Chapter One
Scalar Conservation Laws

This chapter introduces the basic concepts of the present thesis. We review the general theory of linear and nonlinear scalar conservation laws and introduce the fundamental notations of weak solutions and Rankine-Hugoniot jump conditions. Then, we introduce the entropy condition to pick out the physically relevant solution of equation (1.1.1) below.

Finally, we introduce an application (traffic flow) on a real-life problem relevant to the scalar conservation law.

1.1 Overview

We overview the basic details concerning the simplified model of scalar conservation laws. This means that we are looking for the solution $u(x,t)$ of the Cauchy problem for a single hyperbolic partial differential equation of the type

$$\frac{\partial}{\partial t}u+\frac{\partial}{\partial x}f(u)=0 \qquad x\in R,\quad t>0 \tag{1.1.1}$$

$$u(x,0)=u_0(x) \qquad x\in R \tag{1.1.2}$$

The solution $u(x,t)$ is sought for all nonnegative time values, as a function of the space variable $x\in R$. The function $f(u)$ is assumed to be smooth (namely, continuously differentiable at least as many time as needed in the analysis).

1.2 Conservation Law

The term conservation law stems from the following argument: Integrating equation (1.1.1) over the rectangle $0 \le t \le T$, $x_1 \le x \le x_2$ one gets,

$$\int_{x_1}^{x_2} u(x,T)dx - \int_{x_1}^{x_2} u_0(x)dx = -\int_0^T f(u(x_2,t))dt + \int_0^T f(u(x_1,t))dt \qquad (1.2.1)$$

Proof:

Integrating eq. (1.1.1) over the region R as in the Fig. (1.2.1)

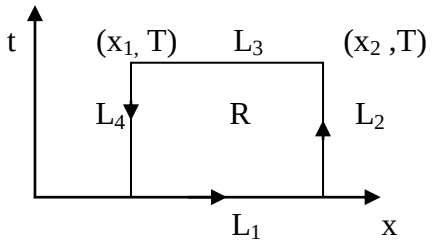


Figure 1.2.1

By Green theorem, we get

$$0 = \iint_R f_x - (-u_t)dxdt = \int_{bdy} -udx + \int_{bdy} f(u)dt$$

$$on \quad L_1: \quad \int_{L_1} = -\int_{x_1}^{x_2} udx + \int f(u)dt = -\int_{x_1}^{x_2} u(x,0)dx$$

$$on \quad L_2: \quad \int_{L_2} = \int_0^T -udx + \int_0^T f(u)dt = \int_0^T f(u(x_2,t))dt$$

$$on \quad L_3: \quad \int_{L_3} = \int_{x_2}^{x_1} -udx + \int f(u)dt = -\int_{x_2}^{x_1} u(x,T)dx$$

$$on \quad L_4: \quad \int_{L_4} = -\int_T^0 udx + \int_T^0 f(u)dt = \int_T^0 f(u(x_1,t))dt$$

Hence,

$$0 = \iint_R [u_t + f(u)]dxdt = \int_{L_1} + \int_{L_2} + \int_{L_3} + \int_{L_4}$$

$$\int_{x_1}^{x_2} u(x,T)dx - \int_{x_1}^{x_2} u_0(x)dx = -\int_0^T f(u(x_2,t))dt + \int_0^T f(u(x_1,t))dt.$$

Thinking of $u(x,t)$ in eq. (1.2.1) as mass density per unit length the integral $\int_{x_1}^{x_2} u(x,t)dx$ expresses the total mass in $[x_1, x_2]$ at time t, while $\int_0^T f(u(x,t))dt$ for any fixed x, can be interpreted as mass flux to the right, at the point x, over the time interval [0, T]. Thus, eq. (1.2.1) may be viewed as a balance equation stating that the gain in total mass in $[x_1, x_2]$ equals the net flux into the interval, through its boundary (end points) x_1 *and* x_2 .

Example (1.2.1):

The simplest conservation law (or nonlinear hyperbolic differential equation) is the Burgers' equation

$$\frac{\partial}{\partial t}u(x,t) + \frac{\partial}{\partial x}(\frac{u^2}{2}) = 0 \qquad in\ R \times R^+ \qquad (1.2.2)$$

1.3 Discontinuous Solution

We begin with the simplest example which leads to discontinuous solution. Consider the single equation (1.2.2) Burgers' equation,

$$u_t + (\frac{u^2}{2})_x = 0$$

Which can be written in the form

$$u_t + uu_x = 0 \qquad (1.3.1)$$

This equation has the rather remarkable property that the only C^1 functions which satisfy this equation for $t > 0,$ are those which are monotonically non decreasing in x, for each fixed $t > 0.$

Suppose (1.3.1) has a smooth solution u $\in C^1$(Rx [0, T]), consider the solution $x(t) \in C^1(0,T) \bigcap C^0[0,T]$ of the initial value problem for the following ordinary differential equation:

$$x'(t) = \frac{dx}{dt} = u(x(t),t), \quad in\ (0,T),\ x(0) = a \tag{1.3.2}$$

For a given constant $a \in R.$ then we have

$$\frac{d}{dt}u(x(t),t) = \frac{\partial u}{\partial t}\frac{dt}{dt} + \frac{\partial u}{\partial x}\frac{dx}{dt} = u_t + uu_x = 0.$$

For all $t \in (0,T).$ this means that $u(x(t),t)$ is constant for all $t \in [0,T],$ or u is constant along the curves $\{(x(t),t) : t \in [0,T]\}.$ these curves are called characteristics. From (1.3.2) we also get that

$$x'(t) = u(x(t),t) = \text{ constant} \tag{1.3.3}$$

For all $t \in (0,T).$ this means that the characteristics are straight lines. Altogether we have shown that the solution u is constant along straight lines. The slopes of these lines are given by (1.3.2):

$$x'(t) = u(x(t),t) = u(x(0),0),$$

(i.e. by the initial values for u). Consider now the initial value problem

$$u_t + (\frac{u^2}{2})_x = 0 \quad in\ R \times R^+,$$
$$u(x,0) = u_0(x) \quad in\ R,$$

Where u_0 is given as in figure (1.3.1)

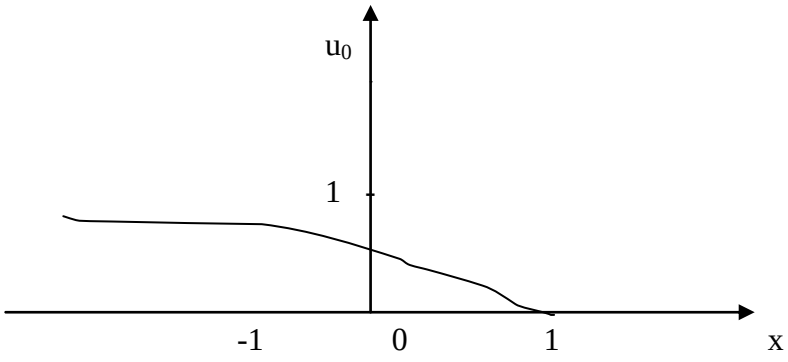


Figure 1.3.1

Then we obtain that u is constant along the characteristic curve (x (t), t), with

$$x'(t) = u(x(0),0) = u_0(x(0))) = \begin{cases} 1 & if \ \ x(0) \le -1 \\ 0 & if \ \ x(0) \ge 1 \end{cases}$$

$$u(x(t),t) = u(x(0),0) = u_0(x(0)) = \begin{cases} 1 & if \ \ x(0) \le -1 \\ 0 & if \ \ x(0) \ge 1 \end{cases}$$

This means, if T is sufficiently large and finite, that the characteristics can meet each other (see figure 1.3.2), and therefore u cannot be a classical solution up to this time. Then a new definition of solutions for conservation laws of type (1.3.1) will be introduced, namely (weak solutions). This new type of solution can have discontinuities, as we shall see in the coming section.

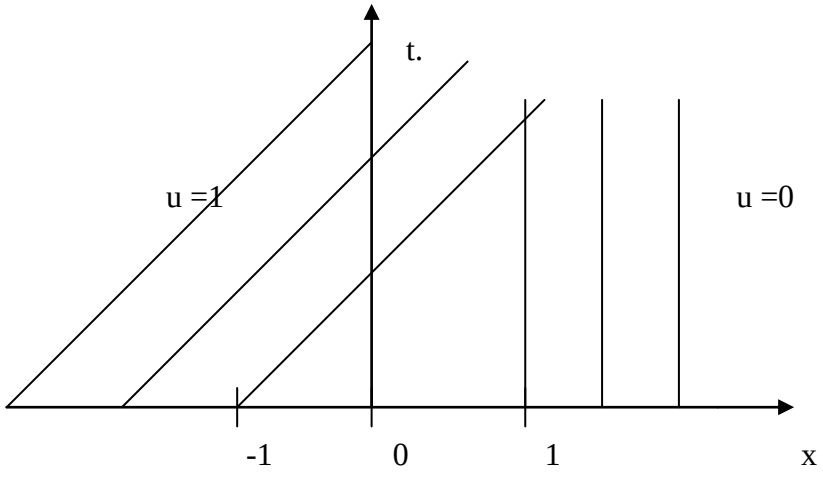


Figure 1.3.2

More generally, consider the initial value problem for the scalar u,

$$u_t + f(u)_x = 0, \quad t > 0$$

$$u(x,0) = u_0(x), \quad x \in R \tag{1.3.4}$$

We can write the equation as $u_t + f'(u)u_x = 0$, and consider the characteristics

$$\frac{dt}{ds} = 1, \qquad \frac{dx}{ds} = f'(u)$$

Along such a curve,

$$\frac{du}{ds} = u_t \frac{dt}{ds} + u_x \frac{dx}{ds} = u_t + f'(u)u_x = 0$$

Thus, again u(x, t) is constant along the characteristics. Since the slope of the characteristics is $\frac{ds}{dx} = \frac{1}{f'(u)}$, (u is constant so $f'(u) = cons\tan t$), so the characteristics are straight lines, having slope determined by their values at $t = 0$: i.e., by $u_0(x)$. so, if there are points $x_1 < x_2$ with

$$m_1 = \frac{1}{f'(u_0(x_1))} < \frac{1}{f'(u_0(x_2)} = m_2,$$

Then the characteristics starting at $(x_1,0)$ and $(x_2,0)$ will cross in $t > 0$;

(see fig 1.3.3). Along l_i, $u(x,t) = u_0(x_i), i = 1,2.$ thus at p the solution must be discontinuous and a shock occurs. Note that this conclusion is independent of the smoothness properties of f *and* u_0; they can each be analytic, and still we cannot obtain a globally defined solution. The phenomenon is a purely nonlinear one.

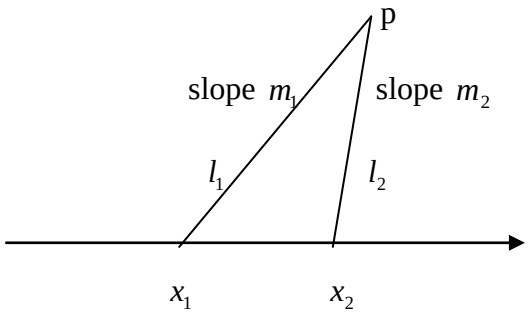


Figure 1.3.3

We can be a bit more explicit and see analytically that discontinuities must form if u_0' is negative at some point. Thus, consider (1.3.4), and assume that $f'' > 0.$ since the characteristics are straight lines, if (x,t) is any point with $t > 0$ we let $(x^*,0)$ denote the unique point on the x-axis which lies on the characteristic through (x, t) since u is constant along characteristics, and $tf'(x^*) = x - x^*,$ we see that u must implicitly be given by

$$u(x,t) = u_0(x - tf'(u(x,t))$$

Now if u_0 is a differentiable function, then we can invoke the implicit function theorem and solve this last equation for u , provided that t is sufficiently small. We find

$$u_t = -\frac{f'(u)u_0'}{1+u_0'f''(u)t} \text{ and } u_x = \frac{u_0'}{1+u_0'f''(u)t}$$

now if $u_0'(x) \geq 0$ for all x, then these formulas show that u_x, u_t stays bounded for all t >0, and the solution u exist for all time. On the other hand, if $u_0'(x) < 0$ at some point , both u_x, u_t becomes unbounded when $1 + u_0'f''(u)t$ tends to zero.

Thus, if we adhere to the notion that a solution must be smooth, then we must content ourselves with solutions which exist for only a finite time.

1.4 The Mathematical Model, Euler Equations of Gas Dynamics

To see how the following conservation laws arise from physical principles

$$
\begin{aligned}
&(i)\quad \rho_t + (\rho u)_x = 0, \quad \text{where } \rho_t = \frac{\partial \rho}{\partial t}. && (\textit{conservation of mass})\\
&(ii)\quad (\rho u)_t + (\rho u^2 + p)_x = 0, && (\textit{conservation of momentum}) \quad (1.4.1)\\
&(iii)\quad [\rho(\frac{u^2}{2} + e)]_t + [\rho u(\frac{u^2}{2} + i)]_x = 0, && (\textit{conservation of energy})
\end{aligned}
$$

Equation (1.4.1) is known as the Euler equations for fluid dynamics. Where ρ is the density of the fluid, u the velocity, p the pressure, e the internal energy, and $i = e + \frac{p}{\rho}$.

We will begin by deriving the equation for conservation of mass in a one-dimensional gas dynamics problem, for example flow in a tube where properties of the gas such as density and velocity are assumed to be constant across each cross section of the tube. Let x represent the distance along the tube and let $\rho(x,t)$ be the density of the gas at point x and t this density is defined in such a way that the total mass of gas in any given section from x_1 to x_2, say, is given by the integral of the density:

$$\text{mass in } [x_1, x_2] \text{ at time t} = \int_{x_1}^{x_2} \rho(x,t)dx$$

If we assume that the walls of the tube are impermeable and that mass is neither created nor destroyed, then the mass in this one section can change only because of gas flowing across the endpoint $x_1, or\ x_2$.

Now let $\nu(x,t)$ be the velocity of the gas at the point x at time t. then the rate of flow, or flux of gas past this point is given by

$$\text{Mass flux at(x, t)} = \rho(x,t)\nu(x,t).$$

By our comments above, the rate of change of mass in $[x_1, x_2]$ is given by the difference in fluxes at x_1, *and* x_2 :

$$\frac{d}{dt}\int_{x_1}^{x_2} \rho(x,t)dx = \rho(x_1,t)\nu(x_1,t) - \rho(x_2,t)\nu(x_2,t).$$

This is one integral form of the conservation law. Another form is obtained by integrating this in time from t_1 *to* t_2, giving an expression for the mass in $[x_1, x_2]$ at time $t_2 > t_1$ in terms of the mass at time t_1 and the total (integrated) flux at each boundary during this time period:

$$\int_{x_1}^{x_2} \rho(x,t_2)dx = \int_{x_1}^{x_2} \rho(x,t_1)dx + \int_{t_1}^{t_2} \rho(x_1,t)\nu(x_1,t)dt - \int_{t_1}^{t_2} \rho(x_2,t)\nu(x_2,t)dt \quad (1.4.2)$$

to derive the differential form of the conservation law, we must now assume that $\rho(x,t)$ and $\nu(x,t)$ are differentiable functions. Then using

$$\rho(x,t_2) - \rho(x,t_1) = \int_{t_1}^{t_2} \frac{\partial}{\partial t}\rho(x,t)dt \quad (1.4.3)$$

and

$$\rho(x_2,t)\nu(x_2,t) - \rho(x_1,t)\nu(x_1,t) = \int_{x_1}^{x_2} \frac{\partial}{\partial x}((\rho(x,t)\nu(x,t))dx \quad (1.4.4)$$

in (1.4.2) gives

$$\int_{t_1}^{t_2}\int_{x_1}^{x_2} \left\{ \frac{\partial}{\partial t}\rho(x,t) + \frac{\partial}{\partial x}(\rho(x,t)\nu(x,t)) \right\} dxdt = 0 \quad (1.4.5)$$

since this must hold for any section $[x_1, x_2]$ and over any time interval $[t_1, t_2]$, we conclude that in fact the integrand in (1.4.5) must be identically zero, i.e.,

$$\rho_t + (\rho\nu)_x = 0 \qquad \text{Conservation of mass} \quad (1.4.6)$$

the conservation law (1.4.6) can be solved in isolation only if the velocity $\nu(x,t)$ is known as a function of $\rho(x,t)$. if it is, then $\rho\nu$ is a function of ρ alone, say

$\rho v = f(\rho)$, and the conservation of mass equation (1.4.6) becomes a scalar conservation law for ρ,

$$\rho_t + f(\rho)_x = 0 \tag{1.4.7}$$

Moreover, if the velocity is constant, v(x, t) =a, then $f(\rho) = a\rho$ and (1.4.6) reduces to

$$\rho_t + a\rho_x = 0 \tag{1.4.8}$$

this equation is known as the linear advection equation or sometimes the one-way wave equation . if this equation is solved for $t \geq 0$ with the initial data

$$\rho(x,0) = \rho_0(x) \qquad -\infty < x < \infty \tag{1.4.9}$$

Then it easy to check (assuming ρ_0 is differentiable) that the solution is simply

$$\rho(x,t) = \rho_0(x - at) \tag{1.4.10}$$

We can also define flux function in a way such that

$$f(\rho, \rho_x) = a\rho - D\rho_{x,} \quad D > 0 \tag{1.4.11}$$

and the conservation law from (1.4.7) becomes

$$\rho_t + (a\rho - D\rho_x)_x = 0 \tag{1.4.12}$$

or assuming D is constant,

$$\rho_t + a\rho_x = D\rho_{xx} \tag{1.4.13}$$

Equation (1.4.13) is called the advection-diffusion equation and $(-D\rho_x)$ is called diffusive flux. This flux is determined by “Fourier’s law of heat conduction”(heat diffuses in much the same way as the chemical concentration). The advection-diffusion equation (1.4.13) is a parabolic equation while (1.4.7) is hyperbolic. One major difference is that (1.4.13) always has smooth solutions for t >0 even if the

initial data $\rho_0(x)$ is discontinuous. We can view (1.4.7) as an approximation to (1.4.13) valid for D very small, but we may need to consider the effect of D in order to properly interpret discontinuous solution to (1.4.7).

1.5 The Linear Advection Equation

We first consider the linear advection equation, derived before, which we now write as

$$u_t + au_x = 0 \qquad (1.5.1)$$

The Cauchy problem is defined by this equation on the domain $-\infty < x < \infty, t \geq 0$ together with the initial condition

$$u(x,0) = u_0(x) \qquad (1.5.2)$$

As noted previously, the solution is simply

$$u(x,t) = u_0(x - at) \qquad (1.5.3)$$

for $t \geq 0.$ as time evolves , the initial data simply propagates unchanged to the right (if $a > 0)$ or left (if $a < 0)$ with velocity $a.$ The solution u(x, t) constant along each ray $x - at = x_0,$ which are known as the characteristics of the equation. (see fig.1.5.1 for the case $a > 0$)

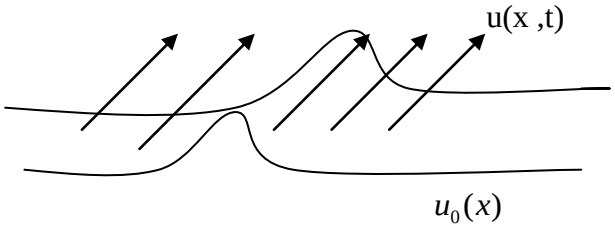


Figure 1.5.1 characteristics and solution for the advection equation.

Note that the characteristics are curves in the x-t plane satisfying the ordinary differential equations $x'(t) = a,\ x(0) = x_0.$ if we differentiate u(x, t) along one of these curves to find the rate of change of u along the characteristic, we find that

$$\frac{d}{dt}u(x(t),t) = \frac{\partial u}{\partial t}\frac{dt}{dt} + \frac{\partial u}{\partial x}\frac{dx}{dt}$$

$$= u_t + au_x \qquad (1.5.4)$$

$$= 0$$

Confirming that u is constant along these characteristics.

More generally, we might consider a variable coefficient advection equation of the form

$$u_t + (a(x)u)_x = 0, \qquad (1.5.5)$$

Where $a(x)$ is smooth function. Recalling the derivation of the advection equation before, this models the evolution of a chemical concentration u(x, t) in a stream with variable velocity $a(x)$.

We can write (1.5.5) as

$$u_t + a(x)u_x = -a'(x)u \qquad (1.5.6)$$

$$(\frac{\partial}{\partial t} + a(x)\frac{\partial}{\partial x})u(x,t) = -a'(x)u(x,t). \qquad (1.5.7)$$

It follows that the evolution of u along any curve x(t) satisfying

$$x'(t) = a(x(t))$$
$$x(0) = x_0 \tag{1.5.8}$$

satisfies a simple ordinary differential equation (ODE):

$$\frac{d}{dt}u(x(t),t) = -a'(x(t))u(x(t),t) \tag{1.5.9}$$

The curves determined by (1.5.8) are again called characteristics. In this case the solution u(x, t) is not constant along these curves, but can be easily determined by solving two sets of ODEs $\left(\frac{dt}{ds} = 1, \frac{dx}{ds} = a(x)\right)$.

Thus, if $u_0(x)$ is a smooth function, say $u_0 \in C^k(-\infty,\infty)$, then the solution u(x, t) is equally smooth in space and time, $u \in C^k((-\infty,\infty)\times(0,\infty))$.

Remark: Domain of dependence

Note that solution to the linear advections (1.5.1) and (1.5.5) have the following property: the solution u(x, t) at any point $(\bar{x},\bar{t})$ depends on the initial data u_0 only at a single point, namely the point $\bar{x}_0$ such that $(\bar{x},\bar{t})$ lies on the characteristic through $\bar{x}_0$. We could change the initial data at any points other than $\bar{x}_0$ with out effecting the solution $u(\bar{x},\bar{t})$. The set $D(\bar{x},\bar{t}) = \{\bar{x}_0\}$ is called the domain of dependence of the point $(\bar{x},\bar{t})$. Here this domain consists of a single point. Conversely, initial data at any given point x_0 can influence the solution only within some cone $\{x : |x - x_0| \le a_{\max} t\}$ of the x-t plane. This region is called the range of influence of the point x_0. See Figure 1.5.2 for an illustration.

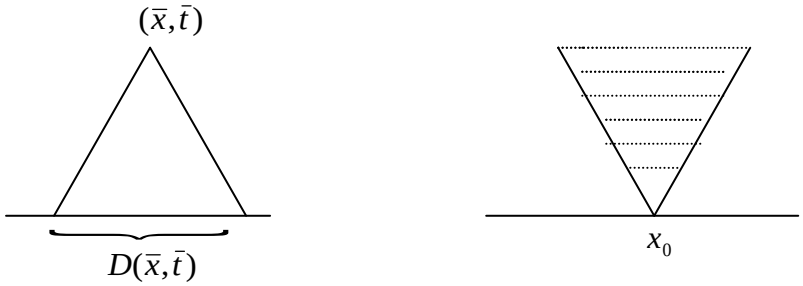


Figure 1.5.2: Domain of dependence and range of influence.

Remarks: Non-Smooth Data

In the manipulations performed above, we have assumed differentiability of u(x, t). however, from our observation that the solution along a characteristics curve depends only on the one value $u_0(x_0)$, it is clear that spatial smoothness is not required for this construction of u(x, t) from $u_0(x)$. we can thus define a solution to the PDE even if $u_0(x)$ is not a smooth function. Note that if $u_0(x)$ has a singularity at some point x_0 (a discontinuity in u_0 or some derivative), then the resulting u(x, t) will have a singularity of the same order along the characteristic curve through x_0, but will remain smooth along characteristics through smooth portions of the data. This is a fundamental property of linear hyperbolic equations: singularities propagate only along characteristics.

1.6 Burgers' Equation

Again consider the nonlinear scalar equation

$$u_t + f(u)_x = 0 \qquad (1.6.1)$$

Where f (u) is a nonlinear function of u .We will assume for the most part that f (u) is a convex function, $f''(u) > 0$ for all u. The convexity assumption corresponds to a "genuine nonlinearity" assumption for system of equations that holds in many important cases, such as Euler equations.

By far the most famous model problem in this field is Burgers' equation, in which $f(u) = \frac{1}{2}u^2$, so (1.6.1) becomes

$$u_t + uu_x = 0 \qquad (1.6.2)$$

Actually this should be called the "inviscid Burgers' equation", since the equation studied by Burgers also includes a viscous term:

$$u_t + uu_x = \varepsilon u_{xx} \qquad (1.6.3)$$

Example 1.6.1

We want to solve the following problem

$$u_t + uu_x = 0$$

With the initial condition

$$u(x,0) = u_0(x) = \begin{cases} 1 \ , & x \le 0 \\ 1-x, & 0 \le x \le 1 \\ 0 \ , & x > 1 \end{cases}$$

By using the method of characteristics, we can solve this problem up to the time when the characteristic intersect. We already know that the characteristic passing through the point $(x_0,0)$ is given by

$$\frac{x-x_0}{t}=\frac{dx}{dt}=u(x,t)=u_0(x_0) \quad \Rightarrow \quad x=x(x_0,t)=x_0+tu_0(x_0)$$

so that

$$x=\begin{cases} x_0+t, & x_0\le 0 \\ x_0+t(1-x_0), & 0\le x_0\le 1 \\ x_0, & x_0\ge 1 \end{cases}$$

For $t<1$, the characteristic do not intersect (see figure 1.6.1). hence , given a point (x, t) with t < 1 , we draw the (backward) characteristic passing through this point and we determine the corresponding point x_0.

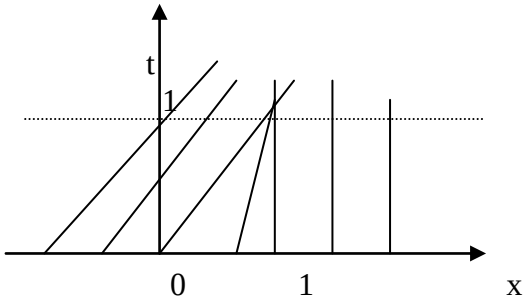


Figure 1.6.1

$$x_0=\begin{cases} x-t & if \quad x\le t<1 \\ (x-t)/(1-t) & if \quad t\le x\le 1 \\ x & if \quad x\ge 1,\ t<1 \end{cases}$$

We obtain the following continuous solution for t < 1

$$u(x,t)=u_0(x_0)=\begin{cases} 1, & x\le t \\ (1-x)/(1-t), & t\le x\le 1 \\ 0, & x\ge 1 \end{cases}$$

Since in general we can only prove local existence , we have to generalize the definition of solution of conservation laws.

Remark: Shock formulation

We consider again Burgers' equation $(u_t + uu_x = 0)$ with the initial data $u(x,0) = u_0(x)$. It's easy to show that $u(x,t) = u_0(x-ut)$ to Burgers' equation. Since $u(x,t) = u(\xi,0) = u_0(\xi) = \frac{dx}{dt} =$ constant, where $\xi = x - ut$.

For each (x, t), the characteristic line passes through $(\xi,0)$ satisfies

$$\frac{dx}{dt} = \frac{x-\xi}{t} = u_0(\xi)$$

or

$$x = \xi + u(\xi)t \qquad (1.6.4)$$

For large t the equation 1.6.4 may not have a unique solution. This happens when the characteristic cross, as will eventually happen if $u_x(x,0)$ is negative at any point. At the time T_b where the characteristic first cross, the function u(x, t) has infinite slope the wave "breaks" and a shock forms. Figure 1.6.2 shows an extreme example where the initial data is piecewise linear and many characteristics comes together at once. More generally an infinite slope in the solution may appear first at just one point x, corresponding via (1.6.4) to the point ξ where the slope of the initial data is most negative. At this point the wave is said to "break" by analogy with waves on a beach. Mathematically speaking, where a shock wave occurs, the solution u(x, t) has a jump discontinuity. This usually occurs along a curve in the x-t plane.

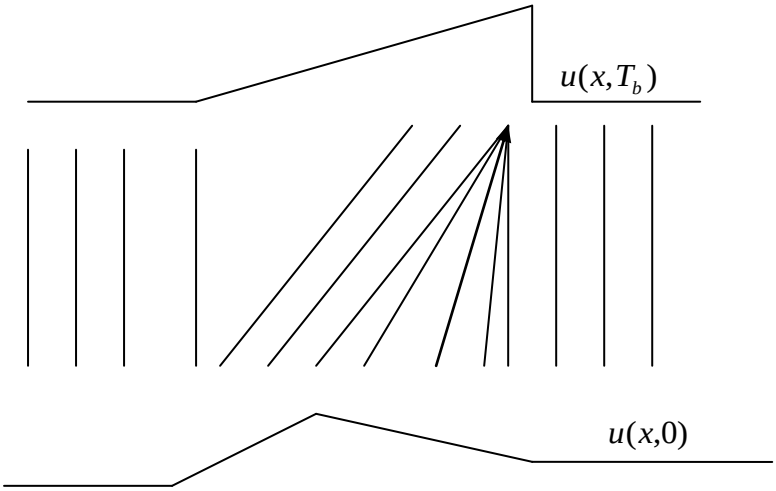


Figure 1.6.2 *shock formation in Burgers' eq.*

If we solve Burgers' equation with smooth initial data $u_0(x)$ for which $u_0'(x)$ is somewhere negative, then the wave will break at time

$$T_b = \frac{-1}{\min u_0'(x)}. \qquad (1.6.5)$$

1.7 Weak Solution

A Natural way to define a generalized solution of the inviscid equation $(u_t + uu_x = 0)$ that does not require differentiability is to go back to the integral form of the conservation law , and say that u(x, t) is a generalized solution if (1.4.5) is satisfied for all x_1, x_2, t_1, t_2.

There is another approach that results in a different integral formulation that is more convenient to work with. This is a mathematical technique that can be applied more generally to write a differential equation in a form where less smoothness is required to define a "solution". The basic idea is to take the PDE, multiply by a smooth "test function", integrate one or more times over some domain, and then use integration by

part to move derivatives off the function u and onto the smooth test function. The result is an equation involving fewer derivatives on u, and hence requiring less smoothness.

In our case we will use test function $\Phi \in C_0^1(R \times R)$. Here C_0^1 is the space of function that are continuously differentiable with "compact support". The latter requirement means that $\Phi(x,t)$ is identically zero outside of bounded set, and so the support of the function lies in a compact set.

If we multiply $u_t + f_x = 0$ *by* $\Phi(x,t)$ and then integrate over space and time, we obtain

$$\int_0^\infty \int_{-\infty}^{\infty} [\Phi u_t + \Phi f(u)_x] dx dt = 0. \qquad (1.7.1)$$

Now integrate by parts, yielding

$$\int_0^\infty \int_{-\infty}^{\infty} [\Phi_t u + \Phi_x f(u)] dx dt = -\int_{-\infty}^{\infty} \Phi(x,0) u(x,0) dx. \qquad (1.7.2)$$

Note that nearly all the boundary terms which normally arise through integration by parts drop out due to the requirement that Φ have compact support, and hence vanishes at infinity. The remaining boundary term brings in the initial condition of the PDE, which must still play a role in our weak formulation.

Definition 1.7.1

The function u(x, t) is called a weak solution of the conservation law if (1.7.2) holds for all function $\Phi \in C_0^1(R \times R)$.

The advantage of this formulation over the original integral form (1.4.2) is that the integration in (1.7.2) is over a fixed domain, all of $R \times R.^{+}$ the integral form (1.4.2) is over an arbitrary rectangle, and to check that u(x, t) is a solution we must verify that this holds for all choices of x_1, x_2, t_1 and t_2. Of course, our new form (1.7.2) has a similar feature, we must check that it holds for all $\Phi \in C_0^1,$ but this turns out to be an easier task.

Mathematically, the two integral forms are equivalent and we should rightly expect a more direct connection between the two that does not involve the differential equation.

This can be achieved by considering special test functions $\Phi(x,t)$ with the property that

$$\Phi(x,t) = \begin{cases} 1 & for\ (x,t) \in [x_1, x_2] \times [t_1, t_2] \\ 0 & for\ (x,t) \notin [x_1 + \in, x_2 + \in] \times [t_1 + \in, t_2 + \in] \end{cases} \qquad (1.7.3)$$

proof

Equation 1.7.1 can be written

$$0 = \int_0^{\infty} \int_{-\infty}^{\infty} [\Phi u_t + \Phi f(u)_x] dx dt = \int_0^{\infty} \int_{-\infty}^{\infty} \Phi u_t dx dt + \int_0^{\infty} \int_{-\infty}^{\infty} \Phi f(u)_x dx dt \qquad (*)$$

$$\int_{-\infty}^{\infty} \int_0^{\infty} \Phi u_t dt dx = \int_{-\infty}^{\infty} [\int_0^{\infty} \Phi u_t dt] dx = \int_{-\infty}^{\infty} ([\Phi u]_0^{\infty}) dx - \int_{-\infty}^{\infty} (\int_0^{\infty} \Phi_t u dt) dx$$

$$= \int_{-\infty}^{\infty} -\Phi(x,0) u(x,0) - \int_{-\infty}^{\infty} \int_0^{\infty} \Phi_t u dt dx, \quad \text{since} \quad \Phi(x,\infty) = 0 \qquad (1)$$

also

$$\int_0^\infty \int_{-\infty}^{\infty} \Phi f_x dx dt = -\int_0^\infty \int_{-\infty}^{\infty} \Phi_x f(u) dx dt \,, \qquad \text{since} \quad \Phi(\pm\infty, t) = 0 \qquad (2)$$

Substitute (1) and (2) in (*) we obtain eq. (1.7.2)

Unfortunately, weak solution is often not unique, and so an additional problem is often to identify which weak solution is the physically correct solution. There are other conditions one can impose on weak solutions that are easier to check and will also pick out the correct solution. These are usually called entropy conditions by analogy with the gas dynamics case. The solution is also called the entropy solution.

1.8 The Riemann Problem

The conservation law together with piecewise constant data having a single discontinuity is known as the Riemann problem. As an example, consider Burgers' equation $u_t + uu_x = 0$ with piecewise constant data

$$u(x,0) = \begin{cases} u_l & x < 0 \\ u_r & x > 0 \end{cases} \qquad (1.8.1)$$

The form of the solution depends on the relation between u_l *and* u_r.

Case I. $u_l > u_r$.

In this case there is a unique weak solution,

$$u(x,t) = \begin{cases} u_l & x < st \\ u_r & x > st \end{cases} \qquad (1.8.2)$$

where

$$s = \frac{u_l + u_r}{2} \qquad (1.8.3)$$

is the shock speed, the speed at which the discontinuity travels. A general expression for the shock speed will be derived below. Note that characteristics in each of the regions where u is constant go into the shock (see fig.1.8.1) as time advances.

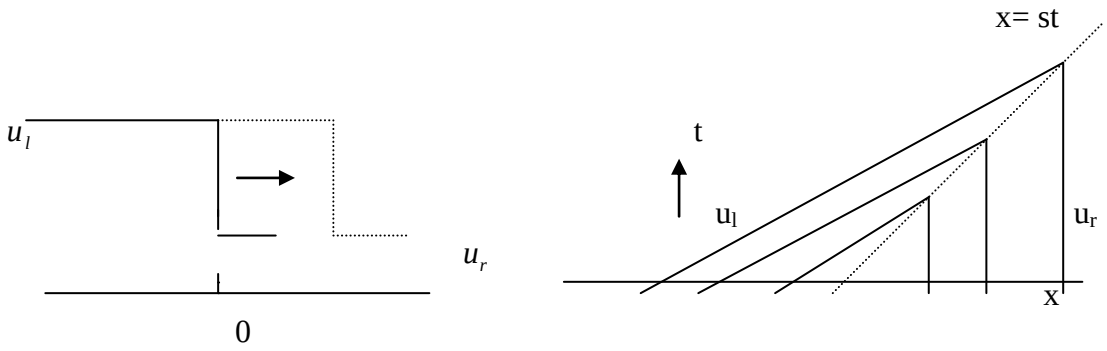


Figure 1.8.1. shockwave

Case II. $u_l < u_r$

In this case there are infinitely many weak solutions. One of these is again (1.8.2), (1.8.3) in which the discontinuity propagates with speed *s*. Note that characteristics now go out of the shock (fig.1.8.2) and that this solution is not stable to perturbations. If the data is smeared out slightly, the solution changes completely.

Another weak solution is the rarefaction wave

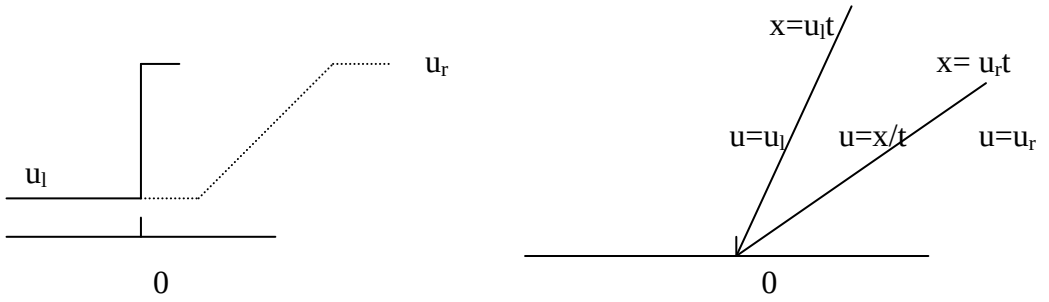


Fig. 1.8.2. Rarefaction wave

$$u(x,t)=\begin{cases} u_l & x < u_l t \\ \dfrac{x}{t} & u_l t \le x \le u_r t \\ u_r & x > u_r t \end{cases} \tag{1.8.4}$$

This solution is stable to perturbations. There are infinitely many other weak solutions of equation $u_t + uu_x = 0$ when $u_l < u_r$. for example,

$$u(x,t)=\begin{cases} u_l & x < s_m t \\ u_m & s_m t \le x \le u_m t \\ x/t & u_m t \le x \le u_r t \\ u_r & x > u_r t \end{cases}$$

is a weak solution for any u_m with $u_l \le u_m \le u_r$ and $s_m = u_l + u_m / 2.$

Another example is the general convex problem

$$u_t + f(u)_x = 0 \tag{1.8.5}$$

with the data (1.8.1) and $u_l < u_r$, since $f'' \ge 0, \textit{then}\ f'$ is an increasing function $f'(u_l) < f'(u_r)$ the rarefaction wave solution is given by

$$u(x,t)=\begin{cases} u_l & x < f'(u_l)t \\ v(x/t) & f'(u_l)t \le x \le f'(u_r)t \\ u_r & x > f'(u_r)t \end{cases} \tag{1.8.6}$$

where $v(\xi)$ is the solution to $f'(v(\xi)) = \xi.$

1.9 Shock Speed

The propagating shock solution (1.8.2) is a weak solution to Burgers' equation only if the speed of propagation is given by (1.8.3). the same discontinuity propagating at a different speed would not be a weak solution.

A gain, the form (1.8.5) is the differential form of the conservation laws, which holds in the usual sense only where u is smooth. More generally, the integral form for a system of equations says that

$$\frac{d}{dt}\int_{x_1}^{x_2} u(x,t)dx = f(u(x_1,t)) - f(u(x_2,t)) \tag{1.9.1}$$

for all $x_1, x_2, t.$

The speed of propagation can be determined by conservation. If M is large compared to (st) then by (1.9.1), $\int_{-M}^{M} u(x,t)dx$ must increase at the rate

$$\begin{aligned}\frac{d}{dt}\int_{-M}^{M} u(x,t)dx &= f(u_l) - f(u_r) \\ &= \frac{1}{2}(u_l^2 - u_r^2)\end{aligned} \tag{1.9.2}$$

for Burgers' equation. On the other hand, the solution (1.8.2) clearly has

$$\int_{-M}^{M} u(x,t)dx = \int_{-M}^{st} u_l dx + \int_{st}^{M} u_r dx = (M+st)u_l + (M-st)u_r$$

so that

$$\frac{d}{dt}\int_{-M}^{M} u(x,t)dx = s(u_l - u_r) \tag{1.9.3}$$

comparing (1.9.2) and (1.9.3) gives (1.8.3)

More generally, for arbitrary flux function f(x) this same argument gives the following relation between the shock speed s and the states $u_l \quad and \quad u_r,$ called the Rankine-Hugoniot jump condition:

$$f(u_l) - f(u_r) = s(u_l - u_r) \tag{1.9.4}$$

For scalar problems this gives simply

$$s = \frac{f(u_l) - f(u_r)}{u_l - u_r} = \frac{[f]}{[u]} \tag{1.9.5}$$

where [.] indicates the jump in some quantity across the discontinuity. Note that any jump is allowed, provided the speed is related via (1.9.5).

the Rankine-Hugoniot (R-H) conditions (1.9.5) hold more generally across any propagating shock, where now u_l *and* u_r denote the values immediately to the left and right of the discontinuity and s is the corresponding instantaneous speed, which varies along with u_l *and* u_r. to verify that the R-H condition must be instantaneously satisfied when u_l *and* u_r vary, we apply the same conservation argument as before but now to a small rectangle as shown in figure (1.9.1), with $x_2 = x_1 + \Delta x$ *and* $t_2 = t_1 + \Delta t.$ assuming that u is smoothly varying on each side of the shock, and that the shock speed s(t) is consequently also smoothly varying, we have the following relation between Δx *and* Δt :

$$\Delta x = s(t_1)\Delta t + o(\Delta t^2) \tag{1.9.6}$$

From the integral form of the conservation law, we have

$$\int_{x_1}^{x_1+\Delta x} u(x, t_1 + \Delta t)dx = \int_{x_1}^{x_1+\Delta x} u(x, t_1)dx + \int_{t_1}^{t_1+\Delta t} f(u(x_1, t))dt - \int_{t_1}^{t_1+\Delta t} f(u(x_1 + \Delta x, t))dt \tag{1.9.7}$$

In the triangular portion of the infinitesimal rectangle that lies to the left of the shock $u(x,t) = u_l(t_1) + o(\Delta t)$, while in the complementary triangle, $u(x,t) = u_r(t_1) + o(\Delta t)$.

Using this in (1.9.7) gives

$$\Delta x u_l = \Delta x u_r + \Delta t f(u_l) - \Delta t f(u_r) + o(\Delta t^2)$$

Using the relation (1.9.6) in the above equation and then dividing by Δt gives

$$s(u_l - u_r) = f(u_l) - f(u_r) + o(\Delta t)$$

where s, u_l, *and* u_r are all evaluated at t_1. Letting $\Delta t \to 0$ gives the R-H condition (1.9.4)

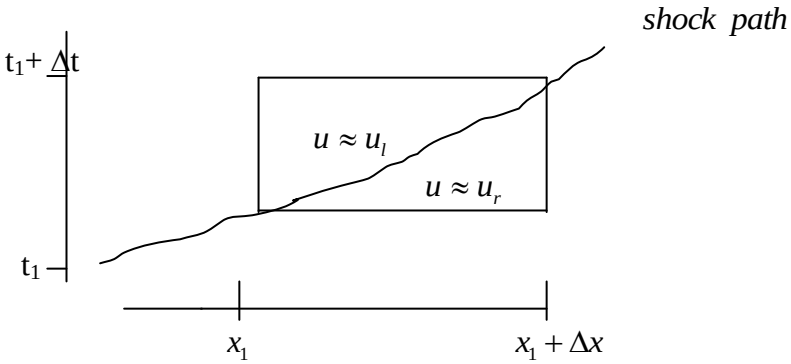


Figure 1.9.1 region of integration for shock speed calculation.

1.10 Entropy Condition

As demonstrated above, there are situations in which the weak solution is not unique and an additional condition is required to pick out the physically relevant solution.

For scalar equations there is an obvious condition suggested by figures (1.8.1) and (1.8.2). A shock should have characteristics going into the shock, as time advance. A propagating discontinuity with characteristics coming out of it is unstable to perturbations. Ether smearing out the initial profiles a little will cause this to be replaced by a rarefaction fan of characteristics. This gives our first version of the entropy condition:

ENTROPY CONDITION (VERSION I):

A discontinuity propagating with speed s given by (1.9.4) satisfies the entropy condition if

$$f'(u_l) > s > f'(u_r) \ . \qquad (1.10.1)$$

Note that $f'(u)$ is the characteristic speed. For convex f, the R-H speed s from (1.9.5) must lie between $f'(u_l)$ *and* $f'(u_r)$ so (1.10.1) reduces to simply the requirement that $f'(u_l) > f'(u_r)$, which again by convexity requires $u_l > u_r$.

ENTROPY CONDITION (VERSION II):

u(x, t) is the entropy solution if all discontinuities have the property that

$$\frac{f(u) - f(u_l)}{u - u_l} \geq s \geq \frac{f(u) - f(u_r)}{u - u_r} \tag{1.10.2}$$

for all u between u_l *and* u_r.

For convex f, this requirement reduces to (1.10.1).

Another form of entropy condition is based on the spreading of characteristics in a rarefaction fan. If u(x,t) is an increasing function of x in some region, then characteristics spread out if $f'' > 0$. the rate of spreading can be quantified, and gives the following condition.

ENTROPY CONDITION (VERSION III):

u(x, t) is the entropy solution if there is a constant $E > 0$ such that for all $a > 0, t > 0$ and $x \in R$,

$$\frac{u(x+a,t) - u(x,t)}{a} < \frac{E}{t} \tag{1.10.3}$$

Note that for a discontinuity propagating with constant left and right states u_l *and* u_r, this can be satisfied only if $u_r - u_l \leq 0$, so this agrees with (1.10.1). The form of (1.10.3) has advantages in studying numerical methods.

1.11 ENTROPY FUNCTIONS

Yet another approach to the entropy condition is to define an entropy function $\eta(u)$ for which an additional conservation law holds for smooth solution that becomes an inequality for discontinuous solutions. In gas dynamics, there exists a physical quantity called entropy that is known to be constant along particle paths in smooth flow and to jump to a higher value as the gas crosses a shock. It can never jump to a lower value, and this gives the physical entropy condition that picks out the correct weak solution in gas dynamics.

Suppose some function $\eta(u)$ satisfies a conservation law of the form

$$\eta(u)_t + \Psi(u)_x = 0 \qquad (1.11.1)$$

for some entropy flux $\Psi(u)$. Then we can obtain from this, for smooth u ,

$$\eta'(u)u_t + \Psi'(u)u_x = 0 \ . \qquad (1.11.2)$$

Recall that the conservation law (1.1.1) can be written as $u_t + f'(u)u_x = 0.$ multiply this by

$\eta'(u)$ and compare with (1.11.2) to obtain

$$\Psi'(u) = \eta'(u) f'(u) \qquad (1.11.3)$$

For a scalar conservation law this equation admits many solutions $\eta(u), \Psi(u).$ an additional condition we place on the entropy function is that it be convex, $\eta''(u) > 0,$ for reasons that will be seen below.

The entropy $\eta(u)$ is conserved for smooth flows by its definition. For discontinuous solutions, however, the manipulations performed above are not valid. Since we are particularly interested in how the entropy behaves for the vanishing viscosity weak

solution, we look at the related viscous problem and will then let the viscosity tends to zero. The viscous equation is

$$u_t + f(u)_x = \in u_{xx}. \tag{1.11.4}$$

Since solution to this problem is always smooth, we can derive the corresponding evolution equation for the entropy following the same manipulations we used for smooth solutions of the inviscid equation, multiplying (1.11.4) by $\eta'(u)$ to obtain

$$\eta(u)_t + \Psi(u)_x = \in \eta'(u)u_{xx}. \tag{1.11.5}$$

We can now rewrite the right hand side to obtain

$$\eta(u)_t + \Psi(u)_x = \in (\eta'(u)u_x)_x - \in \eta''(u)u_x^2 \ . \tag{1.11.6}$$

Integrating this equation over the rectangle $[x_1, x_2]\times[t_1, t_2]$ gives

$$\int_{t_1}^{t_2}\int_{x_1}^{x_2} \eta(u)_t + \Psi(u)_x dxdt = \in \int_{t_1}^{t_2} [\eta'(u(x_2,t))u_x(x_2,t) - \eta'(u(x_1,t))u_x(x_1,t)]dt$$
$$- \in \int_{t_1}^{t_2}\int_{x_1}^{x_2} \eta''(u)u_x^2 dxdt \ .$$

As $\in \to 0,$ the first term on the right hand side vanishes. (This is clearly true if u is smooth at x_1, *and* x_2. the other term, involves integrating over the$[x_1, x_2]\times[t_1, t_2]$. If the limiting weak solution is discontinuous along a curve in this rectangle, then this term will not vanish in the limit. However, since $\in > 0, u_x^2 > 0$ *and* $\eta'' > 0$ (by our convexity assumption), we can conclude that the right hand side is non-positive in the limit and hence the vanishing viscosity weak solution satisfies

$$\int_{t_1}^{t_2}\int_{x_1}^{x_2} \eta(u)_t + \Psi(u)_x dxdt \le 0 \tag{1.11.7}$$

for all $x_1, x_2, t_1,$ *and* t_2. Alternatively, in integral form ,

$$\int_{x_1}^{x_2} \eta(u(x,t))dx]_{t_1}^{t_2} + \int_{t_1}^{t_2} \Psi(u(x,t))dt]_{x_1}^{x_2} \le 0 \qquad (1.11.8)$$

i.e.

$$\int_{x_1}^{x_2} \eta(u(x,t_2))dx \le \int_{x_1}^{x_2} \eta(u(x,t_1))dx - (\int_{t_1}^{t_2} \Psi(u(x_2,t))dt - \int_{t_1}^{t_2} \Psi(u(x_1,t))dt \qquad (1.11.9)$$

Consequently, the total integral of η is not necessarily conserved, but can only decrease.

(Note that our mathematical assumption of convexity leads to an "entropy function" that decreases, whereas the physical entropy in gas dynamics increases.) The fact that (1.11.7) holds for all $x_1, x_2, t_1, \text{and } t_2$ is summarized by saying that $\eta(u)_t + \Psi(u)_x \le 0$ in the weak sense. This gives our final form of the entropy condition, called the entropy inequality.

ENTROPY CONDITION (VERSION IV):

The function u(x, t) is the entropy solution of (1.1.1) if, for all convex entropy functions and corresponding entropy fluxes, the inequality

$$\eta(u)_t + \Psi(u)_x \le 0 \qquad (1.11.10)$$

is satisfied in the weak sense.

This formulation is also useful in analyzing numerical methods. If a discrete form of this entropy inequality is known to hold for some numerical methods, then it can be shown that the method converges to the entropy solution.

Just as for the conservation law, alternative weak form of the entropy condition can be formulated by integrating against smooth test function Φ, now required to be nonnegative since the entropy condition involves an inequality. The weak form of the entropy inequality is

$$\int_0^\infty \int_{-\infty}^\infty \Phi_t(x,t)\eta(x,t) + \Phi_x(x,t)\Psi(x,t)dxdt \le -\int_{-\infty}^\infty \Phi(x,0)\eta(x,0)dx \qquad (1.11.11)$$

for all $\Phi \in C_0^1(R \times R)$ *with* $\Phi(x,t) \ge 0$ for all x , t.

Example:

Consider Burgers' equation with $f(u) = \frac{1}{2}u^2$ and take $\eta(u) = u^2$.

Then (1.11.3) gives $\Psi'(u) = 2u^2$ and hence $\Psi(u) = \frac{2}{3}u^3$. then entropy condition (1.11.10) reads

$$(u^2)_t + (\frac{2}{3}u^3)_x \le 0 \qquad (1.11.12)$$

For smooth solution of Burgers' equation this should hold with equality. If a discontinuity is present, then integrating $(u^2)_t + (\frac{2}{3}u^3)_x$ over an infinitesimal rectangle as in Figure (1.9.1) gives

$$\begin{aligned}\int_{x_1}^{x_2} u^2(x,t)dx]_{t_1}^{t_2} + \int_{t_1}^{t_2} \frac{2}{3}u^3(x,t)dt]_{x_1}^{x_2} &= \Delta x(u_l^2 - u_r^2) + \frac{2}{3}\Delta t(u_r^3 - u_l^3) \\ &= \Delta t(u_l^2 - u_r^2)(s_1 - s_2) + o(\Delta t^2) \\ &= -\frac{1}{6}(u_l - u_r)^3 \Delta t + o(\Delta t^2)\end{aligned}$$

for small $\Delta t > 0,$ the $o(\Delta t^2)$ term will not affect the sign of this quality and so the weak form (1.8.8) is satisfied if and only if $(u_l - u_r)^3 > 0,$ and the only allowable discontinuities have $u_l > u_r$, as expected.

1.12 Scalar Example (Traffic flow)

In this section we will look of example of scalar conservation law with physical meaning, and apply the theory developed in the previous sections. This application (traffic flow) should also help develop some physical intuition that applicable to the more complicated case of gas dynamics, with gas molecules taking the place of cars. Consider the flow of cars on a highway. Let ρ denote the density of cars (in vehicles per mile) and u the velocity. In this application ρ is restricted to a certain range, $0 \le \rho \le \rho_{\max}$, where $\rho_{\max}$ is the value at which cars are bumper to bumper.

Since cars are conserved, the density and velocity must be related by the continuity equation derived before,

$$\rho_t + (\rho u)_x = 0. \qquad (1.12.1)$$

in order to obtain a scalar conservation law for ρ alone, we now assume that u is a given function of ρ. This makes sense: on a highway we would optimally like to drive at some speed $u_{\max}$ (the speed limit perhaps), but in heavy traffic we slow down, with velocity decreasing as density increases. The simplest model is the linear relation

$$u(\rho) = u_{\max}(1 - \rho / \rho_{\max}) \qquad (1.12.2)$$

At zero density (empty road) the speed is $u_{\max}$, but decreases to zero as ρ approaches $\rho_{\max}$. using this in (1.12.1) gives

$$\rho_t + f(\rho)_x = 0 \qquad (1.12.3)$$

where

$$f(\rho) = \rho u_{\max}(1 - \rho / \rho_{\max}) \qquad (1.12.4)$$

Whitham notes that a good fit to data for Lincoln tunnel was found by Greenberg in 1959 by

$$f(\rho) = a\rho \log(\rho_{\max} / \rho),$$

a function shaped similar to (1.12.4).

The characteristic speeds for (1.12.3) with flux (1.12.4) are

$$f'(\rho) = u_{\max}(1 - 2\rho / \rho_{\max}), \tag{1.12.5}$$

while the shock speed for a jump from ρ_l *to* ρ_r is

$$s = \frac{f(\rho_l) - f(\rho_r)}{\rho_l - \rho_r} = u_{\max}(1 - (\rho_l + \rho_r)/\rho_{\max}). \tag{1.12.6}$$

The entropy condition ($f'(u_l) > s > f'(u_r)$ says that a propagating shock must satisfy $f'(\rho_l) > f'(\rho_r)$ which requires $\rho_l < \rho_r$. Note this is the opposite inequality as in Burgers' equation since here f is concave rather than convex.

Example 1.12.1.

Take initial data

$$\rho(x,0) = \begin{cases} \rho_l & x < 0 \\ \rho_r & x > 0 \end{cases} \tag{1.12.7}$$

where $0 < \rho_l < \rho_r < \rho_{\max}$. Then the solution is a shock wave traveling with speed s given by (1.12.6). Note that although $u(\rho) \geq 0$ the shock speed s can be either positive or negative depending on ρ_l *and* ρ_r.

Consider the case $\rho_r = \rho_{\max}$ and $\rho_l < \rho_{\max}$. Then $s < 0$ and the shock propagates to the left. This models the situation in which cars moving at speed $u_l > 0$

unexpectedly encounter a bumper-to-bumper traffic jam and slam on their brakes, instantaneously reducing their velocity to 0 while the density jumps from ρ_l to $\rho_{\max}$. This discontinuity occurs at the shock, and clearly the shock location moves to the left as more cars join the traffic jam. This is illustrated in Figure1.12.1, where the vehicle trajectories ("particle paths") are sketched. Note that the distance between vehicles is inversely proportional to density. In gas dynamics, $1/\rho$ is called the specific volume.

The particle paths should not be confused with the characteristics, which are shown in Figure 1.12.2 for the case $\rho_l = \frac{1}{2}\rho_{\max}$ ($so\ u_l = \frac{1}{2}u_{\max}$), as is the case in figure 1.12.1 also. In this case, $f'(\rho) = 0$. If $\rho_l > \frac{1}{2}\rho_{\max}$ then $f'(\rho_l) < 0$ and all the characteristics go to the left, while if $\rho < \frac{1}{2}\rho_{\max}$ then $f'(\rho) > 0$ and the characteristics to the left of the shock are rightward going.

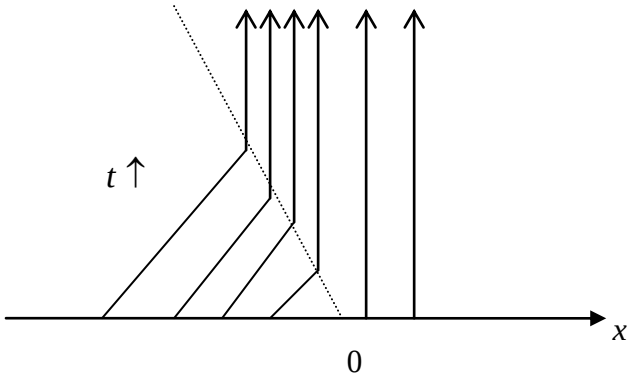


*Figure*1.12.1 *Traffic jam shock wave* (*car pathes*)

with data $\rho_l = \frac{1}{2}\rho_{\max}$, $\rho_r = \rho_{\max}$

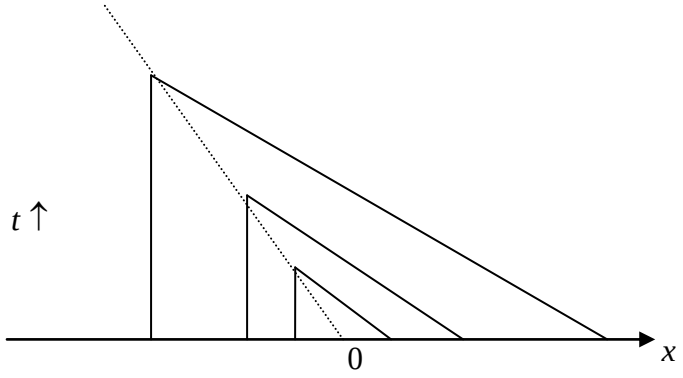


Figure 1.12.2 *characteristics*

Example 1.12.2.

Again consider a Riemann problem with data of the form (1.12.7) but now take $0 < \rho_r < \rho_l < \rho_{\max}$ so that the solution is a rarefaction wave. This might model the start up of cars after a light turns green. The cars to the left are initially stationary but can begin to accelerate once the cars in front of them begin to move. Since the velocity is related to the density by (1.12.2), each driver can speed up only by allowing the distance between her and the previous car to increase, and so we see a gradual acceleration and spreading out of cars.

As cars go through the rarefaction wave, the density decreases. Cars spread out or become “rarified” in the terminology used for gas molecules. Of course n this case there is another weak solution to (1.12.3), the entropy-violating shock. This would correspond to drivers accelerating instantaneously from $u_l = 0$ to $u_r > 0$ as the preceding car moves out of the way.

Chapter two

Analysis of Numerical Methods for Scalar Conservation Laws

This chapter introduces the basic ideas of discrete approximations, such as accuracy and convergence. In section (2.1) we introduce an example to show how finite difference method works. Then we introduce the most important scheme to illustrate the basic principles of finite-difference method. Finally, we analyze the accuracy and convergence of these schemes.

2.1 Solution by Finite Difference Method

Any second order partial differential equation has the form,

$$Au_{xx} + Bu_{xt} + Cu_{tt} = F(x,t,u,u_x,u_t). \qquad (2.1.1)$$

where A,B and C are constant. There are three types of equations:

if $B^2 - 4AC > 0,$ the equation is called hyperbolic. (2.1.2)

If $B^2 - 4AC = 0,$ the equation is called parabolic. (2.1.3)

If $B^2 - 4AC < 0,$ the equation is called elliptic. (2.1.4)

As an example of a hyperbolic partial differential equation, we consider the wave equation

$$u_{tt}(x,t) = c^2 u_{xx}(x,t), \quad \text{for } 0 < x < a \quad and \quad 0 < t < b, \qquad (2.1.5)$$

(where $u_{xx} = \frac{\partial^2 u}{\partial x^2}$ *and* $u_{tt} = \frac{\partial^2 u}{\partial t^2}$).

with the boundary and initial conditions

$$u(0,t) = 0 \quad and \quad u(a,t) = 0 \quad for \quad 0 \le t \le b,$$

$$u(x,0) = f(x) \qquad \text{for} \quad 0 \le x \le a, \qquad (2.1.6)$$

$$u_t(x,0) = g(x) \qquad \text{for} \quad 0 < x < a,$$

The wave equation models the displacement *u* of a vibrating elastic string with fixed ends at *x* = 0 and *x* = *a*. although analytic solution to the wave equation can be obtained with Fourier series; we use the problem as a prototype of a hyperbolic equation.

Derivation of the difference equation

Partition the rectangle R= $\{(x,t) : 0 \le x \le a, \quad 0 \le t \le b \}$ into a grid consisting of (p-1) by (m-1) rectangles with sides $\Delta x = h \quad and \ \Delta t = k,$ as shown in Figure (2.1.1) start at the bottom row, where t = 0 and the solution is known to be $U_j^0 = u(x_j,0) = f(x_j).$ we shall use a difference-equation method to compute $\{U_j^n : j = 1,2,....., p\}$ in successive rows for n = 2, 3,……, m.

The true solution at the grid points is $U_j^n = u(x_j,t_n)$

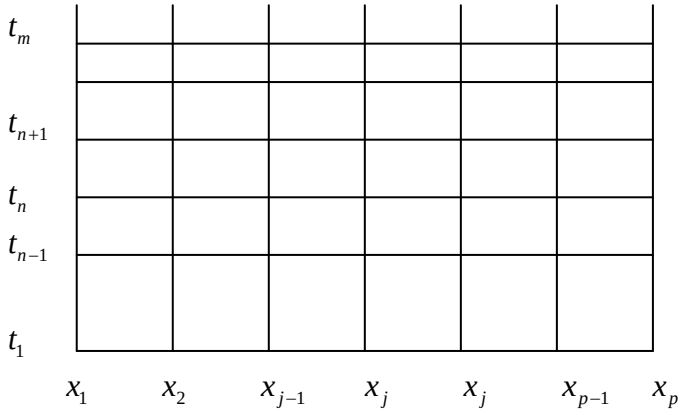


Figure 2.1.1: The grid for solving $u_{tt} = c^2 u_{xx}$ over R.

The central-difference formulas for approximating u_{tt} *and* u_{xx} are

$$u_{tt}(x,t) = \frac{u(x,t+k) - 2u(x,t) + u(x,t-k)}{k^2} + O(k^2) \tag{2.1.7}$$

and

$$u_{xx}(x,t) = \frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2} + O(h^2) \tag{2.1.8}$$

The grid spacing is uniform in every row $x_{j+1} = x_j + h, x_{j-1} = x_j - h$ and it is uniform in every column $t_{n+1} = t_n + k,\ t_{n-1} = t_n - k$, next we drop the terms $O(h^2)$ and $O(k^2)$ and use the approximation $U_j^n = u(x_j, t_n)$ in equation (2.1.7) and 2.1.8, which in turn are substituted into (2.1.5); this produces the difference equation.

$$\frac{U_j^{n+1} - 2U_j^n + U_j^{n-1}}{k^2} = c^2 \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2} \tag{2.1.9}$$

Which approximates the solution to (2.1.5.) For convenience, the substitution $r = ck/h$ is introduced in (2.1.9).

Equation (2.1.9) is employed to find row n+1 across the grid, assuming that approximations in both *n* and *n*-1 are known:

$$U_j^{n+1} = (2-2r^2)U_j^n + r^2(U_{j+1}^n + U_{j-1}^n) - U_j^{n-1} \quad \text{for j=2,3,….,p-1} \qquad (2.1.10)$$

The four known values on the right side of equation (2.1.10), which are used are to create the approximation U_j^{n+1}, are shown in the Figure (2.1.2)

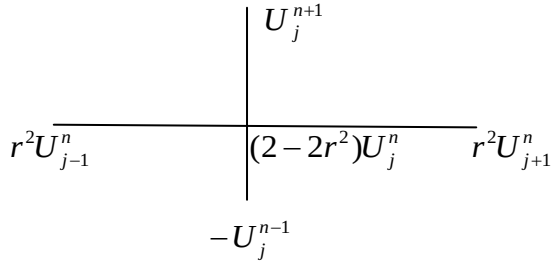


Figure 2.1.2: The wave equation stencil.

Caution must be taken when using formula (2.1.10), if the error made at one stage of the calculation is dampened out, the method is called stable. To guarantee stability in (2.1.10) it is necessary that $r = ck/h \le 1$. There are other schemes, called implicit methods, which are more complicated to implement, but do not have stability restrictions on r.

In order to use formula (2.1.10) to compute the third row two starting values corresponding to n = 1 and n = 2 must be supplied. Since the second row is not usually given, the boundary function g(x) is used to help produce starting approximations in the second row. Fix $x = x_j$ at the boundary and apply Taylor's formula of order 1 for expanding $u(x,t)$ about $(x_j,0)$, the value $u(x_j,k)$ satisfies

$$u(x_j,k) = u(x_j,0) + u_t(x_j,0)k + O(k^2) \qquad (2.1.11)$$

Then use $u(x_j,0) = f(x_j) = f_j$ and $u_t(x_j,0) = g(x_j) = g_j$ in (2.1.11) to produce the formula for computing the numerical approximations in the second row:

$$U_j^2 == f_j + kg_j \qquad for\ j = 2,3,....., p-1 \qquad (2.1.12)$$

We use a very small step size for k so that the values for U_j^2 given in (2.1.11) do not contain a large amount of truncation error.

Often, the boundary function f has a second derivative $f''(x)$ over the interval. In this case we have $u_{xx}(x,0) = f''(x)$ and we use the Taylor formula to help construct the second row,

$$u_{tt}(x_j,0) = c^2 u_{xx}(x_j,0) = c^2 f_{xx}(x_j) = c^2 \frac{f_{j+1} - 2f_j + f_{j-1}}{h^2} + O(h^2) \qquad (2.1.13)$$

Recall the Taylor's formula of order 2 is

$$u(x,k) = u(x,0) + u_t(x,0)k + \frac{u_{tt}(x,0)k^2}{2} + O(k^3)\,. \qquad (2.1.14)$$

Applying formula (2.1.14) at $x = x_j$, together with (2.1.12) and (2.1.13), we get

$$u(x_j,k) = f_j + kg_j + \frac{c^2k^2}{2h^2}(f_{j+1} - 2f_j + f_{j-1}) + O(k^2)O(h^2) + O(k^3)\,. \qquad (2.1.15)$$

Using r = ck/h, formula (2.1.15) can be simplified to obtain a difference formula for the improved numerical approximations in the second row:

$$U_j^2 = (1-r^2)f_j + kg_j + \frac{r^2}{2}(f_{j+1} + f_{j-1}) \qquad \text{for j = 2,3,…..,p-1} \qquad (2.1.16)$$

Remark

Assume that two rows of values $U_j^1 = u(x_j,0)$ and $U_j^2 = u(x_j,k)$ for j = 1,2,…,p are the exact solutions to the wave equation (2.1.5.) if the step size k = h/c is chosen along the t-axis, then r = 1 and formula (2.1.10) becomes

$$U_j^{n+1} = U_{j+1}^n + U_{j-1}^n - U_j^{n-1} \qquad (2.1.17)$$

Example 2.1.1

Consider the equation for a vibrating string

$$u_{tt}(x,t) = 4u_{xx}(x,t) \quad for \quad 0 < x < 1 \; and \; 0 < t < 0.5 \tag{2.1.18}$$

with the boundary conditions

$$\begin{aligned} &u(0,t) = 0 \; and \quad u(1,t) = 0 \qquad for \; 0 \le t \le 0.5, \\ &u(x,0) = f(x) = \sin(\pi x) + \sin(2\pi x) \quad for \; 0 \le x \le 1 \\ &u_t(x,0) = g(x) = 0 \qquad for \; 0 \le x \le 1. \end{aligned} \tag{2.1.19}$$

Now, for convenience we choose h = 0.1 and k = 0.05. Since c = 2 this yields r = ck/h =1. Since g(x) =0 and r = 1, formula (2.1.15) for creating the second row is

$$U_j^2 = \frac{f_{j-1} + f_{j+1}}{2} \quad for \quad j = 2,3,....,9 \tag{2.1.20}$$

Substituting r = 1 into equation (2.1.10) gives the simplified difference equation

$$U_j^{n+1} = U_{j+1}^n + U_{j-1}^n - U_j^{n-1}. \tag{2.1.21}$$

Applying (2.1.20) and (2.1.21) successively to generate rows will produce the approximation to $u(x,t)$ given in table (2.1.1) for $0 < x_j < 1 \; and \; 0 \le t_n \le 0.50$.

Table 2.1.1 solution of wave equation 2.1.17 with the boundary condition 2.1.18

t_j	x_2 =0.1	x_3 =0.2	x_4 =0.3	x_5 =0.4	x_6 =0.5	x_7 =0.6	x_8 =0.7	x_9 =0.8	x_{10} =0.9
0.00	0.896802	1.538842	1.760074	1.538842	1.000000	0.363271	-0.142040	-0.363271	-0.278768
0.05	0.769421	1.328438	1.538842	1.380037	0.951056	0.428980	0.000000	-0.210404	-0.181636
0.10	0.431636	0.769421	0.948401	0.951056	0.809017	0.587785	0.360616	0.181636	0.068364
0.15	0.000000	0.051599	0.181636	0.377381	0.587785	0.740653	0.769421	0.639384	0.363271
0.20	-0.380037	-0.587785	-0.519421	-0.181636	0.309017	0.769421	1.019421	0.951056	0.571020
0.25	-0.587785	-0.951056	-0.951056	-0.587785	0.000000	0.587785	0.951056	0.951056	0.587785
0.30	-0.571020	-0.951056	-1.019421	-0.769421	-0.309017	0.181636	0.519421	0.587785	0.380037
0.35	-0.363271	-0.639384	-0.769421	-0.740653	-0.587785	-0.377381	-0.181636	-0.051599	0.000000
0.40	-0.068364	-0.181636	-0.360616	-0.587785	-0.809017	-0.951056	-0.948401	-0.769421	-0.431636
0.45	0.181636	0.210404	0.000000	-0.428980	-0.951056	-1.380037	-1.538842	-1.328438	-0.769421
0.50	0.278768	0.363271	0.142040	-0.363271	-1.000000	-1.538842	-1.760074	-1.538842	-0.896802

The numerical values in Table (2.1.1) agree to more than six decimal places of accuracy with those obtained with the analytic solution

$$u(x,t) = \frac{1}{2}[f(x+ct) + f(x-ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)ds \tag{2.1.22}$$

Then

$$u(x,t) = \sin(\pi x)\cos(2\pi t) + \sin(2\pi x)\cos(4\pi t)\,. \qquad (2.1.23)$$

For example, by hand calculations, we can find some of U_j^n

* $U_1^n = 0,\ for\ n = 1,2,3,....,11$ the first column, $U_{11}^n = 0$ the last column.

* $U_j^1 = f(x_j)$, j = 2,3,…..,10 the first row

$$U_2^1 = f(x_2) = f(0.1) = \sin(18) + \sin(36) = 0.8968022$$
$$U_3^1 = f(x_3) = f(0.2) = \sin(36) + \sin(72) = 1.5388418$$

e.g. .

.

$$U_{10}^1 = f(x_{10}) = f(0.9) = ... = -0.2787682$$

* for the second row , we use the formula $U_j^2 = \frac{f_{j-1} + f_{j+1}}{2}$, j=2,3,….,10

e.g. for j = 2, $U_2^2 = \frac{f_1 + f_3}{2} = \frac{f(0) + f(0.2)}{2} = 0.7694208$

$$U_3^2 = \frac{f_2 + f_4}{2} = \frac{f(0.1) + f(0.3)}{2} = 1.328438$$

.

.

$$U_{10}^2 = \frac{f_9 + f_{11}}{2} = \frac{f(0.8) + 0}{2} = -0.181636$$

*for the third row we use the formula (j=2, 3, 4,….., 10, and n=3),

$$U_j^{n+1} = U_{j+1}^n + U_{j-1}^n - U_j^{n-1}$$

e.g. $U_2^3 = U_3^2 + U_1^2 - U_2^1 = 1.328438 + 0 - 0.896802 = 0.431636$

$$U_3^3 = U_4^2 + U_2^2 - U_3^1 = 1.538842 + 0.769421 - 1.538842 = 0.769421$$

Similarly, we can find the remaining values of U_j^n, n=4,5,…,11

Algorithm 2.1.1 (finite-difference solution for the wave equation).

To approximate the solution of $u_{tt}(x,t) = \alpha^2 u_{xx}(x,t)$ with

$u(0,t) = u(l,t) = 0, \quad 0 < t < T, \quad$ and $u(x,0) = f(x)$, $u_t(x,0) = g(x)$ *for* $0 \le x \le l$

INPUT endpoint l; maximum time T; constant α; integer $m \ge 2, N \ge 2$.

OUTPUT approximations w_i^j to $u(x_i, t_j)$ *for each* $i = 0,.....,m$ *and* $j = 0,.....,N$.

Step 1 set $h = l/m$;

$k = T/N$;

$\lambda = k\alpha/h$.

Step 2 *For* $j = 1,...,N$ *set* $w_0^j = 0$;

$w_m^j = 0$;

Step 3 set $w_0^0 = f(0)$;

$w_m^0 = f(l)$.

For $i = 1,....,m-1$ (int*ialize for* $t = 0$ *and* $t = k$.)

Step 4 *set* $w_i^0 = f(ih)$;

$w_i^1 = (1-\lambda^2) f(ih) + \frac{\lambda^2}{2}[f(i+1)h) + f((i-1)h)] + kg(ih)$.

For $j = 1,...,N-1$ (*perform matrix multiplica tion*.)

Step 5 *for* $i = 1,....,m-1$

set $w_i^{j+1} = 2(1-\lambda^2)w_i^j + \lambda^2(w_{i+1}^j + w_{i-1}^j) + w_i^{j-1}$.

For $j = 0,.....,N$

set $t = jk$;

Step 6 *for* $i = 0,....,m$

set $x = ih$

OUTPUT (x, t, w_i^j).

Step 7 STOP. (Procedure is complete.)

2. 2 Numerical Approximation for linear Equations

We review some of the basic theory of numerical methods for the linear equation. The emphasis will be on concepts that carry over to the nonlinear case.

We consider the time-dependent Cauchy problem in one space dimension.

$$u_t + au_x = 0 \qquad -\infty < x < \infty, \quad t \geq 0 \tag{2.2.1}$$

$$u(x,0) = u_0(x) \tag{2.2.2}$$

To approximate the model problem (2.2.1) by finite differences we proceed as in section (2.1) and divide the closed domain $\overline{R} \times [0,T]$ by a set of lines parallel to the x- and t-axes to form a grid or mesh. We shall assume, for simplicity, only that the sets of lines are equally spaced, and from now on we shall assume that $\overline{R}$ is the interval [0, 1]. Note that in practice we have to work in a finite time interval [o, T], but T can be as large as we like.

We shall write $\Delta x \quad and \quad \Delta t$ for the line spacing. The crossing points

$$(x_j = j\Delta x, \quad t_n = n\Delta t), \quad \text{j = 0, 1… J, n = 0, 1,….} \tag{2.2.3}$$

where

$$\Delta x = 1/J \tag{2.2.4}$$

are called the grid points or mesh points. We seek approximations of the solution at these mesh points; these approximate values will be denoted by

$$U_j^n \approx u(x_j, t_n) \tag{2.2.5}$$

We shall approximate the derivatives in 2.2.1 by finite differences and then solve the resulting difference equation in an evolutionary manner starting from n = 0.

It will also be useful to define

$$x_{j+1/2} = x_j + \Delta x/2 = (j + \frac{1}{2})\Delta x . \qquad 2.2.6$$

For simplicity we take a uniform mesh. With Δx and Δt constant, although most of the methods discussed can be extended to variable meshes.

The finite difference methods we will develop approximations $U_j^n \in R^m$ to the solution $u(x_j, t_n)$ at the discrete grid points.

This is a standard interpretation of the approximate solution, and will be used at times here, but in developing methods for conservation it is often preferable to view U_j^n as an approximation to a cell average of $u(x,t_n)$ defined by

$$\tilde{u}_j^n \equiv \frac{1}{h}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,t_n), \qquad where\ h = \Delta x \qquad (2.2.7)$$

Rather than as an approximation to the pointwise values u_j^n. this interpretation is natural since the integral form of the conservation law describes precisely the time evolution of integrals such as that appearing in (2.2.7).

As initial data for the numerical method we use $u_0(x)$ to define U^0 either by point-wise values, $U_j^0 = u_j^0 = u_0(x_j)$, j = 1,2,…, J, and $U_0^n = U_J^n = 0$, $n = 0,1,2...$, or preferably by cell averages, $U_j^0 = \tilde{u}_j^0$.

It is also frequently convenient to define a piecewise constant function $U_k(x,t)$, $(where\ k = \Delta t)$ for all x and t from the discrete values U_j^n. we assign this function the value U_j^n in the (j,n) grid cell, i.e.,

$$U_k(x,t) = U_j^n \quad for \quad (x,t) \in [x_{j-1/2}, x_{j+1/2}) \times [t_n, t_{n+1}) . \qquad (2.2.8)$$

We index this function U_k by the time step k, and assume that the mesh width h and step k are related in some fixed way, so that the choice of k defines a unique mesh. For time-dependent hyperbolic equations one generally assume that the mesh ratio k/h is a fixed constant as k, h $\to 0$. This assumption will be made from here on.

From the initial data $u_0(x)$ we have defined data U^0 for our approximation of the solution. We now use a time-marching procedure to construct the approximations

U^1 from U^0, then U^2 *from* U^1 (and possibly also U^0) and so on.

There are a wide variety of finite difference methods that can be used. Many of these are derived simply by replacing the derivatives occurring in (2.2.1) by appropriate finite difference approximations based at the mesh point (x_j, t_n). For example, replacing u_t by a forward-in-time approximation

$$\frac{u(x_j, t_{n+1}) - u(x_j, t_n)}{\Delta t} \approx \frac{\partial u}{\partial t}(x_j, t_n) \tag{2.2.9}$$

and u_x by a spatially centered approximation,

$$\frac{u(x_{j+1}, t_n) - u(x_{j-1}, t_n)}{2\Delta x} \approx \frac{\partial u}{\partial x}(x_j, t_n) \tag{2.2.10}$$

We obtain the following difference equations for U^{n+1}:

$$\frac{U_j^{n+1} - U_j^n}{k} + a\left(\frac{U_{j+1}^n - U_{j-1}^n}{2h}\right) = 0 \tag{2.1.11}$$

This can be solved for U_j^{n+1} to obtain

$$U_j^{n+1} = U_j^n - \frac{k}{2h} a(U_{j+1}^n - U_{j-1}^n). \tag{2.2.12}$$

A far more stable method is obtained by evaluating the centered difference approximations to u_x at time t_{n+1} rather than at time t_n, giving

$$\frac{U_j^{n+1}-U_j^n}{k}+a\left(\frac{U_{j+1}^{n+1}-U_{j-1}^{n+1}}{2h}\right)=0 \qquad (2.2.13)$$

or

$$U_j^{n+1}=U_j^n-\frac{k}{2h}a(U_{j+1}^{n+1}-U_{j-1}^{n+1})\,. \qquad (2.2.14)$$

The method (2.2.12) allows us to determine U^{n+1} explicitly, and is called an explicit method, whereas (2.2.14) is an implicit method. Although the method (2.2.12) is useless due to stability problem, there are other explicit methods which work very satisfactorily (see [RA2], (pp. 100-101)).

If we look at which grid points are involved in the computation of U_j^{n+1} with a given method, we can obtain a diagram that is known as the stencil of the method. The stencils for the methods (2.2.12) and (2.2.14) are shown in figure (2.2.1).

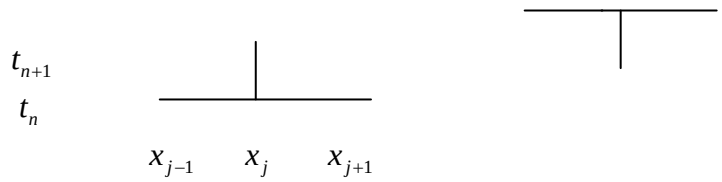


Figure 2.2.1 stencils for the methods (2.2.12) and (2.2.14).

A wide variety of methods for the linear problem can be devised by using different finite difference approximation. Most of these are based directly on finite difference approximations to the PDE. An exception is the **Lax-Wendroff method**, which is based on the Taylor series expansion

$$u(x,t+k)=u(x,t)+ku_t(x,t)+\frac{1}{2}k^2u_{tt}(x,t)+\ldots\ldots \qquad (2.2.15)$$

and the observation that from $u_t = -au_x$ we can compute

$$u_{tt} = -au_{xt} = -au_{tx} = -a(-au_x)_x = a^2 u_{xx} \tag{2.2.16}$$

so that(2.2.15) becomes

$$u(x,t+k) = u(x,t) - kau_x(x,t) + \frac{1}{2}k^2a^2u_{xx}(x,t) + ... \tag{2.2.17}$$

The **Lax-Wendroff method** then results from retaining only the first three terms of (2.2.17) and using centered difference approximations for the derivatives appearing there:

$$u_x(x,t) = \frac{1}{2h}(U^n_{j+1} - U^n_{j-1})$$

$$u_{xx}(x,t) = \frac{U^n_{j+1} - 2U^n_j + U^n_{j-1}}{h^2} \quad , u(x,t+k) = U^{n+1}_j \quad , u(x,t) = U^n_j .$$

We obtain **Lax-Wendroff scheme**

$$U^{n+1}_j = U^n_j - \frac{k}{2h}a(U^n_{j+1} - U^n_{j-1}) + \frac{k^2}{2h^2}a^2(U^n_{j+1} - 2U^n_j + U^n_{j-1}) \tag{2.2.18}$$

where $(k = \Delta t, \text{ and } \quad h = \Delta x)$.

Remark:

Lax-Wendroff Method for the scalar conservation-law **(** $u_t + (f(u))_x = 0$ **).**

The scalar conservation law

$$u_t + f(u)_x = 0 \tag{1}$$

admits the Lax-Wendroff Method (scalar)

$$U^{n+1}_j = U^n_j - \frac{\lambda}{2}(f^n_{j+1} - f^n_{j-1}) + \frac{\lambda^2}{4}[(f'^n_{j+1} + f'^n_j)(f^n_{j+1} - f^n_j) - (f'^n_j + f'^n_{j-1})(f^n_j - f^n_{j-1})]. \tag{2}$$

Here, $f^n_j = f(U^n_j)$, $f'^n_j = f'(U^n_j)$, $f' = \frac{df}{du}$.

Proof

A Taylor expansion in t gives

$$u(x_j,t_{n+1}) = u(x_j,t_n) + ku_t(x_j,t_n) + \frac{k^2}{2}u_{tt}(x_j,t_n) + ... \tag{3}$$

By (1) $u_t = -[f(u)]_x$, and so, using a centered x-difference,

$$u_t(x_j,t_n) \approx -\frac{f(u_{j+1}^n) - f(u_{j-1}^n)}{2h}, \qquad h = \Delta x. \tag{4}$$

Furthermore, $u_{tt} = [f'(u)(f(u))_x]_x$. Now, the usual centered second difference is the forward difference of a backward difference; that is,

$$\delta^2\Phi_n = (\Phi_{n+1} - \Phi_n) - (\Phi_n - \Phi_{n-1}).$$

Hence we approximate the inside x-derivative above as

$$[f(u)]_x \approx \frac{f(u_j^n) - f(u_{j-1}^n)}{h}$$

and represent its multiplier by a mean value:

$$f'(u) = \frac{f'(u_j^n) + f'(u_{j-1}^n)}{2}.$$

The forward differencing corresponding to the outside derivative then gives

$$u_{tt}(x_j,t_n) \approx \frac{1}{h}\left[\frac{f'(u_{j+1}^n) + f'(u_j^n)}{2}\frac{f(u_{j+1}^n) - f(u_j^n)}{h} - \frac{f'(u_j^n) + f'(u_{j-1}^n)}{2}\frac{f(u_j^n) - f(u_{j-1}^n)}{h}\right]$$

Substitution of (4) and the last equation in (3), and replacement of u by U, yields (2).

The program of the Lax-Wendroff method equation (2) is given below.

Example

Use the Lax-Wendroff method (2) to approximate the solution of

$$u_t + (u^2/2)_x = 0 \qquad x > 0,\ t > 0$$
$$u(x,0) = x \qquad x > 0$$
$$u(0,t) = 0 \qquad t > 0$$

At t=1, for $0 \le x \le 1$, compare the numerical solution with the exact solution, $u = x/(t+1)$.

Now by using the preceding program we get

H=0.1 K=0.1 $\lambda = 1$ T=1

	Numerical	Exact
X=0	0.	0.
X=0.1	0.50146	0.0500000
X=0.2	0.100292	0.100000
X=0.3	0.150438	0.150000
X=0.4	0.200585	0.200000
X=0.5	0.250731	0.250000
X=0.6	0.300877	0.300000
X=0.7	0.351023	0.350000
X=0.8	0.401169	0.400000
X=0.9	0.451315	0.450000
X=1.0	0.501461	0.500000 .

For time-dependent conservation laws, 2-level methods are almost exclusively used. We will study explicit 2-level methods almost exclusively, and introduce some special notation for such methods, writing

$$U^{n+1} = H_k(U^n) \tag{2.2.19}$$

here U^{n+1} represents the vector of approximations U_j^{n+1} at time t_{n+1}.

However, to illustrate the basic principles of the underlying finite-difference method, let us first consider the case of the linear equation $u_t + au_x = 0$.

We summarize four different ways in which $\{U_j^{n+1`}\}_{j=-\infty}^{\infty}$ can be derived from the known values $\{U_j^n\}_j$:

(a) The "backward"-difference scheme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = -a\frac{U_j^n - U_{j-1}^n}{\Delta x} \tag{2.2.20}$$

(b) The "forward"-difference scheme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = -a\frac{U_{j+1}^n - U_j^n}{\Delta x} \tag{2.2.21}$$

(c) The Lax-Friedrichs scheme

$$\frac{U_j^{n+1} - \frac{1}{2}[U_{j+1}^n + U_{j-1}^n]}{\Delta t} = -a\frac{U_{j+1}^n - U_{j-1}^n}{2\Delta x} \tag{2.2.22}$$

(d) The Lax-Wendroff scheme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = -a\frac{U_{j+1}^n - U_{j-1}^n}{2\Delta x} + a^2\frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{\Delta x^2}\frac{\Delta t}{2} \tag{2.2.23}$$

First, we set $k = \Delta t$ and take $\Delta x = \frac{\Delta t}{\lambda}$, where $\lambda > 0$ is given (and fixed). Thus, k is the only "small parameter" in the scheme.

Definition 2.2.1

Let $u(x, t)$ be a smooth solution to the conservation law $u_t + f(u)_x = 0$ and let $U^{n+1} = H_k(U^n)$ be an approximating scheme. We say that H_k is accurate of order $p \geq 1$ if, with $u^n = \{u(x_j, t_n)\}_{j=-\infty}^{\infty}$,

$$u^{n+1} - H_k(u^n) = O(k^{p+1}), \qquad k \to 0 \tag{2.2.24}$$

Remark

Observe that the notion of consistency is built into (2.2.24) in the following way:

Define $F_k(u^n) = u^n - H_k(u^n)$. Then (2.2.24) can be rewritten as

$$u^{n+1} - u^n + F_k(u^n) = O(k^{p+1}) \tag{2.2.25}$$

and dividing by k we have

$$\frac{u^{n+1} - u^n}{k} + \frac{1}{k} F_k(u^n) = O(k^p), \qquad k \to 0. \tag{2.2.26}$$

Since $u(x,t)$ is an exact solution, $\frac{u^{n+1} - u^n}{k}$ approximates u_t $(t = t_n)$, and therefore $\frac{1}{k} F_k(u^n)$ should be an approximation for $f(u)_x$ $(at\ t = t_n), as\ k \to 0$. This last conclusion is commonly referred to as the "consistency" of the scheme $H_k(u^n) = u^n - F_k(u^n)$ with the differential equation. Suppose that u(x, t) is a smooth function satisfying (2.2.24) and assume that

$$-u^n + H_k(u^n) + kau_x(x_j, t_n) = O(k^{p+1}) \qquad \text{as } k \to 0. \tag{2.2.27}$$

Inserting (2.2.25) into (2.2.27) we obtain:

$$\frac{u^{n+1}-u^n}{k}+au_x(x_j,t_n)=O(k^p) \qquad as \quad k\to 0, \tag{2.2.28}$$

so that by letting $k\to 0$ we get $u_t+au_x=0.$

Example 2.2.1.

Let us use the backward-difference scheme for equation (2.2.1):

$$\frac{U_j^{n+1}-U_j^n}{k}=-a\frac{U_j^n-U_{j-1}^n}{\Delta x}$$

We replace U_j^n by u(x, t), U_j^{n+1} by u(x, t+k) and U_{j-1}^n by u(x-h, t). Assuming *u* to be a smooth solution of (eq.2.2.1) and using Taylor expansion we get,

$$\begin{aligned} u(x,t+k) &= u(x,t)+ku_t(x,t)+O(k^2)\\ &= u(x,t)-kau_x(x,t)+O(k^2)\\ &= u(x,t)-\frac{ak}{\Delta x}[u(x,t)-u(x-\Delta x,t)]+O(k^2)+O(\Delta x^2)\\ &= H_k u(x,t)+O(k^2) \end{aligned}$$

where

$$H_k u(x,t)=(1-\lambda a)u(x,t)+\lambda au(x-\Delta x,t), \qquad where\ \lambda=\frac{k}{\Delta x} \tag{2.2.29}$$

and we have absorbed $O(\Delta x^2)$ into $O(k^2)$.

We conclude that the scheme is first-order accurate (p= 1). We can prove similarly that both the forward-difference and the Lax-Friedrichs schemes are of first-order accuracy, whereas the Lax-wendroff scheme is of second-order accuracy.

2.3 Convergence

The major question that poses itself in connection with discretized schemes is that of convergence (as $k \to 0$) of the "approximate solution" $\{U_j^n\}$ to a weak solution of the differential equation. To illustrate the situation, we take as before the linear equation

$$u_t + au_x = 0, \qquad u(x,0) = u_0(x). \tag{2.3.1}$$

Theorem 2.3.1

(Non convergence for large $\lambda > 0$) Assume $a > 0$ and take backward-difference scheme (2.2.20). Using (2.2.29) we can verify easily that

$$U_j^n = \sum_{l=0}^{n} \binom{n}{l} (\lambda a)^l (1-\lambda a)^{n-1} U_{j-l}^0 \tag{2.3.2}$$

Proof

from 2.2.20 $U_j^{n+1} - U_j^n + a\lambda(U_j^n - U_{j-1}^n) = 0$

This can be written in the form

$$U_j^{n+1} = (1 - a\lambda + a\lambda H_k)U_j^n,$$

where we have used $H_k(U_j^n) = U_{j-1}^n$ *i.e.* $(H_k^p U_{j-p}^n = U_{j-p}^n)$.

Hence $U_j^n = (1 - a\lambda + a\lambda H_k)U_j^{n-1}$

$$= (1 - a\lambda + a\lambda H_k)(1 - a\lambda + a\lambda H_k)U_j^{n-2}$$

$$= (1 - a\lambda + a\lambda H_k).....(1 - a\lambda + a\lambda H_k)U_j^0$$

$$= (1 - a\lambda + a\lambda H_k)^n U_j^0$$

$$= \sum_{l=0}^{n} \binom{n}{l} (a\lambda H_k)^l (1 - a\lambda)^{n-l} U_j^0$$

$$= \sum_{l=0}^{n} \binom{n}{l} (a\lambda)^l (1-a\lambda)^{n-l} H_k^l U_j^0 = \sum_{l=0}^{n} \binom{n}{l} (a\lambda)^l (1-a\lambda)^{n-l} U_{j-l}^0$$

We can prove by the same way that it true for U_j^{n+1}. □

The initial values $\{U_j^0\}_{j=-\infty}^{\infty}$ are computed from the given initial function $u_0(x)$. A common choice is to define U_j^0 as the average of $u_0(x)$ over the interval of size Δx centered at x_j, that is,

$$U_j^0 = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x)dx, \qquad x_r = r\Delta x. \tag{2.3.3}$$

If we take the initial step function

$$u_0(x) \begin{cases} 1, & x \le 0, \\ 0, & x > 0, \end{cases} \tag{2.3.4}$$

The weak solution of (2.3.1)-(2.3.4) is given by

$$u(x,t) = u_0(x-at). \tag{2.3.5}$$

For the corresponding approximation we get from (2.3.4)

$$U_j^0 = \begin{cases} 1, & j < 0, \\ \frac{1}{2}, & j = 0, \\ 0, & j > 0, \end{cases} \tag{2.3.6}$$

and from (2.3.2), along with simple facts about the binomial coefficients, we get

$$\sum_{j=0}^{n} U_j^n = \sum_{j=0}^{n} \sum_{l=0}^{n} \binom{n}{l} (\lambda a)^l (1-\lambda a)^{n-l} U_{j-l}^0$$

$$= \frac{1}{2} \sum_{j=0}^{n} \binom{n}{j} (\lambda a)^j (1-\lambda a)^{n-j} + \sum_{j=0}^{n} \sum_{l=j+1}^{n} \binom{n}{l} (\lambda a)^l (1-\lambda a)^{n-l}$$

$$= \frac{1}{2} + \sum_{l=0}^{n} l \binom{n}{l} (\lambda a)^l (1-\lambda a)^{n-l} = \frac{1}{2} + \lambda n a,$$

so that

$$\Delta x\sum_{j=0}^{n}U_j^n=\frac{\Delta x}{2}+nka \tag{2.3.7}$$

In analogy with interpretation of U_j^0, we think of U_j^n as approximating the mean value of $u(x,nk)$ in the interval $(x_{j-1/2},x_{j+1/2})$. Thus, the sum in (2.3.7) should be compared with the integral

$$\int_{-\frac{\Delta x}{2}}^{(n+1/2)\Delta x}u(x,nk)dx=\int_{-\frac{\Delta x}{2}}^{(n+1/2)\Delta x}u_0(x-ank)dx$$

$$=\int_{-\frac{\Delta x}{2}-ank}^{(n+\_1/2)\Delta x-ank}u_0(x)dx$$

Now fix $t>0$ and take $nk=t$. As $k\to 0$ and because $\lambda=\frac{k}{\Delta x}$ is fixed, we have $\Delta x\to 0$ and the limit of the last integral can be evaluated as

$$\int_{-\frac{\Delta x}{2}-at}^{\frac{\Delta x}{2}+t(\frac{1}{\lambda}-a)}u_0(x)dx\to\begin{cases}at, & \lambda^{-1}\geq a,\\ \dfrac{1}{\lambda}t, & \lambda^{-1}<a.\end{cases} \tag{2.3.8}$$

However, from (2.3.7), as $k\to 0$,

$$\Delta x\sum_{j=0}^{n}U_j^n\to at\,. \tag{2.3.9}$$

We conclude that

$$\left|\int_{-\frac{\Delta x}{2}}^{(n+1/2)\Delta x}u(x,t)dx-\Delta x\sum_{j=o}^{n}U_j^n\right|\to 0,\qquad as\quad k\to 0 \tag{2.3.10}$$

(with $nk=t\ \ and\ \ \lambda=\frac{k}{\Delta x}$) if and only if

$$\lambda a\leq 1 \tag{2.3.11}$$

Clearly, convergence "in the mean" is a very reasonable way of requesting that the sequence $\{U_j^n\}_{j=-\infty}^{\infty}$ should approximate the exact solution $u(x,nk)$. Since the step function (2.3.4) represents only one possible initial datum, we can only derive a necessary condition for convergence from the foregoing discussion.

First, we formalize the convergence in the mean as follows.

Definition 2.3.1

Fix $T > 0$. We say that $\{U_j^n\}_{j=-\infty}^{\infty}$ converges to the solution $u(x,t)$ *in* $L_{loc}^1(R)$ at fixed time t if, for any $[\alpha,\beta] \in R$,

$$\lim_{\substack{k\to 0 \\ (nk=t)}} \sum_{j=[\frac{\alpha}{\Delta x}]}^{[\frac{\beta}{\Delta x}]} \int_{x_{j-1/2}}^{x_{j+1/2}} \left| U_j^n - u(x,t) \right| dx = 0 \qquad (2.3.12)$$

for any $0 \le t \le T$.

Our conclusion (2.3.10) can be restated as follows.

Corollary 2.3.2

Let $a > 0$. Then (2.3.11) is a necessary condition for the backward-difference scheme to converge in $L_{loc}^1(R)$ to the solution (2.3.1).

Definition 2.3.3

The condition (2.3.11) is called the CFL (Courant-Friedrichs-Lewy)

Condition associated with equation (2.3.1) and scheme (2.2.20). Since $\lambda = \frac{k}{\Delta x}$, the CFL condition can be written as

$$k \le \frac{\Delta x}{a}. \tag{2.3.13}$$

It therefore forces a necessary restriction on the size of the time step $k = \Delta t$, relative to the cell size Δx, for convergence to take place. Later in this section we shall have a geometric interpretation of this condition, in terms of the characteristic lines of the equation. Note that the condition refers not only to the equation but also to the particular scheme used to approximate it. Although it plays a fundamental role in the theory of linear equations, it serves only as a guideline in the nonlinear case (via linearization). Because our primary objective here is the treatment of the nonlinear case, we shall make little use of the general theory related to the CFL condition.

As we shall see throughout this thesis, the backward-difference scheme (2.2.20) (for $a > 0$) plays a fundamental role in the development of accurate high-resolution schemes. The first step in this development is taken in the following theorem, proving the sufficiency of the CFL condition for convergence in $L^1_{loc}(R)$. Some knowledge of the binomial distribution is needed in the proof.

Theorem 2.3.4

Consider the equation $u_t + au_x = 0, a > 0,$ and assume that the initial function $u_0(x) = u(x,0)$ is uniformly bounded. Then, under the CFL condition (2.3.11) the backward-difference scheme (2.2.20) converges in $L^1_{loc}(R)$, that is, in the sense of Definition (2.3.1).

Proof

In view of $u(x,t) = u_0(x-at)$ and (2.3.2) ,(2.3.3) we can write, with $nk = t$,

$$\sum_{j=[\frac{\alpha}{\Delta x}]}^{[\frac{\beta}{\Delta x}]} \int_{x_{j-1/2}}^{x_{j+1/2}} \left| U_j^n - u(x,t) \right| dx$$

$$= \sum_{j=[\frac{\alpha}{\Delta x}]}^{[\frac{\beta}{\Delta x}]} \int_{x_{j-1/2}}^{x_{j+1/2}} \left| \sum_{l=0}^{n} \binom{n}{l} (\lambda a)^l (1-\lambda a)^{n-l} [U_{j-l}^0 - u_0(x-at)] \right| dx$$

$$\leq \sum_{l=0}^{n} \binom{n}{l} (\lambda a)^l (1-\lambda a)^{n-l} p_l, \qquad (2.3.14)$$

where

$$p_l = \sum_{j=[\frac{\alpha}{\Delta x}]}^{[\frac{\beta}{\Delta x}]} \int_{x_{j-1/2}}^{x_{j+1/2}} \left| U_{j-l}^0 - u_0(x-at) \right| dx, \qquad (2.3.15)$$

and where we have used the identity

$$\sum_{l=0}^{n} \binom{n}{l} (\lambda a)^l (1-\lambda a)^{n-l} = 1.$$

We note that $n\Delta x = \frac{nk}{\lambda} = \frac{t}{\lambda} \leq \frac{T}{\lambda},$ so that the numbers $p_l, 0 \leq l \leq n,$ are uniformly bounded by

$$|p_l| \leq 2 \int_{\alpha - \frac{T}{\lambda} - \Delta x}^{\beta + \Delta x} |u_0(x)| dx, \qquad 0 \leq l \leq n .$$

Recall that the "Law of large numbers" states that the binomial distribution $b_{n,l} = \binom{n}{l} (\lambda a)^l (1-\lambda a)^{n-l}$ is concentrated around $l = \lambda a n,$ or, more precisely, that for any $\varepsilon > 0$

$$\lim_{n\to\infty} \sum_{|l-\lambda an|>n_\epsilon} b_{n,l}\, p_l = 0\,. \tag{2.3.16}$$

Thus, going back to (2.3.14), (2.3.15), we obtain for any $\epsilon > 0$

$$\lim_{\substack{k\to 0 \\ (nk=t)}} \sum_{j=[\frac{\alpha}{\Delta x}]}^{[\frac{\beta}{\Delta x}]} \int_{x_{j-1/2}}^{x_{+1/2}} \left|U_j^n - u(x,t)\right| dx \le \lim_{\substack{n\to\infty \\ (nk=t)}} \sum_{|l-\lambda an|\le n_\epsilon} b_{n,l}\, p_l\,. \tag{2.3.17}$$

But if $|l-\lambda an| \le n_\epsilon$, we have $\left|(j-l)\Delta x - (x_j - at)\right| \le n_\epsilon \Delta x \le \dfrac{T}{\lambda}\epsilon,$ *so*

$$p_l \le \sum_{j=[\frac{\alpha}{\Delta x}]}^{[\frac{\beta}{\Delta x}]} \int_{x_{j-l-1/2}}^{x_{j-l+1/2}} \left|U_{j-l}^0 - u_0(x)\right| dx$$

$$+ \sup_{0<h\le \frac{T}{\lambda}\epsilon} \int_{\alpha-\frac{T}{\lambda}}^{\beta} \left|u_0(y+h) - u_0(y)\right| dy\,. \tag{2.3.18}$$

Given $\delta > 0$ we can choose $\varepsilon > 0$ sufficiently small so that the second term on the right-hand side of (2.3.18) is smaller than $\dfrac{\delta}{2}$. This follows from elementary properties of functions in $L^1(R)$; simply approximate u_0 in L^1 by a smooth function. As for the first term in the right-hand side of (2.3.18), recall from (2.3.3) that U_{j-l}^0 is the average value of $u_0(x)$ over $[x_{j-l-1/2}, x_{j-l-1/2}]$. Thus, if $u_0(x)$ is smooth, $U_{j-l}^0 = u_0(y_{j-l})$ for some y_{j-l} in the interval and, since

$$u_0(y_{j-l}) - u_0(x) = \int_x^{y_{j-l}} u_0'(\xi)\, d\xi$$

We get

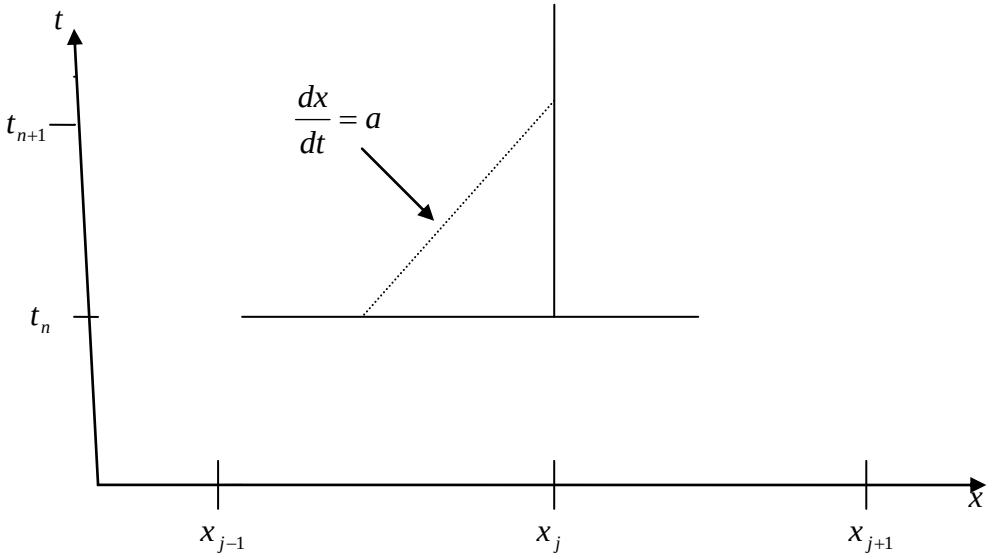
$$\sum_{j=[\frac{\alpha}{\Delta x}]}^{[\frac{\beta}{\Delta x}]} \int_{x_{j-l-1/2}}^{x_{j-l+1/2}} \left|U_{j-l}^0 - u_0(x)\right| dx \le \int_{\alpha-\frac{T}{\lambda}}^{\beta} \left|u_0'(y)\right| dy\,,$$

which is smaller than $\frac{\delta}{2}$ (*for* $0 \le l \le n$) *if* Δx is small. If u_0 is not smooth, it can be approximated (in L^1) by a smooth function, so that the same result holds.

We conclude that $\sup_{|l-\lambda an|<n_\varepsilon} p_l$ can be made arbitrarily small by taking $\varepsilon, \Delta x$ sufficiently small. From (2.3.17) we now get

$$\lim_{\substack{k\to 0 \\ (nk=t)}} \sum_{j=[\frac{\alpha}{\Delta x}]}^{[\frac{\beta}{\Delta x}]} \int_{x_{j-1/2}}^{x_{j+1/2}} \left| U_j^n - u(x,t) \right| dx = 0 ,$$

which proves our theorem. □



. Figure 2.3.1 Geometric (characteristic) interpretation.

The backward-difference scheme has a simple geometric interpretation. Consider the grid (x_j, t_n) as in Figure 2.3.1. As mentioned earlier, the approximating values $\{U_j^n\}$ are associated with the points (x_j, t_n). If the CFL condition (2.3.11) holds, then the characteristic line $x'(x) = a$, issuing from (x_j, t_{n+1}), intersects the line $t = t_n$ at the point $\bar{x} = \lambda a x_{j-1} + (1-\lambda a) x_j \in [x_{j-1}, x_j]$. If we use the linear interpolation

$$U^n(\bar{x}) = \lambda a U^n_{j-1} + (1-\lambda a)U^n_j \tag{2.3.19}$$

then the backward-difference scheme (2.2.20) states simply that

$$U^{n+1}_j = U^n(\bar{x}),$$

which just expresses the fact that the corresponding exact solution is constant along a characteristic line. We can summarize this discussion as follows.

Summary:

The values $\{U^{n+1}_j\}_{j=-\infty}^{\infty}$ as obtained by the scheme (2.2.20) ($a > 0$), subject to the CFL condition (2.3.11), are the exact values $\tilde{u}(x_j, t_{n+1})$, where $\tilde{u}(x,t)$ satisfies the equation $\tilde{u}_t + a\tilde{u}_x = 0$, subject to the initial condition $\tilde{u}(x,t_n) = U^n(x)$. The function $U^n(x)$ is the piecewise linear (continuous) function obtained by interpolating the values $\{U^n_j\}_{j=-\infty}^{\infty}$ at the grid points $\{(x_j, t_n)\}_{j=-\infty}^{\infty}$.

Definition 2.3.5 (Upwinding)

We say that the backward-differences scheme (2.2.20) with $a > 0$, is an "upwind scheme," meaning that the values $\{U^{n+1}_j\}_{j=-\infty}^{\infty}$ are obtained from $\{U^n_j\}_{j=-\infty}^{\infty}$ by following the characteristic lines of the equation.

We now suggest yet another interpretation of the backward-difference scheme (see [BF], (pp. 33-35)). This one, as in the preceding discussion, will also be based on an exact solution of the equation, subject to approximate initial data. However, now we take $U^n(x)$ as the piecewise-constant function defined by

$$U^n(x) = U^n_j, \qquad x_{j-1/2} < x < x_{j+1/2}, \qquad -\infty < j < \infty. \tag{2.3.20}$$

We can make the following claim.

Claim 2.3.6

If we solve the equation $\tilde{\tilde{u}}_t + a\tilde{\tilde{u}}_x = 0,$ subject to the initial condition $\tilde{\tilde{u}}(x,t_n) = U^n(x)$ as in (2.3.20). Then the values $U_j^{n+1},$ as determined by the backward-difference scheme (2.2.20), satisfy

$$U_j^{n+1} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{\tilde{u}}(x,t_{n+1})dx, \tag{2.3.21}$$

Provided the CFL condition (2.3.11) holds.

Proof

The CFL condition implies that the "moving step" solution $[u(x,t) = u_0(x-at)]$ satisfies

$$\tilde{\tilde{u}}(x_{j+1/2},t) = U_j^n, \qquad t_n < t < t_{n+1}, \qquad -\infty < j < \infty. \tag{2.3.22}$$

It follows from the balance equation (1.2.1) that

$$\int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{\tilde{u}}(x,t_{n+1})dx = U_j^n\Delta x - a[U_j^n - U_j^n]k,$$

and, by $k = \lambda\Delta x,$

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{\tilde{u}}(x,t_{n+1})dx = (1-\lambda a)U_j^n + \lambda a U_{j-1}^n = U_j^{n+1}. \quad \square \tag{2.3.23}$$

Observe that although $\tilde{u}(x,t)$ in summary (2.3.5) and $\tilde{\tilde{u}}(x,t)$ in claim (2.3.7) satisfy the same differential equation, they are actually different since the initial data $U^n(x)$,

used to interpolate the discrete values $\{U_j^n\}_{j=-\infty}^{\infty}$, are different for the two cases. In the case of $\tilde{u}(x,t)$ the initial function $U^n(x)$, and hence $\tilde{u}(x,t)$, are continuous, and U_j^{n+1} is taken as the approximate (pointwise) "upwind" value. In contrast, in the case of $\tilde{\tilde{u}}(x,t)$, the initial function $U^n(x)$ is piecewise constant, and hence is in general discontinuous, and the value U_j^{n+1} is taken as the average of the ensuing solution $\tilde{\tilde{u}}(x,t_{n+1})$ *over* $(x_{j-1/2}, x_{j+1/2})$.

Recall that, for the nonlinear conservation law $u_t + f(u)_x = 0$, the solution can develop discontinuities even when subject to very smooth initial data. In this case, therefore, the pointwise upwinding approach expressed by $\tilde{u}(x,t)$, based on continuous interpolation, does not seem appropriate. In contrast, the "averaging" approach, based on the balance equation (1.2.1) applied to piecewise-constant initial data, can be readily generalized to the nonlinear case. It is this approach, first suggested by Godunov, which will serve as the basis of the GRP discussed in the next chapter.

Remark

Note that none of the schemes (2.2.21)-(2.2.23) (i.e., the forward difference, Lax-Friedrichs, and Lax-Wendroff schemes) are amenable to an interpretation based on characteristic values ("upwinding" as in definition 2.3.6) or averaging in the sense of Godunov (as in claim 2.3.7). Nonetheless, all these scheme [including (2.2.20) are conservative in the sense that

$$\sum_{j=-\infty}^{\infty} U_j^{n+1} = \sum_{j=-\infty}^{\infty} U_j^n$$

(when the values U_j^n vanish sufficiently fast as $|j| \to \infty$. This is of course consistent with the conservation property ($\int_{-\infty}^{\infty} u(x,t)dx = \int_{-\infty}^{\infty} u_0(x)dx, \quad 0 \le t \le T$). However, in this thesis we shall not make much use of this conservation property.

We will use the 1-norm almost exclusively, and so a norm with no subscript will generally refer to the 1-norm. for the discrete grid function U^n we use the discrete 1-norm defined by

$$\|U^n\|_1 = h\sum_j |U_j^n|. \tag{2.3.24}$$

note that this is consistent with the function version in the sense that

$$\|U^n\|_1 = \|U_k(.,t_n)\|_1$$

2.4 Local Truncation Error

The local truncation error $L_k(x,t)$ is a measure of how well the difference equation models the differential equation locally. It is defined by replacing the approximation solution U_j^n in the difference equations by the true solution $u(x_j,t_n)$. of course this true solution of the PDE is only an approximate solution of the difference equations, and how well it satisfies the difference equations gives an indication of how well the exact solution of the difference equations satisfies the differential equation.

As an example, consider the Lax-Friedrichs method. This method is similar to the unstable method (2.2.12) but replaces U_j^n *by* $\frac{1}{2}(U_{j-1}^n - U_{j+1}^n)$ and is stable provided k/h is sufficiently small, as we will see later.

We first write this method in the form

$$\frac{1}{k}[U_j^{n+1} - \frac{1}{2}(U_{j-1}^n + U_{j+1}^n)] + \frac{1}{2h}a[U_{j+1}^n - U_{j-1}^n] = 0,$$

so that it appears to be a direct discretization of the PDE. If we now replace each U_j^n by the exact solution at the corresponding point, we will not get zero exactly. What we get instead is defined to be the local truncation error,

$$L_k(x,t) = \frac{1}{k}[u(x,t+k) - \frac{1}{2}(u(x-h,t) + u(x+h,t))] \qquad (2.4.1)$$

$$+ \frac{1}{2h}a[u(x+h,t) - u(x-h,t)]$$

in computing the local truncation error we always assume smooth solutions, and so we can expand each term on the right hand side of 2.3.1 in a Taylor series about *u*(*x*, t). Doing this and collecting terms gives (with $u \equiv u(x,t)$:

$$L_k(x,t) = \frac{1}{k}\left[(u + ku_t + \frac{1}{2}k^2u_{tt} + ...) - (u + \frac{1}{2}h^2u_{xx} + ...)\right]$$

$$+ \frac{1}{2h}a\left[2hu_x + \frac{1}{3}h^3u_{xxx} + ...\right]$$

$$= u_t + au_x + \frac{1}{2}\left(ku_{tt} - \frac{h^2}{k}u_{xx}\right) + O(h^2) \qquad (2.4.2)$$

Since we assume that $u(x, t)$ is the exact solution, $u_t + au_x = 0$ in 2.4.2. Using this and also (2.2.16), we find that

$$L_k(x,t) = \frac{1}{2}k(a^2 - \frac{h^2}{k^2})u_{xx}(x,t) + O(k^2)$$

$$= O(k) \quad as \quad k \to 0. \tag{2.4.3}$$

Recall that we assume that a fixed relation between k and h, k/h= constant, so that h^2/k^2 is constant as the mesh is refined. (This also justifies indexing L_k by k alone rather than by both k and h.)

by being more careful in this analysis, using Taylor's theorem with remainder and assuming uniform bounds on the appropriate derivatives of u(x, t), we can in fact show a sharp bound of the form

$$\left|L_k(x,t)\right| \le Ck \ \ for\ all\ k < k_0 \tag{2.4.4}$$

The constant C depends only on the initial data u_0. If we assume moreover that u_0 has compact support, then $L_k(x,t)$ will have finite 1-norm at each time t and we can obtain a bound of the form

$$\left\|L_k(x,t)\right\| \le C_L k \ \ for\ all \quad k < k_0 \tag{2.4.5}$$

for some constant C_L again depending on u_0.

The Lax-Friedrichs method is said to be first order accurate since the local error (2.4.5) depends linearly on k.

We now extend these notions to 2-level methods.

Definition 2.4.1.

For a general 2-level method, we defined the local truncation error by

$$L_k(x,t) = \frac{1}{k}[u(x,t+k) - H_k(u(.,t);x)]. \qquad (2.4.6)$$

Definition 2.4.2

The method is consistent if

$$\|L_k(.,t)\| \to 0 \quad as\ k \to 0. \qquad (2.4.7)$$

Definition 2.4.3

The method is of order p if for all sufficiently smooth initial data with compact support, there is some constant C_L such that

$$\|L_k(.,t)\| \le C_L k^p \ \ for\ all \quad k < k_0,\ t \le T \qquad (2.4.8)$$

this is the local order of the method, but it turns out that for smooth solution, the global error will be of the same order provided the method is stable.

2.5. Stability

Any two level method can be written (by 2.3.19) in the compact form

$$U^{n+1} = H_k(U^n) \qquad (2.5.1)$$

where U^n represents the vector of approximations $\{U_j^n : j \in Z\}$ at time t_n.

componentwise we have

$$U_j^{n+1} = H_k(U^n; j) \tag{2.5.2}$$

For instance, for the forward EULER METHOD $\left\{ U_j^{n+1} = U_j^n - \frac{\lambda a}{2}(U_{j+1}^n - U_{j-1}^n), \quad where\ \lambda = \frac{k}{h}, \quad k = \Delta t \quad h = \Delta x \right\}$ the operator H_k takes the form

$$U_j^{n+1} = H_k(U^n; j) = U_j^n - \frac{\lambda a}{2}(U_{j+1}^n - U_{j-1}^n) \tag{2.5.3}$$

Definition 2.5.1

A method is said to be stable if for each time T there is a constant $C_T > 0$ (possibly depending on T) and a value $\delta_0 > 0$ such that

$$\left\|U^n\right\|_h \le C_T \left\|U^0\right\|_h \tag{2.5.4}$$

for each $nk \le T$ and $0 < k \le \delta_0$. here,

$$\left\|U\right\|_h := h \sum_{j=-\infty}^{\infty} \left|U_j\right| \tag{2.5.5}$$

is an approximation of the norm of $L^1(R)$.

Since $U^n = H_k(U^{n-1}) = H_k H_k .....H_k(U^0) = H_k^n(U^0)$

stability holds if there exists $\beta > 0$ such that for each

$0 < \Delta t \le \delta_0 \quad and \quad 0 < \Delta x \le \delta_0$

$\left\|H_k V\right\|_h \le (1 + \beta k)\left\|V\right\|_h \quad \forall V.$

As a matter of fact

$$\left\|U^n\right\|_h = \left\|H_k^n(U^0)\right\| \le \left\|H_k^n\right\|\left\|U^0\right\|_h \le (1+\beta k)^n \left\|U^0\right\|_h \le e^{\beta T}\left\|U^0\right\|_h$$

for all k *and* n such that $nk \le T,$ hence (2.4.4) would follow.

Note in particular that the method is stable if $\left\|H_k\right\| \le 1,$ for then

$\left\|H_k^n\right\| \le \left\|H_k\right\|^n \le 1$ *for all* $n,k.$ More generally, some growth is allowed. For example, if

$$\left\|H_k\right\| \le 1+\alpha k \quad \textit{for all } k < k_0 \tag{2.5.6}$$

Then

$$\left\|H_k^n\right\| \le (1+\alpha k)^n \le e^{\alpha k n} \le e^{\alpha T}$$

for all k, n with $nk \le T.$

2.6 Convergence (definitions & examples)

Definition 2.6.1

A difference scheme is said to be convergent if

$$\max_{0 \le n \le T/k} \left\|u(.,t_n) - U^n\right\|_h \to 0 \quad \textit{as } k,h \to 0.$$

We illustrate some results of stability here below.

Example 2.6.1

Consider the Lax- Friedrichs method applied to the scalar advection equation $u_t + au_x = 0.$ we will show that the method is stable provided that *k* and *h* are related in such a way that

$$\left|\frac{ak}{h}\right| \le 1 \tag{2.6.1}$$

this is the stability restriction for the method. For the discrete operator H_k, we will show that $\|U^{n+1}\| \le \|U^n\|$ and hence $\|H_k\| \le 1$. exactly the same proof carries over to obtain the same bound for the continuous operator as well. (Take the norm and the triangle inequality):

We have

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{ak}{2h}(U_{j+1}^n - U_{j-1}^n) \tag{2.6.2}$$

and hence

$$\|U^{n+1}\| = h\sum_j |U_j^{n+1}|$$

$$\le \frac{h}{2}\left[\sum_j \left|(1-\frac{ak}{h})U_{j+1}^n\right| + \sum_j \left|(1+\frac{ak}{h})U_{j-1}^n\right|\right].$$

but the restriction (2.6.5) guarantees that

$$1-\frac{ak}{h} \ge 0, \quad 1+\frac{ak}{h} \ge 0$$

and so these can be pulled out of the absolute values, leaving

$$\|U^{n+1}\| \le \frac{h}{2}\left[(1-\frac{ak}{h})\sum_j |U_{j+1}^n| + (1+\frac{ak}{h})\sum_j |U_{j-1}^n|\right]$$

$$= \frac{1}{2}(1-\frac{ak}{h})\|U^n\| + \frac{1}{2}(1+\frac{ak}{h})\|U^n\|$$

$$= \|U^n\|$$

Hence $\|U^{n+1}\|_h \le \|U^n\|$ as desired.

This shows that (2.6.1) is sufficient for stability. In fact, it is also necessary.

Example 2.6.2

Next consider the upwind scheme (for a > 0) which reads

$$U_j^{n+1} = U_j^n - \lambda a(U_j^n - U_{j-1}^n) \tag{2.6.3}$$

then

$$\left\|U^{n+1}\right\|_h = \left\|(1-\lambda a)U_j^n + \lambda a U_{j-1}^n\right\|_h \le \left\|(1-\lambda a)U_j^n\right\|_h + \left\|(\lambda a)U_{j-1}^n\right\|_h$$

$$\le h\sum_j \left|(1-\frac{k}{h}a)U_j^n\right| + h\sum_j \left|\frac{k}{h}aU_{j-1}^n\right|$$

if we now assume that

$$0 \le \frac{ak}{h} \le 1 \tag{2.6.4}$$

then the coefficients of U_j^n *and* U_{j-1}^n are both nonnegative, therefore

$$\left\|U^{n+1}\right\|_h \le (1-\frac{ak}{h})\left\|U^n\right\|_h + (\frac{ak}{h})\left\|U^n\right\|_h = \left\|U^n\right\|$$

2.7 Conservative Methods for Nonlinear Problems

When we attempt to solve nonlinear conservation laws numerically we run into additional difficulties not seen in the linear equation. Moreover, the nonlinearity makes every thing harder to analyze. In spite of this, a great deal of progress has been made in recent years (see [RA2], (pp. 122-123)).

For smooth solutions to nonlinear problems, the numerical method can often be linearized and results from the theory of linear finite difference methods applied to obtain convergence results for nonlinear problems.

We have already seen some of the difficulties caused by discontinuous solutions even in the linear case. For nonlinear problems there are additional difficulties that can arise:

(a) The method might be "nonlinearly unstable", i.e., unstable on the nonlinear problem even though linearized versions appear to be stable.

(b) The method might converge to a function that is not a weak solution of our original equation (i.e., does not satisfy the entropy condition).

The fact that we might converge to a function that is not a weak solution at all is more puzzling, but goes back to the fact that it is possible to derive a variety of conservation laws that are equivalent for smooth solutions but have different weak solutions. For example, the PDEs

$$u_t + (\frac{1}{2}u^2)_x = 0 \tag{2.7.1}$$

and

$$(u^2)_t + (\frac{2}{3}u^3)_x = 0 \tag{2.7.2}$$

have exactly the same smooth solutions, but the Rankine-Hugoniot condition gives different shock speeds, and hence different weak solutions.

Consider a finite difference method that is consistent with one of these equations, say (2.7.1), using the same definition of consistency as for linear problems (expand in Taylor series). Then the method is also consistent with (2.7.2) since the Taylor series expansion gives the same result in either case. So the method is consistent with both (2.7.1) and (2.7.2) and while we might then expect the method to converge to a function that is a weak solution of both, that is impossible when the two weak solutions differ.

Example 2.7.1.

If we write Burgers' equation (2.6.1) in the quasilinear form

$$u_t + uu_x = 0 \tag{2.7.3}$$

Then a natural finite difference method, obtained by a minor modification of the upwind method for $u_t + au_x = 0$ (and assuming $U_j^n \geq 0$ for all j , n) is

$$U_j^{n+1} = U_j^n - \frac{k}{h} U_j^n (U_j^n - U_{j-1}^n) \tag{2.7.4}$$

The method (2.7.3) is adequate for smooth solutions but will not, in general, converge to a discontinuous weak solution of Burgers' equation (2.6.1) as the grid is refined. Consider, for example, the data which in discrete form gives

$$U_j^0 = \begin{cases} 1 & j < 0 \\ 0 & j \geq 0. \end{cases} \tag{2.7.5}$$

Then it is easy to verify from (2.7.4) and (2.7.5) that $U_j^1 = U_j^0$ for all j and n = 0. this happens in every successive step as well and so $U_j^n = U_j^0$ for all j, regardless of the step size k and h. as the grid is refined, the numerical solution thus converges very nicely to the function $u(x,t) = u_0(x).$ this is not a weak solution of (2.7.1) (or of (2.7.2) either).

In this example the solution is obviously wrong, but similar behavior is seen with other initial data that may give reasonable looking results that are incorrect. Figure (2.7.1) shows the true and computed solutions at time t=1 with Riemann data $u_l = 1.2$ and $u_r = 0.4.$ we get a nice looking solution propagating at entirely the wrong speed.

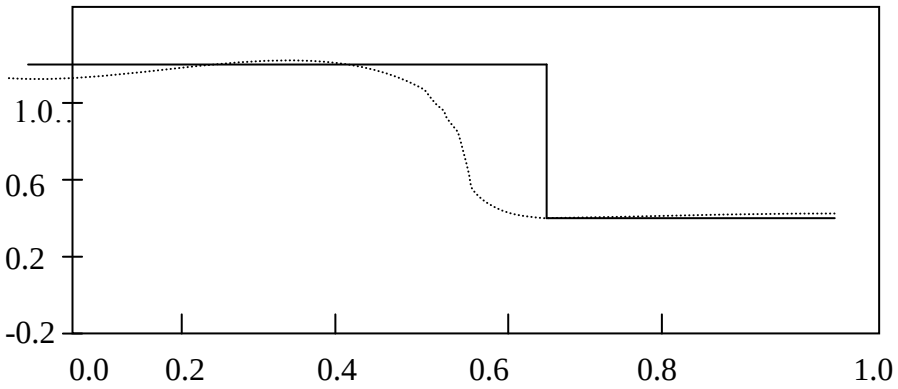


Figure 2.7.1 true and computed solutions to burgers' equation using a non-conservative method.

Chapter Three
The Generalized Riemann Problems Method

This chapter introduces the GRP method in the context of the scalar conservation law $u_t + f(u)_x = 0$. We start in section 3.1 with the classical first- order (conservative) "Godunov Scheme," which leads (sections 3.2, 3.3, 3.4) naturally to its second-order GRP extension. Section 3.5 contains a number of numerical (one dimensional) examples, for linear and nonlinear equations.

3.1 Godunov's Method

In 1959, Godunov proposed a way to make use of the characteristic information within the framework of a conservative method. Rather than attempting to follow characteristics backwards in time, Godunov suggested solving Riemann problems forward in time. Solution to Riemann problems are relatively easy to compute, give substantial information about the characteristic structure, and lead to conservative methods since they are themselves exact solutions of the conservation laws and hence conservative. The basic idea of Godunov scheme is to compose the global solution by the exact solution of the local Riemann problems. For given initial values $u_0 \in L^1(R)$ we define

$$U_j^0 := \frac{1}{h}\int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x)dx \quad , \qquad x_r = rh \tag{3.1.1}$$

Now let us assume that we have already computed the approximation $(U_j^n)\ n \in N$ for the time t_n, where U_j^n is also constant on (x_{j-1}, x_{j+1}) for all $j \in N$.

On each cell (x_{j-1}, x_j) for all $j \in N$ we determine the exact solution of the Riemann problem for

$$u_t + f(u)_x = 0, \quad \text{on} \quad R \times [t_n, t_{n+1}] \tag{3.1.2}$$

with respect to the initial condition

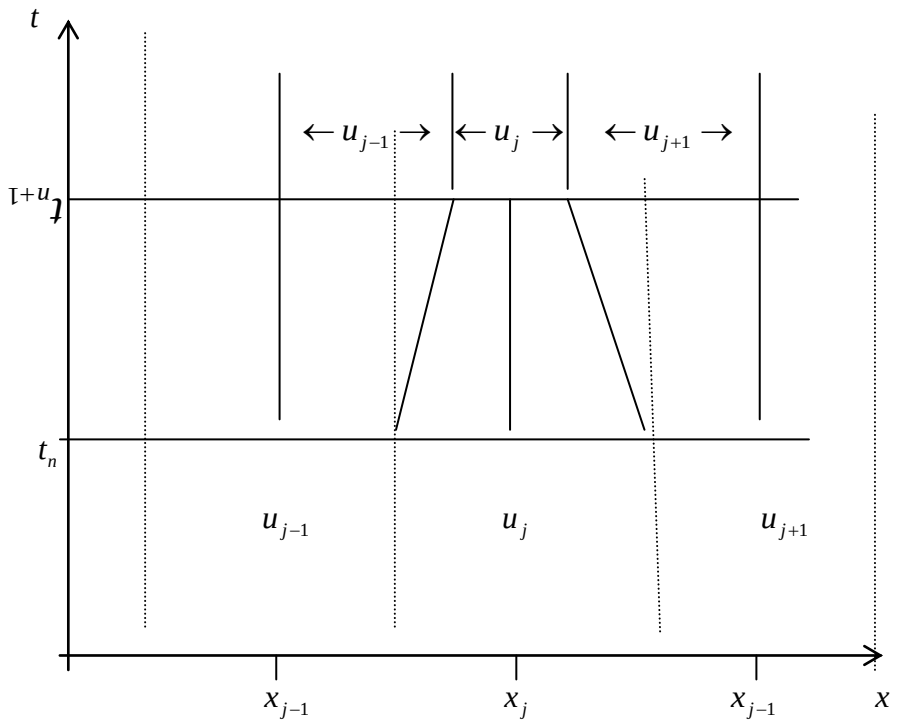
$$u(x, t_n) = \begin{cases} u_{j-1} & \text{if } x < x_{j-\frac{1}{2}} \\ u_j & \text{if } x > x_{j-\frac{1}{2}} \end{cases} \tag{3.1.3}$$

We denote this solution by $u(x,t;u_{j-1},u_j)$. In order to ensure that the neighboring solution $u(x,t;u_{j-1},u_j)$ *and* $u(x,t;u_j,u_{j+1})$ cannot influence each other, we have to assume that the shocks with

$$S_{j-\frac{1}{2}} = \frac{f(u_j) - f(u_{j-1})}{u_i - u_{i-1}} \quad \textit{and} \quad S_{j+\frac{1}{2}} = \frac{f(u_{j+1}) - f(u_j)}{u_{j+1} - u_j}$$

must not intersect. This can be obtained if (see Figure 3.1.1)

$$\left| S_{j-\frac{1}{2}} \right| \frac{k}{h} \le \frac{1}{2}, \left| S_{j+\frac{1}{2}} \right| \frac{k}{h} \le \frac{1}{2}, \quad \text{where} \quad k = \Delta t,\ h = \Delta x.$$



(Figure 3.1.1)

or

$$\frac{k}{h}\sup_{u\in R}|f'(u)| \le \frac{1}{2}. \tag{3.1.4}$$

The condition (3.1.4) is again the Courant, Friedrichs, Lewy condition or CFL condition.

If (3.1.4) is satisfied, the solution $u(x,t;u_{j-1},u_j)$ of the local Riemann problem (3.1.2), (3.1.3) uniquely defines a function v on $R\times[t_n,t_{n+1}]$ such that

$$v(x,t) := \begin{cases} u(x,t;u_{j-1},u_j) & \text{if } x_{j-\frac{1}{2}} < x < x_j,\ t_n \le t \le t_{n+1} \\ u(x,t;u_j,u_{j+1}) & \text{if } x_j < x < x_{j+\frac{1}{2}},\ t_n \le t \le t_{n+1} \end{cases}$$

As for the initial values, we have to ensure that the approximation U^{n+1} at time t_{n+1} is constant on $(x_{j-1/2}, x_{j+1/2})$ for all $j \in N$. Therefore we define

$$U_j^{n+1} := \frac{1}{h} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} v(x, t_{n+1}) dx . \tag{3.1.5}$$

This means U_j^{n+1} is the mean value of v on $(x_{j-1/2}, x_{j+1/2})$ and therefore contains parts of $u(x,t;u_{j-1},u_j)$ *and* $u(x,t;u_j,u_{j+1})$. since v is an exact solution on $(x_{j-1/2}, x_{j+1/2})$, we get (see 1.2.1)

$$\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} v(x, t_{n+1}) dx = \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} v(x, t_n) dx - \int_{t_n}^{t_{n+1}} f(v(x_{j+\frac{1}{2}}, s)) ds + \int_{t_n}^{t_{n+1}} f(v(x_{j-\frac{1}{2}}, s)) ds .$$

Using

$$\begin{aligned} v(x_{j+1/2}, t) &= u(x_{j+1/2}, t; u_j, u_{j+1}) =: u_{j+\frac{1}{2}} \\ v(x_{j-1/2}, t) &= u(x_{j-1/2}, t, u_{j-1}, u_j) =: u_{j-\frac{1}{2}} \end{aligned},$$

We obtain

$$U_j^{n+1} = U_j^n - \lambda (f_{j+\frac{1}{2}}^{n+\frac{1}{2}} - f_{j-\frac{1}{2}}^{n+\frac{1}{2}}) \tag{3.1.6}$$

where $\lambda = \frac{k}{h}$ and $f_{j+1/2}^{n+1/2}$ is an approximation for the average flux $\frac{1}{k} \int_{t_n}^{t_{n+1}} f(u(x_{j+1/2}, t)) dt.$

The scheme (3.1.6) is called the GODUNOV scheme.

3.2 Introduction to Generalized Riemann Problems

The GRP method is a high-resolution numerical approximation of the solution to a conservation law of the form

$$\frac{\partial u}{\partial t}+\frac{\partial}{\partial x} f(u)=0 \qquad x\in R,\ t>0, \tag{3.2.1}$$

$$u(x,0)=u_0(x) \qquad x\in R, \tag{3.2.2}$$

we always assume that $f(u)$ is strictly convex, $f''(u)\geq \mu>0.$ This method is a natural extension to the Godunov (upwind) scheme which we studied in section 3.1. As before, we take a uniform spatial grid $x_j = j\Delta x, \ -\infty < j < \infty,$ and uniformly spaced time levels $t_{n+1}=(n+1)k=nk+k=t_n+k, \quad t_0=0$. We refer to the interval $(x_{j-1/2}, x_{j+1/2})$ as "cell j", and $x_{j\pm 1/2}$ as its "cell boundaries".

Given the approximating functions $U^1(x),.....,U^n(x),$ a numerical scheme consists in constructing $U^{n+1}(x)$, approximating $u(x,t_{n+1})$, the functions $U^n(x)$ are piecewise-constant (Godunov). Where $U^n(x)$ are piecewise-linear for the (GRP). Their averages over cell j are denoted by U_j^n.

Our starting point is the balance equation (1.2.1), to be used over the rectangle $[x_{j-1/2}, x_{j+1/2}]\times[t_n, t_{n+1}]$. Since U_j^n, U_j^{n+1} are supposed to be the average values of the approximating function over bottom and top respectively, the discrete version of (1.2.1) should be

$$U_j^{n+1}=U_j^n-\lambda(f_{j+1/2}^{n+1/2}-f_{j\_1/2}^{n+1/2}) \tag{3.2.3}$$

where $\lambda = \frac{k}{\Delta x}$ and $f_{j+1/2}^{n+1/2}$ is an approximation for the average flux $\frac{1}{k}\int_{t_n}^{t_{n+1}} f(u(x_{j+1/2},t)dt.$

Definition 3.2.1:

The term $f_{j-1/2}^{n+1/2}$ is called the "numerical flux" at the boundary $x_{j+1/2}$ over the time interval $[t_{n,}t_{n+1}]$.

Clearly, once the numerical fluxes are known, the numerical scheme is fully determined.

3.3 Godunov Scheme For Nonlinear Equations

In this thesis we adapt the approach suggested by Godunov as mentioned before. In the present nonlinear case, it can be described as follows:

Take the function $U^n(x)$ as a piecewise constant, with

$$U^n(x) = U_j^n \qquad x_{j-1/2} < x < x_{j+1/2} \tag{3.3.1}$$

Let $\tilde{u}(x,t)$ be the weak solution 3.2.1 for $t \geq t_n$, subject to the initial data $U^n(x)$ at $t = t_{n.}$. Now evaluate the numerical flux as

$$f_{j+1/2}^{G,n+1/2} = \frac{1}{k}\int_{t_n}^{t_{n+1}} f(\tilde{u}(x_{j+1/2},t)dt, \qquad -\infty < j < \infty \tag{3.3.2}$$

The numerical flux associated with Godunov method. The main idea in the application of (3.2.3) with $(f_{j+1/2}^{n+1/2} = f_{j+1/2}^{G,n+1/2})$ is that if k is sufficiently small then

$$\tilde{u}(x_{j+1/2},t) = \text{constant}, \qquad t \in [t_n, t_{n+1}] \tag{3.3.3}$$

so that (3.3.2)is easily evaluated. This is in full agreement with the linear case

[$f(u) = au,\ \ a > 0$]

$$f_{j+1/2}^{G,j+1/2} = \frac{1}{k}\int_{t_n}^{t_{n+1}} f(\tilde{u}(x_{j+1/2},t)dt = \frac{1}{k}\int_{t_n}^{t_{n+1}} a\tilde{u}\ dt = \frac{a\tilde{u}}{k}[t_{n+1} - t_n] = a\tilde{u} = aU_j^n$$

in order to give a more precise meaning to (3.3.3), we set

$$M_n = \sup_{-\infty < j < \infty} \left|U_j^n\right| < \infty, \tag{3.3.4}$$

and let $k = \Delta t$ satisfy the CFL condition ($f'(u) = a$)

$$k \max_{|u| \le M_N} \left| f'(u)\right| < \Delta x. \tag{3.3.5}$$

Observe that near every cell-boundary $x_{j+1/2}$ the solution $\tilde{u}(x,t)$ is a "Riemann solution" $R(\frac{x - x_{j+1/2}}{t - t_n}; U_j^n, U_{j+1}^n)$ associated with the initial data U_j^n, U_{j+1}^n(See Fig 3.3.1)

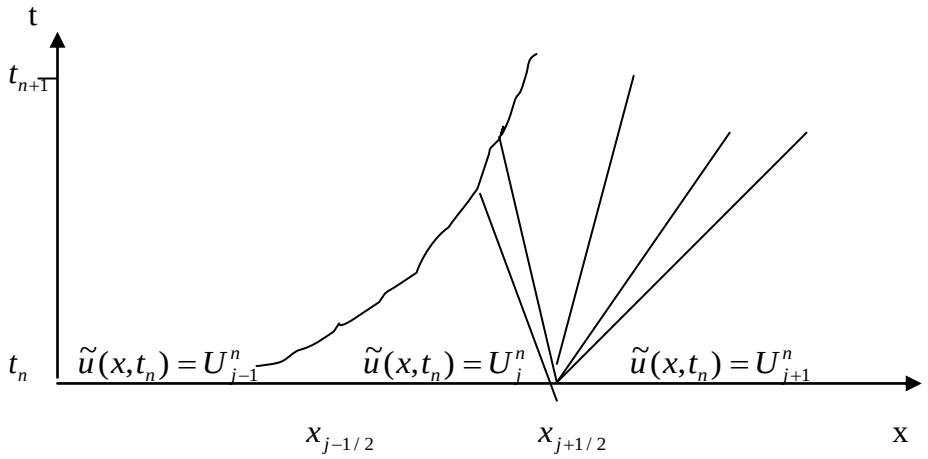


Figure 3.3.1: Wave pattern for piecewise-constant initial data.

The speed of all waves emanating from the points $x_{j+1/2}$, $-\infty < j < \infty$ are bounded by $S_n = \max_{|u| < M_n} \left| f'(u)\right|$ (3.3.6)

The CFL condition (3.3.5) entails the following important conclusion.

Remark

Under the CFL condition (3.3.5), a wave issuing from $(x_{j+1/2},t_n)$ does not reach any other cell-boundary $(x_{j+1/2},t)$ within the time interval $[t_n,t_{n+1}]$.

So the solution $\tilde{u}$ satisfies, for every $-\infty < j < \infty$,

$$\tilde{u}(x_{j+1/2},t) = R(0,U_j^n,U_{j+1}^n) \qquad t_n \le t \le t_{n+1} \tag{3.3.7}$$

and that the numerical flux is given by

$$f_{j+1/2}^{G,n+1/2} = f(R(0,U_j^n,U_{j+1}^n)), \tag{3.3.8}$$

and the balance equation (1.2.1) reads

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{u}(x,t_{n+1}) = U_j^n - \lambda(f_{j+1/2}^{n+1/2} - f_{j-1/2}^{n+1/2}). \tag{3.3.9}$$

In the linear case $f(u) = au, \quad a > 0$ we get $f_{j+1/2}^{G,n+1/2} = aU_j^n$ so that equation (3.2.3) yields the upwind scheme

$[U_j^{n+1} = U_j^n - a\lambda(U_j^n - U_{j-1}^n)]$.

Definition 3.3.1: (The Godunov Scheme).

The scheme given by

$$U_j^{n+1} = U_j^n - \lambda[f(R(0;U_j^n;U_{j+1}^n)) - f(R(0;U_{j-1}^n;U_j^n))]$$
$$U_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{u}_0(x)dx \tag{3.3.10}$$
$$\lambda = \frac{k}{\Delta x}$$

is called the "Godunov scheme" for the approximation of the conservation law (3.2.1).

The first and most fundamental question to be asked about the Godunov scheme (as well as any other approximating scheme) concerns its convergence to the exact weak solution of (3.2.1), (3.2.2). In theorem (2.3.1) and theorem (2.3.4) we have seen that, in the linear case $f(u) = au$, the CFL condition (3.3.5) is a necessary and sufficient condition for convergence in $L^1_{loc}(R)$. The idea of convergence in $L^1_{loc}(R)$ is very reasonable, especially when dealing with discontinuous solutions. It allows for phenomena common to numerical approximation, such as oscillations or "spurious waves," as long as they tend to zero in the mean as the grid is refined $(k = \Delta t \to 0)$, over any fixed finite interval.

Considering the convergence properties of the Godunov scheme in the case of a nonlinear flux function $f(u)$, we can cite the following theorem.

Theorem 3.3.1

Let $u_0(x) \in L^1(R) \cap L^\infty(R)$ and assume further that $u_0(x)$ is a function of finite total variation. Let $u(x.t)$ be the unique entropy solution to (3.2.1), (3.2.2), and let $(\{U_j^n\}, -\infty < j < \infty, n \geq 0)$ be obtained by the Godunov scheme (3.3.10). Then, under the CFL condition (3.3.5), $\{U_j^n\}$ converges to $u(x,t)$ in $L^1_{loc}(R)$ (see definition 2.3.1).

Proof: (see [BF], (pp. 320-329))

Claim 3.3.2 ("Maximum principle for the Godunov scheme")

Given the scheme (3.3.10), and using the notation $M_n = \sup_{-\infty<j<\infty} \left|U_j^n\right| < \infty,$ we have

$$M_{n+1} \le M_n \le ..... \le M_0.$$

Proof

According to 3.2.3 and 3.3.2 U_j^{n+1} is an average (over cell j) of the exact solution $\tilde{u}(x,t_{n+1}),$ subject to the initial data $\tilde{u}(x,t_n) = U^n(x).$ Thus, by the maximum principle for an exact solution theorem

$(\sup_{x\in R} u(x,t) \le \sup_{x\in R} u_0(x), \quad \inf u(x,t) \ge \inf u_0(x))$

$$M_{n+1} = \sup_j \left|U_j^{n+1}\right| \le \sup_x \left|\tilde{u}(x,t_{n+1})\right| \le \sup_x \left|\tilde{u}(x,t_n)\right| = \sup_j \left|U_j^n\right| = M_n. \qquad \square$$

3.4. Second-Order Accuracy Methods

The GRP may be introduced as follows: We consider the balance eq. (1.2.1). However, we assume that at $t = t_n$ the initial distribution is linear in each cell j. This proposed by Van Leer. Retaining the notation U_j^n for the average cell values, we therefore assume that

$$U^n(x) = U_j^n + (x - x_j) S_j^n \qquad x_{j-1/2} < x < x_{j+1/2}, \tag{3.4.1}$$

where S_j^n is the slope of the linear segment $U^n(x)$ in cell j . Note that at cell-boundaries $x_{j+1/2}$ we have in general a jump discontinuity in the values of $U^n(x)$ (namely, between $U_j^n + \frac{\Delta x}{2} S_j^n$ and $U_{j+1}^n - \frac{\Delta x}{2} S_{j+1}^n$), and also in the values of the slopes (S_j^n, S_{j+1}^n).

Let $\tilde{u}(x,t),\quad t_n \le t \le t_{n+1},$ be the weak solution to (3.2.1), subject to the initial data $\tilde{u}(x,t_n) = U^n(x)(as\ in(3.4.1))$. The values $\tilde{u}(x_{j+1/2},t)$ at cell-boundaries now depend on t, even for $t - t_n$ small, in contrast to the previous (Godunov) case, as given (3.3.7). This is of course because now $U^n(x)$ is not constant on either side of $x_{j+1/2},$ so we cannot expect a Riemann solution there. It follows that in the present case the difference scheme (3.2.3) can only be written with numerical fluxes $f_{j+1/2}^{n+1/2}$ which are only approximately equal to $\frac{1}{k}\int_{t_n}^{t_{n+1}} f(\tilde{u}(x_{j+1/2},t)dt.$

Specifically, we assume now that the numerical fluxes $f_{j+1/2}^{n+1/2}$ satisfy,

$$f_{j+1/2}^{n+1/2} - f_{j-1/2}^{n+1/2} = \frac{1}{k}\int_{t_n}^{t_{n+1}} [f(\tilde{u}(x_{j+1/2},t)) - f(\tilde{u}(x_{j-1/2},t))dt + O(k^3)\,. \tag{3.4.2}$$

We now define the new averages $U_j^{n+1},\ -\infty < j < \infty,$ by

$$U_j^{n+1} = U_j^n - \lambda(f_{j+1/2}^{n+1/2} - f_{j-1/2}^{n+1/2}) \tag{3.4.3}$$

combining (3.4. 1)-(3.4. 2) with the balance equation 1.2.1 over $[x_{j-1/2,}x_{j+1/2}]\times[t_n,t_{n+1}]$ we get from 3.4.13,

$$U_j^{n+1} - \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{u}(x,t_n)dx = -\int_{t_n}^{t_{n+1}} [f(\tilde{u}(x_{j+1/2},t)) - f(\tilde{u}(x_{j-1/2},t))]dt + O(k^3)$$

$$U_j^{n+1} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{u}(x,t_{n+1})dx + O(k^3)\,,\quad -\infty < j < \infty\,. \tag{3.4.4}$$

Note that for the special initial data 3.3.1 equation (3.4.2), (3.4.4) were satisfied with no truncation error. In terms of definition (2.2.1) we now conclude that the scheme is of second order accuracy (p =2).

The foregoing derivation was based on the hypothesis (3.4.2). To study its validity we prove the following claim.

Claim 3.4.1

Let $\tilde{u}(x,t)$ be smooth in $x \in [x_{j-1/2}, x_{j+1/2}]$ and $t \geq t_n$. Then (3.4.2) is satisfied with

$$f_{j+1/2}^{n+1/2} = f(\tilde{u}(x_{j+1/2}, t_n)) + \frac{k}{2}\frac{\partial}{\partial t} f(\tilde{u}(x_{j+1/2}, t_n)) \tag{3.4.5}$$

(Namely, $f_{j+1/2}^{n+1/2}$ = the linear approximation (in t) of $f(\tilde{u}(x_{j+1/2}, t))$ evaluated at the midpoint $(t_{n+1/2} = t_n + \frac{k}{2})$.

Proof

This is a direct consequence of Taylor's theorem and the fact that $\lambda = k/\Delta x = constant$. To simplify notation, we introduce the functions:

$$g_{j+1/2}(t) = f(\tilde{u}(x_{j+1/2}, t), \qquad -\infty < j < \infty$$

Writing

$$g_{j+1/2}(t) = g_{j+1/2}(t_{n+1/2}) + g'_{j+1/2}(t_{n+1/2})(t - t_{n+1/2}) + \frac{1}{2} g''_{j+1/2}(t_{n+1/2})(t - t_{n+1/2})^2 + O(k^3) \qquad t_n \leq t \leq t_{n+1}$$

We obtain by integration,

$$\int_{t_n}^{t_{n+1}} [g_{j+1/2}(t) - g_{j-1/2}(t)]dt = [g_{j+1/2}(t_{n+1/2}) - g_{j-1/2}(t_{n+1/2})k + \frac{1}{24}[g''_{j+1/2}(t_{n+1/2}) - g''_{j-1/2}(t_{n+1/2})]k^3 + O(k^4) \tag{3.4.6}$$

However, $g''_{j+1/2}(t_{n+1/2}) - g''_{j-1/2}(t_{n+1/2}) = O(k)$ and by (3.4.5)

$$g_{j+1/2}(t_{n+1/2}) - g_{j-1/2}(t_{n+1/2}) = f_{j+1/2}^{n+1/2} - f_{j-1/2}^{n+1/2} +$$
$$\frac{1}{8}[g''_{j+1/2}(t_n) - g''_{j-1/2}(t_n)]k^2 + O(k^3)$$
$$= f_{j+1/2}^{n+1/2} - f_{j-1/2}^{n+1/2} + O(k^3)$$

Inserting these relations in 3.4.6 yields 3.4.2 □

Remark

It is clear that from the proof that we could replace 3.4.5 by any other expression which approximates, up to $O(k^2)$, the value $f(\tilde{u}(x_{j+1/2}, t_{n+1/2}))$, such as

$$f_{j+1/2}^{n+1/2} = f\left(\tilde{u}(x_{j+1/2}, t_n) + \frac{k}{2}\frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t_n)\right) \qquad (3.4.7)$$

Although claim 3.4.1is of a formal value (as the solution $\tilde{u}(x,t)$ is generally not smooth), it provides the guideline to the construction of the GRP numerical fluxes. Because of the fundamental importance of this construction in the present work, we shall first list the technical steps, and then follow by a detailed discussion.

Construction 3.4.1 (GRP Algorithm)

Given the piecewise-linear distribution $U^n(x)$ (3.4.1) and $\Delta t = k$ such that $\lambda = \frac{k}{\Delta x}$ satisfies the CFL condition (3.3.5) (with $M_n = \sup_{x \in R}\left|U^n(x)\right|$), construct $U^{n+1}(x)$ (which should approximate $\tilde{u}(x, t_n + k)$) as follows.

Step 1. At every cell-boundary $x_{j+1/2}$ evaluate $U^n(x)$ on the two sides by

$$U^n_{j+1/2,\pm} = \lim_{\delta \to 0^+} U^n(x_{j+1/2} \pm \delta) = \begin{cases} U^n_{j+1} - \dfrac{\Delta x}{2} s^n_{j+1}, & \text{"+"} \\ U^n_j + \dfrac{\Delta x}{2} s^n_j, & \text{"–"} \end{cases}$$

Then determine the Riemann solution

$$U^n_{j+1/2} = R(0; U^n_{j+1/2,-}, U^n_{j+1/2,+}). \tag{3.4.8}$$

Note that,

$$U^n_{j+1/2} = \begin{cases} U^n_{j+1/2,-} & \textit{wave moves right, } f'(U^n_{j+1/2}) > 0, \\ U_{j+1/2,+} & \textit{wave moves left, } \quad f'(U_{j+1/2,+}) < 0, \\ u_{\min} & \textit{if } x_{j+1/2} \textit{ is a sonic point,} \\ & f'(U^n_{j+1/2,-}) \le 0 \le f'(U_{j+1/2,+}). \end{cases} \tag{3.4. 9}$$

Step 2. Determine the instantaneous time derivatives $\dfrac{\partial \tilde{u}}{\partial t}(x_{j+1/2,} t_n)$ by,

$$\frac{\partial \tilde{u}}{\partial t}(x_{j+1/2,} t_n) = \begin{cases} -f'(U^n_{j+1/2})\, s^n_j & \textit{if } f'(U^n_{j+1/2}) > 0, \\ -f'(U^n_{j+1/2})\, s^n_{j+1/2} & \textit{if } f'(U^n_{j+1/2}) < 0, \\ 0 & \textit{if } U^n_{j+1/2} = u_{\min}. \end{cases} \tag{3.4.10}$$

Then compute the approximate solution and numerical flux (see 3.4.5) at the midpoint $(x_{j+1/2}, t_{n+1/2})$ by,

$$\begin{aligned} U^{n+1/2}_{j+1/2} &= U^n_{j+1/2} + \frac{k}{2} \frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t_n), \\ f^{n+1/2}_{j+1/2} &= f(U^n_{j+1/2}) + \frac{k}{2} f'(U^n_{j+1/2}) \frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t_n) \end{aligned} \quad . \tag{3.4.11}$$

Step 3. Evaluate the new cell averages as in (3.4.3),

$$U^{n+1}_j = U^n_j - \lambda (f^{n+1/2}_{j+1/2} - f^{n+1/2}_{j-1/2}), \quad -\infty < j < \infty, \tag{3.4.12}$$

And the new slopes by,

$$U_{j+1/2}^{n+1} = U_{j+1/2}^{n} + k\frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t_n), \quad -\infty < j < \infty, \tag{3.4.13}$$

$$s_j^{n+1} = \frac{1}{\Delta x}(U_{j+1/2}^{n+1} - U_{j-1/2}^{n+1}).$$

The construction of the algorithm is not yet complete. We shall later supplement it (see 3.4.5) by a suitable "slope limiter", which ensures certain monotonicity properties of the new profile $U^{n+1}(x)$. However, we shall first make a few comments concerning this algorithm.

The basic hypothesis underlying the GRP construction is that the wave pattern associated with the solution $\tilde{u}(x,t)$ can be fully determined (for sufficiently small $k = \Delta t$) by the Riemann solutions of Step 1 (see (3.4.8). Of course, as has already been observed, a shock wave issuing from $x_{j+1/2}$ will not be (in general) self-similar. In other words, its trajectory will not be of constant slope. This is in contrast to the characteristic lines (comprising a centered rarefaction wave) which in the SCALAR case are always straight lines. However, the assumption here is that at each cell boundary $x_{j+1/2}$ the solution $\tilde{u}(x,t)$ consists of a single wave (shock for $U_{j+1/2,-}^{n} > U_{j+1/2,+}^{n}$, centered rarefaction otherwise), where instantaneous features at $x = x_{j+1/2,}$ $t = t_n$, (i.e. slopes of a shock trajectory or head and tail characteristics of a rarefaction) are completely determined by the Riemann solution

$$\mathrm{R}\left(\frac{x - x_{j+1/2}}{t - t_n}; U_{j+1/2,-}^{n}, U_{j+1/2,+}^{n}\right).$$

Also, the solution $\tilde{u}$ on the two sides of the shock, or inside and outside a centered rarefaction wave, is smooth, with a jump discontinuity across a shock trajectory or

jump discontinuities of the derivatives across the head and tail characteristics of a centered rarefaction wave. As a matter of fact, in the present case (a scalar conservation law with a strictly convex flux function) this assumption can be proved for the unique entropy solution. the CFL condition implies, as in the case of the Godunov scheme, that a wave issuing from a cell-boundary $x_{j+1/2}$ is limited (for $t \in [t_n, t_n + k]$) to the neighboring cells j ,j+1, not reaching their opposite boundaries $x_{j+3/2}, x_{j-1/2}$. In particular, the solution $\tilde{u}(x_{j+1/2}, t),\ t_n \le t \le t_{n+1}$, is not affected by the discontinuities at $x_{l+1/2},\ l \neq j$, and is therefore a smooth function of t. Its derivative $\frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t)$ should be interpreted as the limiting value of $\frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t),\ t > t_n\ as\ t \to t_n$. if the wave moves to the right, the segment $(x_{j+1/2}, t),\ t_n < t < t_{n+1}$ is contained, along with $(x, t_n),\ x_j < x < x_{j+1/2}$, in the same domain of smoothness of $\tilde{u}(x,t)$, hence $\tilde{u}(x,t)$ is a classical solution there, satisfying $\tilde{u}_t(x,t) = -f'(\tilde{u}(x,t))\tilde{u}_x(x,t)$. a similar consideration applies to the case where the wave moves to left. If $x_{j+1/2}$ is a sonic point, the line $x = x_{j+1/2}$ is characteristic, carrying the constant value $\tilde{u}(x_{j+1/2}, t) = u_{\min}$. We obtain therefore all three cases of Eq (3.4.10)

The evaluation of the numerical fluxes (3.4.11) follows the second- order approximation given by (3.4.5), where $\tilde{u}(x_{j+1/2}, t_n) = U^n_{j+1/2}$ is the limiting value (as $t \to t_n$) of $\tilde{u}(x_{j+1/2}, t),\ t > t_n$. The same linear approximation of $\tilde{u}(x_{j+1/2}, t)$ serves to determine the new slopes in (3.4.13)

Remark: (Accuracy of the slope computation)

As in claim (3.4.1), assume that $\tilde{u}(x_{j+1/2},t)$ is smooth in $[x_{j-1/2},x_{j+1/2}]\times[t_n,t_{n+1}]$, then

$$U_{j\pm1/2}^{n+1}=\tilde{u}(x_{j\pm1/2},t_{n+1})-\frac{1}{2}\tilde{u}_{tt}(x_{j\pm1/2},t_n).k^2+O(k^3)$$

$$=\tilde{u}(x_j,t_{n+1})\pm\tilde{u}_x(x_j,t_{n+1}).\frac{\Delta x}{2}$$

$$+\frac{1}{8}\tilde{u}_{xx}(x_j,t_{n+1}).\Delta x^2-\frac{1}{2}\tilde{u}_{tt}(x_{j\pm1/2},t_n).k^2+O(k^3).$$

Thus

$$s_j^{n+1}=\frac{1}{\Delta x}(U_{j+1/2}^{n+1/2}-U_{j-1/2}^{n+1/2})=\tilde{u}_x(x_j,t_{n+1})+O(k^2),$$

which is, naturally, less accurate than the computation of the cell averages U_j^{n+1}

Remark (zero slopes in GRP computation)

Observe that when the slopes s_j^n are set to zero for all cells j and at every time level t_n, the GRP computational scheme naturally reduces to Godunov scheme.

Remark (stationary shocks)

If the Riemann solution $R(\frac{x-x_{j+1/2}}{t-t_n};U_{j+1/2,-}^n,U_{j+1/2,+})$ yields a stationary shock along $x=x_{j+1/2}$, it means (by the Rankine-Hugoniot jump condition) that $f(U_{j+1/2,-})=f(U_{j+1/2,+}),U_{j+1/2,-}>U_{j+1/2,+}^n$. the shock speed is given by

$$\sigma(t)=\frac{f(\tilde{u}^+(x(t),t))-f(\tilde{u}^-(x(t),t)}{\tilde{u}^+(x(t),t)-\tilde{u}^-(x(t),t)},$$

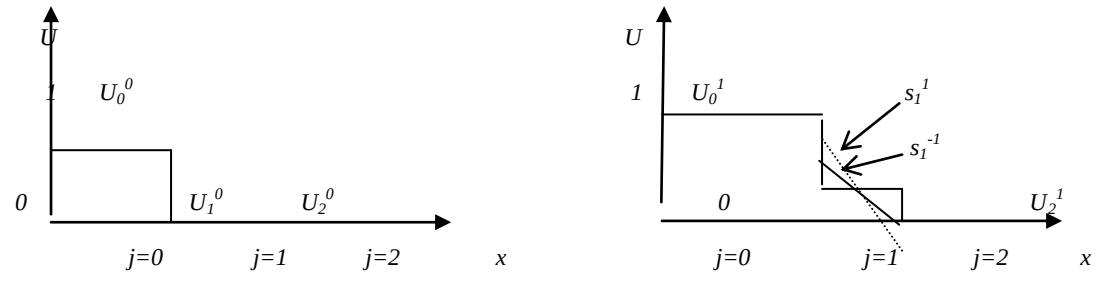
where x(t) is the shock trajectory $(x(t_n) = x_{j+1/2,}\ x'(t) = \sigma(t))$ and $\tilde{u}^{-}(resp.\ \tilde{u}^{+})$ is the value behind (resp. ahead of) the shock , $u^{\pm}(x(t_n), t_n) = U_{j+1/2,\pm}$.. Thus,

$$\sigma'(t)\Big|_{t=t_n} = \frac{-f'(U^n_{j+1/2,+})^2 s^n_{j+1} + f'(U_{j+1/2,-})^2 s^n_j}{U^n_{j+1/2,+} - U^n_{j+1/2,-}}$$

and the value of $\frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t_n)$ is determined according to whether $\pm\sigma'(t_n) > 0$.

The last technical step in the description of the GRP algorithm is concerned with a modification of the slope s_j^{n+1}. In the language common to numerical schemes, it is a "postprocessing" step applied to the new results $\left\{U_j^{n+1}, s_j^{n+1}\right\}_{-\infty<j<\infty}$.

It is a basic rule in all GRP calculations that the new averages U_j^{n+1}, as determined by (3.4.12), are never modified. Their values are obtained by the approximate implementation of the balance equation, which is viewed here as the basis of our methodology. On the other hand, the slopes are less accurately computed, using a discrete differentiation procedure (3.4.13). We can illustrate the need for a "postprocessing intervention" in their values by the following example (see Fig. 3.4.1)



(a) Initial data (unit step function)

(b) Integration by one time-step

Figure 3.4.1: First GRP time-integration cycle of a moving step.

Example 3.4.1

Let the initial data be $U_j^0 = 1$ (resp. $U_j^0 = 0$) for $j \le 0$ (resp. $j > 0$), and let

$f(u) = \frac{1}{2}u^2$, so that the solution is a shock wave moving at speed $\frac{1}{2}$, $\tilde{u}(x,t) = u_0(x - \frac{1}{2}t)$. If we use a time step $k = \Delta t$, it is easy to see that

$$U_j^1 = \begin{cases} 1, & j \le 0, \\ \frac{1}{2}\lambda, & j = 1, \\ 0, & j > 1 \end{cases}$$

The corresponding computed slopes s_j^1. satisfy

$$s_j^1 = \begin{cases} 0, & j \le 0 \\ -\frac{1}{\Delta x}, & j = 1 \\ 0, & j > 1 \end{cases}$$

Thus, if we are to retain all slopes s_j^1 the approximating function $U^1(x)$ in the cell j = 1, should be

$$U^1(x) = \frac{1}{2}\lambda - \frac{x - \Delta x}{\Delta x} = 1 + \frac{1}{2}\lambda - \frac{x}{\Delta x} \qquad \frac{\Delta x}{2} < x < \frac{3}{2}\Delta x. \tag{3.4.14}$$

Thus, $U^1(\frac{3}{2}\Delta x-) = \frac{1}{2}(\lambda - 1)$, which is negative. This is in contradiction with the "moving step" character of the exact (weak) solution.

From the mathematical point of view, the modification of the slopes $\{s_j^n\}_{j=-\infty}^{\infty}$ is needed for the control of the total variation of the approximating solution, in analogy with the total variation properties of the exact (weak) solution.

The modification of the slopes used in our GRP methodology is implemented as follows.

Construction 3.4.2 (GRP "Slope Limiter"). Given the computed slopes s_j^{n+1} (as in (3.4.13), set the final slope values $\bar{s}_j^{n+1}$ to be

$$\bar{s}_j^{n+1} = \frac{1}{\Delta x}\min\text{mod}[2(U_{j+1}^{n+1} - U_j^{n+1}), 2(U_j^{n+1} - U_{j-1}^{n+1}), \Delta x s_j^{n+1}], \tag{3.4.15}$$

where, for any three real numbers a, b, c,

$$\text{minmod[a,b,c]} = \begin{cases} \sigma \min(|a|,|b|,|c|) & \text{if } \sigma = \text{sgn}(a) = \text{sgn}(b) = \text{sgn}(c), \\ 0, & \text{otherwise.} \end{cases}$$

Geometrically speaking, our limiter reflects the minimal change (of s_j^{n+1}) needed to obtain the following "5-point monotonicity" (see Fig. 3.4.2):

form a monotonic sequence, then so do the five values $\{U_{j-1}^{n+1}, U_j^{n+1}, U_{j+1}^{n+1}\}$ If

$$\cdot\left\{U_{j-1}^{n+1}, U_j^{n+1} - \frac{\Delta x}{2}\bar{s}_j^{n+1}, U_j^{n+1}, U_j^{n+1} + \frac{\Delta x}{2}\bar{s}_j^{n+1}, U_{j+1}^{n+1}\right\}$$

If $U_j^{n+1} \geq \max(U_{j\pm1}^{n+1})$ *or* $U_j^{n+1} \leq \min(U_{j\pm1}^{n+1})$ *we set* $\bar{s}_j^{n+1} = 0$. Thus, at external points the slopes are set to zero, whereas elsewhere it is ensured that

(in the case of $U_{j-1}^{n+1} \leq U_j^{n+1} \leq U_{j+1}^{n+1}$)

$$U_j^{n+1} \geq \max(U_{j-1/2,\pm}^{n+1}), \qquad U_j^{n+1} \leq \min(U_{j+1/2,\pm}^{n+1}).$$

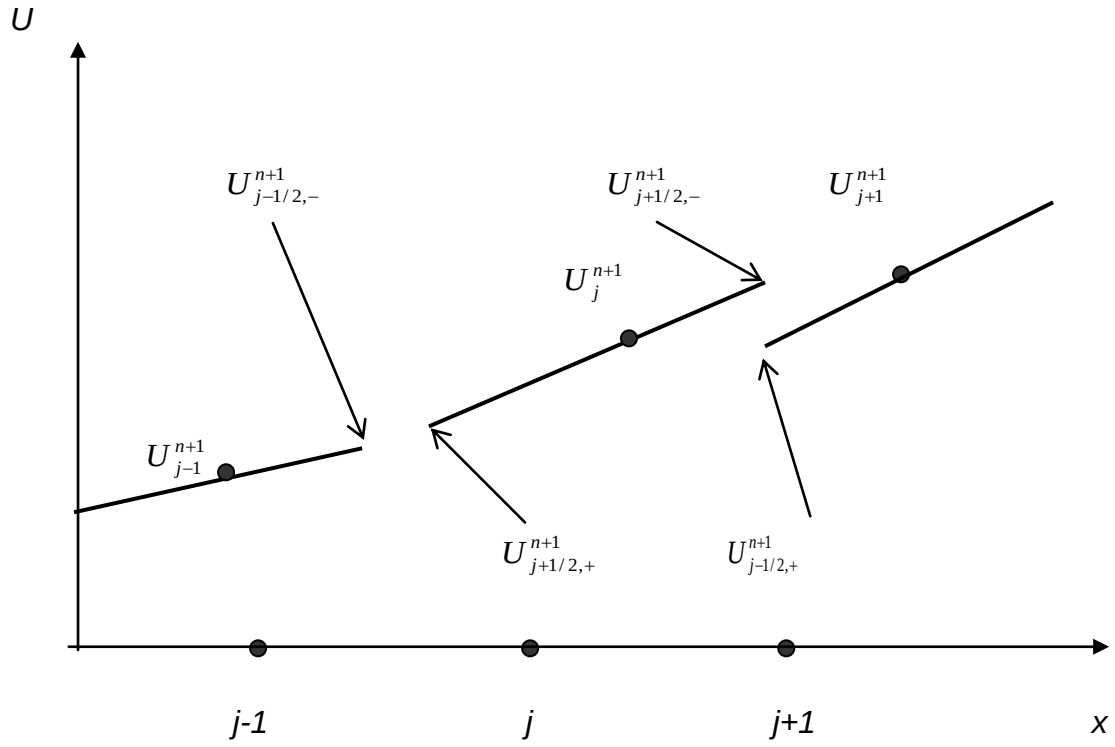


Figure 3.4.2: The "slope limiter" is a 5-point rule

Remark (Convergence of the GRP scheme)

The convergence of the first-order Godunov scheme was stated in theorem 3.3.1. At the time of writing this work, a similar convergence result has not yet been established for the GRP scheme. The main obstacle for convergence proof lies in the rather weak slope limiter as given in construction (3.4.2) (open problem).

3.5 1-D Sample problem

In this section we present numerical solution to scalar conservation laws, linear and nonlinear, in one space dimension. The initial data considered are sufficiently simple, so that the exact solution can be computed and compared to the finite-difference approximations. Two pairs of schemes were chosen for sample problems, one pair of first-order schemes and one pair of second-order schemes. The idea is to demonstrate the difference between the first-order Godunov scheme and its natural second-order extension – GRP scheme (construction 3.4.1). Then, for the sake of comparison, we use another typical scheme in each class. We selected the (first-order) Lax-Friedrichs scheme and the (second-order) Lax-Wendroff scheme.

The Linear Conservation Law

The equation to be considered here is

$$u_t + u_x = 0, \qquad u(x,0) = u_0(x). \tag{3.5.1}$$

The exact solution is given by the traveling wave $u(x,t) = u_0(x-t)$.

First-Order Schemes

We shall use the following pair of first-order schemes:

(a) The Godunov scheme, which in this case (since $a > 0$) is identical to the backward- difference scheme (2.2.20), as explained in claim (2.3.6)

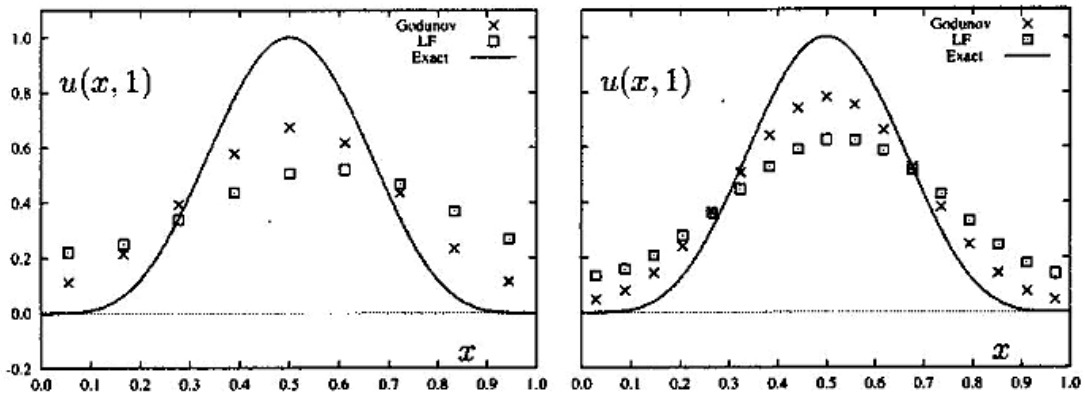
(b) The Lax-Friedrichs (LF) scheme, which in this case is given by (2.2.22).

In all the computations of this section we take constant (fixed) space and time steps, Δx and $k = \Delta t,$ respectively. Their ratio $\lambda = \frac{k}{\Delta x}$ satisfies the CFL condition, namely, $\lambda \leq 1.$

Two initial profiles $u_0(x)$ are considered, the first having smooth periodic data, and the second having step function data. These problems have been chosen for two reasons: (i) one of them has smooth data, and the other has discontinuous data (i.e., only a weak solution exists). (ii) Both problems are defined on R, yet can be solved numerically on some finite interval $x_1 < x < x_2,$ producing the same finite-difference solution that would have been obtained on an unbounded interval of x. The smooth initial data are

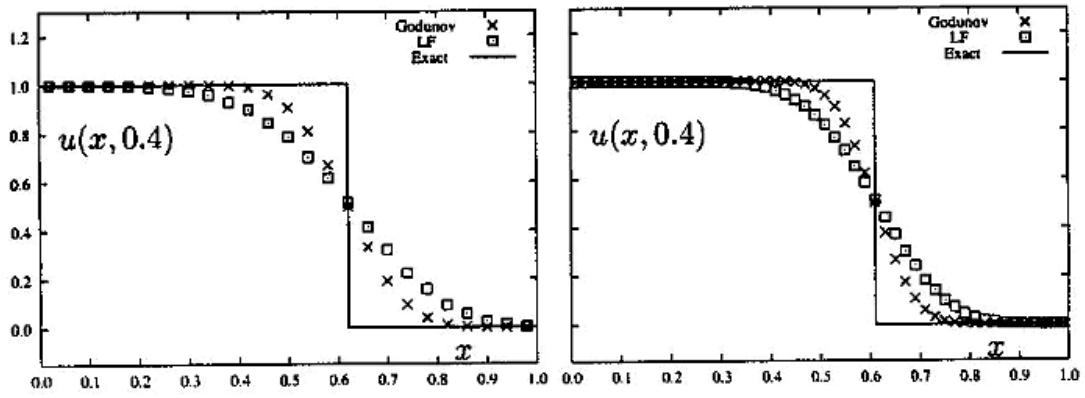
$$u_0(x) = \sin^4(\pi x). \tag{3.5.2}$$

This is a periodic function with a period L=1, so that at time t=1, $u_0(x)$ has propagated exactly through one period. The numerical solution is performed with periodic boundary conditions. Figure (3.5.1) shows the results of such computation, using a coarse grid of $\Delta x = 1/9$ and a refined grid with $\Delta x = 1/17.$ The constant ratios are $\lambda = 0.7500$ *and* $\lambda = 0.7391,$ respectively (corresponding to integration by 12 and 23 time steps, respectively).



(a) $\Delta x = 1/9$. (b) $\Delta x = 1/17$.

Figure 3.5.1 First-order integration of $u_t + u_x = 0$, with initial data $u_0(x) = \sin^4(\pi x)$.



(a) $\Delta x = 1/25$. (b) $\Delta x = 1/50$.

Figure 3.5.2 First-order integration of $u_t + u_x = 0$, with unit step-function initial data.

As is evident from Figure 3.5.1, both finite-difference approximations are rather far from the exact solution, with a smaller error in the finer grid computation. Furthermore, the Godunov scheme clearly produces more accurate results than the LF scheme. This can be interpreted as indicating that although both schemes are

first-order accurate, the Godunov solution is less “smearing” than the LF one and therefore more accurate.

For the step function case, the function is $u_0(x) = 1 \quad for \quad x < x_0$ and $u_0(x) = 0$ for $x > x_0$. The numerical integration is performed in the rang $0 < x < 1$ until a time t =0.4, with $\lambda = 0.5$. The boundary conditions for the time interval $0 < t < 0.4 \quad are\, u(0) = 1, u(1) = 0$. Two grids were used a coarse grid with $\Delta x = 0.04$ (20 time step) and $x_0 = 0.22$ and a fine grid with $\Delta x = 0.02$ (40 time steps) and $x_0 = 0.21$. Referring to the exact solution, we note that the discontinuity is positioned at a mid-cell point at the initial time, as well as at the final time. The datum in cell $x_j = x_0$ is $U_j^0 = 0.5$, in accordance with definition 3.3.1 of Godunov scheme. Again, we observe in Figure 3.5.2 that the Godunov scheme produces more accurate results than the LF scheme. It is also noted that both coarse and fine grid solutions seem to approximate the moving step quite accurately in the mean; i.e., the numerical values are symmetrically distributed about the step, and, moreover, the sharp step is “spread” over about 8 cells in the first grid and over about 12 cells in the second. The width of the “step-spreading” appears to be proportional to $\sqrt{N}$, where N is the number of time integration cycles. This spreading effect is typical of a linear conservation law.

Second-Order Schemes

Turning to second-order schemes, our primary interest is GRP, but for comparison we also consider the Lax-Wendroff (LW) scheme (2.2.23). It is readily verified from

equation (2.2.23) that if $u_0(x)$ is of compact support in R, then $\sum_j U_j^{n+1} = \sum_j U_j^n$. This means that the LW scheme is conservative, although not upwind.

The GRP scheme, by contrast, is both upwind and conservative. It is given by adapting construction 3.4.1 to the case $f(u) = u$. Thus, the Riemann solution is simply a moving step solution, so that

$$U_{j+1/2}^n = R(0; U_{j+1/2,-}^n, U_{j+1/2,+}^n) = U_{j+1/2,-}^n. \tag{3.5.3}$$

It follows from equation (3.4.10) that

$$\frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t_n) = -s_j^n = -\frac{1}{\Delta x}(U_{j+1/2,-} - U_{j+1/2,+}^n); \tag{3.5.4}$$

Hence, as in equation (3.4.11),

$$U_{j+1/2}^{n+1/2} = U_{j+1/2}^n - \frac{k}{2} s_j^n, \tag{3.5.5}$$

$$f_{j+1/2}^{n+1/2} = U_{j+1/2}^{n+1/2}. \tag{3.5.6}$$

The resulting GRP scheme is, as in equation (3.4.12)

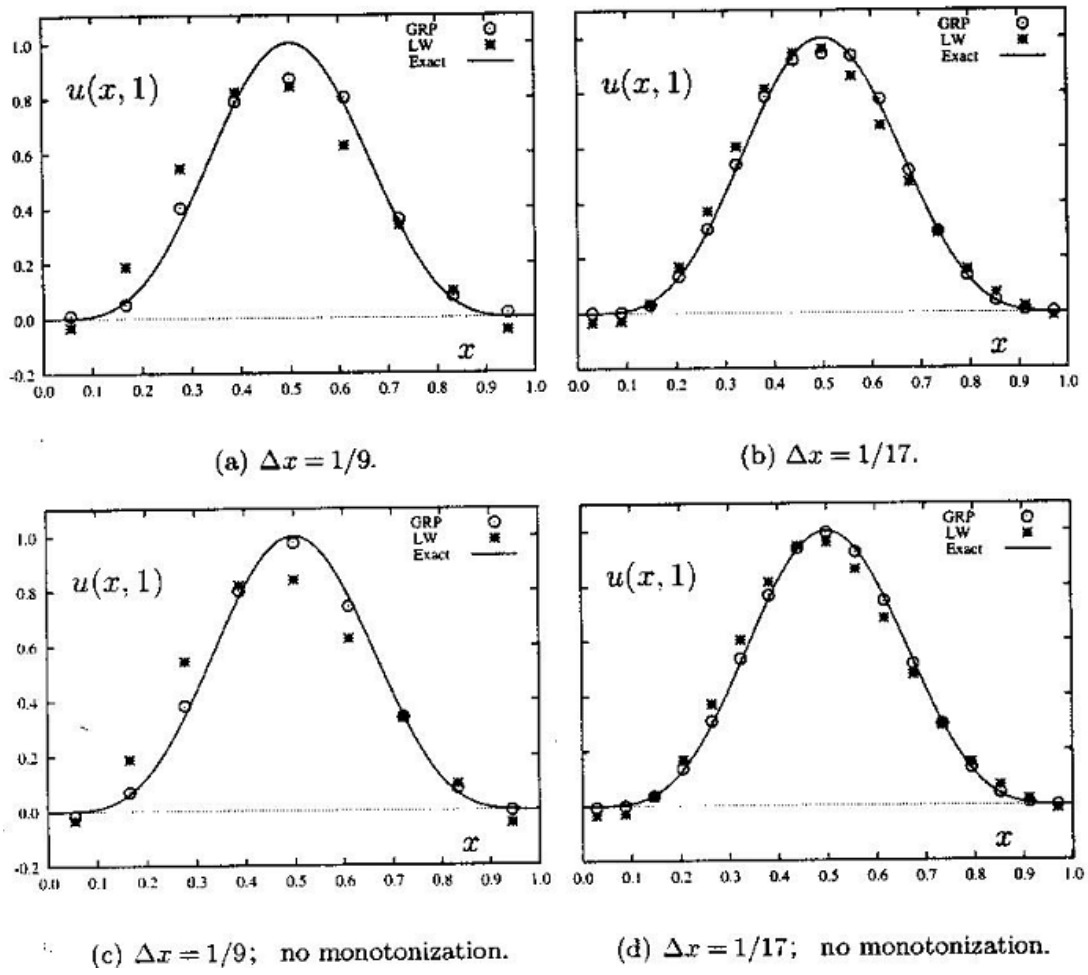
$$U_j^{n+1} = U_j^n - \lambda(U_{j+1/2}^{n+1/2} - U_{j-1/2}^{n+1/2}). \tag{3.5.7}$$

Finally, the new slopes s_j^{n+1} are obtained as follows [see eq. (3.4.13)]:

$$U_{j+1/2}^{n+1} = U_{j+1/2}^n - k s_j^n, \tag{3.5.8}$$

$$s_j^{n+1} = \frac{1}{\Delta x}(U_{j+1/2}^{n+1} - U_{j-1/2}^{n+1}). \tag{3.5.9}$$

The slopes s_j^{n+1} are further subjected to the monotonicity algorithm given by construction 3.4.2.



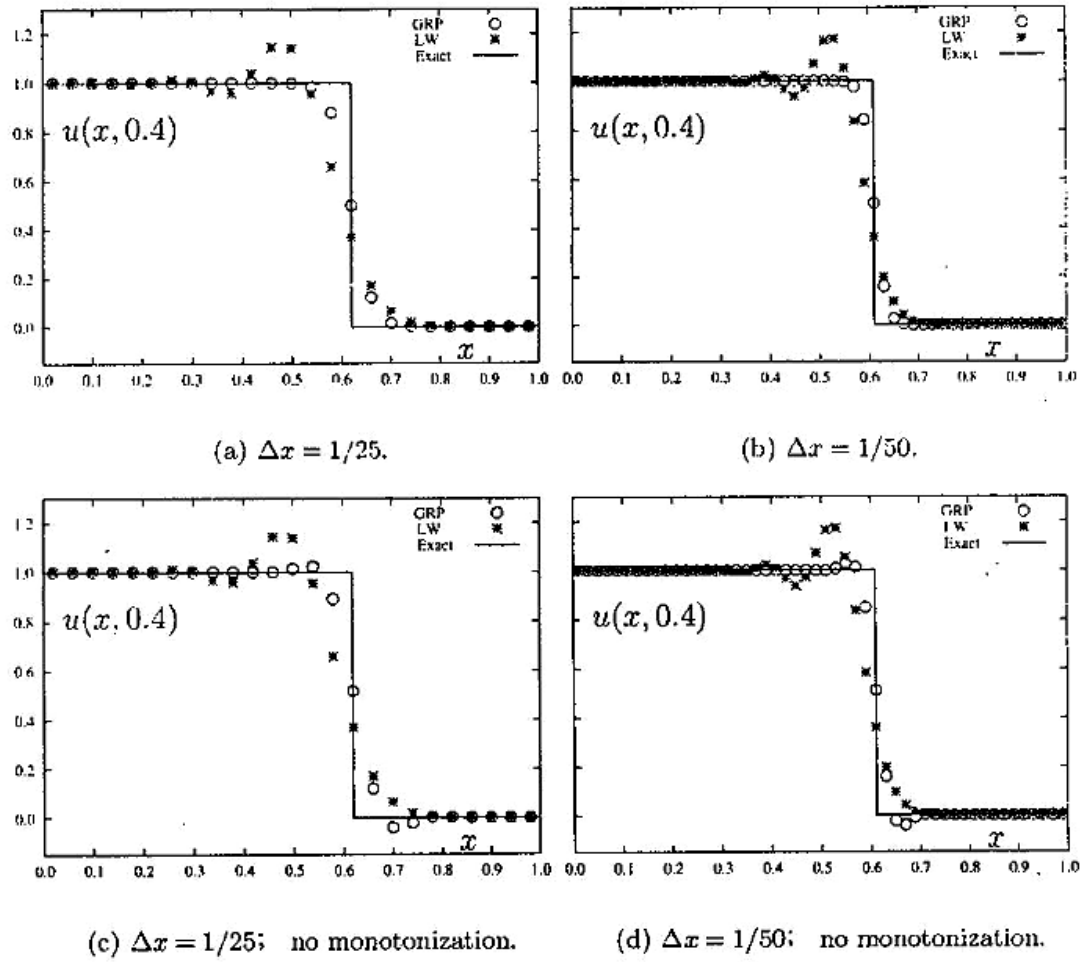
(a) $\Delta x = 1/9$. (b) $\Delta x = 1/17$.

(c) $\Delta x = 1/9$; no monotonization. (d) $\Delta x = 1/17$; no monotonization.

Figure 3.5.3.Second order integration of $u_t + u_x = 0$, with initial data $u_0(x) = \sin^4(\pi x)$.

The sample problems considered here are the same two problems previously used for the first-order schemes, including the same grids and time step specifications. The second order results for the periodic case are given in Figure3.5.3, where a comparison between GRP and LW schemes is shown. We notice a significant improvement relative to the first order results in Figure 3.5.1, and it is also evident that the convergence with grid refinement is faster in the second order case than in the first order one. On the whole, the GRP values are closer to the exact solution than are the LW values. Furthermore, the LW results have a significant phase-shift error, whereas the GRP results do not. How does the monotonization algorithm affect the GRP results? In Figure 3.5.3(a) and 3.5.3(b) we show that the GRP results that were subject to the slope limiter given in construction 3.4.2. The LW scheme, however, does not include any monotonization or slope limiting algorithm. For comparison, we therefore repeated the GRP computation without applying the monotonization algorithm, and the results are shown in Figure 3.5.3(c) and 3.5.3(d). Clearly, the GRP points near the peak (where slope limiting is most effective) are now higher, including that indeed the limiting algorithm is required to suppress peak-forming tendencies. We also note on Figures 3.5.3(c) and 3.5.3(d) that some GRP and LW points have $u < 0$. These "undershoot" values are in violation of the maximum-minimum principle. Slope limiting eliminates such violation by a second order scheme and is hence mandatory to comply with the maximum-minimum principle. We note that the Godunov scheme is in agreement with that principle (as stated before), and the results in Figure 3.5.1 are evidence to that property. We now turn to the step-function problem, identical to that considered in the first order scheme. In particular, we use the same $\lambda, \Delta x,$ and final time. As is

clearly visible in Figure 3.5.4(a) and 3.5.4(b), the "shock-captured" solution obtained here is similar to that of the first order scheme already discussed (Figure 3.5.2). However, the discontinuity is more sharply resolved by the second order schemes, with the sharpest (and most accurate) results obtained by the GRP. Here the jump in U_j^n is spread over about three cells, both in the coarse grid and in the fine-grid computations.



(a) $\Delta x = 1/25$. (b) $\Delta x = 1/50$.

(c) $\Delta x = 1/25$; no monotonization. (d) $\Delta x = 1/50$; no monotonization.

Figure 3.5.4 second-order integration of $u_t + u_x = 0,$ with unit step-function initial data.

Again, for comparison we repeated the two cases without the GRP monotonization constraint, and the results are shown in Figures 3.5.4(c) and 3.5.4(d). The GRP now produces some "overshoot" and "undershoot" values near the step. The indispensability of monotonization has thus been amply demonstrated, and in subsequent GRP computations we shall no longer consider the nonmonotonization option.

It is also interesting to notice the nature of the LW solution. No monotonization is applied in this scheme, and indeed the numerical solution develops pronounced oscillations behind the shock, which is also a typical feature for this scheme when it is extended to the fluid dynamical equations.

The Burgers Nonlinear Conservation Law

Here we consider the Burgers equation,

$$u_t + (\frac{u^2}{2})_x = 0, \qquad u(x,0) = u_0(x). \tag{3.5.10}$$

As explained before, in the case of smooth initial data the solution to this equation is obtained by the invariance of $u(x,t)$ along characteristic lines. When characteristic lines intersect, a smooth solution no longer exists, and from that time on only a (weak) solution, with shocks that obey the jump condition [$s = \frac{f(u_2) - f(u_1)}{u_2 - u_1}$], is possible. In the case of the Burgers equation the characteristic speed is $\frac{dx}{dt} = u,$ and the speed of a shock wave is given by $s = \frac{1}{2}(u_L + u_R),$ where the left and the right

values at the shock discontinuity u_L, u_R must obey the inequality $u_L \geq u_R$.

Two initial value problems are considered. The first has the smooth periodic data

$$u_0(x) = \sin(2\pi x), \tag{3.5.11}$$

and the second is a moving step problem, having the initial data $u_0(x) = 1 \quad for \quad x \leq x_0 \quad and \; u_0(x) = 0 \quad for \quad x > x_0$. Both problems have exact solutions, which, for the simple initial data considered here, are readily calculated by using the previously mentioned characteristic construction. Both problem defined on R, yet, with appropriate boundary condition, they can be solved numerically on some bounded interval $[x_1, x_2]$ of R, yielding the same finite-difference solution that would have been obtained on R.

First-Order Computation

Here we use the same two first-order schemes (Godunov and Lax-Friedrichs) previously considered in the context of the linear sample problems. The Godunov scheme is given by 3.3.10. The Lax-Fridrichs scheme, however, for a general flux function $f(u)$ is given by

$$U_j^{n+1} = \frac{1}{2}(U_j^n + U_{j-1}^n) - \frac{\lambda}{2}(f(U_{j+1}^n) - f(U_{j-1}^n)), \tag{3.5.12}$$

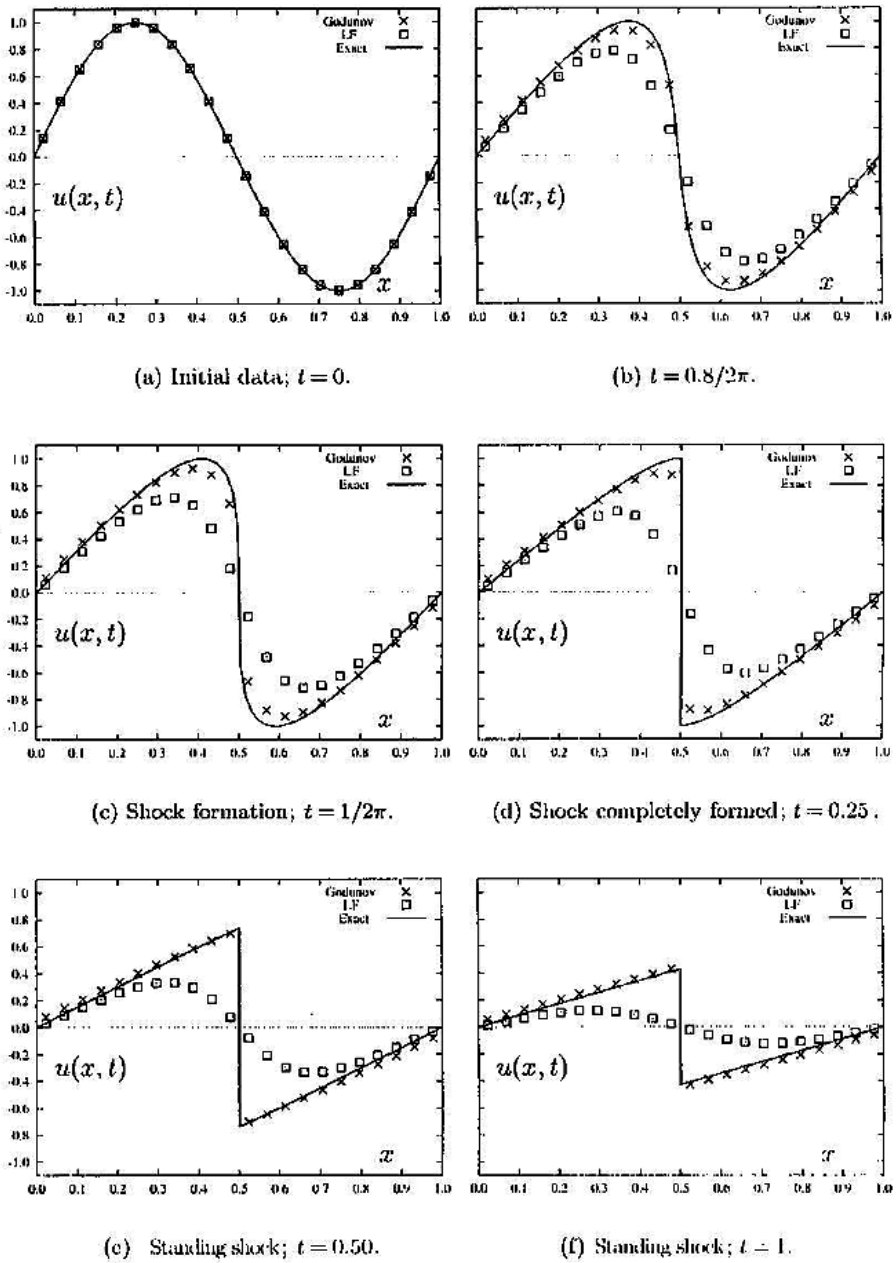
where the Burgers equation scheme is obtained by taking $f(u) = \frac{1}{2}u^2$. In our computations (both first and second order) we take fixed values for k and Δx, so

that the ratio $\lambda = \frac{k}{\Delta x}$ satisfies the CFL condition ($k.\max_{|u|\le M_n} |f'(u)| < \Delta x$). In fact, we take k such that the left-hand side in CFL condition is approximately equal to $\frac{1}{2}\Delta x$. In the case of smooth initial data, the equation is solved in the domain [0,1] with periodic boundary conditions. The computational cell size is $\Delta x = \frac{1}{22}$. The result are displayed as a time sequence in Figure 3.5.5, which compares the finite-difference solutions with the exact solution obtained by the method of characteristics. Prior to shock formation, in Figure 3.5.5 (b), the solution is smooth and displays the expected steepening in the interval where $\frac{\partial}{\partial x}u(x,t) < 0.$ The smooth solution breaks down at the moment $t = 1/2\pi$, where the slope at $x = 0.5$ becomes unbounded, as readily derived by taking the limit

$$\lim_{\varepsilon\to 0}\frac{\varepsilon}{\sin(2\pi(0.5-\varepsilon))} = \lim_{\varepsilon\to 0}\frac{\varepsilon}{2\pi\varepsilon} = \frac{1}{2\pi},$$

which corresponds to the point where the characteristic line emanating from $(x,t) = (0.5-\varepsilon,0)$ intersects the line $x = 0.5$. The solution at the breakdown time is shown in Figure 3.5.5(c). Beginning at this time the jump discontinuity at $x = 0.5$ gradually increases, reaching a maximal value (between $u = -1$ *and* $u = 1$) at $t = 0.25$ [see Figure 3.5.5(d). This is the moment at which the characteristic lines emanating from the external points $(x,t) = (0.5 \pm 0.25, 0)$ reach the discontinuity point $(x,t) = (0.5, 0.25)$. We also observe that by jump condition for the Burgers equation the speed of propagation of a shock discontinuity

$[u_L, u_R]$ *is* $s = 0.5(u_L + u_R)$, which vanishes owing to the symmetry of (u, t) about $x = 0.5$. The shock discontinuity at $x = 0.5$ is thus a standing shock.



(a) Initial data; $t = 0$. (b) $t = 0.8/2\pi$.

(c) Shock formation; $t = 1/2\pi$. (d) Shock completely formed; $t = 0.25$.

(e) Standing shock; $t = 0.50$. (f) Standing shock; $t = 1$.

Figure 3.5.5. First-order integration of $u_t + (\frac{u^2}{2})_x = 0$, with $u_0(x) = \sin(2\pi x)$.

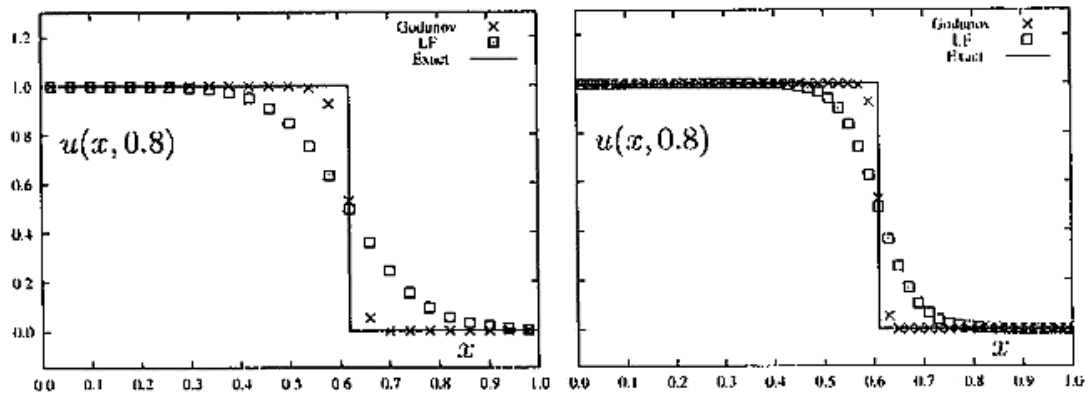
At later times $(t > 0.25)$, the jump $[u]$ at the shock discontinuity decreases progressively from its maximal value of $[u] = u_L - u_R = 2$, as clearly visible in Figures 3.5.5(d) – 3.5.5(f). Can this observation be explained by theoretical consideration? Indeed, it can be explained using the concepts of energy and dissipation. Let the "energy measure" of $u(x,t)$ be the finite integral $E(t) = \int_a^{a+1} \frac{1}{2}[u(x,t)]^2 dx$, *where* $0 \le a \le 1$ is a constant. The periodicity of $u(x,t)$ in x implies that $E(t)$ is independent of a. Now, multiply the Burgers equation by $u(x,t)$, obtaining $(\frac{1}{2}u^2)_t + (\frac{1}{3}u^3)_x = 0$, and integrate the resulting equation over an interval $[a, a+1]$. Interchanging the order of x integration and time derivative, we obtain $\frac{\partial}{\partial t}E(t) = 0$ [since $u(x,t)$ is periodic in x]. This means that the Burgers equation preserves the energy measure over time, which seems to agree with the finite-difference solution prior to the shock formation. After shock formation, however, this result is clearly in disagreement with both the exact and the numerical solution. Indeed, this equation may not be integrated over an interval containing a shock discontinuity, while disregarding the jump in $u(x,t)$ there. A correct way to perform the integration when a shock is present is to choose $a = -0.5$, so that the integration will extend from the right side of the shock at $x = -0.5$ to the left side of the shock at $x = 0.5$. The resulting rate of dissipation is then given by

$$\frac{\partial}{\partial t}E(t) = \frac{2}{3}\left[u_L^3 - u_R^3\right] , \qquad (3.5.13)$$

and since $u_L = -u_R > 0$ the energy $E(t)$ is decreasing in time, as is also evident by observing the evolution of the solution from $t = 0.25$ [Figure 3.5.5(d)] to $t = 1$

[Figure 3.5.5(f)]. This clearly demonstrates the dissipation effect of a shock wave in the solution to the Burgers equation.

We now turn to the step-function example, the initial data being $u_0(x) = 1$ *for* $x < x_0$ *and* $u_0(x) = 0$ *for* $x > x_0$, where x_0 is the initial position of the discontinuity. The numerical integration is performed in the rang $0 < x < 1$, with boundary conditions $u(0) = 1, u(1) = 0$. Two grids were used: (i) a coarse grid with $\Delta x = 0.04$ and an initial (mid-cell) position $x_0 = 0.22$ and (ii)a fine grid with $\Delta x = 0.02$ and an initial (mid-cell) position $x_0 = 0.11$. In both grids we chose a constant $\lambda = k / \Delta x,$ having the value $\lambda = 0.5,$ and the integration was performed to time $t = 0.8,$ so that the step propagates through the same distance as in the linear case.



(a) $\Delta x = 1/25.$ (b) $\Delta x = 1/50.$

Figure 3.5.6 First-order integration of $u_t + (\frac{1}{2}u^2)_x = 0$, with unit step-function initial data.

As is clearly observed in Figure 3.5.6, the Godunov scheme approximates the moving step to a considerably higher level of accuracy and resolution than the LF scheme. In particular, the Godunov scheme captures the shock over about three cells, versus nine cells in the LF scheme. It is also noted that these cell numbers are virtually unchanged by grid refinement (for constant time).

Comparing this feature to the linear case (Figure3.5.2), where the "discontinuity spreading" increases with grid refinement, we interpret this as a "stabilization" effect typical of shock capturing in the nonlinear case.

Second-Order Computation

Our primary interest here is the GRP scheme given by construction 3.4.1 while taking $f(u) = \frac{1}{2}u^2$. The slopes s_j^{n+1} are further subjected to the monotonicity algorithm given by construction 3.4.2. For comparison we take the Lax-wendroff scheme, which in the case of equation $[\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0]$ generalizes 2.2.23 as

$$U_j^{n+1} = U_j^n - \frac{\lambda}{2}\left[f_{j+1}^n - f_{j-1}^n\right] + \frac{\lambda^2}{2}\left[g_{j+1/2}^n - g_{j-1/2}^n\right], \qquad (3.5.14)$$

where $\lambda = \frac{k}{\Delta x}$. The first-order term (in λ) in 3.5.14 approximates $k(U_t)_j^n$ with

$$f_{j\pm1}^n = f(U_{j\pm1}^n),$$

and the second-order term approximates $\frac{k^2}{2}(U_{tt})_j^n$. The second time-derivative is based on the identity $u_{tt} = g(u)_x$, $g(u) = f'(u)f(u)_x$, obtained by differentiating the scalar conservation law $u_t + f(u)_x = 0$. Its finite-difference approximation is then given by:

$$(U_{tt})_j^n = \left[g_{j+1/2}^n - g_{j-1/2}^n\right]/\Delta x,$$
$$g_{j\pm1/2}^n = \pm f'\left(\tfrac{1}{2}(U_j^n + U_{j\pm1}^n)\right)\left[f_{j\pm1}^n - f_j^n\right] \tag{3.5.15}$$

We recall GRP algorithm (construction 3.4.1) in the special case $f(u) = \frac{1}{2}u^2$.

(I) Given the values

$$U_{j+1/2,\pm}^n = \begin{cases} U_{j+1}^n - \frac{\Delta x}{2} s_{j+1}^n, & \text{"+"} \\ U_j^n + \frac{\Delta x}{2} s_j^n, & \text{"–"} \end{cases} \tag{3.5.16}$$

We let $U_{j+1/2}^n$ be the solution to the corresponding Riemann problem.

Explicitly, it is given by

$$U_{j+1/2}^n = \begin{cases} \max\left(\left|U_{j+1/2,-}^n\right|, \left|U_{j+1/2,+}^n\right|\right).\operatorname{sgn}(U_{j+1/2,-}^n + U_{j+1/2,+}^n), \\ \qquad\qquad \text{if} \quad U_{j+1/2,-} > U_{j+1/2,+}^n \\ \min\left(\left|U_{j+1/2,-}^n\right|, \left|U_{j+1/2,+}^n\right|\right)\operatorname{sgn}(U_{j+1/2,-}^n), \\ \qquad\qquad \text{if} \ \operatorname{sgn}(U_{j+1/2,-}^n).\operatorname{sgn}(U_{j+1/2,+}^n) \geq 0, \\ 0, \qquad\qquad \text{if} \ U_{j+1/2,-} < 0 < U_{j+1/2,+}^n, \end{cases} \tag{3.5.17}$$

and of course $U_{j+1/2}^n = U_{j+1/2,-}^n$ if $U_{j+1/2,-}^n = U_{j+1/2,+}^n$.

(II) The instantaneous time derivatives $\left(\frac{\partial \tilde{u}}{\partial t}\right)_{j+1/2}^n = \frac{\partial \tilde{u}}{\partial t}(x_{j+1/2}, t_n)$ are now given by

$$\left(\frac{\partial \tilde{u}}{\partial t}\right)_{j+1/2}^n = \begin{cases} -U_{j+1/2}^n.s_j^n, & \text{if} \ U_{j+1/2}^n > 0, \\ -U_{j+1/2}^n.s_{j+1}^n, & \text{if} \ U_{j+1/2}^n < 0, \\ 0, & \text{if} \ U_{j+1/2}^n = 0, \end{cases} \tag{3.5.18}$$

and numerical fluxes are given by

$$f_{j+1/2}^{n+1/2} = f(U_{j+1/2}^n) + \frac{k}{2} U_{j+1/2}^n . \left(\frac{\partial \tilde{u}}{\partial t}\right)_{j+1/2}^n,$$

so that

$$U_j^{n+1} = U_j^n - \lambda\left(f_{j+1/2}^{n+1/2} - f_{j-1/2}^{n+1/2}\right).$$

(III) The new values at the cell boundaries are given by

$$U_{j+1/2}^{n+1} = U_{j+1/2}^{n} + k\left(\frac{\partial \tilde{u}}{\partial t}\right)_{j+1/2}^{n},$$

and the new slopes updated by

$$s_j^{n+1} = \frac{1}{\Delta x}\left(U_{j+1/2}^{n+1} - U_{j-1/2}^{n+1}\right).$$

(IV) Finally, the computed slopes s_j^{n+1} are modified by the slope limiter as in construction 3.4.2.

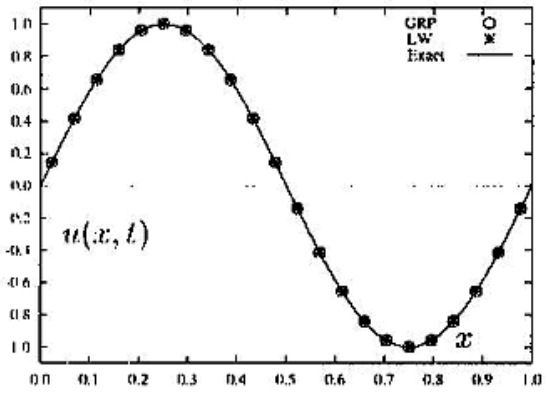
The same two sample problems are considered here as for the first-order schemes already discussed, with identical data, boundary conditions, grids, and final time. The time sequence results for the periodic case are shown in Figure3.5.7.

How do the second-order results compare to the first-order ones (Figure 3.5.5)?

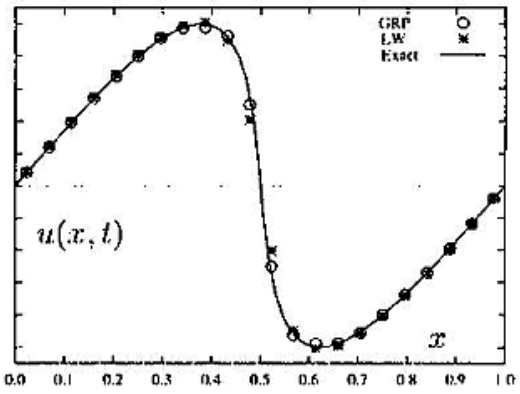
The GRP approximation is generally very close to the exact solution, which is a considerable improvement relative to the Godunov approximation. The LW scheme produces fairly accurate results, although less accurate than those of the GRP. As in the smooth linear case (Figure3.5.3), the LW scheme is characterized by overshoots near extremal points, notably in Figures 3.5.7(c) and 3.5.7(d). It is particularly interesting to compare the GRP formation of the "N-wave" [Figure 3.5.7(d)-3.5.7(f) with the corresponding Godunov results {Figure 3.5.5(d)-3.5.5(f). The GRP points are considerably closer to the exact solution than those of the Godunov scheme, notably near the shock for still higher order schemes (not considered in this thesis).

Turning to the step-function example, having the same data as in the first-order case, we show the results in Figure 3.5.8 at $t = 0.8$. As in the linear case (Figure 3.5.4), the LW scheme produces significant oscillations behind the step, indicating that this feature of the scheme is not suppressed by the nonlinearity of the

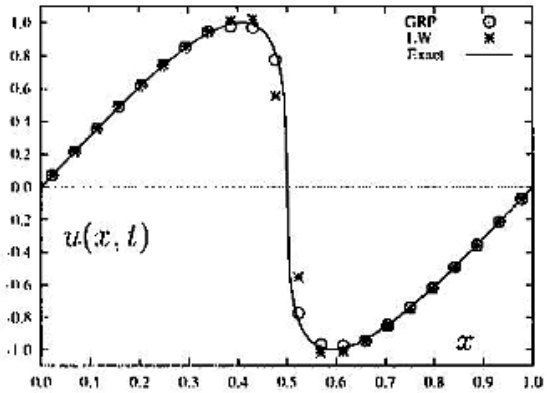
scalar conservation law. By comparing the GRP and the Godunov approximations (Figure 3.5.8), it is evident that the improvement in accuracy takes place only near the shock discontinuity. The GRP results are a near-perfect approximation to the step, with only (the inevitable) single point representing the average value of 0.5 at the mid-cell $x = 0.62 \; or \; x = 0.61$, where the exact jump is positioned. This demonstrates the higher-resolution feature of the GRP, which also characterizes the fluid dynamical GRP scheme.



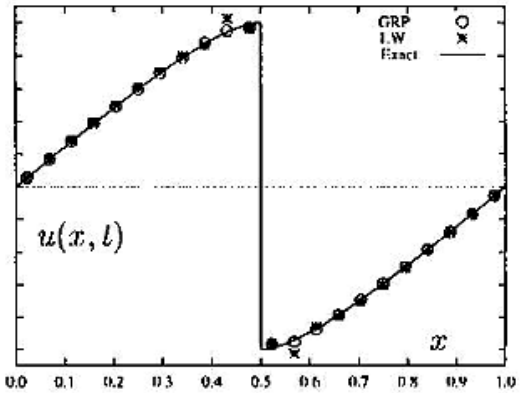
(a) Initial data; $t = 0$.



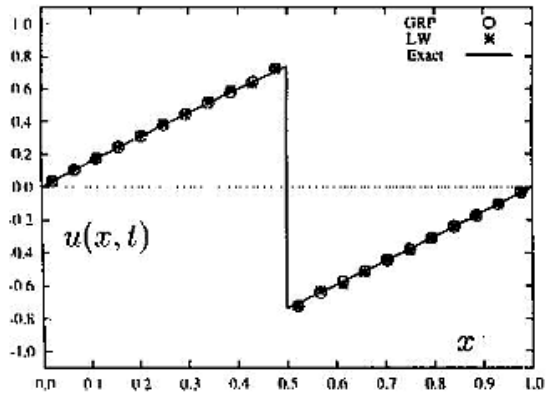
(b) $t = 0.8/2\pi$.



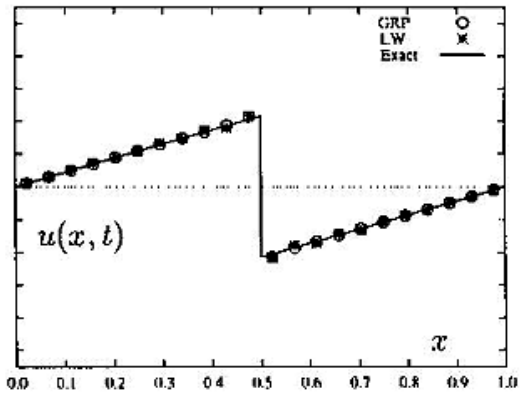
(c) Shock formation; $t = 1/2\pi$.



(d) Shock completely formed; $t = 0.25$.

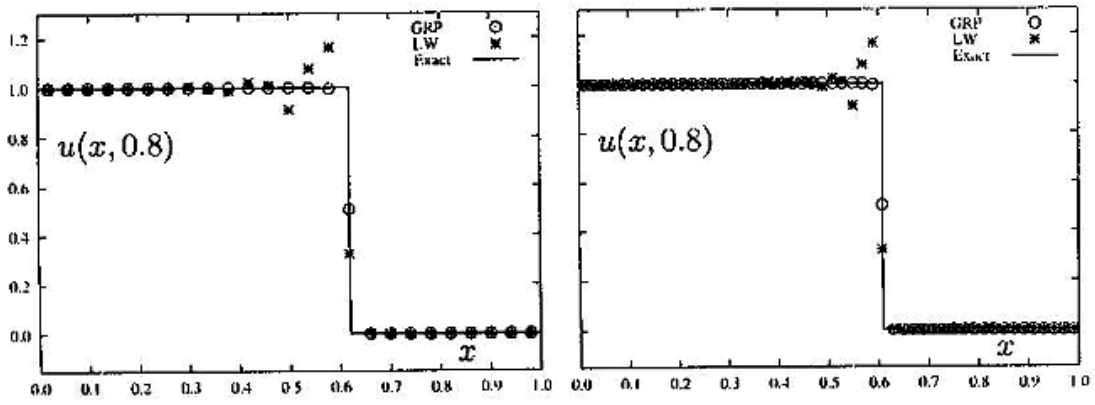


(e) Standing shock; $t = 0.50$.



(f) Standing shock; $t = 1$.

Figure 3.5.7. Second-order integration of $u_t + (\frac{1}{2}u^2)_x = 0$, with initial data $u_0(x) = \sin(2\pi x)$.



(a) $\Delta x = 1/25$. (b) $\Delta x = 1/50$.

Figure 3.5.8. Second-order integration of $u_t + (\frac{1}{2}u^2)_x = 0$, with unit step-function initial data.

References:

[AD G]	F. Angrand, A. Dervieux, J.A. desideri and R. Glowinski, Numerical Methods for Euler Equations of Fluid Dynamics, Siam Philadelphia. 1985.
[AJ]	Alexandre J. Chorin, and Jerrold E. Marsden, A Mathematical Introduction to Fluid Mechanics, third edition, Springer-Verlag, New York, 1993.
[ATP]	D.A.Anderson, J.C.Tannehill, and R.H.Pletcher, Computational Fluid Mechanics and Heat Transfer, McGraw-Hill, 1984.
[BF]	M.Ben-Artzi, J.Falcovitz, Generalized Rieman Problems in computational Fluid Dynamics, Cambridge University press 2003.
[BR]	L.Brillouin, Wave Propagation and Group Velocity, Acadimic Press, 1960.
[DA]	S.F.Davis, A rotationally biased upwind difference scheme for the Euler equation, J. Comput. Phys.,1984.
[EO]	B.Engquist and S.Osher, Stable and entropy satisfying approximations for transonic flow calculations, Math. Comp.,1980.
[GR]	Edwige Goldlewski and Pierre-Arnaud Raviart, Mathematiques and Application, copyright edition marketing, ellipses, 1991.
[HA]	W. Hackbusch, Multi-grid methods and applications. Springer, Berlin, 1985.
[HF]	M. Hinatsu, and J.H. Ferziger, Numerical computation of unsteady incompressible flow in complex geometry using a composite multigrid

	technique. Int. J. Numer. Methods Fluids, 1991.
[JO1]	F. John, Partial Differential Equations, Springer, 1971.
[JO2]	John H. Mathews, Numerical Methods for Mathematics, Science, and Engineering. Second Edition, Prentice-Hall International, 1992.
[JO3]	John C. Strikwerda, Finite Difference Schemes and Partial Differential Equations. Wadsworth & Books/cole, 1989.
[KR]	Rainer Kress, Numerical Analysis, Springer-Verlag, New York. 1998.
[LW]	P.D. Lax, and B. Wendroff, System of conservation laws.1960.
[MM]	K.W. Morton, and D.F. Mayers, Numerical Solution of Partial Differential Equations, Cambridge University Press, 1994.
[RA1]	Randall J. Leveque, Time-Split Methods for Partial Differential Equations, PHD thesis, Stanford, 1982.
[RA2]	Randall J. Leveque, Numerical Methods for Conservation Laws, Birkhauser Verlag, 1990.
[RF]	Richard L. Burden., and Douglas Fairs, Numerical Analysis, Sixth Edition, Books Cole Publishing Company, 1997.
[SM1]	Joel Smoller, Shock Waves and Reaction-Diffusion Equations, Springer-Verlag, 1983.
[SM2]	G.D. Smith, Numerical Solution of Partial Differential Equations. Third Edition, Clarendon Press.Oxford. 1985.