HYPERBOLIC EQUATIONS

MARIA CAMERON

Contents

1. Linear advection equation

Please read Chapter 10 and Appendix E.3 in R. LeVeque "Finite difference methods for ordinary and partial differential equations". These notes are complimentary.

We will consider the linear advection equation on the interval $x \in [0,1]$ with the periodic boundary condition. the resulting initial and boundary value problem (IBVP) is

$$
\begin{cases} u_t + au_x = 0, & x \in [0,1], \ t \geq 0, \\ u(0,t) = u(1,t), & t \geq 0 \\ u(x,0) = \eta(x). \end{cases} \tag{1}
$$

The exact solution to (1) is $\eta(x - at)$ periodically extended, i.e., $\eta(x - at \mod 1)$.

1.1. **Useful matrices and their spectra.** It is convenient to conduct the stability analysis of methods for linear advection equation by thinking of the *method of lines* and and the forward Euler time discretization. Below we will list matrices that arise the right-hand side of the method of lines and find their eigenvalues and eigenvectors.

Throughout this section, we will assume that the interval $[0,1]$ is partitioned to m subintervals of length $h = 1/m$. The periodic boundary conditions make the points $x_0 = 0$ and $x_m = 1$ identical. Therefore, we need to find the numerical solutions only at the points x_0, x_1, ..., x_{m-1} at each $t_n = kn$. Hence, the matrices in the right-hand side of the MOL will be $m \times m$.

- Matrix A_1 arises whenever the first derivative in space approximated using the central difference
$$\frac{U_{j+1} - U_{j-1}}{2h}$$
and the periodic boundary conditions are imposed:
$$
A_1 = \begin{bmatrix} & 1 & & & -1 \\ -1 & & \ddots & & \\ & \ddots & & \ddots & \\ & & \ddots & & 1 \\ 1 & & & -1 & \end{bmatrix}. \tag{2}
$$
In words, A_1 has 1s along its first superdiagonal, -1s along is first subdiagonal, -1 in the top right corner, and 1 in the bottom left corner. All other entries of A_1 are zero. We can guess the form of the eigenvector of A_1 taking into account that its jth entry must be proportional to the difference between its nearest neighbors. Moreover, the eigenvectors of A_1 should be extendable periodically. Hence, we try the vector v with entries $v_j = e^{ijb}$ as the candidate for the eigenvector. Periodicity requires
$$1 = v_0 = v_m = e^{ijbm}.$$
Hence $jbm = 2\pi p$ where $p \in \mathbb{Z}$. Therefore, $b = \frac{2\pi p}{m}$. For all $0 \leq j \leq m-1$ we have:
$$[Av]_j = v_{j+1} - v_{j-1} = e^{\frac{2\pi p i}{m}(j+1)} - e^{\frac{2\pi p i}{m}(j-1)} = e^{\frac{2\pi p i}{m}j} 2i \sin\left(\frac{2\pi p}{m}\right) = v_j 2i \sin\left(\frac{2\pi p}{m}\right).$$

Furthermore, the m distinct eigenpairs correspond to $p = 0, 1, \ldots, m-1$. Therefore, the eigenvectors and eigenvalues of A_1 are

$$v^p = \begin{bmatrix} 1 \\ e^{\frac{2\pi p i}{m}} \\ e^{\frac{2\pi p i}{m}2} \\ \vdots \\ e^{\frac{2\pi p i}{m}(m-1)} \end{bmatrix}, \quad \lambda_p = 2i\sin\left(\frac{2\pi p}{m}\right), \quad p = 0, 1, \ldots, m-1. \tag{3}$$

- Matrix A_2 arises when the artificial viscosity, i.e. a term with

$$\frac{U_{j+1} - 2U_j + U_{j-1}}{h^2},$$

is added to the finite difference scheme and the periodic boundary conditions are imposed. It is added in the Lax-Friedrichs scheme for stability purposes. In the Lax-Wendroff scheme, it compensates the error term proportional to u_{xx} resulting from the finite difference approximation of the first derivative and makes the method stable. The matrix A_2 is:

$$A_2 = \begin{bmatrix} -2 & 1 & & & 1 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ 1 & & & 1 & -2 \end{bmatrix}. \tag{4}$$

Its eigenvectors are the same as those of A_1 (see (3)). The eigenvalues of A_2 are

$$\lambda_p = 2\cos\left(\frac{2\pi p}{m}\right) - 2, \quad p = 0, 1, \ldots, m-1. \tag{5}$$

- Matrices A_3 and A_4 arise in left and right upwind schemes, in which the first derivatives in x are approximated using

$$\frac{U_j - U_{j-1}}{h} \quad \text{and} \quad \frac{U_{j+1} - U_j}{h},$$

respectively:

$$A_3 = \begin{bmatrix} 1 & & & & -1 \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}, \quad A_4 = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 1 & & & & -1 \end{bmatrix}. \tag{6}$$

Their eigenvectors are the same as those of A_1 (see (3)). Their eigenvalues are

$$\lambda_p(A_3) = 1 - e^{-\frac{2\pi p i}{m}}, \quad \lambda_p(A_4) = e^{\frac{2\pi p i}{m}} - 1, \quad p = 0, 1, \ldots, m-1. \tag{7}$$

1.2. Finite difference schemes and their spectra.

1.2.1. *Lax-Friedrichs.* The scheme:

$$U_j^{n+1} = \frac{U_{j+1}^n + U_{j-1}^n}{2} - \frac{ak}{2h}\left(U_{j+1}^n - U_{j-1}^n\right). \tag{8}$$

The local truncation error for Lax-Friedrichs is $O(h^2)+O(k)$ making it first-order accurate. The corresponding MOL discretization gives:

$$\frac{dU}{dt} = \frac{1}{2k}A_2U - \frac{a}{2h}A_1U. \tag{9}$$

The eigenvalues of the matrix $M := \frac{1}{2k}A_2 - \frac{a}{2h}A_1$ are:

$$\lambda_p = \frac{1}{k}\left(\cos\left(\frac{2\pi p}{m}\right) - 1\right) - i\frac{a}{h}\sin\left(\frac{2\pi p}{m}\right), \quad p = 0, 1, \dots, m-1. \tag{10}$$

The Lax-Friedrichs method is obtained from the MOL equation (9) by using the forward Euler time stepping. Hence, the stability condition requires that $|k\lambda_p + 1| \le 1$, i.e.,

$$\left|\cos\left(\frac{2\pi p}{m}\right) - i\frac{ka}{h}\sin\left(\frac{2\pi p}{m}\right)\right| = \left[\cos^2\left(\frac{2\pi p}{m}\right) + \left(\frac{ka}{h}\right)^2\sin^2\left(\frac{2\pi p}{m}\right)\right]^{1/2} < 1. \tag{11}$$

This condition holds iff

$$\left|\frac{ka}{h}\right| \le 1. \tag{12}$$

2. Modified equations

Consideration of modified equations is a convenient tool for the analysis of finite difference schemes for PDEs. A modified equation is an equation that is satisfied by the numerical solution more exactly than the original PDE. The extra terms in a modified equation explain the behavior of the numerical error.

2.1. A derivation of the modified equation for the Lax-Wendroff scheme.

We consider the linear advection equation

$$u_t + au_x = 0. \tag{13}$$

The Lax-Wendroff scheme for it is

$$U_j^{n+1} = U_j^n - \frac{ak}{2h}\left(U_{j+1}^n - U_{j-1}^n\right) + \frac{a^2k^2}{2h^2}\left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right). \tag{14}$$

Assume that a smooth function $v(t,x)$ satisfies the Lax-Wendroff scheme, i.e.,

$$\begin{aligned} v(t+k,x) = v(t,x) &- \frac{ak}{2h}\left(v(t,x+h) - v(t,x-h)\right) \\ &+ \frac{a^2k^2}{2h^2}\left(v(t,x+h) - 2v(t,x) + v(t,x-h)\right). \end{aligned} \tag{15}$$

To find a PDE for v we use Taylor expansions around (t, x):

$$\begin{aligned} & v + kv_t + \frac{k^2}{2}v_{tt} + \frac{k^3}{6}v_{ttt} + O(k^4) \\ =& v - \frac{ak}{2h}\left(2hv_x + 2\frac{h^3}{6}v_{xxx} + O(h^5)\right) + \frac{a^2k^2}{2h^2}\left(2\frac{h^2}{2}v_{xx} + 2\frac{h^4}{24}v_{xxxx} + O(h^6)\right) \end{aligned}$$

Canceling v and dividing by k we get:

$$v_t + av_x = \frac{k}{2}\left(a^2v_{xx} - v_{tt}\right) - \frac{k^2}{6}\left(v_{ttt} + \frac{ah^2}{k^2}v_{xxx}\right) + O(k^3). \tag{16}$$

Here we have taken into account that the Courant number $\nu := \frac{ak}{h}$ is constant.

Equation (16) implies that $v_t + av_x$ is at most $O(k)$. Let us show that it is actually $O(k^2)$ and that $v_{tt} - a^2v_{xx}$ is also $O(k^2)$. The next two equations are obtained by differentiating (16) with respect to t and x, respectively:

$$v_{tt} + av_{xt} = \frac{k}{2}\left(a^2v_{xxt} - v_{ttt}\right) + O(k^2), \tag{17}$$

$$v_{tx} + av_{xx} = \frac{k}{2}\left(a^2v_{xxx} - v_{ttx}\right) + O(k^2). \tag{18}$$

Multiplying (18) by a and subtracting it from (17) we obtain:

$$v_{tt} - a^2v_{xx} = \frac{k}{2}\left[\left(a^2v_{xx} - v_{tt}\right)_t - a\left(a^2v_{xx} - v_{tt}\right)_x\right] + O(k^2). \tag{19}$$

Equation (19) implies that $v_{tt} - a^2v_{xx}$ is at most $O(k)$. But this means that

$$v_{tt} - a^2v_{xx} = \frac{k}{2}\left([O(k)]_t - a[O(k)]_x\right) + O(k^2) = O(k^2). \tag{20}$$

Therefore, $v_{tt} - a^2v_{xx} = O(k^2)$ which, together with (16) implied that $v_t + av_x = O(k^2)$.

Plugging the relationship $v_{tt} - a^2v_{xx} = O(k^2)$ into (16) we get

$$v_t + av_x = -\frac{k^2}{6}\left(v_{ttt} + \frac{ah^2}{k^2}v_{xxx}\right) + O(k^3). \tag{21}$$

We would like to express the third derivative in t via the third derivative in x on the right-hand side in (21). We have:

$$(\partial_t + a\partial_x)v = O(k^2).$$

We apply the differential operator $(\partial_{tt} - a\partial_{tx} + a^2\partial_{xx})$ to both parts of the last equality and obtain:

$$(\partial_{tt} - a\partial_{tx} + a^2\partial_{xx})(\partial_t + a\partial_x)v = (\partial_{ttt} + a^3\partial_{xxx})v = O(k^2). \tag{22}$$

We still have $O(k^2)$ in the right-hand side because the differential operator $(\partial_{tt} - a\partial_{tx} + a^2\partial_{xx})$ does not contain k and hence at least cannot lower the order in k. Hence

$$\frac{\partial^3}{\partial t^3} = -a^3\frac{\partial^3}{\partial x^3} + O(k^2). \tag{23}$$

This means that, in (21), the term in the large parensethes on the right-hand side is:

$$v_{ttt} + \frac{ah^2}{k^2} v_{xxx} = a\left(-a^2 + \frac{h^2}{k^2}\right) v_{xxx}.$$

From the fact that $\nu = \frac{ak}{h}$ is constant we get $k = \frac{\nu h}{a}$ and $\frac{h}{k} = \frac{a}{\nu}$. Therefore, (21) can be rewritten as

$$v_t + av_x = -\frac{\nu^2 h^2}{6a^2} a\left(-a^2 + \frac{a^2}{\nu^2}\right) v_{xxx} + O(h^3) = -\frac{ah^2}{6}\left(-\nu^2 + 1\right) v_{xxx} + O(h^3). \tag{24}$$

Equation (24) is the modified equation whose solution satisfies the Lax-Wendroff scheme exactly. If we are willing to truncate the higher-order terms on the right-hand side of (24), we obtain a simpler modified equation that the numerical solution satisfies approximately, but more exactly than the original PDE (13):

$$\boxed{v_t + av_x + \frac{ah^2}{6}\left(1 - \nu^2\right) v_{xxx} = 0.} \tag{25}$$

The modified equations (24) and (25) have a leading-order extra term in comparison with the original advection equation (13) which is of the order $O(h^2)$ and dispersive, i.e., proportional to v_{xxx}. This indicates that the Lax-Wendroff scheme is second-order accurate and the numerical error in its numerical solution will be oscillatory.

2.2. **Fourier analysis of the modified equation for Lax-Wendroff.** The numerical solution to the advection equation $u_t + \sqrt{2}u_x = 0$ with $\nu = 0.8$ at times $t = 25$, $t = 50$, $t = 75$, and $t = 100$ is displayed in Fig. 1. The initial condition is $u(x, 0) = \exp(-20(x - 5)^2)$. The space interval is $0 \le x \le 25$. The boundary conditions are periodic. We observe that the numerical error is oscillatory and the oscillatory tail is traveling behind the main hump. Furthermore, the numerical solution broadens with time and gets more and more behind the exact one. These phenomena can be explained using Fourier analysis tools.

The modified equation (25) contains only odd derivatives in x. Hence the error is primarily dispersive. Subjecting the modified equation for Lax-Wendroff (25) to the Fourier transform we obtain:

$$\hat{v}_t = \left[-ai\xi + i\xi^3 \frac{ah^2}{6}(1 - \nu^2)\right] \hat{v}. \tag{26}$$

Hence, the Fourier mode coefficient with index ξ evolves in time as

$$\hat{v}(t) = e^{-ai\xi\left(1 - \frac{h^2}{6}(1-\nu^2)\xi^2\right)t} \hat{v} \equiv e^{-i\omega(\xi)t} \hat{v}. \tag{27}$$

The function $\omega(\xi)$ in (27) is called *the dispersion relation* (see Appendix E3 in R. LeVeque "Finite difference methods for ordinary and partial differential equations" for more details). The *phase velocity* for the numerical solution by Lax-Wendroff, i.e., the velocity with which the ξ's Fourier mode is traveling, is

$$c_p \equiv \frac{\omega(\xi)}{\xi} = a\left(1 - \frac{h^2}{6}(1 - \nu^2)\xi^2\right). \tag{28}$$

The *group velocity* for the numerical solution by Lax-Wendroff, i.e., the velocity with which the wave packet is traveling as a whole, is

$$c_p \equiv \frac{d\omega(\xi)}{d\xi} = a\left(1 - \frac{h^2}{2}(1-\nu^2)\xi^2\right). \tag{29}$$

Therefore, we have

$$c_g < c_p < a. \tag{30}$$

Hence, the wave packet, i.e., the numerical solution, is traveling slower than the exact solution and gradually is getting behind it. At the same time, the peaks of the numerical solution travel faster than the numerical solution as a whole but slower than the exact solution. This is what we see in Fig. 1.

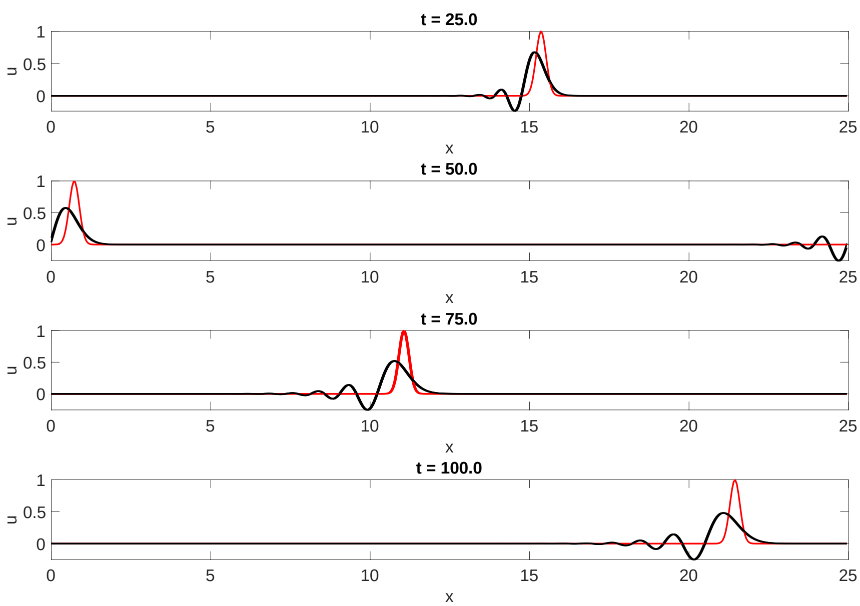


FIGURE 1. The numerical solution (black) by Lax-Wendroff and the exact solution (red) to the advection equation $u_t + \sqrt{2}u_x = 0$ on the time interval $0 \le t \le 100$.

3. Burgers's equation

Ref. R. LeVeque, "Numerical Methods for Conservation Laws", Birkhauser, 1992.
Burgers's equation

$$u_t + uu_x = \nu u_{xx} \tag{31}$$

is a successful, though rather simplified, mathematical model of the motion of a viscous compressible gas, where

- u = the speed of the gas,
- ν = the kinematic viscosity,
- x = the spatial coordinate,
- t = the time.

3.1. **Solution of the Burgers equation with nonzero viscosity.** Let us look for a solution of Eq. (31) of the form of traveling wave [1], i.e.,

$$u(x,t) = w((x - x_0) - st) \equiv w(y).$$

Then $u_t = -sw'$, $u_x = w'$, and $u_{xx} = w''$. Plugging this into Eq. (31) we obtain

$$\begin{aligned}
-sw' + ww' &= \nu w'' \\
-sw' + \left(\frac{w^2}{2}\right)' &= \nu w'' \\
-sw + \frac{w^2}{2} &= \nu w' + C.
\end{aligned}$$

We also impose conditions at the $\pm\infty$: $w(-\infty) = u_L$, $w(\infty) = u_R$, where $u_L > u_R$, and $w'(\pm\infty) = 0$. Then we have

$$-su_L + \frac{u_L^2}{2} = C = -su_R + \frac{u_R^2}{2}.$$

Therefore, s must be $(u_L + u_R)/2$. Thus, the shock speed is the same as in the case of zero viscosity. Hence $C = -u_L u_R/2$. Then we continue.

$$\begin{aligned}
\nu w' &= \frac{w^2}{2} - \frac{u_L + u_R}{2} w + \frac{u_L u_R}{2} \\
\frac{dy}{2\nu} &= \frac{dw}{w^2 - (u_L + u_R)w + u_L u_R} \\
\frac{dy}{2\nu} &= \frac{dw}{w^2 - (u_L + u_R)w + \frac{(u_L+u_R)^2}{4} - \frac{(u_L-u_R)^2}{4}} \\
\frac{dy}{2\nu} &= \frac{dw}{\left(w - \frac{u_L+u_R}{2}\right)^2 - \frac{(u_L-u_R)^2}{4}}
\end{aligned}$$

Integrating the both parts using

$$\int \frac{dw}{(w-a)^2 - b^2} = \frac{1}{2b}\log\left|\frac{w-a-b}{w-a+b}\right|$$

we obtain

$$\frac{y}{2\nu} + C = \frac{1}{u_L - u_R}\log\left|\frac{w - \frac{u_L+u_R}{2} - \frac{u_L-u_R}{2}}{w - \frac{u_L+u_R}{2} + \frac{u_L-u_R}{2}}\right| = \frac{1}{u_L - u_R}\log\left|\frac{w - u_L}{w - u_R}\right| = \frac{1}{u_L - u_R}\log\frac{u_L - w}{w - u_R}.$$

In the last equality we used the fact that $u_L > w > u_R$. Hence,

$$\begin{aligned}
\frac{u_L - w}{w - u_R} &= e^{\frac{y(u_L-u_R)}{2\nu}+C} \\
u_L - w &= we^A - u_R e^A, \quad \text{where} \quad A = y(u_L-u_R)/(2\nu) + C \\
w(e^A+1) &= (u_L + u_R e^A) \\
w = \frac{u_L + u_R e^A}{e^A+1} &= u_R + \frac{u_L - u_R}{2}\frac{2}{e^A+1}.
\end{aligned}$$

Multiplying and dividing by $\exp(-A/2)$ and using the identity

$$\frac{2e^{-A/2}}{e^{A/2}+e^{-A/2}} = 1 - \frac{e^{A/2}-e^{-A/2}}{e^{A/2}+e^{-A/2}} = 1 - \tanh\frac{A}{2}$$

we get

$$w(y) = \frac{u_R+u_L}{2} - \frac{u_L-u_R}{2}\tanh\left(\frac{y(u_L-u_R)}{4\nu} + C\right).$$

Hence,

$$u(x,t) = \frac{u_R+u_L}{2} - \frac{u_L-u_R}{2}\tanh\left(\frac{([x-x_0]-st)(u_L-u_R)}{4\nu}\right). \tag{32}$$

The profiles $w(y)$ for various values of ν are shown in Fig. 2. As $\nu \to 0$, $u(x,t)$ tends to a step function of the argument $x - st$.

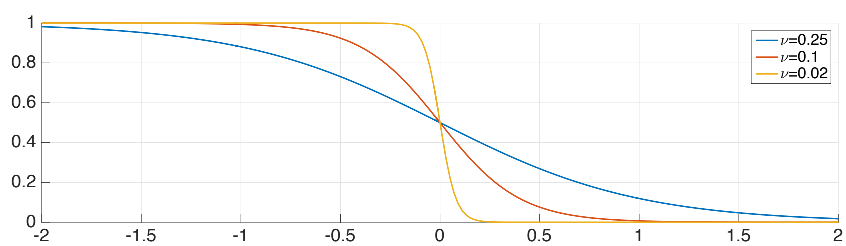


FIGURE 2. The profiles of the solution of the viscous Burgers equation for $u_R = 0$, $u_L = 1$, $x_0 = 0$, and ν equal to 0.25, 0.1, and 0.02. Note that $w(y)$ tends to a step function as $\nu \to 0$.

3.2. **Shock speed.** If the viscosity $\nu = 0$, or neglected, Eq. (31) can be rewritten as

$$u_t + [\tfrac{1}{2}u^2]_x = 0 \tag{33}$$

Eq. (33) is easier to study theoretically and numerically than Eq. (31). From now on, unless indicated otherwise, we will refer to Eq. (33) as the Burgers equation.

Equation (33) has a solution in the form of the traveling wave [2]

$$u(x,t) = V(x - st), \tag{34}$$

where $V(y)$ is a step function:

$$V(y) = \left\{ \begin{array}{ll} u_L & y < 0 \\ u_R & y > 0 \end{array} \right\}, \tag{35}$$

where $u_L > u_R$. This wave is called the *shock wave* and s is the speed of propagation of the shock wave. It can be obtained from the following reasoning. Let M be some large number. Consider the integral

$$\int_{-M}^{M} u(x,t)dx.$$

Then

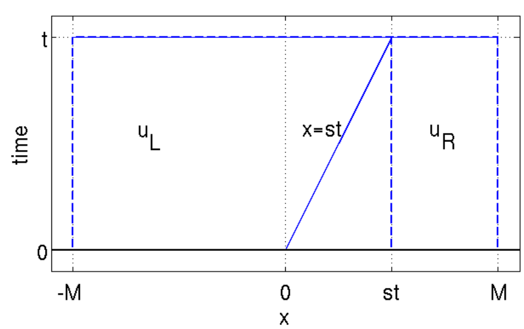


FIGURE 3. Finding the shock speed using the law of conservation of momentum.

$$\frac{d}{dt}\int_{-M}^{M} u(x,t)dx = \int_{-M}^{M} -uu_x dx = -\frac{u^2}{2}\bigg|_{-M}^{M} = \frac{u_L^2}{2} - \frac{u_R^2}{2}.$$

On the other hand,

$$\int_{-M}^{M} u(x,t)dx = (M + st)u_L + (M - st)u_R.$$

Therefore,

$$\frac{d}{dt}\int_{-M}^{M} u(x,t)dx = s(u_L - u_R).$$

Hence, the speed of propagation of the wave is

$$s = \left(\frac{u_L^2}{2} - \frac{u_R^2}{2}\right)/(u_L - u_R) = \frac{u_L + u_R}{2}. \tag{36}$$

Remark The argument above is valid for a more general equation of the form

$$u_t + [f(u)]_x = 0. \tag{37}$$

Such equations are called *hyperbolic conservation laws.* The shock speed is given by

$$s = \frac{f(u_L) - f(u_R)}{u_L - u_R} = \frac{\text{jump in } f(u)}{\text{jump in } u}. \tag{38}$$

This equation is called the *Rankine-Hugoniot condition.*

3.3. **Characteristics of the Burgers equation.** The characteristics of Eq. (33) are given by

$$\frac{dx}{dt} = u(x,t). \tag{39}$$

Let us show that u is constant along the characteristics. Let $(x(t),t)$ be a characteristic. Then

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

Therefore, the solution of Eq. (39) is given by

$$x(t) = u(x(0),0)t + x(0) = u_0(x_0)t + x_0, \quad \text{where} \quad x_0 = x(0), \quad u_0(x) = u(x,0). \tag{40}$$

Eq. (40) shows that

- the characteristics are straight lines,
- they may intersect,
- they do not necessarily cover the entire (x,t) space.

This is a new phenomenon in comparison with the linear advection equation $u_t + au_x = 0$. For the linear advection equation, there is a unique characteristic passing through every point of the (x,t) space. Thus, its characteristics never intersect and cover the entire space.

Moreover, even for a smooth initial speed distribution $u_0(x)$ the solution of the Burgers equation may become discontinuous in a finite time. This happens when $u_0'(x)$ is negative somewhere. Then the characteristics intersect, i.e., the wave breaks. Let us find the break time. Consider two characteristics $x(t) = u_0(x_1)t + x_1$ and $x(t) = u_0(x_2)t + x_2$. Then we equate

$$x(t) = u_0(x_1)t + x_1 = u_0(x_2)t + x_2.$$

Then the time at which they intersect is

$$t = -\frac{x_2 - x_1}{u_0(x_2) - u_0(x_1)}.$$

Therefore, the break time is

$$T_b = \min_{x_1,x_2\in\mathbb{R}}\left(-\frac{x_2 - x_1}{u_0(x_2) - u_0(x_1)}\right) = \frac{1}{\max_{x_1,x_2\in\mathbb{R}}\left(\frac{u_0(x_1)-u_0(x_2)}{x_2-x_1}\right)}$$
$$= \frac{1}{\max_{x_1,x_2\in\mathbb{R}}\left(-\frac{1}{x_2-x_1}\int_{x_1}^{x_2} u_0'(x)dx\right)} = \frac{1}{\max_{x\in\mathbb{R}}[-u_0'(x)]}.$$

In words, the *breaking time is the reciprocal of the maximal negative slope of* u_0'. An example is shown in Figure 4. The initial data $u_0(x) = \exp(-16x^2)$ and the corresponding characteristics of the Burgers equation are shown in Fig. 4 (a) and (b) respectively. The breaking time is found by taking the derivative of $u_0(x)$ and finding a point at which it has the most negative slope which is $x^* = 1/\sqrt{32}$. The breaking time is

$$T_b = \left[32x^* \exp(-16(x^*)^2)\right]^{-1} \approx 0.2915.$$

The solution at times $t = 0.5$ and $t = 0.8$ obtained by the method of characteristics is shown in Fig. 4 (c) and (e). The numerical solution computed by Godunov's method (see Section 5) is shown in Fig. 4 (d) and (f). The solution obtained by the method of characteristics is triple-valued at some values of x and non-physical in the sense that it is not the vanishing viscosity solution (see Section 3.1). In contrast, the solution computed by Godunov's method tends to the vanishing viscosity solution as we refine the mesh.

3.4. **Weak solutions.** We have seen in the previous section that a solution to the Burgers equation can become discontinuous even if the initial data are smooth. Then the discontinuity travels with a certain speed, the shock speed s, given by Eq. (36). In Section 3.2 we found s using the underlying integral conservation law. However, at this point, we can only call the step function given by Eqs (34), (35), and (36) a solution to the integral conservation law rather than the solution to Eq. (33). In order to validate discontinuous solutions for differential equations, the concept of the weak solutions was introduced (see e.g. [3]). This extension of the concept of the solution must satisfy the following requirements:

- a smooth function is a weak solution if and only if it is a regular solution,
- a discontinuous function can be a weak solution,
- only those discontinuous functions that satisfy the associated integral equation can be weak solutions.

Motivation. Let $\phi(x,t)$ be an infinitely smooth function with a compact support, i.e., it is different from zero only within some compact subset of the space $(x,t) = \mathbb{R}\times[0,+\infty)$. Let $u(x,t)$ be a smooth solution of a hyperbolic conservation law given by Eq. (37). Then

$$(41) \qquad 0 = \int_0^\infty \int_{-\infty}^\infty (u_t + [f(u)]_x)\phi dxdt = \left(\int_{-\infty}^\infty \phi u dx\right)\Big|_0^\infty - \int_0^\infty \int_{-\infty}^\infty \phi_t u + \phi_x f(u) dxdt.$$

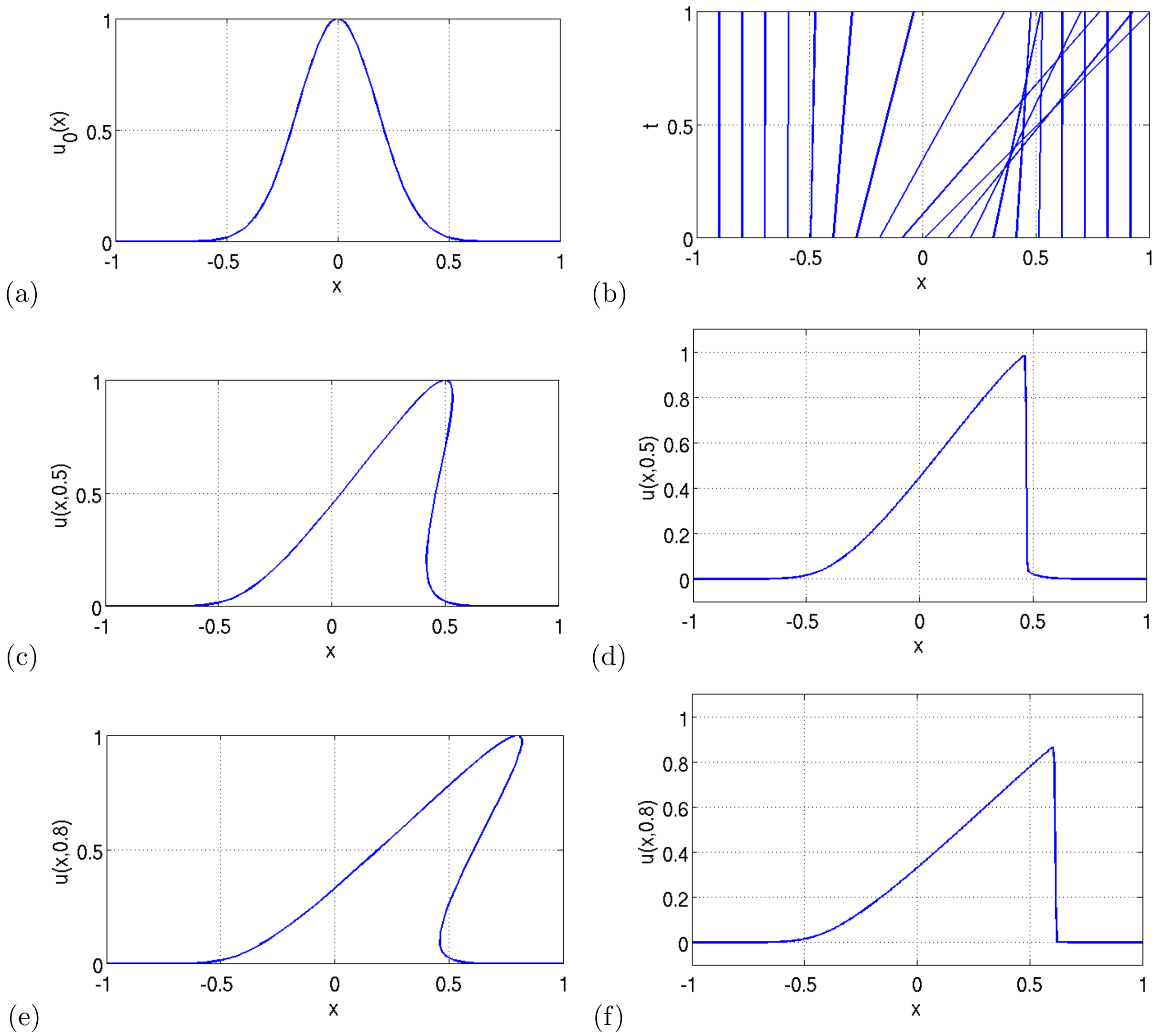


FIGURE 4. (a) The initial data $u_0(x) = \exp(-16x^2)$. (b) The corresponding characteristics of the Burgers equation.(c-d) The solution at time $t = 0.5$ obtained by the method of characteristics (c) and computed numerically (by Godunov's method) (d). (e-f) The solution at time $t = 0.8$ obtained by the method of characteristics (e) and computed numerically (by Godunov's method) (f).

Therefore, here is the definition [3].

Definition 1. *$u(x,t)$ is a weak solution of the conservation law $u_t + [f(u)]_x = 0$ if for any infinitely differentiable function $\phi(x,t)$ with a compact support*

$$\int_0^{\infty}\int_{-\infty}^{\infty}\phi_t u + \phi_x f(u)dxdt = \left(\int_{-\infty}^{\infty}\phi u dx\right)\Big|_0^{\infty} = -\int_{-\infty}^{\infty}\phi(x,0)u(x,0)dx. \tag{42}$$

Such a function $\phi(x,t)$ is called a *test function*.

4. The Riemann problem

In this section, we consider the following initial value problem for the hyperbolic conservation law $u_t + [f(u)]_x = 0$:

$$u(x,0) = \left\{\begin{array}{ll} u_L & x < 0 \\ u_R & x > 0 \end{array}\right\}.$$

This problem is called the *Riemann problem*. We will consider two cases.

4.1. Case 1: $u_L > u_R$.

In this case, the characteristics cover the entire (x,t) space but also cross. Hence the construction of the solution using only the characteristics is ambiguous. Let us show that in this case,

$$u(x,t) = \left\{\begin{array}{ll} u_L & x < st \\ u_R & x > st \end{array}\right\} \tag{43}$$

is a weak solution if and only if the shock speed is given by the *Rankine-Hugoniot condition*

$$s = \frac{f(u_L) - f(u_R)}{u_L - u_R}. \tag{44}$$

The characteristics of this solution are shown in Fig. 5(a).

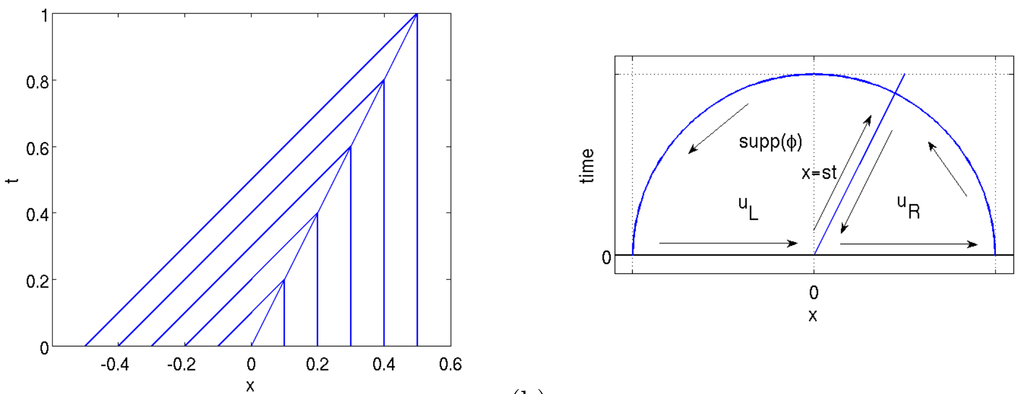
Let $\phi(x,t)$ be a test function. First, suppose that support U lies entirely in one of the sets $\{x < st\}$ or $\{x > st\}$. Then since $u(x,t)$ is constant in each of these sets, it satisfies the Burgers equation on the support of ϕ. Then using Eq. (41) we conclude that Eq. (42) holds.

Now suppose that the support U of ϕ is divided by the line $x = st$ into two sets U_L and U_R (Fig. 5(b)). Then we have

$$\int_0^{\infty}\int_{-\infty}^{\infty}(\phi_t u + \phi_x f(u))\,dxdt = \iint_{U_L}[\phi_t u + \phi_x f(u)]dxdt + \iint_{U_R}[\phi_t u + \phi_x f(u)]dxdt,$$

where $f(u) = u^2/2$. Applying the Green identity

$$\iint_D (P_x - Q_t)dxdt = \int_{\partial D} Pdt + Qdx \tag{45}$$



(a) (b)

FIGURE 5. (a) The characteristics for the shock wave. (b) Illustration for the proof that the shock wave is the unique weak solution.

and noting that u is constant within U_L and within U_R, hence $(\phi u)_t = \phi_t u$ and $(\phi f(u))_x = \phi_x f(u)$, we continue

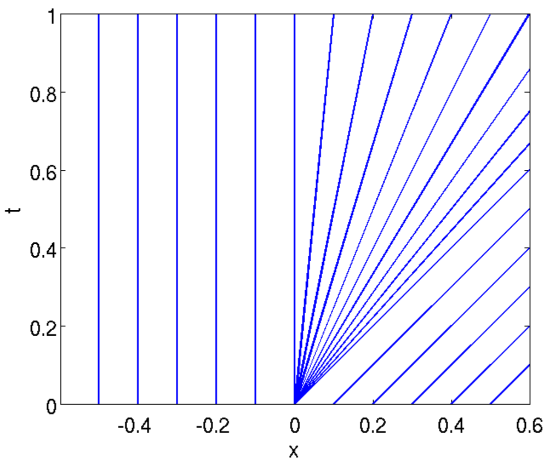
$$
\begin{aligned}
&= \int_{\partial U_L} \phi\left(f(u_L)dt - u_L dx\right) + \int_{\partial U_R} \phi\left(f(u_R)dt - u_R dx\right) \\
&= \int_{x=st} \phi\left(\frac{f(u_L)}{s} - u_L\right)dx - \int_{x=st} \phi\left(\frac{f(u_R)}{s} - u_R\right)dx - \int_{-\infty}^{0} \phi u_L dx - \int_{0}^{\infty} \phi u_R dx \\
&= \int_{x=st} \left(\frac{f(u_L) - f(u_R)}{s} - (u_L - u_R)\right)\phi dx - \int_{-\infty}^{\infty} \phi(x,0)u(x,0)dx
\end{aligned}
$$

The first integral in the last equality is zero for any test function ϕ iff s is given by (44) Furthermore, the discussion in Section 3.1 indicates that the solution given by Eq. (43) is the *vanishing viscosity solution*, i.e. the limit of the solutions of Eq. (31) as $\nu \to 0$.

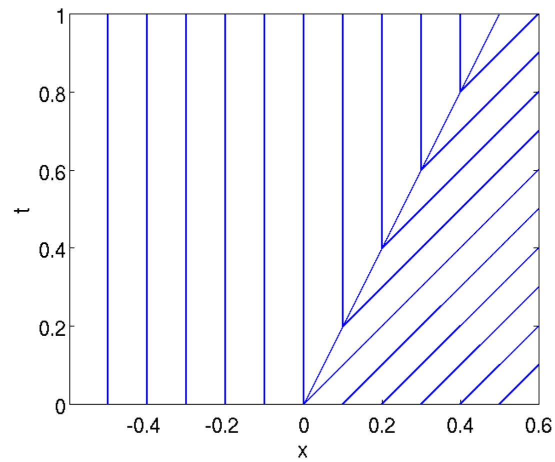
Remark Note that we did not use the fact that $u_L > u_R$ is while checking that (43) is a weak solution. This means that this is also a weak solution if $u_L < u_R$ which is nonphysical!

4.2. **Case 2:** $u_L < u_R$**.** Then the characteristics do not cross but do not cover the entire space (x,t). There are many weak solutions. Two of them are shown in Fig. 6. The one in Fig. 6(a) called the *rarefaction fan* or the *transonic rarefaction* is given by

$$
(46) \qquad u(x,t) = \left\{ \begin{array}{ll} u_L, & x < u_L t \\ x/t, & u_L t \le x \le u_R t \\ u_R, & x > u_R t \end{array} \right\}.
$$



(a)



(a)

FIGURE 6. The characteristics for the rarefaction fan (a) and the rarefaction shock (b). Both of these are weak solutions. However, the rarefaction wave is the physical vanishing viscosity solution, while the rarefaction shock is not.

Another weak solution given by (43), shown in Fig. 6 (b), is called the *rarefaction shock*. A family of weak solutions with rarefaction fans starting at some point of the shock line can be constructed out of these two weak solutions.

Despite there being many weak solutions, only one of them is "physical", i.e., the vanishing viscosity solution. It is the rarefaction fan solution. One can reject all of the nonphysical weak solutions by analyzing Eq. (31). However, the analysis of the equation with nonzero viscosity is harder than the analysis of the one with zero viscosity. Then an additional simpler-to-verify condition, the so-called *entropy condition* was introduced to eliminate nonphysical weak solutions. There are several variations of the entropy condition. We will state only the simplest one.

Definition 2. *A discontinuity propagating with speed s given by Eq. (38) satisfies entropy condition if $f'(u_L) > s > f'(u_R)$.*

For the Burgers equation this entropy condition reduces to the requirement that if a discontinuity is propagating with speed s then $u_L > u_R$.

4.3. **Finding the exact solution to the Burgers equation.** In this section, we work out an exact solution to the Burgers equation $u_t + [0.5u^2]_x = 0$ for the initial condition

$$u(x,0) = \begin{cases} 0, & x < 0 \\ 1, & 0 < x < 1 \\ 0, & x > 1 \end{cases}. \tag{47}$$

The solution is constructed using the fact that the shock speed is given by the Rankine-

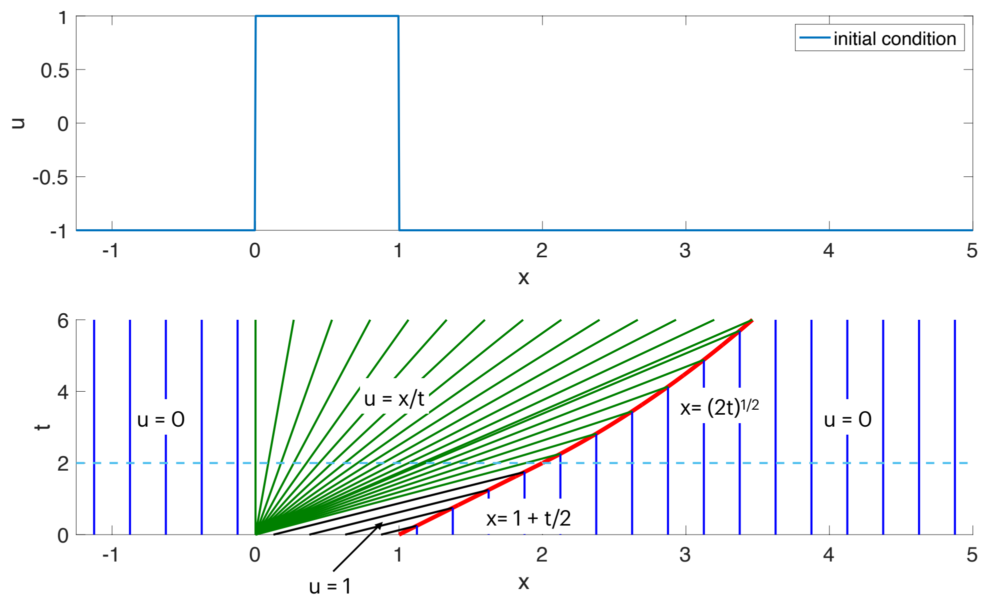


Figure 7. The exact solution to the Burgers equation (bottom) for the initial condition given by a pulse (top).

Hugoniot condition whenever there is a discontinuity in u, and that the gaps between the characteristics are filled with rarefaction fans. A rarefaction fan originating at a point x_0 at time 0 gives $u(x,t) = (x - x_0)/t$.

It is evident from the initial condition (47) that there will be a rarefaction fan emanating from the origin and a shock line emanating from $(1,0)$ on the (x,t)-plane – see Fig. 7. The shock line separates the region with $u = u_L = 1$ and $u = u_R = 0$. Hence the shock speed is $0.5(u_L + u_R) = 0.5$. The shock line is given by

$$x(t) = 1 + {}^{t}\!/_{2}.$$

However, we note that the shock line collides with the rarefaction fan. This happens at x found by equating the shock line $x(t) = 1 + {}^{t}\!/_{2}$ and the boundary characteristic of the rarefaction fan $x(t) = t$:

$$t = 1 + {}^{t}\!/_{2}.$$

Hence, the collision point is $x = 2$, $t = 2$. Therefore, for time $t \geq 2$, the shock line will result from the collision of $u = x/t$ on the left and $u = 0$ on the right. The Rankine-Hugoniot condition gives:

$$s(t) = \frac{\frac{1}{2}(x/t)^2 - 0}{x/t - 0} = \frac{x}{2t}.$$

Hence, to find the shock line, we need to solve the following ODE:

$$\frac{dx}{dt} = s(t) = \frac{x}{2t}, \quad x(2) = 2.$$

Its solution is given by

$$x = \sqrt{2t}.$$

The resulting shock diagram is depicted in Fig. 7(bottom). The exact solution is given by the following formulas. If $t \leq 2$,

$$u(x,t) = \begin{cases} 0, & x < 0 \\ x/t, & 0 \leq x \leq t \\ 1, & t < x < 1 + {}^{t}\!/\!_{2}, \\ 0, & x > {}^{t}\!/\!_{2} \end{cases}. \tag{48}$$

If $t > 2$,

$$u(x,t) = \begin{cases} 0, & x < 0 \\ x/t, & 0 \leq x \leq \sqrt{2t} \\ 0, & x > \sqrt{2t} \end{cases}. \tag{49}$$

5. Numerical methods for hyperbolic conservation laws

Numerical solution of the Burgers equation is a challenging problem because perfectly consistent and stable schemes might propagate discontinuities with wrong speeds and hence fail to converge to the physically correct vanishing viscosity solution with the mesh refinement. In order to address this issue, additional requirements have been imposed on numerical schemes for conservation laws that guarantee that they propagate discontinuities with the right speeds [2].

5.1. **Conservative methods for nonlinear problems.**

A specific difficulty in computing discontinuous solutions of hyperbolic conservation laws can be illustrated by the following simple example.

Example Consider the Burgers equation written in the quasi-linear form

$$u_t + uu_x = 0.$$

Let

$$u(x,0) = u_0(x) = \begin{cases} 1, & x < 0, \\ 0, & x \geq 0. \end{cases}$$

A generalization of the left upwind method for the case where the speed $a = u$ gives

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

The initial condition gives $U_j^0 = 1$ for $j < 0$ and $U_j^0 = 0$ for $j \ge 0$. Then

$$U_j^1 = \begin{cases} 1 - \frac{k}{h} 1(1-1) = 1, & j < 0, \\ 0 - \frac{k}{h} 0(0 - U_{j-1}^0) = 0, & j \ge 0. \end{cases}$$

Hence, $U_j^1 = U_j^0$. Therefore, $U_j^n = U_j^0$ for all j. This means that the method propagates the discontinuity with a wrong speed $s = 0$.

How can we guarantee that a method propagates the discontinuity at the correct speed? Let us recall what a conservation law is. A hyperbolic conservation law

$$u_t + [f(u)]_x = 0 \tag{50}$$

means that the conserved quantity

$$\int_L^R u(x,t)dx$$

can only change due to the flux through the boundaries, i.e.

$$\int_L^R u(x,t+k)dx - \int_L^R u(x,t)dx = \int_0^k f(u(L,t+\tau))d\tau - \int_0^k f(u(R,t+\tau))d\tau.$$

Definition 3. *A numerical method is in a conservation form if it can be rewritten in the form*

$$U_j^{n+1} = U_j^n - \frac{k}{h}\left[F(U_{j-p}^n, \dots, U_{j+q}^n) - F(U_{j-p-1}^n, \dots, U_{j+q-1}^n)\right], \tag{51}$$

for some function F which is called the numerical flux. A method that can be written in a conservation form is called conservative.

5.2. **Discrete conservation.** The basic principle of a conservation law is that the conserved quantity in a given interval $[L,R]$ can change only due to the flux through the boundaries, i.e.

$$\int_L^R u(x,t_2)dx = \int_L^R u(x,t_1)dx - \left(\int_{t_1}^{t_2} f(u(R,t))dt - \int_{t_1}^{t_2} f(u(L,t))dt\right).$$

A similar identity holds for conservative methods due to telescoping the sums. If we sum a conservative method (51) from $j = j_L$ to $j = j_R$ we get

$$h\sum_{j=j_L+1}^{j_R} U_j^{n+1} = h\sum_{j=j_L+1}^{j_R} U_j^n - k\sum_{j=j_L+1}^{j_R}\left[F(U^n;j) - F(U^n;j-1)\right].$$

Here, for brevity, we have used the notation $F(U^n;j) \equiv F(U_{j-p}^n, \dots, U_{j+q}^n)$. The sum telescopes and only boundary fluxes remain as a result:

$$h\sum_{j=j_L+1}^{j_R} U_j^{n+1} = h\sum_{j=j_L+1}^{j_R} U_j^n - k\left[F(U^n;j_R) - F(U^n;j_L)\right].$$

Therefore, the numerical solution, like the exact solution, allows the conserved quantity $h\sum_{j=j_L+1}^{j_R} U_j^n$ to change only due to the flux through the boundaries.

5.3. **Consistency.** A method of the form (51) is consistent with the original conservation law (50) if the numerical flux function F reduces to the true flux function f (the one in the equation $u_t + [f(u)]_x = 0$) for the case of constant flow. I.e., if $u(x,t) \equiv \bar{u}$ then

$$F(\bar{u},\dots,\bar{u}) = f(\bar{u}). \tag{52}$$

Furthermore, if the arguments of F approach some constant value $\bar{u}$, F should approach $f(\bar{u})$ smoothly, i.e.

$$\lim_{v_1,\dots,v_r\to\bar{u}} F(v_1,\dots,v_r) = f(\bar{u}). \tag{53}$$

It is sufficient to require the Lipschitz continuity of F whenever f is Lipschitz-continuous in order to satisfy the smoothness condition given by Eq. (53). Recall that $F(v_1,\dots,v_r)$ is *Lipschitz continuous* if

$$|F(v_1,\dots,v_r) - F(w_1,\dots,w_r)| \le K\max\{|v_1-w_1|,\dots,|v_r-w_r|\}, \tag{54}$$

where K is a constant depending on F but not on its arguments called the *Lipschitz constant.*

Therefore, sufficient consistency conditions are given by

$$F(\bar{u},\dots,\bar{u}) = f(\bar{u}), \tag{55}$$

$$|F(v_1,\dots,v_r) - F(\bar{u},\dots,\bar{u})| \le K\max\{|v_1-\bar{u}|,\dots,|v_r-\bar{u}|\}. \tag{56}$$

5.4. **Generalization of methods developed for the advection equation.** Now we generalize some methods developed for the advection equation $u_t + au_x = 0$ for hyperbolic conservation laws (50) making sure that these generalizations are conservative and consistent.

Lax-Friedrichs

The generalization of the Lax-Friedrichs method to Eq. (50) takes the form

$$U_j^{n+1} = \frac{1}{2}\left(U_{j-1}^n + U_{j+1}^n\right) - \frac{k}{2h}\left(f(U_{j+1}^n) - f(U_{j-1}^n)\right). \tag{57}$$

Observe that this method can be rewritten in the conservative form as

$$U_j^{n+1} = U_j^n - \frac{k}{h}\left[F(U_{j+1}^n, U_j^n) - F(U_j^n, U_{j-1}^n)\right], \quad \text{where}$$

$$F(U_{j+1}, U_j) = \frac{h}{2k}\left(U_j - U_{j+1}\right) + \frac{1}{2}\left(f(U_j) + f(U_{j+1})\right). \tag{58}$$

Consistency check:

$$F(U,U) = \frac{h}{2k}(U-U) + \frac{1}{2}(f(U)+f(U)) = f(U),$$

$$|F(V,W) - F(U,U)| = \left|\frac{h}{2k}(V-U) - \frac{h}{2k}(W-U) + \frac{1}{2}(f(V)-f(U)) + \frac{1}{2}(f(W)-f(U))\right|$$

$$\le \left(\frac{h}{k} + K_f\right)\max\{|V-U|, |W-U|\},$$

where K_f is the Lipschitz constant for the true flux f (we have assumed that f is Lipschitz-continuous).

Lax-Wendroff

Recall that the Lax-Wendroff method for the advection equation $u_t + au_x = 0$ is given by

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

The most straightforward generalization requires the evaluation of the Jacobian $f'(u)$ (in the multidimensional case this is especially unpleasant). Two alternative extensions were developed by Richtmyer and MacCormack. Both of them are two-step procedures.

Richtmyer:

$$U_{j+1/2}^{n+1/2} = \frac{1}{2}\left(U_j^n + U_{j+1}^n\right) - \frac{k}{2h}\left[f(U_{j+1}^n) - f(U_j^n)\right],$$

(59) $$U_j^{n+1} = U_j^n - \frac{k}{h}\left[f(U_{j+1/2}^{n+1/2}) - f(U_{j-1/2}^{n+1/2})\right],$$

MacCormack:

$$U_j^* = U_j^n - \frac{k}{h}\left[f(U_{j+1}^n) - f(U_j^n)\right],$$

(60) $$U_j^{n+1} = \frac{1}{2}\left(U_j^n + U_j^*\right) - \frac{k}{2h}\left[f(U_j^*) - f(U_{j-1}^*)\right],$$

Exercise Show that the methods (59) and (60) reduce to the Lax-Wendroff method for $f(u) \equiv au$. Show that methods (59) and (60) are second-order consistent on smooth solutions. Determine numerical flux functions for the methods (59) and (60) and show that they are conservative.

5.5. **Convergence.** *Lax and Wendroff have proven that a consistent and conservative method converges to a weak solution of the conservation law almost everywhere as $k,h \to 0$ and k/h satisfies stability conditions.*

However, as we know from Section 4, a weak solution might be non-unique in the case where $u_L < u_R$ at the discontinuity. The entropy condition (see Definition 2) allows us to reject all physically irrelevant weak solutions and select the physically correct one, which is consistent with the limit of the solutions of the viscous equation $u_t + [f(u)]_x = \epsilon u_{xx}$ as $\epsilon \to 0$. There is a danger that the numerical solution of $u_t + [f(u)]_x = 0$ by a conservative method converges to a weak but physically irrelevant solution. The following example demonstrates that a seemingly reasonable method can fall into this trap.

Example Consider the Burgers equation $u_t + [u^2/2]_x = 0$ with the initial data

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

The physically relevant solution to this problem (the vanishing viscosity solution) is *a transonic rarefaction* given by

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

The stationary discontinuity $u(x,t) = u_0(x)$ is another weak solution (note, the shock speed is zero: $s = (u_L + u_R)/2 = (1 + (-1))/2 = 0$).

Let us set the numerical initial velocity to

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

and consider a conservative method obtained by a generalization of the upwind scheme:

$$U_j^{n+1} = U_j^n - \frac{k}{h}\left[F(U_{j+1}^n, U_j^n) - F(U_j^n, U_{j-1}^n)\right], \quad \text{where}$$

$$F(v,w) = \begin{cases} f(v), & \text{if} \quad (f(v) - f(w))/(v-w) \ge 0, \\ f(w), & \text{if} \quad (f(v) - f(w))/(v-w) < 0, \end{cases}.$$

For the problem above, we have $f(v) = v^2/2$ and

$$F(1,1) = f(1) = 1/2, \quad F(-1,-1) = f(-1) = 1/2, \quad F(1,-1) = f(1) = 1/2.$$

Therefore,

$$U_j^1 = U_j^0 - \frac{k}{h}[1/2 - 1/2] = U_j^0.$$

Hence $U_j^n = U_j^0$ for all n and j. Thus, the numerical solution will converge to the physically irrelevant "stationary discontinuity" weak solution. On the other hand, if we set

$$U_j^0 = \begin{cases} -1, & j < 0, \\ 0, & j = 0, \\ 1, & j > 0, \end{cases}$$

then the numerical solution by the same method will be the transonic rarefaction. Check this.

In order to outlaw conservative methods that might produce discrete approximations to physically irrelevant weak solutions, we need a discrete analog of the entropy condition (Definition 2). Unfortunately, the simple version of the entropy condition in Definition 2 is not extended to the discrete case. Another, more involved version, requiring an introduction of an entropy function, has a discrete analog. I will not discuss this topic here. Instead, I refer curious students to Ref. [2], Sections 3.8.1, 12.5, and 13.4.

In the next section, we will discuss a different class of methods for conservation laws for which the version of the entropy condition involving an entropy function can be readily verified. (We will not verify it here. You are welcome to read Section 13.4 in [2]).

5.6. **Godunov's method.** The idea of Godunov's method (1959) is the following. Let U_j^n be a numerical solution at time $t_n = kn$. We define a function $\tilde{u}(x,t)$ for $t_n \le t \le t_{n+1}$ as follows. At $t = t_n$,

$$\tilde{u}(x,t_n) = U_j^n, \quad x_j - h/2 < x < x_j + h/2, \quad j = 2,\dots,n-1.$$

Then $\tilde{u}(x,t)$ is the solution to the collection of the Riemann problems on the interval $[t_n, t_{n+1}]$. If the time step k is small enough so that the characteristics starting at the points $x_j + h/2$ do not intersect within this interval (i.e., the CFL condition is satisfied), then $\tilde{u}(x,t_{n+1})$ is determined unambiguously. Then the numerical solution on the next layer, U_j^{n+1} is defined by averaging $\tilde{u}(x,t_{n+1})$ over the intervals $x_j - h/2 < x < x_j + h/2$:

$$U_j^{n+1} = \frac{1}{h}\int_{x_j-h/2}^{x_j+h/2} \tilde{u}(t_{n+1})dx. \tag{61}$$

This idea is illustrated in Fig. 8.

In practice, the cell averages (61) can be easily calculated using the integral form of conservation law:

$$\begin{aligned}\frac{1}{h}\int_{x_j-h/2}^{x_j+h/2} \tilde{u}(x,t_{n+1})dx &= \frac{1}{h}\int_{x_j-h/2}^{x_j+h/2} \tilde{u}(x,t_n)dx \\ &- \frac{1}{h}\left[\int_{t_n}^{t_{n+1}} f\left(\tilde{u}(x_j+h/2,t)\right)dt - \int_{t_n}^{t_{n+1}} f\left(\tilde{u}(x_j-h/2,t)\right)dt\right].\end{aligned}$$

Observing that $\tilde{u}(x_j+h/2,t)$ and $\tilde{u}(x_j-h/2,t)$ are constant over the time interval $[t_n,t_{n+1}]$ we obtain

$$\begin{aligned}&U_j^{n+1} = U_j^n - \frac{k}{h}\left[F(U_j^n,U_{j+1}^n) - F(U_{j-1}^n,U_j^n)\right],\\ &\quad\text{where}\quad F(u_L,u_R) = f\left(u^*(u_L,u_R)\right).\end{aligned}\tag{62}$$

Therefore, Godunov's method is conservative.

The analysis below will be done for a scalar conservation law, i.e., for the case where u is a scalar function. The value $u^*(u_L,u_R)$ in the numerical flux function $F(u_L,u_R) = f\left(u^*(u_L,u_R)\right)$ is defined so that the entropy condition is satisfied and hence the weak solution, to which the numerical solution converges, is the vanishing viscosity solution. If $f(u)$ is convex (if f is twice differentiable then $f''(u) > 0$), the following four cases must be considered:

(1) $f'(u_L) \ge 0$ and $f'(u_R) \ge 0$. Then $u^* = u_L$.
(2) $f'(u_L) \le 0$ and $f'(u_R) \le 0$. Then $u^* = u_R$.

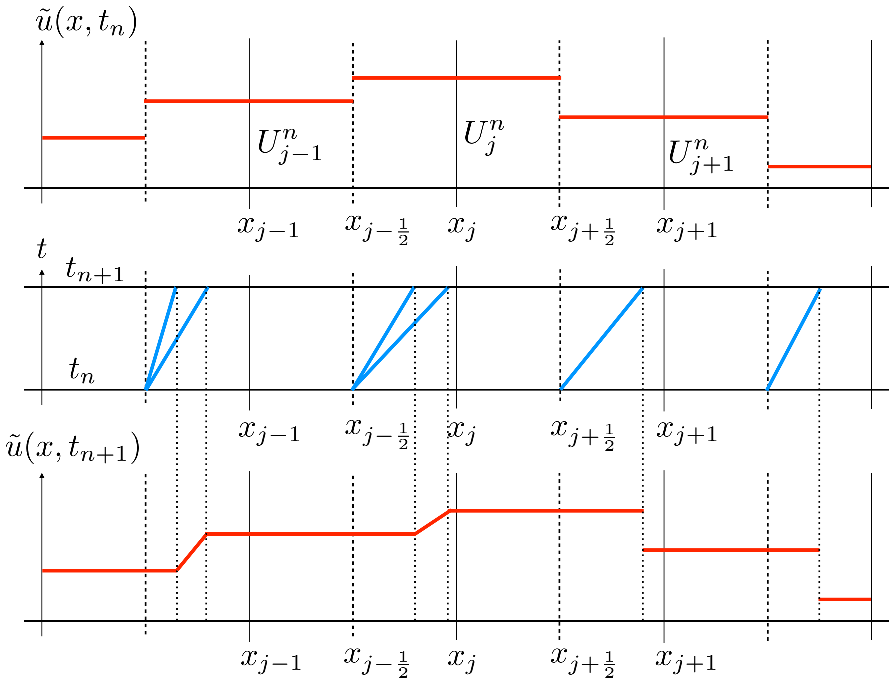


FIGURE 8. Top: The function $\tilde{u}(t_n)$ obtained by setting its value in the intervals $x_j - h/2 < x < x_j + h/2$ to U_j^n. Middle: The diagram on the (x,t)-plane showing lines of discontinuity and the critical characteristics. It allows us to obtain $\tilde{u}(t_{n+1})$. Bottom: $\tilde{u}(t_{n+1})$ obtained using the diagram on the (x,t)-plane.

(3) $f'(u_L) \ge 0 \ge f'(u_R)$. Then

$$u^* = \begin{cases} u_L, & \text{if } \frac{f(u_L)-f(u_R)}{u_L-u_R} > 0, \\ u_R, & \text{if } \frac{f(u_L)-f(u_R)}{u_L-u_R} < 0. \end{cases} \tag{63}$$

(4) $f'(u_L) < 0 < f'(u_R)$. Then $u^* = u_s$ (transonic rarefaction), where the value u_s is such that $f'(u_s) = 0$. It is called the *sonic point*. For example, for the Burgers equation $u_t + [u^2/2]_x = 0$, $u_s = 0$.

In the first three cases, the value u^* is either u_L and u_R, and it can be simply determined by Eq. (63). Note that in Cases 1 and 2, u^* is the same whether the physically correct weak solution to the Riemann problem is a shock wave or a rarefaction. Only in Case 4, the transonic rarefaction, the value of u^* differs from the one determined by Eq. (63). This is the value of u for which the characteristic speed is zero.

Note that the numerical flux determined by Cases 1 - 4 can be rewritten more compactly as

$$F(u_L, u_R) = \begin{cases} \min_{u_L \le u \le u_R} f(u), & \text{if} \quad u_L \le u_R, \\ \max_{u_R \le u \le u_L} f(u), & \text{if} \quad u_L > u_R. \end{cases} \tag{64}$$

It was proven that the numerical flux given by Eq. (64) gives the physically correct flux for scalar conservation laws even if $f(u)$ is non-convex.

Here we considered only the methods for scalar conservation laws. In applications relevant to gas and fluid dynamics the conservation laws are non-scalar. Further reading: R. LeVeque, Finite Volume Methods for Hyperbolic Problems, Cambridge University Press, 2002.

References

[1] G. I. Barenblatt, Scaling, self-similarity, and intermediate asymptotics, Cambridge University Press, 1996
[2] R. J. LeVeque, Numerical Methods for Conservation Laws, Birkhauser, Basel, Boston, Berlin, 1992
[3] Lawrence C. Evans, Partial Differential Equations, AMS, Providence, RI, 1998