb{R})}}{\epsilon}\|u^{n-1}(t)\|_{\mathrm{BV}(I_i)}\\
&\leq\sum_{i\in\mathbb{Z}}\frac{h^2}{6}\frac{\|\eta'\|_{L^1(\mathbb{R})}}{\epsilon}\|u^{n-1}(t)\|_{\mathrm{BV}(I_i)}\\
&=\frac{h^2}{6}\frac{\|\eta'\|_{L^1(\mathbb{R})}}{\epsilon}\|u^{n-1}(t)\|_{\mathrm{BV}(\mathbb{R})}
\end{aligned}$$

The inequality $\|u^{n-1}(t)\|_{\mathrm{BV}(\mathbb{R})}\leq\|u_0\|_{\mathrm{BV}(\mathbb{R})}$ is clear, because the choice of the initial data (3.1), the evolution of u^{n-1} through (2.1) (3.2), and the random choice process (3.3) are all variation diminishing. Thus,

$$E^n(\rho_\epsilon(u^n(t),u(t))-\rho_\epsilon(u^{n-1}(t),u(t)))\leq\frac{h^2}{6}\frac{\|\eta'\|_{L^1(\mathbb{R})}}{\epsilon}\|u_0\|_{\mathrm{BV}(\mathbb{R})}$$

uniformly with respect to the other random variables X^i, implying that $E(\rho_\epsilon(u^n(t),u(t))-\rho_\epsilon(u^{n-1}(t),u(t)))$ is bounded by the same quantity. Therefore, if $T=(N+1)\Delta t$, by using an obvious bound for the initial error, we have

$$\begin{aligned}
(3.5)\quad &E(\|u(\,\cdot\,,T)-u^N(\,\cdot\,,T)\|_{L^1(\mathbb{R})})\\
&\qquad\leq h\|u_0\|_{\mathrm{BV}(\mathbb{R})}+2\epsilon\|u_0\|_{\mathrm{BV}(\mathbb{R})}+\frac{T}{\Delta t}\frac{h^2}{6}\frac{\|\eta'\|_{L^1(\mathbb{R})}}{\epsilon}\|u_0\|_{\mathrm{BV}(\mathbb{R})}.
\end{aligned}$$

By letting $\eta\to\frac{1}{2}\chi_{[-1,1]}$, $\|\eta'\|_{L^1(\mathbb{R})}$ may be chosen arbitrarily close to 1. Minimizing (3.5) with respect to ϵ gives

$$\begin{aligned}
(3.6)\quad &E(\|u(\,\cdot\,,T)-u^N(\,\cdot\,,T)\|_{L^1(\mathbb{R})})\\
&\qquad\leq\left(h+\frac{2}{\sqrt{3}}\left(\frac{h}{\Delta t}\right)^{1/2}(hT)^{1/2}\right)\|u_0\|_{\mathrm{BV}(\mathbb{R})}.
\end{aligned}$$

The theorem is proved. We remark that if one chooses to interpret Glimm's method as providing that the approximate solution is equal to U_i^n on $[ih,(i+1)h)\times[t^n,t^{n+1})$, then the above inequality still holds with a small change for the error incurred in at most one time step.

Extensive numerical evidence, and some theoretical work, shows that monotone finite difference schemes, such as the Engquist-Osher scheme, perform better for problems with uniformly convex fluxes than for problems with linear fluxes; in fact, the Engquist-Osher method is $O(h)$ accurate for the problem

$$\begin{aligned}
(3.7)\qquad &u_t+(\frac{u^2}{2})_x=0, && x\in\mathbb{R},\ t>0,\\
&u(x,0)=\chi_{(-\infty,0]}(x), && x\in\mathbb{R}.
\end{aligned}$$

For problems, such as this one, whose solution consists of a single shock of height one, the expected error in Glimm's scheme can be estimated directly by applying the Central Limit Theorem. If the shock speed is s, and $p = s\Delta t/h$, then after N time steps, the probability distribution of the shock location (measured in spatial intervals) is binomial with parameters N and p; therefore, for large N, the shock location error is approximately normal with mean 0 and variance $\sigma^2 = Np(1-p)h^2$. Asymptotically, the expected value of the $L^1(\mathbb{R})$ error, which is the absolute value of the shock location error, is

$$\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}|\xi|e^{-\xi^2/2\sigma^2}\,d\xi = \sqrt{\frac{2}{\pi}}\,\sigma\int_0^{\infty}e^{-\xi^2/2\sigma^2}\,d\frac{\xi^2}{2\sigma^2} = \sqrt{\frac{2}{\pi}}\,\sigma$$

or $(\frac{2}{\pi}p(1-p)\frac{h}{\Delta t})^{1/2}(Th)^{1/2}$, where $T = N\Delta t$. Our bound on the ratio $\Delta t/h$ implies that $0 \le p \le 1/2$. For example, when Glimm's scheme is applied to (3.7) with $\Delta t = h/2$, $p = 1/4$, the expected value of the error is about $0.3455(\frac{h}{\Delta t})^{1/2}(hT)^{1/2}$. Theorem 1 gives a bound of $1.1547(\frac{h}{\Delta t})^{1/2}(hT)^{1/2}$ independently of the value of p, a fairly close result. Note that Glimm's scheme does not converge for this problem when $h^2/\Delta t$ does not tend to zero as h and Δt tend to zero.

§4. Godunov's Method

GODUNOV'S method was one of the first numerical methods for hyperbolic systems of nonlinear conservation laws. We shall give two different bounds for the error in Godunov's method applied to the scalar equation (2.1). The first bound is derived by showing that Godunov's method falls into the general class of monotone scheme. The second bound, which we shall use later, can be derived much as our previous bound for the random choice method.

Godunov's scheme differs from Glimm's scheme only in that Godunov determines U_i^{n+1} by averaging $u^n(\,\cdot\,, t^{n+1})$ over I_i:

$$U_i^{n+1} = \frac{1}{h}\int_0^h u^n(ih + X, t^{n+1})dX. \tag{4.1}$$

We first determine how to calculate these quantities efficiently when Δt and h satisfy the CFL condition.

We shall consider specifically how to calculate U_0^1. Because $u^0(x,t)$ depends only on $u_0(\xi)$ for $|\xi - x| \le Lt$, where $L = \sup_\xi |f'(\xi)|$, if we restrict $L\Delta t \le h$ then the value of $u^0(0,t)$ for $t < \Delta t$ depends only on U_{-1}^0 and U_0^0.

OK, I don't really understand yet what's going on, but I want to hand out these notes tomorrow when I'll try to offer an explanation. So let me just define

$$F(u_L, u_R) := \begin{cases} \min_{u_L<u<u_R} f(u), & u_L < u_R, \\ \max_{u_R<u<u_L} f(u), & u_R < u_L. \end{cases}$$

Then the Godunov scheme can be written

$$\frac{U_i^{k+1} - U_i^k}{\Delta t} + \frac{F(U_i^k, U_{i+1}^k) - F(U_{i-1}^k, U_i^k)}{h} = 0, \quad i \in \mathbb{Z},\ k \geq 0.$$

This scheme is monotone because the underlying differential equation and the averaging process (4.1) are monotone. Thus, it can be analyzed using the techniques of Chapter 3, and it has an error bound of $C(\sqrt{\Delta t T} + \sqrt{hT})\|u_0\|_{\mathrm{BV}(\mathbb{R})}$. This bound decreases as Δt decreases. Our next bound, based on Lemma 2.1, decreases as Δt *increases*. This makes sense in a way, because the only entropy "violation" occurs when we average $u^k(x, t^{k+1})$ on each interval I_i to obtain U_i^{k+1}; the approximation can be considered an exact entropy solution between time steps.

The following theorem shows that the error in Godunov's method is bounded by the same expression that bounds the expected error in Glimm's scheme.

Theorem 4.1. *If $u^n(x,t)$ is the solution of Godunov's method for $t^n \leq t < t^{n+1}$, $u(x,t)$ is the entropy solution of (2.1), and $T = (N+1)\Delta t$, then*

(4.2)

$$\|u(\,\cdot\,, T) - u^N(\,\cdot\,, T-0)\|_{L^1(\mathbb{R})} \leq \left(h + \frac{2}{\sqrt{3}} \left(\frac{h}{\Delta t} \right)^{1/2} (hT)^{1/2} \right) \|u_0\|_{\mathrm{BV}(\mathbb{R})}.$$

Proof. We proceed as in Theorem 3.1. Again, with $t = t^n$,

$$\begin{aligned}
&\rho_\epsilon(u^n(t), u(t)) - \rho_\epsilon(u^{n-1}(t), u(t)) \\
&= \int_{\mathbb{R}} \sum_{i \in \mathbb{Z}} \int_{I_i} \eta_\epsilon(x-y) \left\{ |U_i^n - u(y,t)| - |u^{n-1}(x,t) - u(y,t)| \right\} dx\, dy \\
&= \int_{\mathbb{R}} \sum_{i \in \mathbb{Z}} \int_{I_i} \eta_\epsilon(x-y) \\
&\quad \times \{ |\frac{1}{h} \int_0^h u^{n-1}(ih + X, t)\, dX - u(y,t)| - |u^{n-1}(x,t) - u(y,t)| \} dx\, dy \\
&\leq \int_0^h \int_{\mathbb{R}} \sum_{i \in \mathbb{Z}} \int_{I_i} \eta_\epsilon(x-y) \\
&\quad \times \frac{1}{h} \left\{ |u^{n-1}(ih + X, t) - u(y,t)| - |u^{n-1}(x,t) - u(y,t)| \right\} dx\, dy\, dX.
\end{aligned}$$

One may now follow the series of inequalities in (3.4) and the subsequent arguments to obtain the estimate in the statement of the theorem. □

The estimate is rather sharp, as can be seen by considering

$$\begin{aligned}
&u_t + u_x = 0, && x \in \mathbb{R},\ t > 0, \\
&u(x,0) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0, \end{cases} && x \in \mathbb{R}.
\end{aligned} \tag{4.3}$$

For this problem, Godunov's method and the Engquist-Osher method are identical. In §3.6 we showed that the Engquist-Osher scheme has an asymptotic error of

$$\left(\frac{2}{\pi}\right)^{1/2}(hT)^{1/2}\left(1-\frac{\Delta t}{h}\right)^{1/2};$$

when $\Delta t/h = \frac{1}{2}$ the error is asymptotically $(\frac{1}{\pi})^{1/2}(Th)^{1/2} = 0.564(Th)^{1/2}$, compared to our estimate of $2\sqrt{\frac{2}{3}}(Th)^{1/2} = 1.633(Th)^{1/2}$.

Note that our analysis applies even if the CFL condition $\Delta t < h/\|f'\|_{L^\infty(\mathbb{R})}$ is violated, as long as the wave interactions in the solution of (3.2) are calculated exactly; this will prove important in the next chapter.

Chapter 5

Stability and Moving Grid Numerical Methods

Beginning in this chapter, we shall study more closely the properties of scalar conservation laws in one space dimension. In §1 we shall study the stability of conservation laws when the flux, as well as the initial data, is perturbed. We shall use this result in §2 to analyze a moving grid method introduced by Dafermos based on piecewise constant approximations that will exhibit a convergence rate of first order in the number of parameters; this is to be contrasted with the convergence rate of order $\frac{1}{2}$ exhibited by monotone finite difference schemes. We go on in §3 to devise a numerical method based on piecewise linear approximations on a moving grid that achieves second order approximation in the number of parameters. In the next chapter we shall study what these new approximation results imply about the regularity of solutions of such scalar conservation laws.

§1. Stability

In this section we consider the stability of the entropy weak solution of the conservation law

$$\begin{aligned} &u_t + \nabla \cdot f(u) = 0, && x \in \mathbb{R}^n,\ t > 0, \\ &u(x,0) = u_0(x), && x \in \mathbb{R}^n, \end{aligned} \tag{1.1}$$

under changes in the flux f as well as the initial data u_0. Specifically, we prove the following theorem.

Theorem 1.1. *Assume that $u(x,t)$ is the entropy weak solution of* (1.1) *and that $v(x,t)$ is the entropy weak solution of*

$$\begin{aligned} &v_t + \nabla \cdot g(v) = 0, && x \in \mathbb{R}^n,\ t > 0, \\ &v(x,0) = v_0(x), && x \in \mathbb{R}^n, \end{aligned}$$

with u_0 and $v_0 \in \mathrm{BV}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ and f and g Lipschitz continuous. Then

$$\begin{aligned} \|u(\,\cdot\,,T) - v(\,\cdot\,,T)\|_{L^1(\mathbb{R}^n)} &\le \|u_0 - v_0\|_{L^1(\mathbb{R}^n)} \\ &\quad + T\| \, |f' - g'| \, \|_{L^\infty(\mathbb{R})} \min(\|u_0\|_{\mathrm{BV}(\mathbb{R}^n)}, \|v_0\|_{\mathrm{BV}(\mathbb{R}^n)}). \end{aligned} \tag{1.2}$$

Proof. For notational convenience we first prove the result in one dimension.

We shall use Kuznetsov's approximation Theorem 1.2 of Chapter 2 to compare u and v, considering v as an approximation to u. Because we are considering two different fluxes f and g, we shall use the slightly expanded notation $\Lambda_\epsilon^{\epsilon_0}(w,z,T,f)$ and $\Lambda_\epsilon^{\epsilon_0}(w,z,T,g)$ for the average entropy inequality for the fluxes f and g respectively.

When we consider v as an approximation to u, we need to bound $\Lambda_\epsilon^{\epsilon_0}(v,u,T,f)$ in Theorem 1.2 of Chapter 2. We set $v := v(x,t)$, $u := u(x',t')$, and $\omega := \omega(x-x',t-t')$ and note that $\Lambda_\epsilon^{\epsilon_0}(v,u,T,g) \le 0$ to calculate

(1.3)

$$
\begin{aligned}
&\Lambda_\epsilon^{\epsilon_0}(v,u,T,f)\\
&= -\int_0^T\int_{\mathbb{R}}\int_0^T\int_{\mathbb{R}} |v-u|\,\omega_t + [f(v\vee u)-f(v\wedge u)]\,\omega_x\,dx\,dt\,dx'dt'\\
&\quad + \int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}} |v(x,T)-u|\,\omega(x-x',T-t')\,dx\,dx'dt'\\
&\quad - \int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}} |v(x,0)-u|\,\omega(x-x',0-t')\,dx\,dx'dt'\\
&= -\int_0^T\int_{\mathbb{R}}\int_0^T\int_{\mathbb{R}} |v-u|\,\omega_t + [g(v\vee u)-g(v\wedge u)]\,\omega_x\,dx\,dt\,dx'dt'\\
&\quad + \int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}} |v(x,T)-u|\,\omega(x-x',T-t')\,dx\,dx'dt'\\
&\quad - \int_0^T\int_{\mathbb{R}}\int_{\mathbb{R}} |v(x,0)-u|\,\omega(x-x',0-t')\,dx\,dx'dt'\\
&\quad - \int_0^T\int_{\mathbb{R}}\int_0^T\int_{\mathbb{R}} [(f-g)(v\vee u)-(f-g)(v\wedge u)]\,\omega_x\,dx\,dt\,dx'dt'\\
&= \Lambda_\epsilon^{\epsilon_0}(v,u,T,g)\\
&\quad - \int_0^T\int_{\mathbb{R}}\int_0^T\int_{\mathbb{R}} [(f-g)(v\vee u)-(f-g)(v\wedge u)]\,\omega_x\,dx\,dt\,dx'dt'\\
&\le -\int_0^T\int_{\mathbb{R}}\int_0^T\int_{\mathbb{R}} [(f-g)(v\vee u)-(f-g)(v\wedge u)]\,\omega_x\,dx\,dt\,dx'dt'.
\end{aligned}
$$

This is the basic inequality of this theorem; we must now bound the right side of (1.3) in a suitable way.

We define $h(v) := (f-g)(v)$ and $\bar{h}(v,u) := h(v\vee u)-h(v\wedge u)$. Note that $|\bar{h}(v_1,u)-\bar{h}(v_2,u)| \le |h(v_1)-h(v_2)|$. Because u, v, and ω_x are bounded, we have by the Lebesgue Dominated Convergence Theorem that the absolute

value of the right side of (1.3) is equal to the limit as $\Delta x \to 0$ of

$$\begin{aligned}
&\left| \frac{1}{\Delta x} \int_0^T \int_{\mathbb{R}} \int_0^T \int_{\mathbb{R}} \bar{h}(v,u)\,[\omega(x+\Delta x - x', t-t') - \omega]\,dx\,dt\,dx'dt' \right| \\
&= \left| \frac{1}{\Delta x} \int_0^T \int_{\mathbb{R}} \int_0^T \int_{\mathbb{R}} [\bar{h}(v,u) - \bar{h}(v(x-\Delta x,t),u)]\,\omega\,dx\,dt\,dx'dt' \right| \\
&\le \frac{1}{|\Delta x|} \int_0^T \int_{\mathbb{R}} \int_0^T \int_{\mathbb{R}} |h(v) - h(v(x-\Delta x,t))|\,\omega\,dx\,dt\,dx'dt' \\
&\le \frac{1}{|\Delta x|} \int_0^T \int_{\mathbb{R}} |h(v) - h(v(x-\Delta x,t))|\,dx\,dt \\
&\le \|h'\|_{L^\infty(\mathbb{R})} \frac{1}{|\Delta x|} \int_0^T \int_{\mathbb{R}} |v - v(x-\Delta x,t)|\,dx\,dt \\
&\le T\|f' - g'\|_{L^\infty(\mathbb{R})} \|v_0\|_{\mathrm{BV}(\mathbb{R})}.
\end{aligned}$$

This bound is independent of Δx, ϵ, or ϵ_0. We let these parameters tend to zero in Theorem 1.2 of Chapter 2, to see that

$$\|u(\,\cdot\,,T) - v(\,\cdot\,,T)\|_{L^1(\mathbb{R})} \le \|u_0 - v_0\|_{L^1(\mathbb{R})} + T\|f' - g'\|_{L^\infty(\mathbb{R})} \|v_0\|_{\mathrm{BV}(\mathbb{R})}.$$

The theorem follows from symmetry in u and v.

In several space dimensions we note that

$$|f'(\xi) - g'(\xi)| := \max_{1\le j\le n} |f_j'(\xi) - g_j'(\xi)|$$

and apply the same argument in each space dimension. □

REMARK 1.1. Just as in Kuznetsov's theorem, the right side of (1.2) can be written in different ways depending on how one bounds (1.3).

§2. Moving Grid Methods I: Dafermos's Method

DAFERMOS came up with a technique for proving the existence of entropy solutions of

$$\begin{aligned}
&u_t + f(u)_x = 0, \qquad x \in \mathbb{R},\ t > 0, \\
&u(x,0) = u_0(x), \qquad x \in \mathbb{R},
\end{aligned} \tag{2.1}$$

that involves solving a perturbed problem. We can interpret his construction as a numerical method for approximating solutions of (2.1). HEDSTROM implemented his technique as a numerical method both for scalar equations and hyperbolic systems of two conservation laws.

We shall assume that u_0 has bounded variation, and additionally that the support of u_0 is contained in a finite interval, $[0,1]$, say. We choose a positive number N and set $h = N^{-1}$ and $I_i := [ih, (i+1)h)$. The initial approximation $\tilde{u}_0$ is given by

$$\tilde{u}_0(x)|_{I_i} = \frac{1}{h} \int_{I_i} u_0(s)\,ds.$$

We now choose an approximation $\tilde{f}$ to f. The function $\tilde{f}$ will be continuous, piecewise linear, with breakpoints at the points ih, such that $f(ih) = \tilde{f}(ih)$ for $i \in \mathbb{Z}$. We solve the following perturbed problem *exactly*:

$$\begin{aligned} &\tilde{u}_t + \tilde{f}(\tilde{u})_x = 0, \qquad && x \in \mathbb{R},\ t > 0, \\ &\tilde{u}(x,0) = \tilde{u}_0(x), && x \in \mathbb{R}. \end{aligned} \tag{2.2}$$

We must show two things: that the error is bounded in a reasonable way, and that we can in fact solve (2.2) in a finite number of steps.

If we wish to apply Theorem 1.1, we need to bound $\|f' - \tilde{f}'\|_{L^\infty(\mathbb{R})}$. On the interval $(0,h)$, we have that

$$\begin{aligned} |f'(\xi) - \tilde{f}'(\xi)| &= |f'(\xi) - \frac{f(h) - f(0)}{h}| \\ &= |f'(\xi) - \frac{1}{h}\int_0^h f'(\eta)\,d\eta| \\ &\le \frac{1}{h}\int_0^h |f'(\xi) - f'(\eta)|\,d\eta \\ &\le \frac{\|f''\|_{L^\infty(\mathbb{R})}}{h}\int_0^h |\xi - \eta|\,d\eta \\ &\le \frac{\|f''\|_{L^\infty(\mathbb{R})}}{2N} \qquad \text{(achieved when } \xi = 0 \text{ or } \xi = h\text{)}. \end{aligned}$$

The same inequality holds for any ξ. We have the by now obvious inequalities $\|u_0 - \tilde{u}_0\|_{L^1(\mathbb{R})} \le h\|u_0\|_{\mathrm{BV}(\mathbb{R})}$ and $\|\tilde{u}_0\|_{\mathrm{BV}(\mathbb{R})} \le \|u_0\|_{\mathrm{BV}(\mathbb{R})}$. Therefore, Theorem 1.1 allows us to conclude that

$$\|u(\,\cdot\,,T) - \tilde{u}(\,\cdot\,,T)\|_{L^1(\mathbb{R})} \le \big(1 + \frac{1}{2}\|f''\|_{L^\infty(\mathbb{R})}T\big)\frac{1}{N}\|u_0\|_{\mathrm{BV}(\mathbb{R})}, \qquad T > 0.$$

We must bound the complexity of solving the perturbed problem (2.2) if we are to consider this a practical numerical method. We shall assume that f, and hence $\tilde{f}$, are convex. First, we bound the number of constant states that can occur immediately after $t = 0$. Because we assumed that the initial data has support in $[0,1]$, we start with $N+2$ constant states in $\tilde{u}_0$ (including the zero values for $x < 0$ and $x > 1$), denoted by U_i, $i = 0, \dots, N+1$. In order to solve (2.2) we must solve the $N+1$ Riemann problems between U_i and U_{i+1}. We have described the solution of the Riemann problem for a piecewise linear flux in §1 of Chapter 4; in our case a constant state in $\tilde{u}(x,t)$ will arise for each j with $U_i < j/N < U_{i+1}$, since the break-points of $\tilde{f}$ are at the points j/N. We can bound the number of these new constant states in terms of the variation of $\tilde{u}_0$ by noting that if there are $\mathcal{N}_i$ constant states introduced between U_i and U_{i+1}, then $|U_i - U_{i+1}| \ge (\mathcal{N}_i - 1)/N$. Therefore, $\mathcal{N}_i \le 1 + N|U_i - U_{i+1}|$ and the

total number of added states is bounded by

$$\sum_{i=-1}^{N-1} \mathcal{N}_i \le N+1+N\sum_{i=-1}^{N-1}|U_i-U_{i+1}|$$
$$= N+1+N\|\tilde{u}_0\|_{\mathrm{BV}(\mathbb{R})} \le N+1+N\|u_0\|_{\mathrm{BV}(\mathbb{R})}.$$

Therefore, for a short time after $t=0$, there are no more than $2N+3+N\|u_0\|_{\mathrm{BV}(\mathbb{R})}$ constant pieces in $\tilde{u}(x,t)$. These pieces are separated by shocks that move at a speed determined by the Rankine-Hugoniot condition. (Here we abuse our previous terminology, and refer to all discontinuities as shocks, even if they can be considered rarefaction waves of the piecewise linear flux.)

These moving shocks could collide, or interact, for large times. If $\tilde{f}$ is convex, then each time two shocks collide, with three constant states given by u_L (to the left of the left shock), u_M (between the two shocks), and u_R (to the right), then only one shock emerges, with left state u_L, right state u_R, and speed given by the Rankine-Hugoniot condition. (This can easily be shown based on a trivial, but tedious, case analysis of the six possible orderings of u_L, u_M, and u_R.) Thus, interaction of waves causes a reduction in the number of constant states. Since there are $O(N)$ constant states initially, there can be at most $O(N)$ wave interactions in the solution of $\tilde{u}(x,t)$ *for all time.* Thus, one can completely determine the solution $\tilde{u}(x,t)$ as constant on $O(N)$ polygonal regions in the x-t plane, with all of the regions determined by at most $O(N)$ line segments.

One would program this numerical method as a simulation, where the events are the intersection of two shocks. We shall briefly describe an algorithm that allows one to completely calculate $\tilde{u}(x,t)$ for all $x\in\mathbb{R}$ and $t>0$ in $O(N\log N)$ steps.

First, one determines, left to right, all the constant states and shocks that will emerge from the initial data. The shocks to the left and right of each constant state are noted, as are the constant states to the left and right of each shock. This takes $O(N)$ time.

For each pair of adjacent shocks, the time of intersection is noted (if the shocks are diverging, then this time is infinite); this information is put on a *heap.* A heap is a binary tree with the following properties. The tree is *complete,* i.e., all levels but the last are full, and the last level has all its vacancies on the right. There is an ordered *key* associated with the heap (in our case it is the time of intersection), and the key of each node in the tree is smaller than the keys of the children (if any) of that node. Consequently, the earliest time of intersection between two shocks will be at the root of the tree.

There are simple algorithms for adding and removing elements from a heap while maintaining these data invariants. Each algorithm takes but $O(\log N)$ time. Therefore, it takes $O(N\log N)$ operations to construct the heap originally.

The root of the tree contains the first wave interaction that will occur

in the solution of $\tilde{u}$. Let's say that it involves the three states u_L, u_M, and u_R, and that there are states u_{LL} to the left of u_L and u_{RR} to the right of u_R. After the u_L/u_M and u_M/u_R waves interact, the state u_M and the u_L/u_M and u_M/u_R waves no longer exists, so these are removed from the list of states and waves, respectively, The new shock between u_L and u_R is calculated, and the root of the tree, which used to represent the wave interaction is removed from the tree. The interaction between the u_{LL}/u_L shock and the u_L/u_M shock, now no longer valid, is removed from the tree; so is the similar interaction on the right. Finally, the interaction between the u_{LL}/u_L shock and the new u_L/u_R shock is added to the tree; again, a similar calculation is performed on the right.

When manipulated in this way, a heap serves as a *priority queue,* a useful tool for programming simulations. The data manipulations for each wave interaction in $\tilde{u}$ requires $O(\log N)$ operations to calculate; since there are at most $O(N)$ wave interactions, the calculation of $\tilde{u}$ requires at most $O(N \log N)$ operations.

Our analysis is summarized in the following theorem.

Theorem 2.1. *Let $u_0 \in \mathrm{BV}(\mathbb{R})$ have support in $[0,1]$ and assume that there exists a constant M such that $0 \le f'' \le M$. Let $u(x,t)$ denote the solution of (2.1). Then for each positive N, there exists a function $\tilde{u}(x,t)$ that is piecewise constant on no more than $2N+3+N\|u_0\|_{\mathrm{BV}(\mathbb{R})}$ polygonal regions in x-t space, and*

$$\|u(\,\cdot\,,t) - \tilde{u}(\,\cdot\,,t)\|_{L^1(\mathbb{R})} \le \big(1 + \frac{1}{2}Mt\big)\frac{1}{N}\|u_0\|_{\mathrm{BV}(\mathbb{R})}, \qquad t > 0.$$

Furthermore, $\tilde{u}$ can be calculated as the solution of (2.2) in $O(N \log N)$ operations.

Dafermos's method is much more computationally efficient than monotone finite difference methods. On an interval in x-t space, $[0,1] \times [0,T]$, say, it takes $O(N^2)$ operations to calculate the solution of a monotone finite difference method (with $h = 1/N$) that achieves an accuracy of $O(N^{-1/2})$. In other words, Error = Work$^{-1/4}$ for monotone finite difference schemes. In contrast, Dafermos's method achieves an error of $O(N^{-1})$ with roughly $O(N)$ operations ($\log_2 N$ is bounded by 20 for $N \le 1,000,000$), achieving Error = Work^{-1}.

§3. Moving Grid Methods II: The Large-Time-Step Godunov Method

LeVeque introduced several interesting techniques in a numerical method for (2.1) that he considered to be a large-time-step generalization of Godunov's method. We shall describe and analyze a variant of his method in this section.

We shall again assume that $0 \le f'' \le M$, that u_0 is of bounded variation and has support in $[0,1]$, and that N is a positive integer parameter

with $h = 1/N$. The approximation to the initial data u_0 is the piecewise constant average of u_0 on each interval $[ih, (i+1)h)$, $i = 0, \ldots, N-1$.

A procedure, called (appropriately enough!) an approximate Riemann solver, is used to approximate the solution of the Riemann problem between adjacent constant states with the flux f. (Approximate Riemann solvers are an integral part of many numerical methods.) This approximate Riemann solver works in the following way. If the solution of the Riemann problem consists of a single shock separating the two constant states of the initial data, then that is also the solution of the approximate Riemann solver. If, however, the solution of the Riemann problem for f consists of a rarefaction wave, then that rarefaction wave is replaced by a "staircase" of small, entropy violating shocks (whose speeds are still determined by the Rankine-Hugoniot condition for f), between the left state and the right state. To be precise, we shall assume that the added, artificial, constant states are located only at the points $u = j/N$ for $u_L < j/N < u_R$. Thus, the height of these entropy-violating shocks is at most $1/N$.

LeVeque wished to have a Godunov-type scheme, where after some time Δt he would project the approximate solution onto the original grid, but one must decide how to deal with the interaction of the waves generated by our approximate Riemann solver when $\|f'\|_{L^\infty(\mathbb{R})}\Delta t \geq h/2$. In later papers LeVeque allowed these waves to pass through each other as though in a linear wave equation, but in his first paper his algorithm strove to calculate their interaction *exactly*; i.e., whenever two waves came together, he would calculate anew the waves that arose from the new Riemann problem. The method continues in this way for a time Δt, at which point the approximate solution is projected onto the original grid and the process starts over again.

LeVeque presented intriguing computations in his paper that showed that as one increased the time step, until eventually $\Delta t = T$, the error in his scheme *decreased.* Our analysis will be able to account for this behavior.

The method as we have described it can (almost) be analyzed as a combination of Dafermos's method and Godunov's method. The introduction of entropy-violating shocks is algebraically identical to Dafermos's piecewise linear approximation $\tilde{f}$ to f; in each case the discontinuity speed between the states $u_L = j/N$ and $u_R = (j+1)/N$ is

$$\frac{f(u_L) - f(u_R)}{u_L - u_R} = \frac{\tilde{f}(u_L) - \tilde{f}(u_R)}{u_L - u_R}.$$

The speed of numerical shocks will differ slightly between the two methods, however, because for general $u \neq j/N$, $f(u) \neq \tilde{f}(u)$. Therefore, let us assume that LeVeque's ingeneous entropy-violating "staircase" arises from the piecewise linear approximation of f by $\tilde{f}$. Thus, the approximate Riemann solver and the exact calculation of wave interactions can be considered an

approximation by problem (2.2), which introduces an error bounded by

$$\|u(\,\cdot\,,T)-\tilde u(\,\cdot\,,T)\|_{L^1(\mathbb{R})}\le \big(1+\frac{1}{2}MT\big)h\|u_0\|_{\mathrm{BV}(\mathbb{R})}.$$

The averaging step can be considered an application of Godunov's method to (2.2), which was analyzed in §4 of Chapter 4; it introduces an error bounded by

$$\frac{2}{\sqrt{3}}\left(\frac{h}{\Delta t}\right)^{1/2}(hT)^{1/2}\|u_0\|_{\mathrm{BV}(\mathbb{R})}.$$

(The $O(h)$ term is missing because the initial data of (2.2) is piecewise constant on the grid of size h.) Thus, we achieve the final bound for the large-time-step Godunov method of

$$\|u(\,\cdot\,,T)-\tilde u(\,\cdot\,,T)\|_{L^1(\mathbb{R})}\le \big(h+\frac{1}{2}MTh+\frac{2}{\sqrt{3}}\left(\frac{h}{\Delta t}\right)^{1/2}(hT)^{1/2}\big)\|u_0\|_{\mathrm{BV}(\mathbb{R})}.$$

Thus, when $\Delta t\approx h$, (e.g., when the CFL number is 1), the error is $O((hT)^{1/2})$ and the method is order–$\frac{1}{2}$ accurate If, however, $\Delta t=T$, and only one time step is taken, then the error is $O(h)$ and the method is first-order accurate. This explains the success of LeVeque's numerical experiments.

§4. Moving Grid Methods III: Piecewise Linear Approximation

In this section we describe a method for approximating solutions of (2.1) by piecewise linear functions, and in so doing we achieve an approximation rate of $O(N^{-2})$ with $O(N)$ linear pieces. Just as for Dafermos's scheme, the computational complexity will be $O(N\log N)$.

This method is based on several observations. The first observation is that if $f(u):=au^2+bu+c$ is quadratic in u, and $u_0(x):=\alpha x+\beta$ is linear in x, then $u(x,t)$ will be linear in x for as long as the solution exists. We calculate from the method of characteristics

$$u=u_0(x-f'(u)t)=\alpha(x-(2au+b)t)+\beta,$$

so that

$$u(x,t)=\frac{\alpha x+\beta-\alpha bt}{1+2\alpha at}. \tag{4.1}$$

The second observation is that the above property is purely local. For example, if

$$u(x,0)=\begin{cases}0, & x<0,\\ x, & 0\le x\le 1,\\ 1, & x>1,\end{cases}$$

and $f(u)$ is quadratic only for $u \in [0,1]$, then between the characteristics emenating from $x = 0$ and $x = 1$, the solution of (2.1) is linear. Thus, one need track only the trajectories of these two points, and the solution is linear between them.

The third and final observation is that one can determine exactly the trajectory of a shock at $x = 0$, say, when immediately to the left of the shock u is linear and $f(u)$ is quadratic, and to the right u is a possibly different linear function and $f(u)$ is a different quadratic function. This follows from the following shock propagation condition, which is equivalent to the Rankine-Hugoniot condition.

Let us consider a shock propagating from $x = 0$ at time $t = 0$ along a curve $x = \xi(t)$. At each point (ξ, t) on the curve, let $(x_L, 0)$ and $(x_R, 0)$ be the starting points of the characteristics passing through the point (ξ, t) from the left and right, respectively. Then $x_L < 0$, $x_R > 0$, and x_L, x_R, ξ, and t satisfy the two equations

$$\xi = x_L + f'(u_0(x_L))\, t \tag{4.2}$$

and

$$\xi = x_R + f'(u_0(x_R))\, t. \tag{4.3}$$

The points $(x_L, 0)$, $(x_R, 0)$, and (ξ, t) form the vertices of a triangle $\mathcal{T}$ in the x-t plane. If we assume that the solution u of (2.1) is smooth to the left and right of the shock, we can integrate (2.1) over $\mathcal{T}$ to see that

$$\begin{aligned} 0 &= \int_{\mathcal{T}} u_t + f(u)_x \, dx \\ &= \int_{\partial\mathcal{T}} (f(u), u) \cdot n \, d\sigma \\ &= -\left[f(u_0(x_L)) - f'(u_0(x_L))u_0(x_L)\right] t \\ &\quad + \left[f(u_0(x_R)) - f'(u_0(x_R))u_0(x_R)\right] t \\ &\quad - \int_{x_L}^{x_R} u_0(x)\, dx. \end{aligned} \tag{4.4}$$

Here we have used the fact that u, and hence $f(u)$ and $f'(u)$, are constant along the top sides of $\mathcal{T}$, and that the length of the top left line segment of $\mathcal{T}$, for example, is

$$\sqrt{t^2 + (\xi - x_L)^2} = \sqrt{t^2 + [f'(u_0(x_L))t]^2} = t\sqrt{1 + [f'(u_0(x_L))]^2}.$$

Thus, the propagation of the shock is totally determined by the system of equations (4.2), (4.3), and (4.4). This is a system of three (nonlinear) equations in ξ, t, and the two auxiliary variables x_L and x_R. In our example, we have that

$$u_0(x) = \begin{cases} \alpha_L x + \beta_L, & x < 0, \\ \alpha_R x + \beta_R, & x \geq 0, \end{cases}$$

with

$$f(u) = \begin{cases} a_L u^2 + b_L u + c_L, & u \text{ near } u_0(0^-) = \beta_L, \\ a_R u^2 + b_R u + c_R, & u \text{ near } u_0(0^+) = \beta_R. \end{cases}$$

In this case, (4.2), (4.3), and (4.4) form a system of *polynomial* equations in ξ, t, x_L and x_R, and the two auxiliary variables can be eliminated through a process called, well, elimination, leaving us a single equation in the two variables of interest, ξ and t. We can use MACSYMA to find this equation, which leads to

$$\begin{aligned}
&(2\alpha_L a_L t + 1)(2\alpha_R a_R t + 1) \\
&(2\alpha_L \alpha_R a_R t \xi^2 - 2\alpha_L \alpha_R a_L t \xi^2 - \alpha_R \xi^2 + \alpha_L \xi^2 \\
&- 4b_L \alpha_L \alpha_R a_R t^2 \xi + 4b_R \alpha_L \alpha_R a_L t^2 \xi + 4\beta_L \alpha_R a_R t \xi \\
&- 4\beta_R \alpha_L a_L t \xi + 2b_R \alpha_R t \xi - 2b_L \alpha_L t \xi - 2\beta_R \xi + 2\beta_L \xi \\
&+ 8c_R \alpha_L \alpha_R a_L a_R t^3 - 8c_L \alpha_L \alpha_R a_L a_R t^3 + 2b_L^2 \alpha_L \alpha_R a_R t^3 \\
&- 2b_R^2 \alpha_L \alpha_R a_L t^3 - 4\beta_L^2 \alpha_R a_L a_R t^2 + 4\beta_R^2 \alpha_L a_L a_R t^2 \\
&- 4\beta_L b_L \alpha_R a_R t^2 + 4c_R \alpha_R a_R t^2 - 4c_L \alpha_R a_R t^2 \\
&+ 4\beta_R b_R \alpha_L a_L t^2 + 4c_R \alpha_L a_L t^2 - 4c_L \alpha_L a_L t^2 - b_R^2 \alpha_R t^2 \\
&+ b_L^2 \alpha_L t^2 + 2\beta_R^2 a_R t - 2\beta_L^2 a_L t + 2\beta_R b_R t \\
&- 2\beta_L b_L t + 2c_R t - 2c_L t) = 0.
\end{aligned}$$

The first two factors are zero when the solution to the left and right of the shock no longer exists; compare with the denominator in (4.1). The third factor is a cubic polynomial in ξ and t, which the shock trajectory must satisfy. Thus, the shock trajectory forms part of the zero set of a cubic polynomial, or a cubic curve. (Newton attempted to classify these curves in *Enumeratio linearum tertii ordinis,* published as an appendix of his *Opticks* in 1704 and excerpted extensively in *A Source Book in Mathematics, 1200–1800,* edited by D. J. Struik. Struik indicates that Newton's paper formed the greatest advance in algebraic geometry since the time of the Greeks.) In fact, since the equation is quadratic in ξ, the shock path can be written

down as a function of t:

$$\begin{aligned}\xi = \{ \pm \big[&(-1)(2\alpha_L a_L t + 1)(2\alpha_R a_R t + 1) \\ &(4\alpha_L c_R \alpha_R a_R t^2 - 4c_L \alpha_L \alpha_R a_R t^2 - \alpha_L {b_R}^2 \alpha_R t^2 + 2b_L \alpha_L b_R \alpha_R t^2 \\ &- 4\alpha_L a_L c_R \alpha_R t^2 + 4c_L \alpha_L a_L \alpha_R t^2 - {b_L}^2 \alpha_L \alpha_R t^2 - 2\beta_L^2 \alpha_R a_R t \\ &+ 2\alpha_L \beta_R^2 a_R t - 2\beta_L b_R \alpha_R t - 2c_R \alpha_R t + 2\beta_L^2 a_L \alpha_R t + 2\beta_L b_L \alpha_R t \\ &+ 2c_L \alpha_R t + 2\alpha_L \beta_R b_R t - 2\alpha_L a_L \beta_R^2 t - 2b_L \alpha_L \beta_R t \\ &+ 2\alpha_L c_R t - 2c_L \alpha_L t - \beta_R^2 + 2\beta_L \beta_R - \beta_L^2)\big]^{1/2} \\ &+ 2\alpha_L \alpha_R (b_L a_R - a_L b_R) t^2 \\ &- (2\beta_L \alpha_R a_R + b_R \alpha_R - 2\alpha_L a_L \beta_R - b_L \alpha_L) t \\ &+ \beta_R - \beta_L \} / (2\alpha_L \alpha_R (a_R - a_L) t - \alpha_R + \alpha_L).\end{aligned}$$

If f is convex then an entropy-satisfying shock has $\beta_L > \beta_R$, so the plus sign is needed to ensure that $\xi = 0$ when $t = 0$. If $\alpha_L = \alpha_R$, then we need to find a more useful expression for the shock trajectory near zero.

To summarize, our new formulation of the shock condition allows us to solve for the shock trajectory *exactly* when u is piecewise linear near the shock and f is piecewise quadratic near the values taken on by u near the shock.

From these observations we can build a numerical method based on piecewise linear approximations that achieves an error of $O(N^{-2})$ when there are $O(N)$ pieces in the approximation. In some sense, this method is the next higher-order generalization of Dafermos's method, in that we solve (2.2) with approximations $\tilde{f}$ to f and $\tilde{u}_0$ to u_0.

We choose a parameter N, set $h = 1/N$, and we assume $\tilde{f}$ is a C^1, piecewise quadratic function with breakpoints at j/N, such that

$$\tilde{f}'(\frac{j}{N}) = f'(\frac{j}{N}), \quad j \in \mathbb{Z}, \quad \text{and} \quad \tilde{f}(0) = f(0).$$

Then for $\xi \in (0, h)$ we have that

$$\begin{aligned} f'(\xi) - \tilde{f}'(\xi) &= f'(\xi) - \frac{(h-\xi) f'(0) + \xi f'(h)}{h} \\ &= \int_0^\xi f''(\eta) - \frac{f'(h) - f'(0)}{h} \, d\eta \\ &= \int_0^\xi f''(\eta) - \frac{1}{h} \int_0^h f''(\zeta) \, d\zeta \, d\eta \\ &= \frac{1}{h} \int_0^h \int_0^\xi f''(\eta) - f''(\zeta) \, d\eta \, d\zeta. \\ &= \frac{1}{h} \int_\xi^h \int_0^\xi f''(\eta) - f''(\zeta) \, d\eta \, d\zeta. \end{aligned}$$

Thus,

$$
\begin{aligned}
|f'(\xi) - \tilde{f}'(\xi)| &\le \frac{\|f'''\|_{L^\infty(\mathbb{R})}}{h} \int_\xi^h \int_0^\xi \zeta - \eta \, d\eta \, d\zeta. \\
&\le \frac{\|f'''\|_{L^\infty(\mathbb{R})}}{h} \frac{1}{2} h(h-\xi)\xi \qquad (4.5) \\
&\le \frac{1}{8} \|f'''\|_{L^\infty(\mathbb{R})} h^2.
\end{aligned}
$$

The same inequality holds for all ξ. Therefore, Theorem 1.1 shows that if $\tilde{u}_0$ is any approximation to u_0 with $\|\tilde{u}_0\|_{\mathrm{BV}(\mathbb{R})} \le \|u_0\|_{\mathrm{BV}(\mathbb{R})}$, then

$$
\|u(\,\cdot\,,t) - \tilde{u}(\,\cdot\,,t)\|_{L^1(\mathbb{R})} \le \|u_0 - \tilde{u}_0\|_{L^1(\mathbb{R})} + \frac{t}{8} \|f'''\|_{L^\infty(\mathbb{R})} \|u_0\|_{\mathrm{BV}(\mathbb{R})} \frac{1}{N^2}.
$$

For example, if u_0 is C^2 and has support in $[0,1]$, then the continuous, piecewise linear interpolant at the points j/N, $j = 0, \ldots, N$, satisfies

$$
\|u_0 - \tilde{u}_0\|_{L^1(\mathbb{R})} \le \|u_0 - \tilde{u}_0\|_{L^\infty(\mathbb{R})} \le \frac{1}{8} \|u_0''\|_{L^\infty(\mathbb{R})} \frac{1}{N^2}.
$$

In this case we have the error bound

$$
\|u(\,\cdot\,,t) - \tilde{u}(\,\cdot\,,t)\|_{L^1(\mathbb{R})} \le \frac{1}{8} [\|u_0''\|_{L^\infty(\mathbb{R})} + t\|f'''\|_{L^\infty(\mathbb{R})} \|u_0\|_{\mathrm{BV}(\mathbb{R})}] \frac{1}{N^2}.
$$

Just as for Dafermos's method, we must now characterize the form of the solution $\tilde{u}(x,t)$ and bound the computational complexity of an algorithm to calculate it.

TO BE CONTINUED