

# Chapter 1

## Lecture

In this chapter, we learn how to solve hydrodynamical equations and magnetohydrodynamical equations. These equations are hyperbolic partial differential equations and describe the propagation of waves. The wave property is key in constructing a numerical scheme for solving hyperbolic partial differential equations. Thus, in the first half of this chapter, we learn how to solve a simple one-dimensional wave equation. The numerical method gives us a basis for the hydrodynamical and magnetohydrodynamical equations, which are multi-dimensional and nonlinear.

### 1.1 Finite Difference Method

Physical quantities such as density and velocity are described as a field, i.e., a function of time and space, in hydrodynamics and magnetohydrodynamics. These quantities are described by arrays having finite numbers of discrete elements in numerical computations. You may already be familiar with this concept, which is also used to solve ordinary differential equations. When we solve partial differential equations, the physical quantities are tensor, because time and space are independent variables.

We consider the simplest case, in which a physical quantity,  $f$ , is a function of time,  $t$ , and the distance,  $x$ . Suppose that we know the value of  $f$  at the points  $(x, t) = (j\Delta x, n\Delta t)$  for any integers  $j$  and  $n$ , where  $\Delta t$  and  $\Delta x$  denote a very short time interval and a very short distance, respectively. Then the array,

$$f_{j,n} \equiv f(x = j\Delta x, t = n\Delta t), \quad (1.1)$$

can be regarded as a function of  $x$  and  $t$  in a practical sense. We call  $\Delta x$  and  $\Delta t$  the grid spacing and time step, respectively.

We can evaluate the partial derivative  $\partial f / \partial x$  from the array  $f_{j,n}$ . However, the evaluation is not unique. Suppose that the function,  $f$ , is a solution of a partial differential equation. Several different expressions converge to  $\partial f / \partial x$  in the limit of  $\Delta x \rightarrow 0$ . For simplicity, we assume that the function  $f$  is smooth and differentiable. Then, the following three expressions:

$$\frac{f_{j+1,n} - f_{j,n}}{\Delta x} = \frac{\partial f}{\partial x} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \Delta x + \mathcal{O}(\Delta x^2), \quad (1.2)$$

$$\frac{f_{j,n} - f_{j-1,n}}{\Delta x} = \frac{\partial f}{\partial x} - \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \Delta x + \mathcal{O}(\Delta x^2), \quad (1.3)$$

$$\frac{f_{j+1,n} - f_{j-1,n}}{2\Delta x} = \frac{\partial f}{\partial x} + \frac{1}{6} \frac{\partial^3 f}{\partial x^3} \Delta x^2 + \mathcal{O}(\Delta x^3), \quad (1.4)$$

denote good approximations to  $\partial f/\partial x$ . Equations (1.2), (1.3), and (1.4), are referred to as the forward difference, backward difference, and central difference, respectively. The forward difference and backward difference are of first-order accuracy because the deviation from the differential is proportional to  $\Delta x$ . The central difference is of second-order accuracy because the deviation is proportional to  $\Delta x^2$ . One might expect that the central difference should always be superior to the forward difference and backward difference because the accuracy is of higher order. This is not true however, and, as shown later, we obtain a better numerical solution when a differential equation is approximated with either the forward difference or the backward difference. An example is presented in the next section.

## 1.2 Linear Wave Equation

We consider the simplest wave equation:

$$\frac{\partial f}{\partial t} + c \frac{\partial f}{\partial x} = 0, \quad (1.5)$$

where  $c$  is a constant and denotes the wave propagation speed. We approximate Equation (1.5) by replacing each term with a finite difference. It is a natural choice to approximate

$$\frac{\partial f}{\partial t} \rightarrow \frac{f_{j,n+1} - f_{j,n}}{\Delta t}, \quad (1.6)$$

and to express  $\partial f/\partial x$  as a function of  $f_{j+1,n}$ ,  $f_{j,n}$ ,  $f_{j-1,n}$ . Then, we can express  $f_{j,n+1}$  as an explicit function of  $f_{j+1,n}$ ,  $f_{j,n}$ ,  $f_{j-1,n}$  and solve the initial value problem of Equation (1.5) easily. If we apply the forward difference to the spatial derivative, we have

$$\frac{f_{j,n+1} - f_{j,n}}{\Delta t} + c \frac{f_{j+1,n} - f_{j,n}}{\Delta x} = 0. \quad (1.7)$$

Equation (1.7) is of forward difference in space and in time. If we apply the backward difference and central difference to the spatial derivative, we have

$$\frac{f_{j,n+1} - f_{j,n}}{\Delta t} + c \frac{f_{j,n} - f_{j-1,n}}{\Delta x} = 0, \quad (1.8)$$

and

$$\frac{f_{j,n+1} - f_{j,n}}{\Delta t} + c \frac{f_{j+1,n} - f_{j-1,n}}{2\Delta x} = 0, \quad (1.9)$$

respectively. Although these three finite difference equations are quite similar in expression, their respective solutions are quite different from one another, as shown below.

## 1.3 Stability Analysis

We next compare the solutions of equations (1.7), (1.8), and (1.9), with the solution of equation (1.5). Suppose that the initial value at  $t = 0$  is given by

$$f = \exp\left(\frac{ikx}{\Delta x}\right), \quad (1.10)$$

where  $i$  denotes  $\sqrt{-1}$ . For later convenience, the function  $f$  is assumed to be complex.<sup>1</sup> Then, the initial value is expressed as

$$f_{j,0} = \exp(ikj), \quad (1.11)$$

in array form. The wave number is assumed to be in the range  $0 < k \leq \pi$  without loss of generality.<sup>2</sup> The solution of the differential equation is

$$f = \exp \left[ \frac{ik(x - ct)}{\Delta x} \right]. \quad (1.12)$$

Thus, ideally, the solution of the finite difference equation is expressed as

$$f_{j,n} = \exp \left[ ik \left( j - \frac{cn\Delta t}{\Delta x} \right) \right]. \quad (1.13)$$

The solution of Equation (1.7) is

$$f_{j,n}^{(F)} = \left\{ 1 - \frac{c\Delta t}{\Delta x} [\exp(ik) - 1] \right\}^n \exp(ikj). \quad (1.14)$$

Although some computations are required to derive the solution, it is easy to confirm the solution by substitution. Similarly, the solutions of (1.8) and (1.9) are

$$f_{j,n}^{(B)} = \left\{ 1 - \frac{c\Delta t}{\Delta x} [1 - \exp(-ik)] \right\}^n \exp(ikj), \quad (1.15)$$

and

$$f_{j,n}^{(C)} = \left\{ 1 - i \frac{c\Delta t}{\Delta x} \sin k \right\}^n \exp(ikj), \quad (1.16)$$

respectively. All of these solutions approach the exact solution given by Equation (1.13) in the limit of very small  $\Delta t$ , because the term in brackets is approximately  $1 - ikc\Delta t/\Delta x$ . The ratios between the approximate and ideal solutions are evaluated to be

$$f_{j,n}^{(F)} = \left[ g^{(F)} \right]^n f_{j,n}, \quad (1.17)$$

$$f_{j,n}^{(B)} = \left[ g^{(B)} \right]^n f_{j,n}, \quad (1.18)$$

$$f_{j,n}^{(C)} = \left[ g^{(C)} \right]^n f_{j,n}, \quad (1.19)$$

where

$$g^{(F)} = \left\{ 1 - \frac{c\Delta t}{\Delta x} [\exp(ik) - 1] \right\} \exp \left( \frac{ick\Delta t}{\Delta x} \right), \quad (1.20)$$

$$g^{(B)} = \left\{ 1 - \frac{c\Delta t}{\Delta x} [1 - \exp(-ik)] \right\} \exp \left( \frac{ick\Delta t}{\Delta x} \right), \quad (1.21)$$

$$g^{(C)} = \left[ 1 - i \frac{c\Delta t}{\Delta x} \sin k \right] \exp \left( \frac{ick\Delta t}{\Delta x} \right). \quad (1.22)$$

---

<sup>1</sup>We assume that the reader is familiar with complex analysis. The same mathematical techniques are used in the analysis of electric circuits.

<sup>2</sup>Since  $j$  is an integer, the value of  $f_{j,0}$  remains unchanged or changes only its sign when we replace  $k$  with  $k' = k + n\pi$ . The restriction,  $0 < k \leq \pi$ , does not cause any loss of generality.

Thus, the error due to the finite difference is estimated from the deviations of  $g^{(F)}$ ,  $g^{(B)}$ , and  $g^{(C)}$  from unity.

The error depends on  $c$ ,  $k$ ,  $\Delta t$ , and  $\Delta x$ . However, we can restrict ourselves to the case of  $c > 0$  without any loss of generality. Note the relation,  $g^{(F)}(-c) = g^{(B)}(c)^*$ , where the asterisk denotes the complex conjugate. The error of the forward difference for negative  $c$  is evaluated from that of the backward difference. This is based on the fact that simultaneous inversion of the propagation speed and coordinate causes no practical change. The error does not depend on the sign of the propagation speed in the central difference because  $g^{(C)}(-c) = g^{(C)}(c)$ . Furthermore, the speed of sound appears only in the form of  $c\Delta t/\Delta x$  in Equations (1.21), (1.22), and (1.22). The term,  $c\Delta t/\Delta x$ , is called the CFL number, where CFL stands for three mathematicians, Courant, Friedrich, and Levy. Thus, the error depends on two nondimensional numbers,  $c\Delta t/\Delta x$  and  $k$ . When the wavenumber is very small ( $0 < k \ll 1$ ), any  $g$  is very close to unity. Thus, we are interested in the case in which the wave number  $k$  is relatively large. Here, we study the dependence of the error on the CFL number for  $k/\pi = 0.25, 0.5$ , and  $0.75$ .

As shown in Figure 1.1, the absolute value  $|g^{(F)}|$  is larger than unity and increases monotonically with  $c\Delta t/\Delta x$ . This causes difficulty in that the wave amplitude increases with time in the forward difference. Even when  $|g^{(F)}|$  is only slightly larger than unity, the wave amplitude increases exponentially with time because we repeat the manipulation.

Next, we examine the backward difference. Figure 1.2 shows the error of the numerical solution obtained by the backward difference. The absolute value of the ratio,  $|g^{(B)}|$ , does not exceed unity in the range  $0 \leq c\Delta t/\Delta x \leq 1$ . Accordingly, the amplitude does not grow unnaturally as long as the CFL number is smaller than or equal to unity. The backward difference gives the best solution among the three differences compared in this section, as shown later. The reader might wonder why this is the case. Why does the wave amplitude diminish in the backward difference solution, whereas it is constant in the exact solution? We shall examine the central difference before answering this question.

Figure 1.3 shows the error of the numerical solution obtained by the central difference. The ratio,  $|g^{(C)}|$ , is the closest to unity than  $|g^{(F)}|$  and  $|g^{(B)}|$  when the CFL number is small. However, the ratio is slightly larger than unity and the wave amplitude grows exponentially with time. Thus, the central difference does not provide a satisfactory solution.

The above analysis, which is referred to as von Neumann stability analysis, is applied to a sinusoidal wave. However, the result is valid for any initial value because any solution is expressed by superpositions of sinusoidal waves for a linear differential equation.<sup>3</sup> If the von Neumann stability analysis does not allow for the possibility of divergence for any sinusoidal wave, the amplitude of the wave never diverges, irrespective of the wave form. We shall confirm this for a rectangular wave.

Figure 1.4 shows the numerical solution of Equation (1.5) for  $c = 1$ . The solution is obtained by the forward difference. The initial value is given by

$$f = \begin{cases} 1 & (x < 0) \\ 0 & (x \geq 0) \end{cases}, \quad (1.23)$$

at  $t = 0$ . The time step and grid spacing are taken to be  $\Delta t = 0.08$  and  $\Delta x = 0.1$ , respectively. The solid curve denotes the numerical solution at  $t = 0.24$ , while the dotted curve denotes the exact solution at  $t = 0.24$ . As expected from the von Neumann stability analysis, the numerical solution shows a large oscillation and is very different from the exact solution. Note that only three time steps spoil the solution completely.

---

<sup>3</sup>This is proven by Fourier analysis.

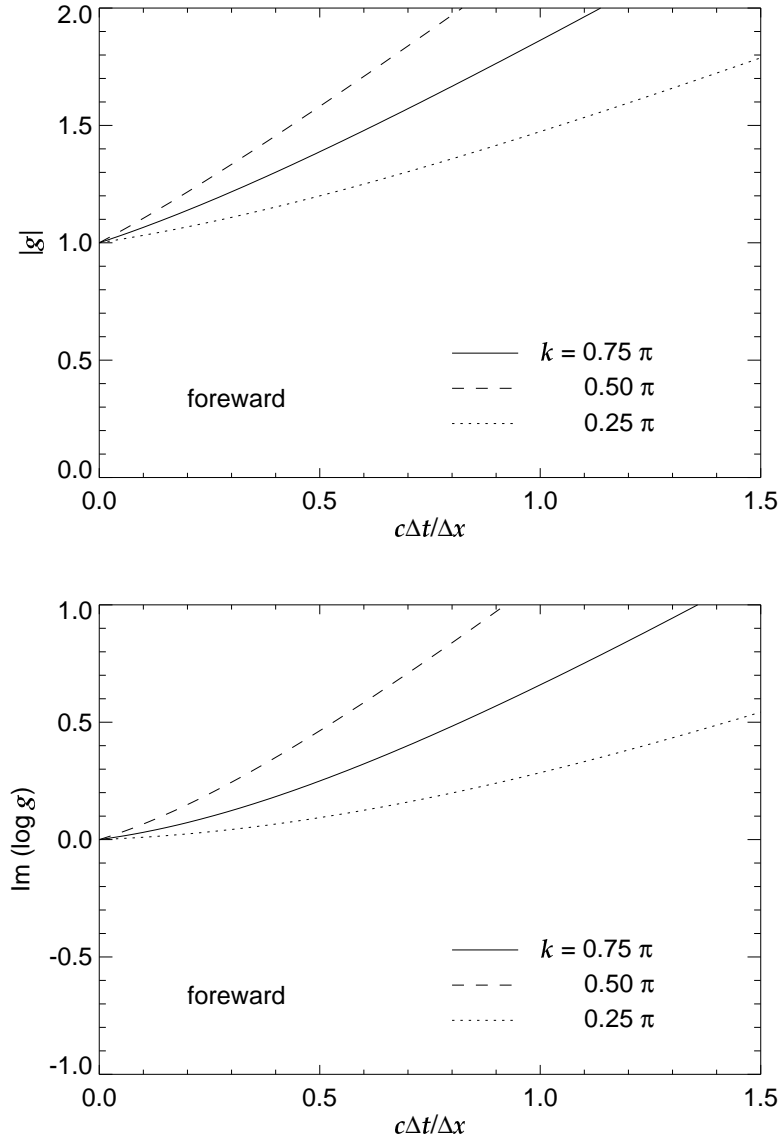


Figure 1.1: The ratio of the numerical solution obtained by the forward difference to the exact solution is shown as a function of the CFL number. The upper panel denotes the absolute value,  $|g^{(\text{F})}|$ , while the lower denotes the phase difference,  $\Im \log g^{(\text{F})}$ .

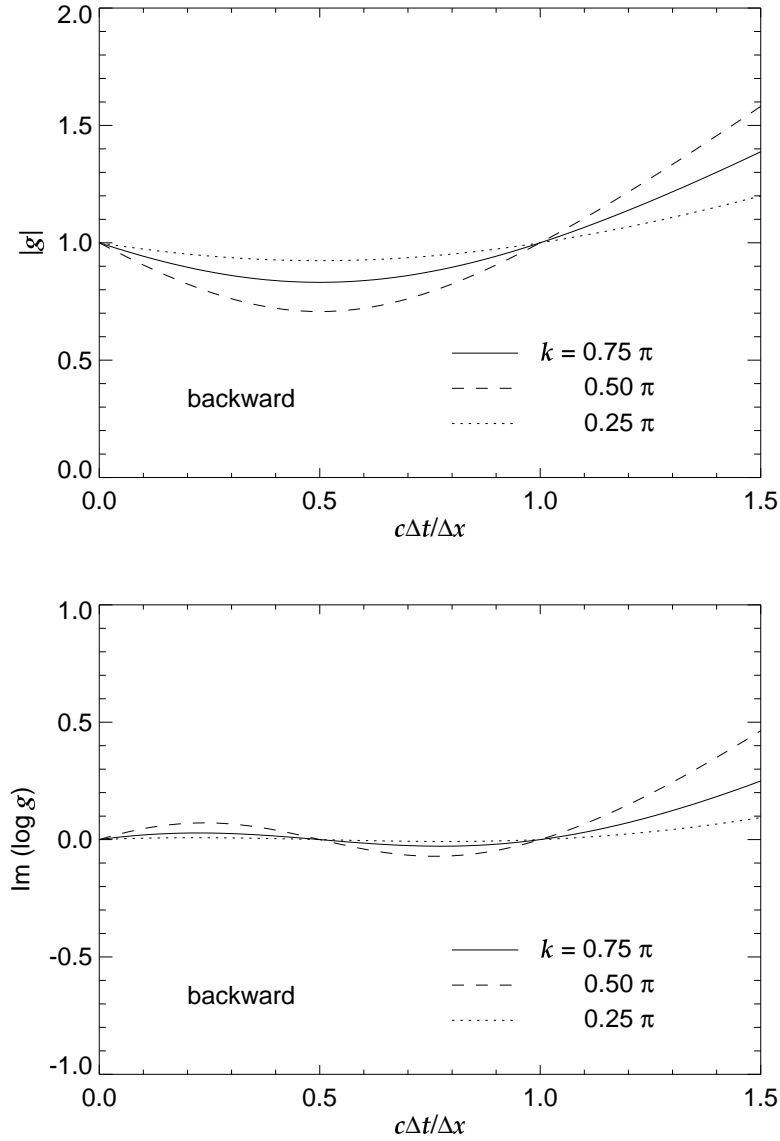


Figure 1.2: The ratio of the numerical solution obtained by the backward difference to the exact solution is shown as a function of the CFL number. The upper panel denotes the absolute value,  $|g^{(F)}|$ , while the lower denotes the phase difference,  $\Im \log g^{(F)}$ .

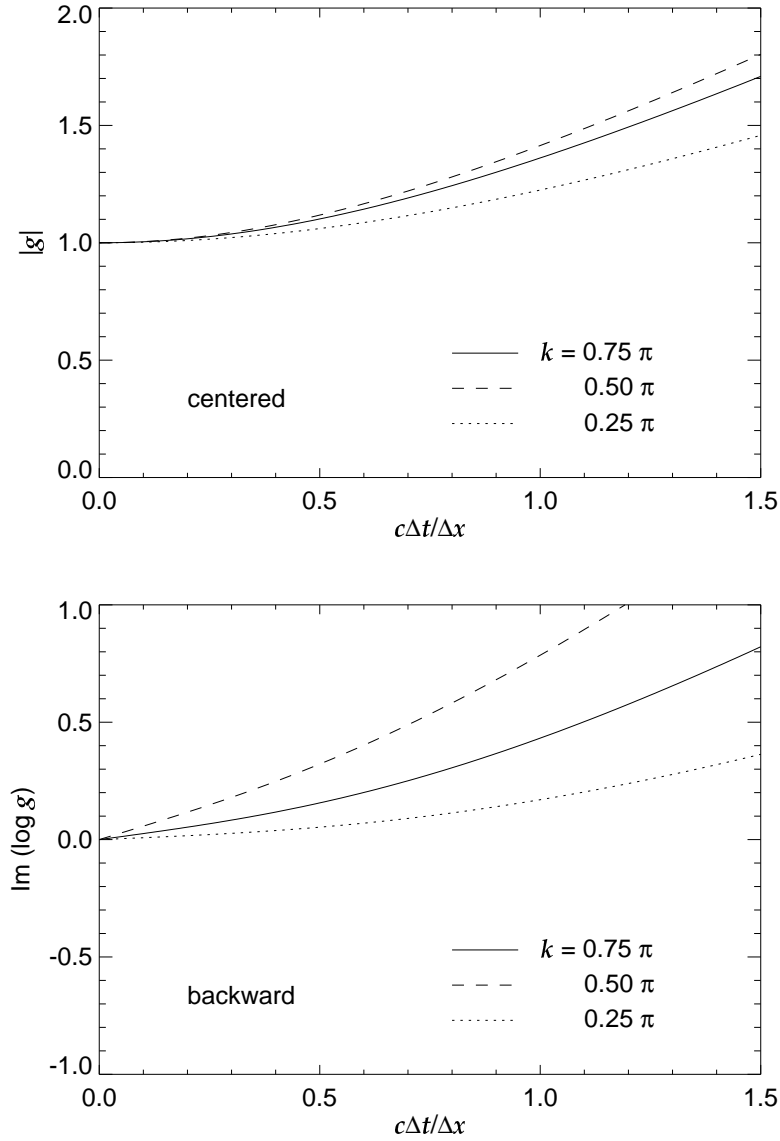


Figure 1.3: The ratio of the numerical solution obtained by the central difference to the exact solution is shown as a function of the CFL number. The upper panel denotes the absolute value,  $|g^{(F)}|$ , while the lower denotes the phase difference,  $\Im \log g^{(F)}$ .

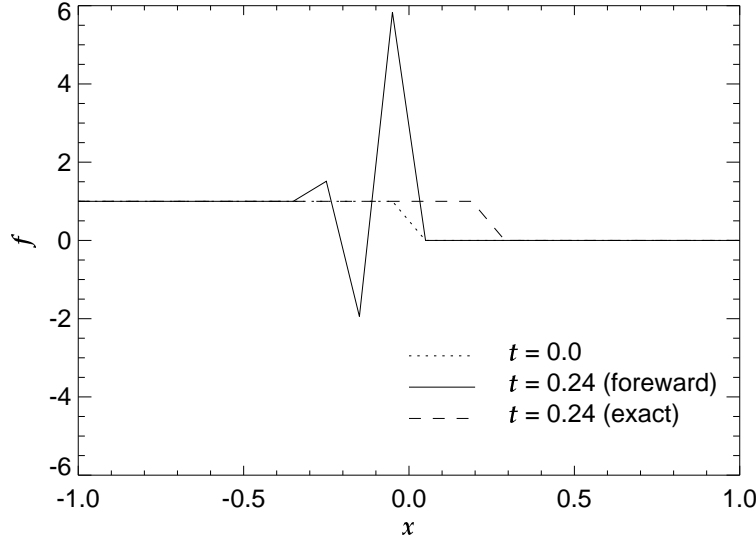


Figure 1.4: Numerical solution obtained by the forward difference. The dotted line denotes the initial condition, while the solid curve denotes the numerical solution for the stage forwarded three time steps. It is very different from the exact solution for the corresponding stage (dashed curve). The CFL number is taken to be 0.8.

The numerical solution shown in Figure 1.4 has another problem in that the wave does not seem to propagate to the right. This problem is also explained by the von Neumann stability analysis. The lower panel of Figure 1.1 shows that the phase error,  $\Im \log g^{(F)}$ , is large for the forward difference. The wave propagation can not be followed properly in the forward difference due to the large phase error.

Figure 1.5 shows the numerical solution of Equation (1.5) for  $c = 1$ . The solution is obtained by the backward difference. The initial condition, wave propagation speed, time step and grid spacing are the same as those for Figure 1.4. The backward difference gives a good approximation of the exact solution, although the numerical solution is appreciably round at the wave front.

Figure 1.6 shows the numerical solution of Equation (1.5) for  $c = 1$ . The solution is obtained by the central difference. The numerical solution is closer to that obtained by the forward difference than that obtained by the backward difference. Equations (1.7)-(1.9) imply that the central difference is something similar to the average of the forward difference and backward difference. The average of the good solution and the bad solution is similar to the bad solution, because oscillation grows exponentially in the bad solution.

Comparison of Figures 1.4, 1.5 and 1.6 show that only the backward difference provides an acceptable solution. Next, we examine whether the backward difference gives a good solution for a later stage. Figure 1.7 shows the solution obtained by the backward difference for  $t = 0.4, 0.8, 1.2, 1.6$ , and  $2.0$ . The time step and grid spacing are again  $\Delta t = 0.08$  and  $\Delta x = 0.1$ , respectively. No artificial oscillation appears in the solution obtained by the backward difference. However, the approximate solution has an appreciably dull wave front at  $t = 20$ , whereas the wave front in the exact solution is sharp. The width of the transition region widens with time due to the damping of short waves, as indicated by the von Neumann stability analysis.

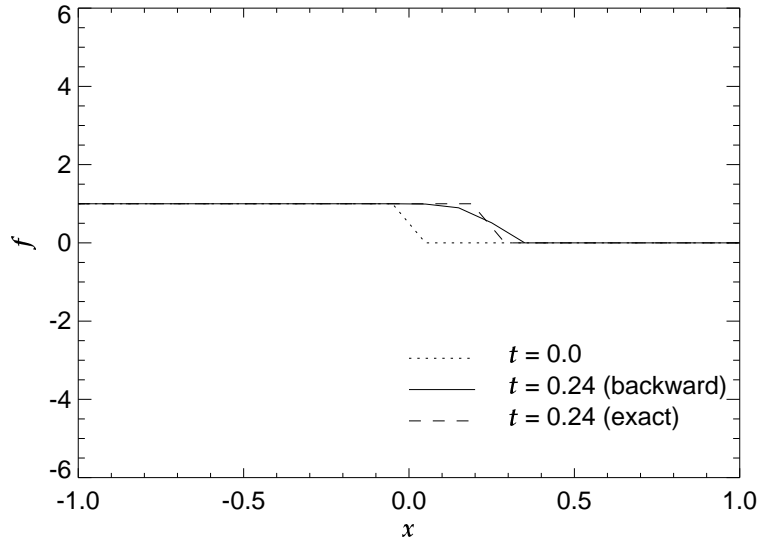


Figure 1.5: Numerical solution obtained by the backward difference. The dotted line denotes the initial condition, while the solid curve denotes the numerical solution for the stage forwarded three time steps. It is very different from the exact solution for the corresponding stage (dashed curve). The CFL number is taken to be 0.8.

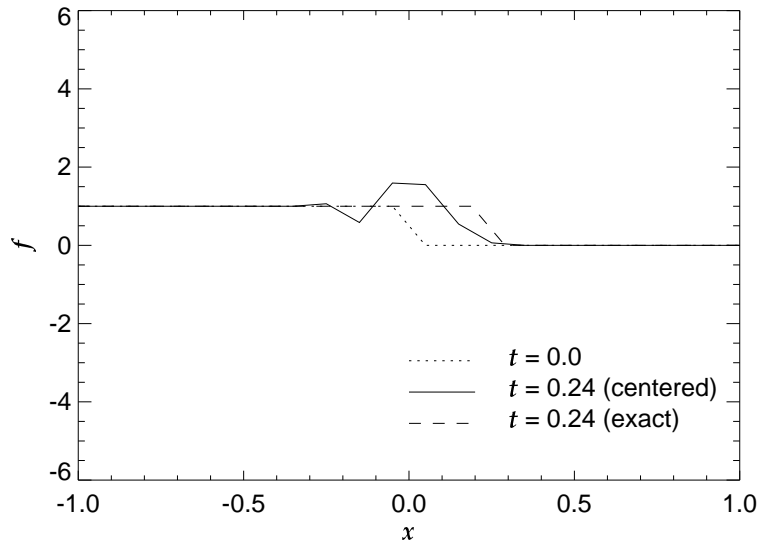


Figure 1.6: Numerical solution obtained by the central difference. The dotted line denotes the initial condition, while the solid curve denotes the numerical solution for the stage forwarded three time steps. It is very different from the exact solution for the corresponding stage (dashed curve). The CFL number is taken to be 0.8.

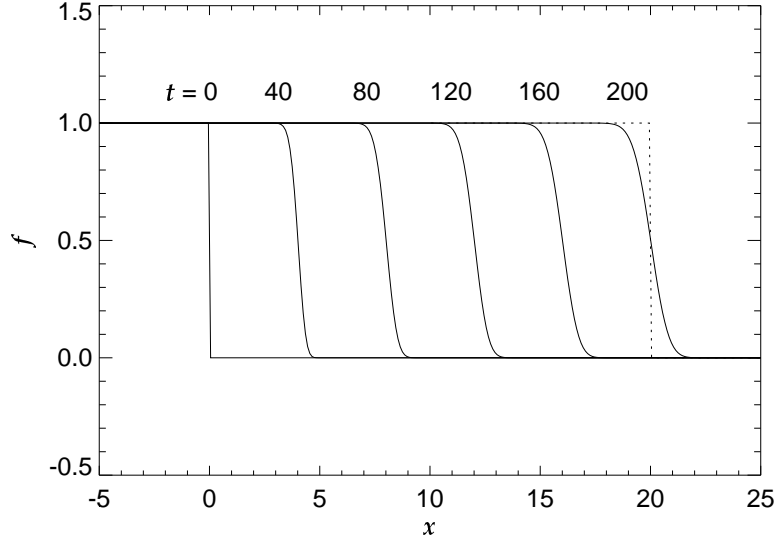


Figure 1.7: Numerical solution obtained by the backward difference. The dotted line denotes the initial condition, while the solid curve denotes the numerical solution for a later stage. It is very different from the exact solution for the corresponding stage (dashed curve). The CFL number is taken to be 0.8.

The numerical solution also depends on the CFL number. Figure 1.8 compares the solutions obtained with different values of  $\Delta t$ . The dashed curve denotes the solution obtained with  $\Delta t = 0.04$ . The wave front is rounder and the transition region is wider. This is due to two factors: the wave amplitude decreases more in each time step (smaller  $|g^{(B)}|$ ) and more time steps are required to reach a given  $t$ . The solution obtained with a smaller  $\Delta t$  is no better, and the time step should be smaller than  $\Delta t \leq 0.1$  in this problem. The solution obtained with  $\Delta t = 0.101$  (dash-dotted curve) shows an unnaturally large oscillation. The von Neumann stability analysis indicates that the wave amplitude grows when the CFL number exceeds unity. When  $\Delta t = 0.101$ , the CFL number is 1.01 and the value of  $|g^{(B)}|$  is only slightly larger than unity. However, the amplification is repeated 99 times at  $t = 10$ . Thus, the solution is not acceptable.

We can now be certain that the backward difference gives a good numerical solution as long as the CFL number is smaller than unity. Next, we consider why the backward difference succeeds while the forward time step fails.

The general solution of the wave equation [Equation (1.5)] is expressed as

$$f(x, t) = F(x - ct), \quad (1.24)$$

where  $F$  denotes an arbitrary function. Thus, we can derive the relation,

$$f(x, t + \Delta t) = f(x - c\Delta t, t). \quad (1.25)$$

On the other hand, Equation (1.8) is equivalent to

$$f(x, t + \Delta t) = \left(1 - \frac{c\Delta t}{\Delta x}\right) f(x, t) + \frac{c\Delta t}{\Delta x} f(x - \Delta x, t). \quad (1.26)$$

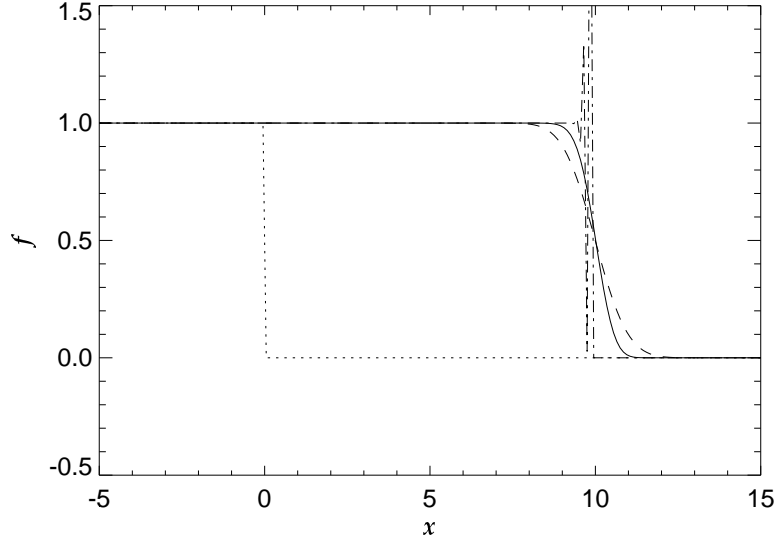


Figure 1.8: Comparison of numerical solutions obtained by the backward time step. The dotted curve denotes the initial condition at  $t = 0$ . The dashed curve denotes the solution at  $t = 10$  obtained with  $\Delta t = 0.08$ . The dashed curve and dash-dotted curves denote the solutions obtained with  $\Delta t = 0.04$  and  $0.101$ , respectively.

Equation (1.26) can be derived from Equation (1.25) by interpolating  $f$  from the values at two adjacent points. When the CFL number is larger than unity, Equation (1.26) is derived not by interpolation but by extrapolation. It is not a coincidence that the condition for stability coincides with the condition for Equation (1.26) to be derived by interpolation.

Similarly, Equation (1.7) is equivalent to

$$f(x, t + \Delta t) = \left(1 + \frac{c\Delta t}{\Delta x}\right) f(x, t) - \frac{c\Delta t}{\Delta x} f(x + \Delta x, t). \quad (1.27)$$

Equation (1.27) can be derived from Equation (1.25) by extrapolation. Interpolation is modest and gives a stable approximate solution, whereas extrapolation occasionally gives an unstable solution.

We have learned that inadequate difference schemes give solutions that show unnatural oscillations. The absence of such an oscillation of numerical origin is often measured by the time evolution of the total variation in numerical hydrodynamics. When the total variation diminishes with time for any initial condition, the solution is ensured to contain no oscillation of the numerical origin. This condition is expressed as

$$\sum_j |f_{j+1,n+1} - f_{j,n+1}| \leq \sum_j |f_{j+1,n} - f_{j,n}|, \quad (1.28)$$

and is referred to as the TVD condition, or total variation diminishing condition. Although it may appear that the total variation should diminish, Equation (1.28) means that

$$\frac{d}{dt} \left[ \int \left| \frac{\partial f}{\partial x} \right| dx \right] \leq 0, \quad (1.29)$$

because

$$\sum_j |f_{j+1,n} - f_{j,n}| \simeq \int \left| \frac{\partial f}{\partial x} \right| dx. \quad (1.30)$$

The left-hand side of Equation (1.29) vanishes in the exact solution.<sup>4</sup> The TVD condition does not require reduction of the total variation. In the same way that the von Neumann stability analysis requires the wave amplitude not to grow, the TVD condition requires the total variation not to grow. The error is smaller when the decrease in the total variation is smaller.

Here, we consider the case in which the propagation speed,  $c$ , is negative. When the propagation speed is reversed, the forward difference and backward difference exchange stabilities, as mentioned in §1.2. When  $c < 0$ , the backward difference is unstable and the forward difference is stable in the range  $-1 \leq c\Delta t/\Delta x \leq 0$ . The stability criterion is then generalized as follows.

The numerical solution is stable when the spatial derivative is replaced by the backward difference with respect to the wave propagation and the CFL number is in the range  $|c|\Delta t/\Delta x$ .

This criterion is valid irrespective of the sign of the propagation speed. The direction opposite to wave propagation corresponds to the upwind direction. Thus, the difference scheme given by Equation (1.8) is called the first-order upwind scheme, rather than the backward difference.

## 1.4 Nonlinear Wave Equation

In this section, as an example of a nonlinear wave equation, we study the Burgers equation,

$$\frac{\partial f}{\partial t} + f \frac{\partial f}{\partial x} = 0. \quad (1.31)$$

The Burgers equation is the same as Equation (1.5) except that the propagation speed  $c$  is replaced with  $f$ . Before discussing the numerical method used to solve the equation, we examine the mathematical properties.

First, we examine the case in which  $f$  is nearly constant and is expressed as follows:

$$f(x, t) = f_0 + f_1(x, t), \quad (1.32)$$

where

$$|f_1| \ll |f_0|. \quad (1.33)$$

Substituting Equation (1.32) into Equation (1.31) and neglecting a small second-order quantity, we obtain the following linearized equation:

$$\frac{\partial f_1}{\partial t} + f_0 \frac{\partial f_1}{\partial x} = 0. \quad (1.34)$$

If we substitute  $f_0$  for  $c$ , we obtain Equation (1.5). The wave amplitude coincides with the propagation speed in the Burgers equation.

Equation (1.31) has a special solution,

$$f = \frac{x - a}{t - b}, \quad (1.35)$$

---

<sup>4</sup>This statement is valid even when  $c$  is not constant.

where  $a$  and  $b$  are arbitrary constants. The value of  $f$  is always zero at  $x = a$  in this special solution, which is not surprising since the propagation speed is zero at this location. Note the change in the spatial derivative,  $\partial f / \partial x$ . In the range  $t > b$ , it is positive and decreases with time, and in the range  $t < b$ , it is negative and increases with time.

Equation (1.35) may not be valid at  $t = b$ , because the value  $f$  diverges. In order to consider the validity, we consider the initial condition

$$f = \begin{cases} -1 & (x \geq 1) \\ -x & (-1 < x < 1) \\ 1 & (x \leq -1), \end{cases} \quad (1.36)$$

at  $t = 0$ . The solution is expressed as

$$f = \begin{cases} -1 & (x \geq 1 - t) \\ \frac{x}{t - 1} & (-1 + t < x < 1 - t) , \\ 1 & (x \leq -1 + t) \end{cases} \quad (1.37)$$

for the period  $t < 1$ . At  $t = 1$ , the gradient,  $\partial f / \partial x$ , diverges at  $x = 0$  and the value of  $f$  becomes indefinite. Thus, the differential equation has no rigorous solution after this because the differential is already infinite. Still we can obtain a unique solution to the equation, as will be shown later. Equation (1.31) is equivalent to

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial x} \left( \frac{f^2}{2} \right) = 0. \quad (1.38)$$

Integrating Equation (1.38), we obtain

$$\frac{\partial}{\partial t} \int_a^b f dx + \left[ \frac{f^2}{2} \right]_a^b = 0. \quad (1.39)$$

When the function  $f$  satisfies Equation (1.39) for an arbitrary interval  $a \leq x \leq b$ , it is a weak solution. While any solution of Equation (1.31) satisfies Equation (1.39), the reverse is not true. A function  $f$  may satisfy Equation (1.39), even when the gradient is indefinite at some points and hence it is not a solution of Equation (1.31). In fact, the function,

$$f = \begin{cases} -1 & (x > 0) \\ 1 & (x < 0) \end{cases}, \quad (1.40)$$

satisfies Equation (1.39) in the period  $t \geq 1$  and continues to Equation (1.32) at  $t = 1$ . Thus, Equation (1.40) is a weak solution of Equation (1.31).

The value of  $f$  jumps from 1 to  $-1$  at  $x = 0$  in Equation (1.40). This jump is similar to the shock wave that appears in hydrodynamics. The density gradient and velocity gradient are very steep at the shock front, and sophisticated mathematics are required to handle an infinitely large quantity. Remember that the divergence apparently disappears in Equation (1.38). A similar feature is seen in Gauss's law of electromagnetism. The divergence seems to disappear in the integral form. In other words, the amount of divergence is evaluated correctly. Thus, the integral form is also used when we derive the numerical solution of Equation (1.31).<sup>5</sup>

---

<sup>5</sup>The integral form is also used in the finite element method.

Equation 1.38 is often referred to as the conservation form of the Burgers equation because it denotes conservation of  $f$ . Similarly, we can rewrite Equation (1.5) in conservation form as follows:

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial x}(cf) = 0. \quad (1.41)$$

Then, we have

$$\frac{f_{j,n+1} - f_{j,n}}{\Delta t} + \frac{\frac{f_{j,n}^2}{2} - \frac{f_{j-1,n}^2}{2}}{\Delta x} = 0, \quad (1.42)$$

which gives us an upwind scheme for  $f > 0$ , and

$$\frac{f_{j,n+1} - f_{j,n}}{\Delta t} + \frac{\frac{f_{j+1,n}^2}{2} - \frac{f_{j,n}^2}{2}}{\Delta x} = 0, \quad (1.43)$$

which gives us an upwind scheme for  $f < 0$ . Since the wave propagation velocity is constant in Equation 1.5, either the forward difference is upwind everywhere or the backward difference is upwind everywhere. However, the propagation speed depends on  $f$  and accordingly changes its sign with time and place in the Burgers equation. Thus, we need to switch the forward difference and backward difference so that the scheme is upwind. Although the **if** clause can manage the switch both in Fortran and in C, we want to avoid the **if** clause for the following two reasons. First, the **if** clause makes a computer program more complex. Second, the **if** clause slows the computation speed appreciably. To avoid this, we rewrite Equations (1.42) and (1.43) as follows:

$$\frac{f_{j,n+1} - f_{j,n}}{\Delta t} + \frac{F_{j+1/2,n}^* - F_{j-1/2,n}^*}{\Delta x} = 0, \quad (1.44)$$

where

$$\begin{aligned} F_{j+1/2,n}^* &= \frac{1}{2} \left( \frac{f_{j+1,n}^2}{2} + \frac{f_{j,n}^2}{2} \right) \\ &+ \frac{1}{2} \frac{|f_{j+1/2,n} + f_{j-1/2,n}|}{2} (f_{j+1,n} - f_{j,n}). \end{aligned} \quad (1.45)$$

The value of  $F_{j+1/2,n}^*$  coincides with either  $f_{j+1,n}^2/2$  or  $f_{j,n}^2/2$  depending on the sign of  $f_{j+1,n} + f_{j,n}$ . When  $f_{j+1,n} + f_{j,n} > 0$ , Equations (1.44) and (1.45) are equivalent to Equation (1.42). When  $f_{j+1,n} + f_{j,n} < 0$ , they are equivalent to Equation (1.43). The variable,  $F_{j+1/2,n}^*$ , is called the numerical flux and denotes the flux at the midpoint,  $x = (x_{j+1} + x_j)/2$ .

Comparison of Equations (1.44) and (1.39) indicates that the variable,  $f_{j,n}$ , should be defined as

$$f_{j,n} \equiv \frac{1}{\Delta x} \int_{x_j - \Delta x/2}^{x_j + \Delta x/2} f(x, t_n) dx. \quad (1.46)$$

This definition states that the variable  $f_{j,n}$  denotes the average of  $f(x, t_n)$  in the cell  $x_j - \Delta x/2 \leq x \leq x_j + \Delta x/2$ . In other words, the variable  $f_{j,n}$  denotes the cell average, and the spatial distance,  $\Delta x$ , denotes the cell width.

Figure 1.9 shows a numerical solution of the Burgers equation obtained by the upwind scheme. This solution has no unnatural oscillation and is an acceptable solution. The error is appreciable near  $x = t - 1$  and  $1 - t$ . The gradient,  $\partial f / \partial x$ , changes continuously, although it should change discontinuously at  $x = t - 1$  and  $1 - t$ . This numerical error is similar to that

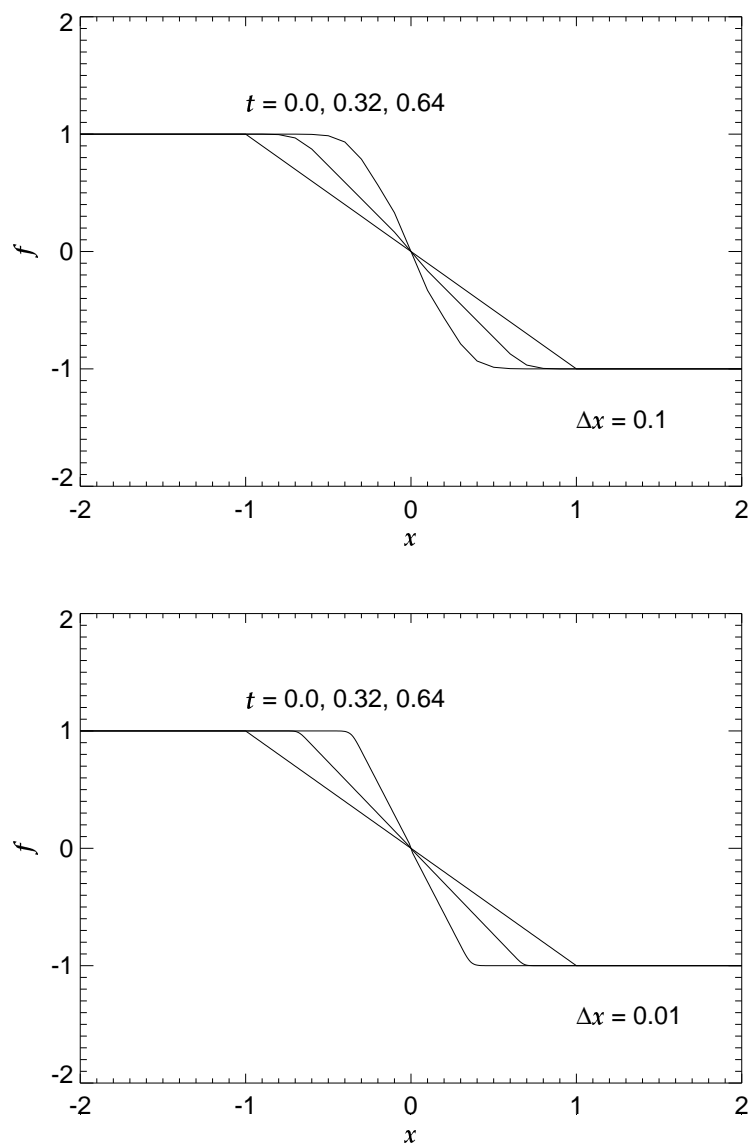


Figure 1.9: A numerical solution of the Burgers equation obtained by the upwind scheme. The initial value is given by Equation (1.40) at  $t = 0$ . The solid curves denote the solutions at  $t = 0, 0.32$ , and  $0.64$ . The grid spacing and time step are taken to be  $\Delta x = 0.1$  and  $\Delta t = 0.08$  in the upper panel and are taken to be  $\Delta x = 0.01$  and  $\Delta t = 0.008$  in the lower panel.

seen in the solution of the linear wave equation. The grid spacing,  $\Delta x$ , is 10 times smaller in the lower panel than in the upper panel. Thus, the resolution is higher and the solution is more accurate.

The next example shows the propagation of a shock wave in the Burgers equation. The initial condition is expressed as

$$f = \begin{cases} 1 & (x \leq -1) \\ -x & (-1 < x < 0) \\ 0 & (x \geq 0) \end{cases} . \quad (1.47)$$

The solution is expressed as

$$f = \begin{cases} 1 & (x \leq -1+t) \\ -\frac{x}{1-t} & (-1+t < x < 0) \\ 0 & (x \geq 0) \end{cases} , \quad (1.48)$$

for  $0 < t < 1$ , and

$$f = \begin{cases} 1 & (x \leq t/2 - 1/2) \\ 0 & (x > t/2 - 1/2) \end{cases} , \quad (1.49)$$

for  $t \geq 1$ . The latter solution is a weak solution, i.e., it satisfies Equation (1.39) rather than Equation (1.31). Figure 1.10 shows the numerical solution of this problem solved by the upwind scheme. The grid spacing is taken to be rather large ( $\Delta x = 0.1$ ) so that the error is easy to read. Thus, the shock front is not very sharp but keeps its form sharp. Remember that the wave front becomes less sharp in the numerical solution of the linear wave equation, as shown in Figure 1.7. On the other hand, the gradient becomes steeper in the Burgers equation when it is negative. The shock front maintains its sharpness due to the physical steepening. The asterisks denote positions at which the value,  $f$ , is evaluated in Figure 1.10. There is only one point in the transition region of  $0.1 < f < 0.9$ . Thus, the numerical solution is regarded as capturing the shock front with one point. Since the solution has no unnatural oscillation near the shock front, we call this scheme a shock capturing scheme.

Next, we examine the case in which the gradient  $\partial f / \partial x$  is positive. Figure 1.11 shows the solution for which the initial condition is given by

$$f = \begin{cases} -1 & (x \leq -1) \\ x & (-1 < x < 1) \\ 1 & (x \geq 1) \end{cases} . \quad (1.50)$$

Although the grid spacing is  $\Delta x = 0.1$  both in the upper and lower panels, the value of  $f$  is evaluated at  $x = j \Delta x$  in the upper panel and at  $x = (j + 1/2) \Delta x$  in the lower panel. Note the appreciable difference between the solutions around  $x = 0$ . The exact solution is

$$f = \begin{cases} -1 & (x \leq -1-t) \\ \frac{x}{t+1} & (-1-t < x < 1+t) \\ 1 & (x \geq 1+t) \end{cases} , \quad (1.51)$$

for  $t > 0$ . The gradient  $\partial f / \partial x$  decreases with time in the exact solution and in the upper panel but maintains its initial value in the lower panel. This odd feature is called expansion shock and

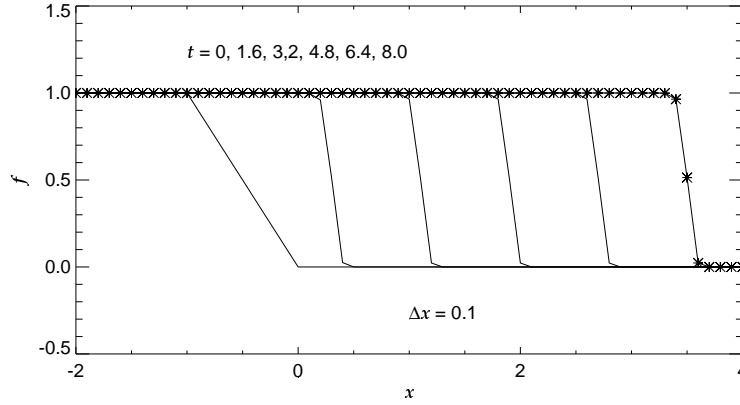


Figure 1.10: A numerical solution of the Burgers equation obtained by the upwind scheme. The grid spacing is taken to be  $\Delta x = 0.1$ .

occurs due to the erroneous evaluation of the numerical flux at  $x = 0$  in the lower panel. The flux is  $F = f^2/2 = 0$  at  $x = 0$  in the exact solution, while the numerical flux is not ( $F^* \neq 0$ ) in the lower panel.

Expansion shock is avoided in the upper panel because the numerical flux is not evaluated at  $x = 0$  but at  $x = \pm\Delta x/2$ . The expansion shock can be avoided by another means. If we replace Equation (1.45) with

$$F_{j+1/2,n}^* = \frac{1}{2} \left( \frac{f_{j+1,n}^2}{2} + \frac{f_{j,n}^2}{2} \right) - \frac{|\lambda|}{2} (f_{j+1,n} - f_{j,n}), \quad (1.52)$$

$$|\lambda| = \begin{cases} \frac{|f_{j,n} + f_{j+1,n}|}{2} & \left( \frac{|f_{j+1,n} + f_{j,n}|}{2} \geq \varepsilon \right) \\ \varepsilon & (\text{otherwise}) \end{cases}, \quad (1.53)$$

$$\varepsilon = \max \left( 0, \frac{f_{j+1,n} - f_{j,n}}{2} \right), \quad (1.54)$$

then the solution has no expansion shock. This modification changes the numerical flux only when  $f_{j+1,n} > f_{j,n}$ . Thus, the solutions shown in Figures 1.9 and 1.10 are not affected by this modification. We often refer to this procedure as the addition of the entropy condition.

## 1.5 Upwind Scheme for the Hydrodynamical Equations

This section deals with the hydrodynamical equations. First, we consider the simplest problem, a one-dimensional flow, and show that the hydrodynamical equations are a system of wave equations.

The mass conservation is described as

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} (\rho v) = 0, \quad (1.55)$$

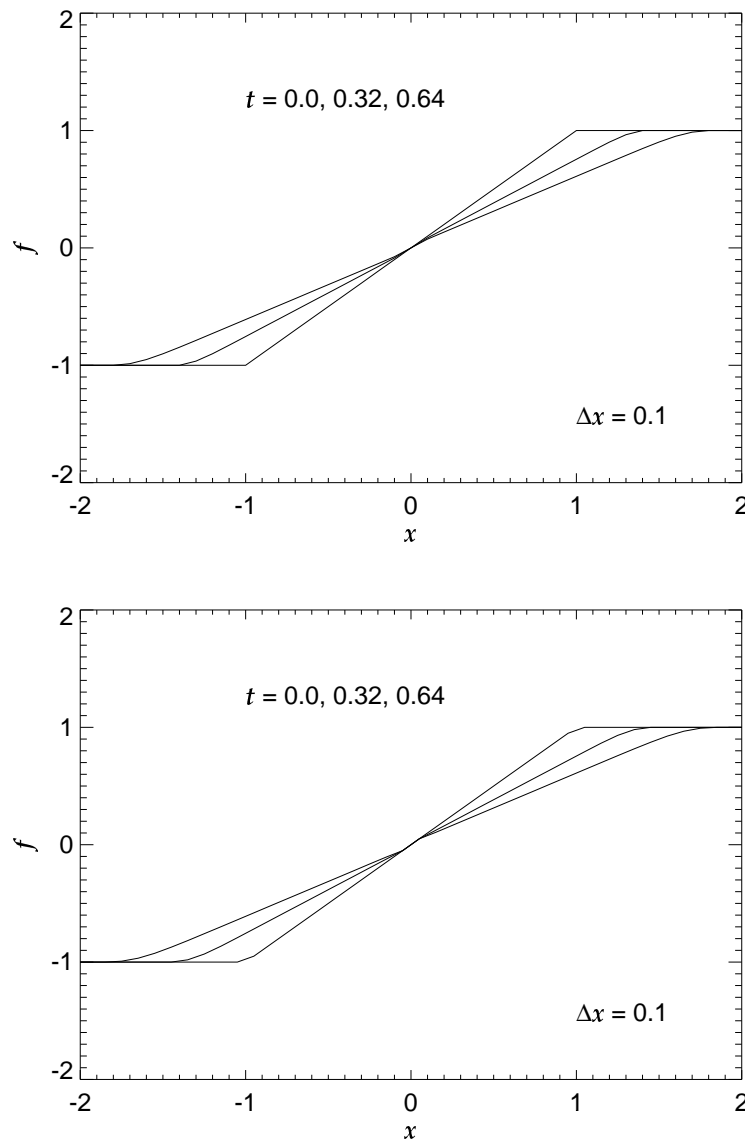


Figure 1.11: Numerical solutions of the Burgers equation. The upwind scheme is used both in the upper and lower panels. The initial condition and the grid spacing are the same. Nevertheless, the solutions are appreciably different. The value of  $f$  is evaluated at  $x_j = j\Delta x$  in the upper panel and is evaluated at  $x_j = (j + 1/2)\Delta x$  in the lower panel.

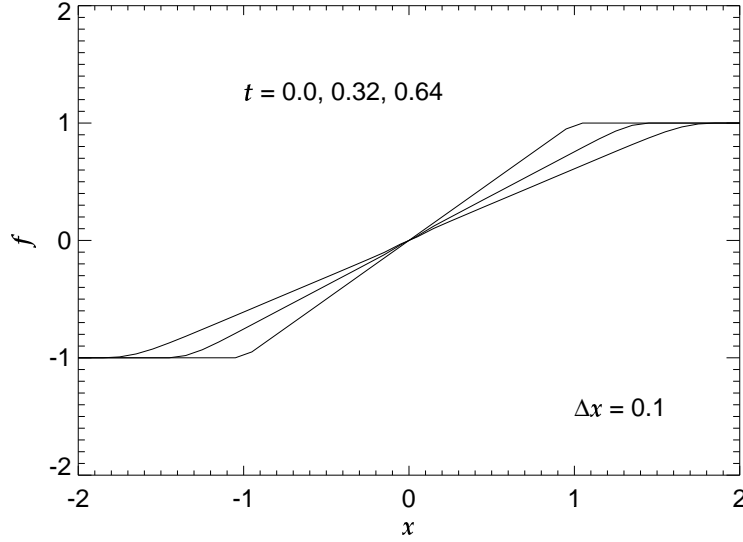


Figure 1.12: Numerical solution obtained with the addition of the entropy condition. The grid spacing and initial condition are the same as those in the lower panel of Figure 1.11. The expansion shock is removed by the addition of the entropy condition.

where  $\rho$  and  $v$  denote the density and velocity in the  $x$ -direction, respectively. The equation of motion is expressed as

$$\frac{Dv}{Dt} + \frac{\partial P}{\partial x} = 0, \quad (1.56)$$

where  $P$  denotes the pressure. Here, the symbol  $D/Dt$  denotes the time Lagrangian time derivative and is given by  $\partial/\partial t + v \partial/\partial x$ . Thus, Equation (1.56) is rewritten as

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + \frac{\partial P}{\partial x} = 0. \quad (1.57)$$

Although the pressure ( $P$ ) is generally a function of density and temperature ( $T$ ), for simplicity, we herein assume that the pressure depends solely on the density.<sup>6</sup> For later convenience, we define the speed of sound as<sup>7</sup>

$$a \equiv \sqrt{\frac{dP}{d\rho}}. \quad (1.58)$$

Then, Equation (1.57) is rewritten as

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + a^2 \frac{\partial \rho}{\partial x} = 0. \quad (1.59)$$

<sup>6</sup>When the pressure is a function of the density, the fluid is defined to be barotropic. This is a good approximation when heating and cooling are negligible and, accordingly, the entropy is constant with respect to time and space. This is also a good approximation when the temperature is maintained constant.

<sup>7</sup>When a gas is thermodynamically stable, the quantity under the square root is positive,  $dP/d\rho$ . If the pressure decreases with increasing density, the gas is further compressed and the density increases further. Thus, the gas is thermodynamically unstable. When  $dP/d\rho < 0$ , the gas separates into two phases, i.e., high-density liquid and low-density gas.

The sum of Equation (1.59) and Equation (1.55) multiplied by  $a/\rho$  yields

$$\frac{a}{\rho} \frac{\partial \rho}{\partial t} + \frac{\partial v}{\partial t} + (v + a) \left( \frac{a}{\rho} \frac{\partial \rho}{\partial x} + \frac{\partial v}{\partial x} \right) = 0. \quad (1.60)$$

This equation is similar to the linear wave equation and the Burgers equation, and is equivalent to

$$\frac{\partial J_+}{\partial t} + (v + a) \frac{\partial J_+}{\partial x} = 0, \quad (1.61)$$

where

$$J_+ \equiv \int \frac{a}{\rho} d\rho + v. \quad (1.62)$$

Equation 1.61 is quite similar to the Burgers equation. Similarly, we obtain

$$\frac{\partial J_-}{\partial t} + (v - a) \frac{\partial J_-}{\partial x} = 0, \quad (1.63)$$

where

$$J_- \equiv \int \frac{a}{\rho} d\rho - v \quad (1.64)$$

from Equations (1.55) and (1.59). The hydrodynamical equations are equivalent to Equations (1.61) and (1.63) when the flow is one-dimensional and barotropic.<sup>8</sup>

Next, we examine Equations (1.61) and (1.63). Equation (1.61) demonstrates that a hydrodynamical wave propagates with  $v + a$ . Since the phase velocity is the sum of the flow velocity ( $v$ ) and speed of sound ( $a$ ), the wave propagates faster than the flow. Thus, the symbol  $J_+$  denotes the amplitude of the wave that propagates toward the right, i.e., in the direction of increasing  $x$ . On the other hand, the symbol  $J_-$  denotes the amplitude of the wave that propagates toward the left. In mathematical terminology, the characteristic speed is identical to the phase velocity. The wave amplitudes,  $J_+$  and  $J_-$ , are referred to as Riemann invariants in mathematics.

The velocity and density are expressed as a function of the two Riemann invariants. The velocity is proportional to the difference of the Riemann invariants,  $v = (J_+ - J_-)/2$ . Thus, the velocity change contains two components that propagate at speeds of  $v + a$  and  $v - a$ . Similarly, the density distribution has two components.

In the previous sections, we learned that we need to use either the backward difference or the forward difference, depending on the sign of the propagation speed, when we solve a wave equation. When the speed of sound ( $a$ ) is larger than the absolute value of the flow velocity ( $|v|$ ), the two characteristic speeds have different signs. Thus, neither the forward difference nor the backward difference can treat one of the two components properly. This means that neither the forward difference nor the backward difference succeeds in solving Equations (1.55) and (1.56). On the other hand, Equations (1.61) and (1.63) describe only one wave component and so can be solved numerically either by the forward difference or by the backward difference. Thus, we have rewritten Equations (1.55) and (1.56) to obtain Equations (1.61) and (1.63).

Roe (1981) proposed a numerical scheme in which hydrodynamical equations are decomposed into components, and in which each component equation is solved with the upwind difference. This scheme is widely used because it is relatively easy to implement and gives a good numerical

---

<sup>8</sup>This argument is presented by Sakashita and Ikeuchi (1996). Their textbook entitled “Uchuu Ryuutai Rikigaku” discusses the mathematical properties of the hydrodynamical equations in greater detail.

solution. This scheme is now applied to the magnetohydrodynamical equations. Although a number of good schemes exist, we herein consider only this scheme.

Roe's original scheme is for an ideal gas. Here, the energy conservation is expressed as

$$\frac{\partial}{\partial t}(\rho E) + v \frac{\partial}{\partial x}(\rho H v) = 0, \quad (1.65)$$

where

$$E = \frac{v^2}{2} + \varepsilon, \quad (1.66)$$

$$H = E + \frac{P}{\rho}. \quad (1.67)$$

where  $E$  and  $\varepsilon$  denote the specific energy of the fluid and the specific energy of the internal energy, respectively. Thus,  $H$  denotes the sum of the specific enthalpy and specific kinetic energy. The specific energy of the fluid and the sum of the specific enthalpy and specific kinetic energy are evaluated to be

$$E = \frac{v^2}{2} + \frac{1}{\gamma - 1} \frac{P}{\rho}, \quad (1.68)$$

and

$$H = \frac{v^2}{2} + \frac{\gamma}{\gamma - 1} \frac{P}{\rho}, \quad (1.69)$$

respectively, for an ideal gas, where  $\gamma$  denotes the ratio of specific heat at constant pressure and constant volume, respectively. The dependent variables are expressed as a function of  $\rho$ ,  $v$ , and  $E$ . The pressure is evaluated to be

$$P = (\gamma - 1) \rho \left( E - \frac{v^2}{2} \right). \quad (1.70)$$

Similarly, we obtain

$$H = \gamma \left( E - \frac{v^2}{2} \right) + \frac{v^2}{2}. \quad (1.71)$$

We have Equations (1.55), (1.57), and (1.65) and three unknowns,  $\rho$ ,  $v$ , and  $E$ . Therefore, we can compute the change in the density and the change in the velocity by solving these equations simultaneously with the appropriate initial and boundary conditions.

For later convenience, we rewrite Equation (1.57) as follows:

$$\frac{\partial}{\partial t}(\rho v) + \frac{\partial}{\partial x}(\rho v^2 + P) = 0, \quad (1.72)$$

using Equation (1.55). Equation (1.72), like Equations (1.55) and (1.65), is written in conservation form. These equations can be expressed in vector form as follows:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0, \quad (1.73)$$

$$\mathbf{U} = \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} = \begin{pmatrix} \rho \\ \rho v \\ \rho E \end{pmatrix}, \quad (1.74)$$

$$\mathbf{F} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} = \begin{pmatrix} \rho v \\ \rho v^2 + P \\ \rho H v \end{pmatrix}. \quad (1.75)$$

In the following, we refer to  $\mathbf{U}$  and  $\mathbf{F}$  as the state vector and the flux vector, respectively. This expression is useful in deriving the phase velocities of hydrodynamical waves. The flux vector is expressed as a function of the state vector, i.e.,

$$\mathbf{F} = \begin{pmatrix} U_2 \\ \frac{3-\gamma}{2} \frac{(U_2)^2}{U_1} + (\gamma-1) U_3 \\ \frac{U_2}{U_1} \left[ \gamma U_3 - (\gamma-1) \frac{(U_2)^2}{2U_1} \right] \end{pmatrix}. \quad (1.76)$$

We can then rewrite Equation (1.73) as

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A} \frac{\partial \mathbf{U}}{\partial x} = 0, \quad (1.77)$$

where

$$\begin{aligned} \mathbf{A} &\equiv \frac{\partial \mathbf{F}}{\partial \mathbf{U}} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma-3) \left( \frac{U_2}{U_1} \right)^2 & -(\gamma-3) \frac{U_2}{U_1} & \gamma-1 \\ -\frac{\gamma U_2 U_3}{(U_1)^2} + (\gamma-1) \left( \frac{U_2}{U_1} \right)^3 & \frac{\gamma U_3}{U_1} - \frac{3}{2}(\gamma-1) \left( \frac{U_2}{U_1} \right)^2 & \frac{\gamma U_2}{U_1} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ -\frac{3-\gamma}{2} v^2 & (3-\gamma)v & \gamma-1 \\ \left( \frac{\gamma-1}{2} v^2 - H \right) v & H - (\gamma-1)u^2 & \gamma v \end{bmatrix}. \end{aligned} \quad (1.79)$$

Comparison of Equations (1.77) and (1.5) reveals that the matrix  $\mathbf{A}$  denotes the velocity of the wave.

From the previous sections, we have

$$\frac{\mathbf{U}_{j,n+1} - \mathbf{U}_{j,n}}{\Delta t} + \frac{\mathbf{F}_{j,n+1/2}^* - \mathbf{F}_{j,n-1/2}^*}{\Delta x} = 0, \quad (1.80)$$

$$\mathbf{F}_{j+1/2,n}^* = \frac{1}{2} [\mathbf{F}_{j+1,n} + \mathbf{F}_{j,n} - |\mathbf{A}| (\mathbf{U}_{j+1,n} - \mathbf{U}_{j,n})]. \quad (1.81)$$

One remaining problem is the evaluation of  $|\mathbf{A}|$ , which is evaluated as follows. The velocity matrix has eigenvalues  $\lambda_i$  and right eigenvectors  $\mathbf{r}_i$ .

$$\mathbf{A} \mathbf{r}_i = \lambda_i \mathbf{r}_i, \quad (1.82)$$

$$\lambda_1 = v - a, \quad (1.83)$$

$$\lambda_2 = v, \quad (1.84)$$

$$\lambda_3 = v + a, \quad (1.85)$$

$$\mathbf{r}_1 = \begin{pmatrix} 1 \\ v - a \\ H - va \end{pmatrix}, \quad (1.86)$$

$$\mathbf{r}_2 = \begin{pmatrix} 1 \\ v \\ \frac{v^2}{2} \end{pmatrix}, \quad (1.87)$$

$$\mathbf{r}_3 = \begin{pmatrix} 1 \\ v + a \\ H + va \end{pmatrix}, \quad (1.88)$$

where

$$a = \sqrt{\frac{\gamma P}{\rho}}, \quad (1.89)$$

$$= (\gamma - 1) \left( H - \frac{v^2}{2} \right). \quad (1.90)$$

Using the matrix,

$$\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \quad (1.91)$$

$$= \begin{pmatrix} 1 & 1 & 1 \\ v - a & v & v + a \\ H - va & \frac{v^2}{2} & H + va \end{pmatrix}, \quad (1.92)$$

we can diagonalize the velocity matrix as follows

$$\mathbf{A} = \mathbf{R} \begin{pmatrix} v - a & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & v + a \end{pmatrix} \mathbf{R}^{-1}. \quad (1.93)$$

The matrix,  $\mathbf{R}^{-1}$ , is expressed as

$$\mathbf{R}^{-1} = \frac{\gamma - 1}{2a^2} \begin{pmatrix} \frac{v^2}{2} + \frac{va}{\gamma - 1} & -v + \frac{a}{\gamma - 1} & 1 \\ -v^2 + \frac{2a^2}{\gamma - 1} & 2v & -2 \\ -\frac{v^2}{2} - \frac{va}{\gamma - 1} & v - \frac{a}{\gamma - 1} & 1 \end{pmatrix} \quad (1.94)$$

$$= \mathbf{L} \quad (1.95)$$

$$= \begin{pmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \end{pmatrix}. \quad (1.96)$$

The vectors  $\ell_1$ ,  $\ell_2$ , and  $\ell_3$  are left eigenvectors because

$$\ell_i \mathbf{A} = \lambda_i \ell_i. \quad (1.97)$$

The eigenvalues, the diagonal elements of the central matrix in Equation (1.93), indicate the propagation speed of the hydrodynamical waves. Thus, we can evaluate  $|\mathbf{A}|$  to be

$$|\mathbf{A}| = \mathbf{R} \begin{pmatrix} |v - a| & 0 & 0 \\ 0 & |v| & 0 \\ 0 & 0 & |v + a| \end{pmatrix} \mathbf{L}. \quad (1.98)$$

The eigenvalues  $v + a$  and  $v - a$  denote the phase velocities of sound waves propagating toward the right and toward the left, respectively. These waves also appear in a barotropic fluid. The other wave, the phase velocity of which coincides with the flow velocity, disappears in a barotropic fluid.

Equation (1.98) should be evaluated from  $\mathbf{U}_{j+1,n}$  and  $\mathbf{U}_{j,n}$ , i.e., from an average state vector, as in the case of the Burgers equation. Roe proposed evaluation of  $\mathbf{R}$ ,  $\mathbf{L}$ ,  $\lambda_i$  ( $i = 1, 2, 3$ ) by the following averages:

$$\bar{v} = \frac{\sqrt{\rho_j}v_j + \sqrt{\rho_{j+1}}v_{j+1}}{\sqrt{\rho_j} + \sqrt{\rho_{j+1}}}, \quad (1.99)$$

$$\bar{H} = \frac{\sqrt{\rho_j}H_j + \sqrt{\rho_{j+1}}H_{j+1}}{\sqrt{\rho_j} + \sqrt{\rho_{j+1}}}, \quad (1.100)$$

$$\bar{a}^2 = (\gamma - 1) \left( \bar{H} - \frac{\bar{v}^2}{2} \right). \quad (1.101)$$

Although computing  $|\mathbf{A}|$  may appear difficult, the right-hand side of Equation (1.81) can be rewritten as

$$\mathbf{F}_{j+1/2,n}^* = \frac{1}{2} \left[ \mathbf{F}_{j+1,n} + \mathbf{F}_{j,n} - \sum_{k=1}^3 |\lambda_k| w_k \mathbf{r}_k \right], \quad (1.102)$$

where

$$w_1 = \frac{1}{2\bar{a}} \left[ \frac{P_{j+1} - P_j}{\bar{a}} - \bar{\rho} (v_{j+1} - v_j) \right], \quad (1.103)$$

$$w_2 = \rho_{j+1} - \rho_j - \frac{P_{j+1} - P_j}{\bar{a}^2}, \quad (1.104)$$

$$w_3 = \frac{1}{2\bar{a}} \left[ \frac{P_{j+1} - P_j}{\bar{a}} + \bar{\rho} (v_{j+1} - v_j) \right], \quad (1.105)$$

because

$$\mathbf{U}_{j+1} - \mathbf{U}_j = \sum_{k=1}^3 w_k \mathbf{r}_k, \quad (1.106)$$

for any  $\mathbf{U}_j$  and  $\mathbf{U}_{j+1}$ .

Here, we consider a wave having a phase velocity of  $v$ . As shown in Equation (1.104), the amplitude is  $w_2$  and is evaluated to be

$$w_2 \simeq \left( \frac{\partial P}{\partial s} \right)_\rho (s_{j+1} - s_j), \quad (1.107)$$

because

$$\delta P = \left( \frac{\partial P}{\partial \rho} \right)_s + \left( \frac{\partial P}{\partial s} \right)_\rho \delta s \quad (1.108)$$

$$= a^2 \delta \rho + \left( \frac{\partial P}{\partial s} \right)_\rho ds, \quad (1.109)$$

where  $s$  denotes the specific entropy. This means that the wave amplitude is proportional to the entropy difference. Thus, the wave is referred to as the entropy wave. When only the entropy wave exists, the pressure and velocity are uniform and the temperature changes. A cold front, in which the temperature drops appreciably, is an example of an entropy wave.<sup>9</sup>

As in the case of the Burgers equation, the characteristic speeds should be modified as

$$|\lambda_1| = \max(|\bar{v} - \bar{a}|, \varepsilon_1), \quad (1.110)$$

$$\varepsilon_1 = \max(\lambda_{j+1,1} - \lambda_{j,1}, 0), \quad (1.111)$$

$$\lambda_{j,1} = v_j - \sqrt{\frac{\gamma P_j}{\rho_j}}, \quad (1.112)$$

$$|\lambda_3| = \max(|\bar{v} + \bar{a}|, \varepsilon_3), \quad (1.113)$$

$$\varepsilon_3 = \max(\lambda_{j+1,3} - \lambda_{j,3}, 0), \quad (1.114)$$

$$\lambda_{j,3} = v_j + \sqrt{\frac{\gamma P_j}{\rho_j}}, \quad (1.115)$$

in order to avoid the expansion shock. Entropy correction is not required for the entropy wave ( $\lambda_2$ ). Entropy waves never evolve into an expansion shock.

Figures 1.13 and 1.14 show numerical examples in which a shock tube problem is solved by the Roe scheme. Figures 1.13 and 1.14 confirm that the Roe scheme can solve the hydrodynamical equations without difficulty, even when the initial density and pressure distributions are discontinuous. The figures show no unnatural oscillation around the shock wave.

Next, we evaluate the numerical error by comparing the solutions for different values of  $\Delta x$ . Figure 1.15 compares solutions obtained with different values of  $\Delta x$ . The initial condition is the same as that of the solution shown in Figures 1.13 and 1.14. The dash-dotted curve denotes the density distribution obtained with  $\Delta x = 0.1$ , while the solid and dashed curves denote those obtained with  $\Delta x = 0.001$  and  $0.01$ , respectively. All of the curves denote the density distribution at  $t = 0.8$ . As  $\Delta x$  decreases, the density decreases gradually in the range  $-1.04 \leq x \leq 0.03$  and approaches the distribution having discontinuities at  $x = 0.78$  and  $1.43$ . These features indicate that the numerical solution obtained by the Roe scheme converges to the exact solution in the limit of infinitesimal  $\Delta x$ . Convergence is an important issue, and a stable scheme is of limited value if the error remains too high.

The speed of convergence depends on the position. The density and temperature change in the rarefaction wave, at the contact discontinuity ( $x = 0.78$ ), and at the shock front ( $x = 1.43$ ). As demonstrated in Figure 1.15, the convergence is fast at the shock front but is very slow at the contact discontinuity. The dashed curve ( $\Delta x = 0.01$ ) is a fairly good solution around the shock front, while the contact discontinuity is not sharp enough, even in the solution with  $\Delta x = 0.001$ . The speed of convergence depends little on the initial condition. The contact discontinuity is always less sharp than the shock front. The change in temperature becomes vague, although it should change sharply at the contact discontinuity.

Figure 1.16 shows the speed of convergence quantitatively. The abscissa denotes the grid spacing  $\Delta x$ , while the ordinate denotes the numerical errors in the density (solid curve) and velocity (dashed curve) at  $(x, t) = (-0.75, 0.8)$ . The time step is taken to be  $\Delta t = 0.5 \Delta x$  so that the CFL number is the same. The errors are proportional to the grid spacing, i.e., they are of the first order. This is because the truncation error is proportional to  $\Delta x^2$  in the first-order

---

<sup>9</sup>The entropy is given by  $s = \log P - \gamma \log \rho$  for an ideal fluid with specific heat ratio  $\gamma$ . The pressure is then expressed as  $P = e^s \rho^\gamma$ .

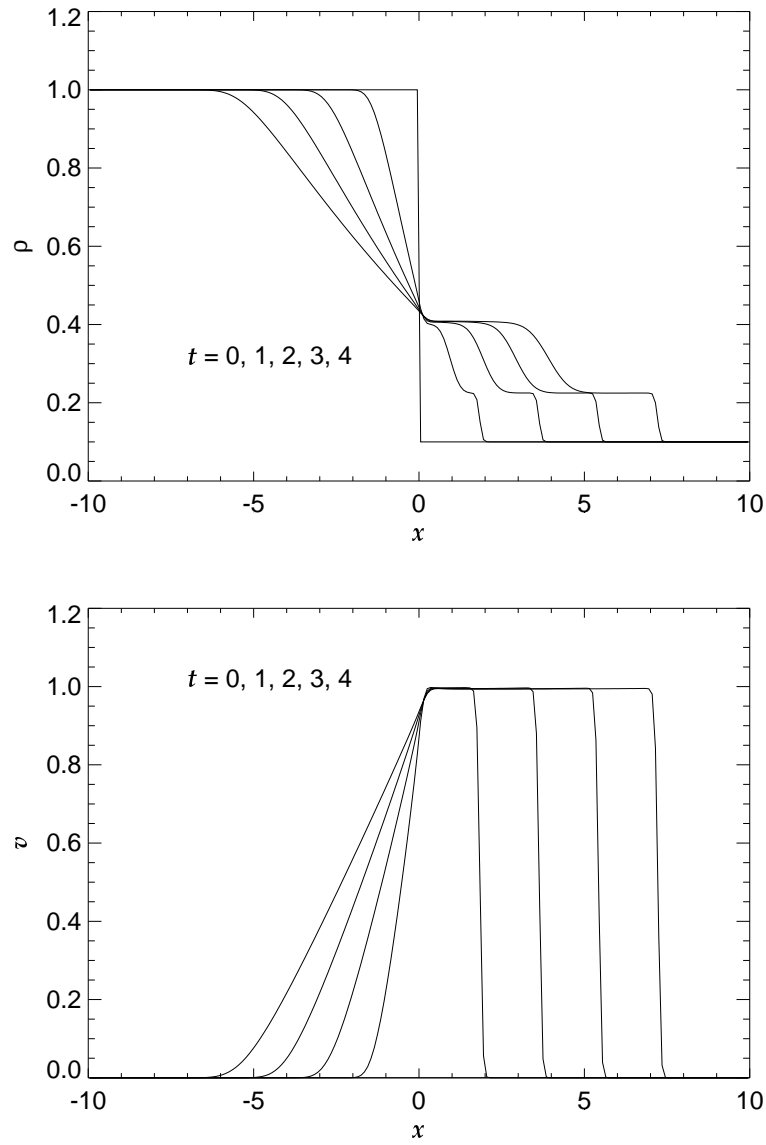


Figure 1.13: A numerical solution obtained with the Roe scheme. The initial density and pressure are  $\rho = 0.1$  and  $P = 0.05$  in the region  $x > 0$  and are  $\rho = 1.0$  and  $P = 1.0$  in the region  $x < 0$ . The initial velocity is  $v = 0$  over the entire region. The specific heat ratio is set to be  $\gamma = 5/3$ . The grid spacing is  $\Delta x = 0.1$  and the time step is  $\Delta t = 0.04$ . The upper panel shows the density, and the lower panel shows the velocity.

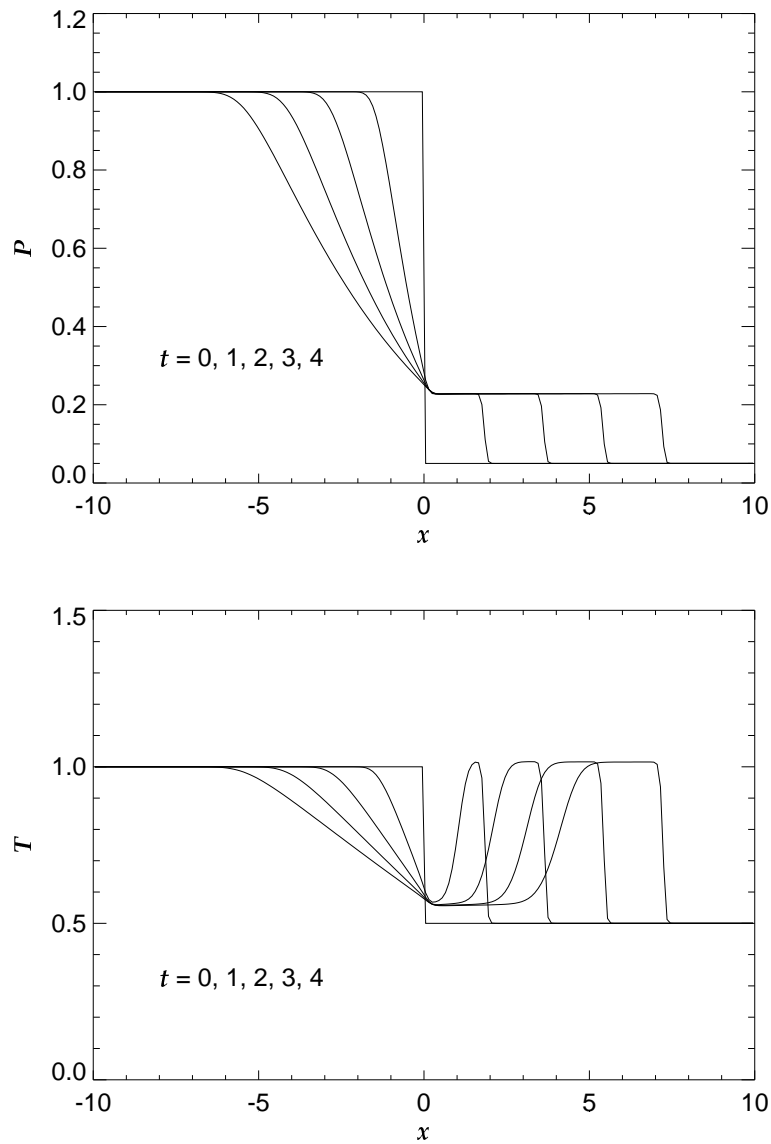


Figure 1.14: A numerical solution obtained with the Roe scheme. The initial density and pressure are  $\rho = 0.1$  and  $P = 0.05$  in the region  $x > 0$  and are  $\rho = 1.0$  and  $P = 1.0$  in the region  $x < 0$ . The initial velocity is  $v = 0$  over the entire region. The grid spacing is  $\Delta x = 0.1$  and the time step is  $\Delta t = 0.04$ . The upper panel shows the pressure, and the lower panel shows the temperature.

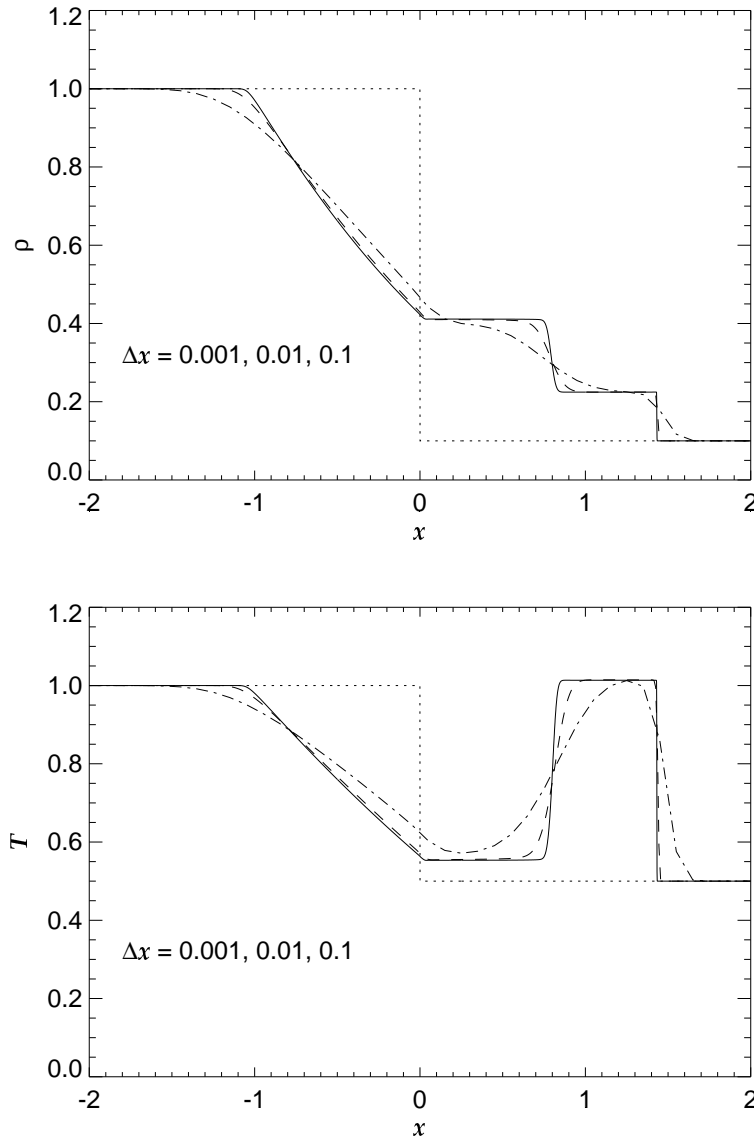


Figure 1.15: Comparison of the numerical solutions with different values of  $\Delta x$ . The initial condition is the same as that of the solution shown in Figures 1.13 and 1.14. The upper and lower panels show the density and temperature, respectively. The solid curve denotes the solution obtained with  $\Delta x = 0.001$ , and the dashed curve shows the solution obtained with  $\Delta x = 0.01$ . The dash-dotted curve denotes the solution obtained with  $\Delta x = 0.1$ . The time step is taken to be  $\Delta t = 0.4 \Delta x$  in all of the solutions.

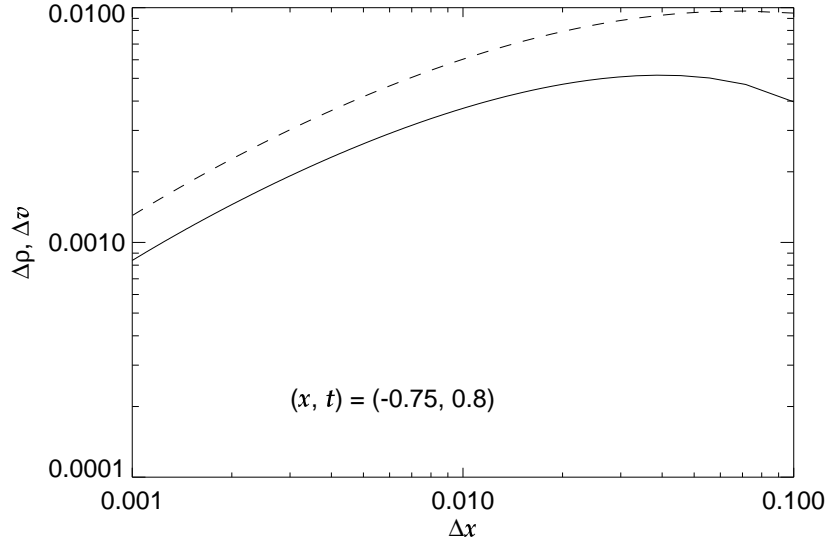


Figure 1.16: The errors in the density and velocity are shown as a function of the grid spacing for the solutions given in Figures 1.13 and 1.14. The errors are evaluated at  $(x, t) = (-0.75, 0.8)$ .

upwind scheme. The number of time steps required to reach a given  $t$  is inversely proportional to  $\Delta x$  because the time step is proportional to  $\Delta x$ . A scheme is of first-order accuracy when the error is proportional to the grid spacing. A higher-order scheme is explained in the next section.

The next example gives the solution for the following initial condition:

$$(\rho, P) = \begin{cases} (0.5, 0.2) & (x \leq -1) \\ (1.0, 1.0) & (-1 < x \leq 1) \\ (0.1, 0.05) & (x > 1) \end{cases} . \quad (1.116)$$

The initial velocity vanishes everywhere ( $v = 0$ ). The upper and lower panels denote the density and pressure distributions, respectively, by gray scale and contour. The density and pressure are higher in the black regions. The contours radiate from  $(x, t) = (0, \pm 1)$  in the early phase in the diagrams. The contours are straight because the waves have constant phase velocities. The two rarefaction waves cross each other at  $t \simeq 1$ . After crossing, the phase velocities change and the slopes of the contours change. Similar results are obtained when the initial velocity is not spatially constant, and also when the changes in initial density, velocity and pressure are more complex. These three types of waves radiate from any point at which the initial density, velocity, or pressure changes. The waves propagate and cross each other. Thus, a scheme can solve the hydrodynamical equations for a given initial condition if it can solve a shock tube problem. Thus, shock tube problems are often used as test problem for numerical simulation code.

Before proceeding further, we consider why the Roe scheme succeeds in solving the shock wave. The primary reason is that we have diagonalized the velocity matrix. Equation (1.77)

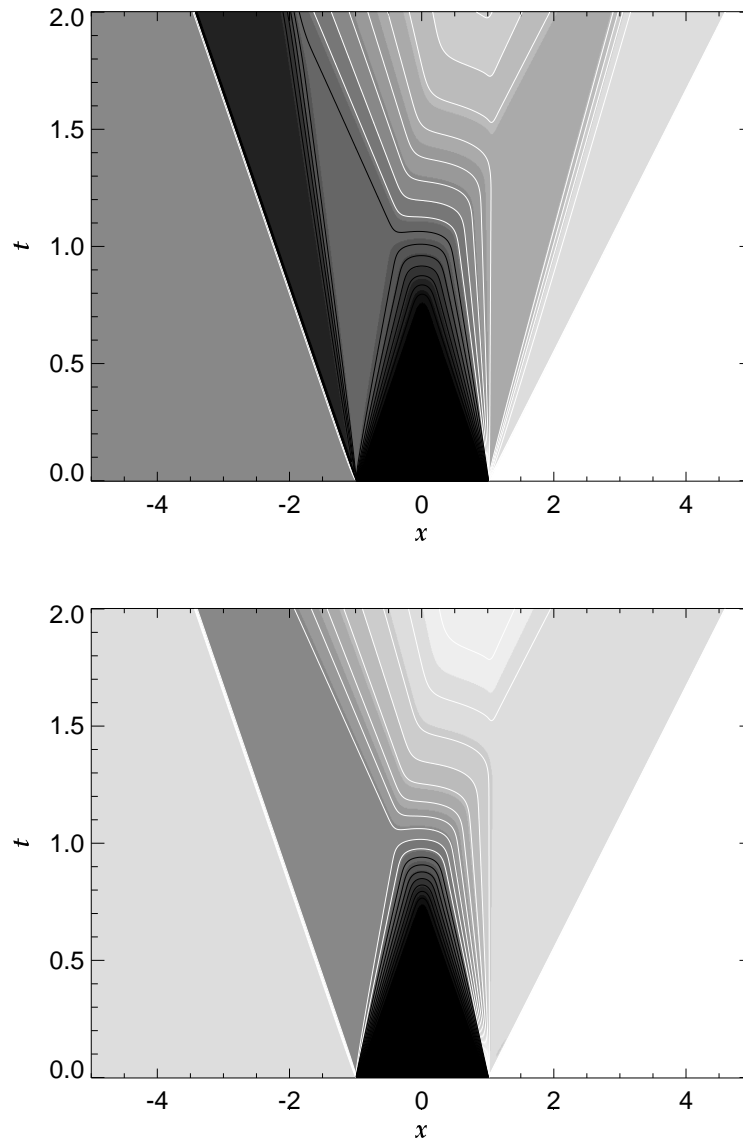


Figure 1.17: A test problem in which the initial density and pressure change at two points. The upper panel denotes the evolution of the density distribution by gray scale and contour, and the lower panel denotes the evolution of the pressure distribution.

can be rewritten as

$$\frac{\partial \mathbf{J}}{\partial t} + \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \frac{\partial \mathbf{J}}{\partial x} = 0, \quad (1.117)$$

where

$$d\mathbf{J} = \mathbf{L} d\mathbf{U}, \quad (1.118)$$

by multiplying the matrix  $\mathbf{L}$ . The newly defined variables,  $J_1$ ,  $J_2$ , and  $J_3$ , are Riemann invariants. The multiplications of  $\mathbf{R}$  and  $\mathbf{L}$  transform the hydrodynamical equations into a set of simple wave equations. However, Equations (1.99) – (1.101) also contribute to the success of the Roe scheme. Thanks to these equations, the relation,

$$\mathbf{F}(\mathbf{U}_{j+1}) - \mathbf{F}(\mathbf{U}_j) = \mathbf{A}(\mathbf{U}_{j+1} - \mathbf{U}_j), \quad (1.119)$$

holds exactly for any given  $\mathbf{U}_j$  and  $\mathbf{U}_{j+1}$ . This is the virtue of Roe's averaging. When  $\mathbf{U}_j = \mathbf{U}_{j+1}$ , the velocity matrix coincides with  $\partial \mathbf{F} / \partial \mathbf{U}$ . Furthermore, Equation (1.101) guarantees that  $a^2$  is always positive, and hence all of the phase velocities are real (not complex). These three conditions are referred to as the U property. If the velocity matrix  $\mathbf{A}$  fulfills the U property in a numerical scheme, then the hyperbolic differential equations can be solved by the upwind scheme and the solution is stable and free from numerical oscillation. Such velocity matrixes have been found for two-dimensional flows, three-dimensional flows and non-ideal-gas flows. When a numerical scheme is based on the velocity matrix satisfying the U property, it is generally referred to as a Roe type scheme, even when it is not derived by Roe. A Roe type scheme is also obtained for the magnetohydrodynamical equations. Some numerical schemes are classified as Roe type schemes by the developer even when they do not satisfy the U property. These schemes provide relatively good solutions, although they are often unstable with respect to strong shock waves.

The velocity matrix that satisfies the U property for an arbitrary equation of state was obtained by Nobuta and Hanawa (1999, ApJ, vol. 510, p. 614). They applied this scheme to a gas in which black body radiation was taken into account, and the scheme was shown to be applicable to degenerate electron gas. Eulerink and Mellema (1993, A&AS, 110, p. 587) obtained the velocity matrix for special relativistic hydrodynamical equations. Their velocity matrix satisfies the U property. A Roe type scheme was derived for the magnetohydrodynamical equations for the first time by Brio and Wu (1988, J. Comput. Phys., 75, p. 400). Their velocity matrix does not satisfy the U property, except for  $\gamma = 2$ . The velocity matrix satisfying the U property for any  $\gamma$  was presented by Cargo and Gallice (1997, J. Comput. Phys., 136, 446).

## 1.6 Higher-order Accuracy

In the previous section, a numerical method by which to solve one-dimensional hydrodynamical equations was described. In this section, we introduce a method by which to obtain a solution of higher-order accuracy. Numerical methods used to solve two-dimensional and three-dimensional hydrodynamical equations are shown in the next section.

The upwind scheme, which was shown in the previous section, is of the first-order in space and time. As the grid spacing and time step are taken to be smaller, the truncation error is smaller. However, an enormous computation time is required to obtain an accurate solution with the first-order scheme. Most numerical simulations employ a second-order or even higher-order scheme to reduce the computation time. In this section, we introduce the Monotone

Upstream-centered Scheme for Conservation Laws (MUSCL) to achieve second-order accuracy. Since MUSCL is rather difficult, easier methods may be desired. Thus, we first understand necessity of a rather complicated method.

First, we consider the one-dimensional wave equation to simplify the problem. Equation (1.9), the central difference, is of second-order accuracy because the error is a small third-order quantity. However, the solution obtained by Equation (1.9) is unstable and shows unnatural oscillations around a discontinuity. Is it possible to realize a higher-order accuracy while maintaining stability and avoiding unnatural oscillations? We shall examine the case in which the solution of Equation (1.5) at step  $n + 1$  should be a linear combination of the solution at step  $n$ , i.e.,

$$f_{j,n+1} = \sum_k B_k f_{j+k,n}. \quad (1.120)$$

This expression includes the forward difference, the backward difference, and the central difference, which are all shown in Section 1.2. If Equation (1.120) denotes the backward difference, then the coefficients,  $B_k$ , are given by

$$B_k = \begin{cases} 1 - \frac{c\Delta t}{\Delta x} & (k = 0) \\ \frac{c\Delta t}{\Delta x} & (k = -1) \\ 0 & (\text{otherwise}) \end{cases}. \quad (1.121)$$

Similarly, for the central difference scheme, they are given by

$$B_k = \begin{cases} 1 & (k = 0) \\ -\frac{c\Delta t}{2\Delta x} & (k = 1) \\ \frac{c\Delta t}{2\Delta x} & (k = -1) \\ 0 & (\text{otherwise}) \end{cases}. \quad (1.122)$$

By arranging  $B_k$ , we can create an infinite number of schemes. The Godunov theorem states that no scheme expressed by Equation (1.120) can achieve second- or higher-order accuracy without unnatural numerical oscillations.

**Godunov theorem:** Any second- or higher-order scheme expressed by Equation (1.120) cannot achieve monotonicity of the solution.

The gradient  $\partial f / \partial x$  changes its sign in the region in which an unnatural numerical oscillation arises. However, the gradient should remain positive if the initial gradient is positive because the wave form should keep the original sign. In other words, the wave amplitude should always increase monotonically with increasing  $x$  if the initial amplitude increases monotonically with increasing  $x$ . It is desired that the monotonicity be preserved in a numerical solution. The monotonicity is preserved if and only if

$$B_k \geq 0. \quad (1.123)$$

The proof of this equation is not easy. First, we prove that the monotonicity is preserved if Equation (1.123) holds. Equation (1.120) yields

$$f_{j+1,n+1} - f_{j,n} = \sum_k B_k (f_{j+1+k,n} - f_{j+k,n}). \quad (1.124)$$

If  $B_k \geq 0$  and  $f_{j+1,n} - f_{j,n} \geq 0$  for any given  $k$  at  $t = n\Delta t$ , then  $f_{j+1,n+1} - f_{j,n+1} \geq 0$ . The same is true for  $f_{j+1,n} - f_{j,n} \leq 0$ . Thus, the monotonicity of function  $f$  is preserved if  $B_k \geq 0$ .

The necessity of  $B_k \geq 0$  for maintaining monotonicity is proven by reduction to absurdity. Suppose  $B_m < 0$  for a given  $m$ . If the initial condition is given by

$$f_{j,n} = \begin{cases} 0 & (j > 0) \\ 1 & (j \leq 0) \end{cases}, \quad (1.125)$$

we obtain

$$f_{-m+1,n+1} - f_{-m,n+1} = \sum_k B_k (f_{k-m+1,n} - f_{k-m,n}) \quad (1.126)$$

$$= B_m (f_{1,n} - f_{0,n}) \quad (1.127)$$

$$> 0. \quad (1.128)$$

This solution violates monotonicity, because the function  $f_{j,n}$  increases monotonically with  $j$  while  $f_{j,n+1}$  does not. Thus, we have proven that the monotonicity is preserved if and only if  $B_k \geq 0$ .

Next, we examine the truncation error using the Taylor series. Taylor series expansion gives

$$f_{j,n+1} = f_{j,n} + \sum_{m=1}^{\infty} \frac{\partial^m f}{\partial t^m} \frac{(\Delta t)^m}{m!}, \quad (1.129)$$

which is rewritten as

$$f_{j,n+1} = f_{j,n} + \sum_{m=1}^{\infty} \frac{\partial^m f}{\partial x^m} \frac{(c\Delta t)^m}{m!}, \quad (1.130)$$

because

$$\frac{\partial^m f}{\partial t^m} = (-c)^m \frac{\partial^m f}{\partial x^m} \quad (1.131)$$

is derived from Equation (1.5). Another Taylor series expansion gives

$$f_{j+k,n} = f_{j,n} + \sum_{k=1}^{\infty} \frac{\partial^k f}{\partial x^k} (k\Delta x)^m. \quad (1.132)$$

Comparison of Equations (1.132) and (1.130) gives

$$\sum_k B_k = 1 \quad (1.133)$$

$$\sum_k k^m B_k = \left( -\frac{c\Delta t}{\Delta x} \right)^m. \quad (1.134)$$

from the condition that Equation (1.134) gives the condition that the scheme is of  $m$ -th order accuracy. We can easily confirm that the backward difference satisfies Equations (1.133) and (1.134) for  $m = 1$  but not for  $m = 2$ . The central difference satisfies Equation (1.134) for  $m \leq 2$ , although it does not guarantee the monotonicity of the solution because  $B_1 < 0$ .

We obtain

$$\sum_k \left[ k^2 - 2k \frac{c\Delta t}{\Delta x} + \left( \frac{c\Delta t}{\Delta x} \right)^2 \right] B_k = 0, \quad (1.135)$$

from the sum of Equation 1.133 multiplied by  $(c\Delta t/\Delta x)^2$ , Equation (1.134) for  $m = 1$  multiplied by  $2(c\Delta t/\Delta x)$ , and Equation (1.134) for  $m = 2$ . Equation (1.135) is rewritten as

$$\sum_k \left( k - \frac{c\Delta t}{\Delta x} \right)^2 B_k = 0. \quad (1.136)$$

If  $B_k \geq 0$  for any  $k$ , Equation (1.136) does not hold except when  $c\Delta t/\Delta x$  is an integer. Thus, Equations (1.133) and (1.134) for  $m = 1$  and 2 do not hold simultaneously as long as  $B_k \geq 0$  for any  $k$  except  $\Delta t$ .

The Godunov theorem tells us that conservation of monotonicity is a tough constraint. However, monotonicity is violated and unnatural oscillations arise only in a limited region. As shown in Section 1.2, the central difference causes numerical oscillation only in the region in which the value of  $f$  changes drastically. Thus, to avoid unnatural oscillations, we need care about monotonicity only in this region. When the value of  $f$  changes drastically, the Taylor series does not provide a good approximation and second-order accuracy is not important in this case. Second-order accuracy has no value in particular at the shock front because the density and velocity are discontinuous there. This allows for a balance between monotonicity and second-order accuracy to be struck. Second-order accuracy is realized almost everywhere, except in small regions in which the physical quantities change drastically.

Next, we reexamine second-order accuracy, paying attention to the numerical flux. Substituting the numerical flux defined as

$$F_{j+1/2,n} = c \left( \frac{f_{j+1,n} + f_{j,n}}{2} \right), \quad (1.137)$$

into Equation (1.44), we obtain the central difference. The numerical flux denotes the flux at  $x = (x_{j+1} + x_j)/2$ , as is indicated by the subscripts. The Taylor series expansion of the right-hand side of Equation (1.137) is of first-order accuracy when expanded in the Taylor series. On the other hand, the numerical flux is evaluated either at  $x = x_j$  or  $x_{j+1}$  in the backward difference and the forward difference. Thus, the truncation error is larger.

We next compare the accuracy of the first-order upwind scheme and that of the central difference using Figure 1.18. Suppose that the initial condition is given by  $f_j$ . The function is implicitly assumed to be a step function (dashed line) in the upwind scheme because the numerical flux at  $x = x_{j+1/2}$  is evaluated by either  $f_j$  or  $f_{j+1}$ . On the other hand, the function is implicitly assumed to be a stepwise linear function (solid curve) in the central difference, because the numerical flux is evaluated by the average of  $f_j$  and  $f_{j+1}$ . If we can approximate the initial condition not by the simple linear interpolation, but by another stepwise linear function, we will be able to achieve first-order accuracy in the numerical flux and second-order accuracy in the solution.

Fortunately, in addition to for the simple linear interpolation, we have several ways to approximate a function by a stepwise function. For example, extrapolating from the left-hand side,

$$f_{j+1/2}^{(L)} = f_j + \frac{1}{2}(f_j - f_{j-1}) \quad (1.138)$$

$$= \frac{3f_j - f_{j-1}}{2}, \quad (1.139)$$

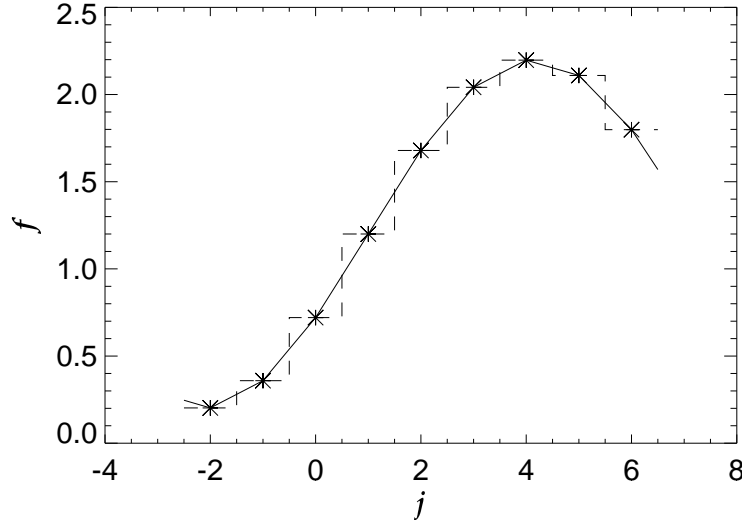


Figure 1.18: Example of simple linear interpolation.

is of first-order accuracy. Similarly, extrapolating from the right-hand side,

$$f_{j+1/2}^{(R)} = f_{j+1} - \frac{1}{2}(f_{j+2} - f_{j+1}) \quad (1.140)$$

$$= \frac{3f_{j+1} - f_{j+2}}{2}, \quad (1.141)$$

is of first-order accuracy. Here,  $f_{j+1/2}^{(L)}$  and  $f_{j+1/2}^{(R)}$  provide a solution of second-order accuracy if used in the evaluation of the numerical flux.

The use of  $f_{j+1/2}^{(L)}$  and  $f_{j+1/2}^{(R)}$  requires careful consideration. Figure 1.19 shows an example of extrapolation by Equations (1.139) and (1.141). The initial data are the same as those shown in Figure 1.18. Unnatural irregularities appear around the local minima and local maxima in the extrapolation.

It is a natural consequence of the Godunov theorem that the extrapolation violates monotonicity. The Godunov theorem states that the monotonicity is preserved only by the upwind scheme, i.e., by the step function.

Next, we modify Equation (1.139) to guarantee monotonicity. If we replace Equation (1.139) with

$$f_{j+1/2}^{*(L)} = \begin{cases} f_j & \Delta_{j+1/2} \Delta_{j-1/2} \leq 0 \\ f_j + \frac{\Delta_{j-1/2}}{2} \min \left( 1, \left| \frac{\Delta_{j+1/2}}{\Delta_{j-1/2}} \right| \right) & \text{(otherwise)} \end{cases}, \quad (1.142)$$

$$\Delta_{j+1/2} = f_{j+1} - f_j, \quad (1.143)$$

$$\Delta_{j-1/2} = f_j - f_{j-1}, \quad (1.144)$$

then the monotonicity problem is resolved. Equation (1.144) is of first-order accuracy only when the sign of  $\Delta_{j+1/2}$  is the same as that of  $\Delta_{j-1/2}$ , i.e., only when the function increases

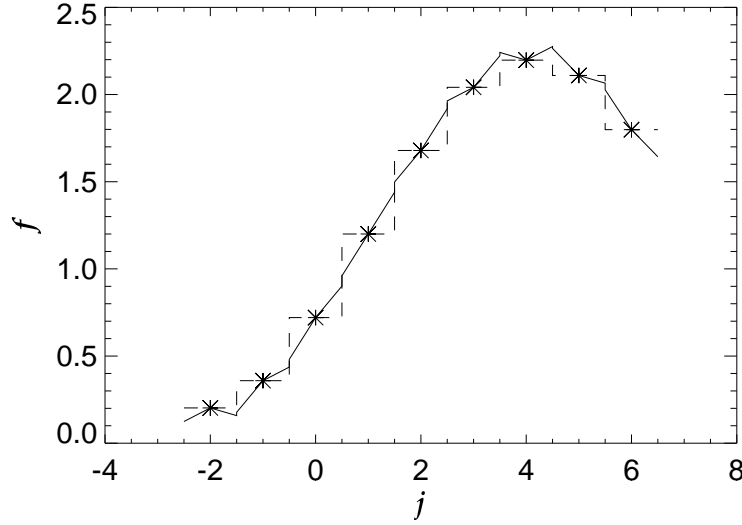


Figure 1.19: An example of extrapolation by Equations (1.139) and (1.141). The data are the same as those of Figure 1.18.

or decreases monotonically in the interval  $x_{j-1} \leq x \leq x_{j+1}$ . Otherwise, Equation (1.144) is approximated by  $f_j$  because the gradient changes and the function has either a local minimum or a local maximum in the interval. When the sign of  $\Delta_{j+1/2}$  is the same as that of  $\Delta_{j-1/2}$ , the minimum absolute gradient is used for the extrapolation. In most textbooks, Equation (1.142) is expressed as

$$f_{j+1/2}^{(L)} = f_j + \frac{\Delta_{j-1/2}}{2} \Psi \left( \frac{\Delta_{j+1/2}}{\Delta_{j-1/2}} \right), \quad (1.145)$$

$$\Psi(r) = \begin{cases} 0 & (r \leq 0) \\ r & (0 < r < 1) \\ 1 & (r \geq 1) \end{cases}, \quad (1.146)$$

and  $\Psi$  is referred to as the minmod limiter. In short, Equation (1.144) provides a modest extrapolation.<sup>10</sup>

Similarly, Equation (1.141) is replaced by

$$f_{j+1/2}^{*(R)} = \begin{cases} f_{j+1} & \Delta_{j+3/2} \Delta_{j+1/2} \leq 0 \\ f_{j+1} - \frac{\Delta_{j+1/2}}{2} \min \left( 1, \left| \frac{\Delta_{j+3/2}}{\Delta_{j+1/2}} \right| \right) & (\text{otherwise}) \end{cases}. \quad (1.147)$$

Equation (1.147) also gives a modest extrapolation from the right-hand side. When  $f_j \leq f_{j+1}$ , we obtain the following inequality:

$$f_j \leq f_{j+1/2}^{*(L)} \leq f_{j+1/2}^{*R} \leq f_{j+1}, \quad (1.148)$$

---

<sup>10</sup>In addition to the minmod limiter, the superbee and other limiters are described by Hirsch (1981).

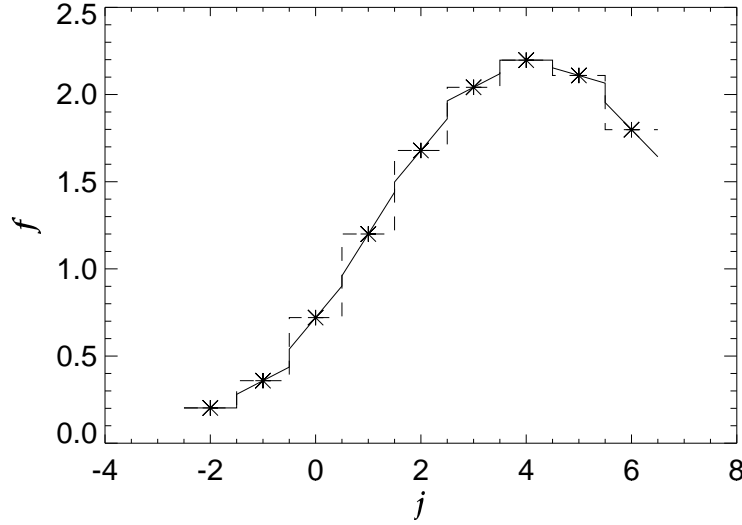


Figure 1.20: An example of extrapolation by Equations (1.144) and (1.147). The data are the same as those shown in Figure 1.18.

On the other hand, we obtain the following inequality:

$$f_j \geq f_{j+1/2}^{*(L)} \geq f_{j+1/2}^{*(R)} \geq f_{j+1}, \quad (1.149)$$

when  $f_j \geq f_{j+1}$ . The monotonicity is preserved in both cases. Figure 1.20 gives an example of the stepwise linear function obtained by Equations (1.144) and (1.147), which demonstrates that this stepwise linear extrapolation preserves the monotonicity.

Using  $f_{j+1/2}^{*(L)}$  and  $f_{j+1/2}^{*(R)}$ , we can construct the numerical flux as follows:

$$F_{j+1/2}^* = \frac{c}{2} \left( f_{j+1/2}^{*(L)} + f_{j+1/2}^{*(R)} \right) - \frac{|c|}{2} \left( \frac{f_{j+1/2}^{*(L)} - f_{j+1/2}^{*(R)}}{2} \right). \quad (1.150)$$

This numerical flux is constructed from the data of the upwind side and is of first-order accuracy. This numerical flux provides second-order accuracy in space.

Second-order accuracy in time is achieved by a two-stage method such as the predictor-corrector method for the ordinary differential equation. Suppose that we want to obtain the function  $f$  at  $t = t_0 + \Delta t$  from that at  $t = t_0$ . In this case, we obtain the function  $f$  at  $t = t_0 + \Delta t/2$  at the first stage. At the second stage, we obtain the function  $f$  at  $t = t_0 + \Delta t$  using the numerical flux at  $t = t_0 + \Delta/2$ .

The following is the procedure used to solve the Burgers equation with second-order accuracy in space and time.

(1) Extrapolate the function while preserving the monotonicity.

$$f_{j+1/2,n}^{*(L)} = \begin{cases} f_{j,n} & (\Delta_{j+1/2,n} \Delta_{j-1/2,n} \leq 0) \\ f_{j,n} + \frac{\Delta_{j-1/2,n}}{2} \min \left( 1, \left| \frac{\Delta_{j+1/2,n}}{\Delta_{j-1/2,n}} \right| \right) & (\text{otherwise}) \end{cases}, \quad (1.151)$$

$$f_{j+1/2,n}^{*(R)} = \begin{cases} f_{j+1,n} & (\Delta_{j+1/2,n} \Delta_{j+3/2,n} \leq 0) \\ f_{j+1,n} - \frac{\Delta_{j+3/2,n}}{2} \min \left( 1, \left| \frac{\Delta_{j+1/2,n}}{\Delta_{j+3/2,n}} \right| \right) & (\text{otherwise}) \end{cases}, \quad (1.152)$$

$$\Delta_{j-1/2,n} = f_{j,n} - f_{j-1,n}, \quad (1.153)$$

$$\Delta_{j+1/2,n} = f_{j+1,n} - f_{j,n}, \quad (1.154)$$

$$\Delta_{j+3/2,n} = f_{j+2,n} - f_{j+1,n}. \quad (1.155)$$

(2) Compute the numerical flux from  $f_{j+1/2,n}^{*(L)}$  and  $f_{j+1/2,n}^{*(R)}$ .

$$F_{j+1/2,n}^* = \frac{1}{2} \left[ \frac{\left( f_{j+1/2,n}^{*(L)} \right)^2}{2} + \frac{\left( f_{j+1/2,n}^{*(R)} \right)^2}{2} \right] - \frac{|\lambda|}{2} \left( f_{j+1/2,n}^{*(L)} - f_{j+1/2,n}^{*(R)} \right), \quad (1.156)$$

$$|\lambda| = \begin{cases} \frac{|f_{j+1/2,n}^{*(L)} + f_{j+1/2,n}^{*(R)}|}{2} & \left( \frac{|f_{j+1/2,n}^{*(L)} + f_{j+1/2,n}^{*(R)}|}{2} \geq \varepsilon \right) \\ \varepsilon & (\text{otherwise}) \end{cases}, \quad (1.157)$$

$$\varepsilon = \max \left( 0, \frac{f_{j+1/2,n}^{*(L)} - f_{j+1/2,n}^{*(R)}}{2} \right), \quad (1.158)$$

(3) Obtain  $f_{j,n+1/2}$  using the numerical flux.

$$f_{j,n+1/2} = f_{j,n} + \frac{\Delta t}{2\Delta x} \left( F_{j+1/2,n}^* - F_{j-1/2,n}^* \right). \quad (1.159)$$

(4) Extrapolate  $f_{j,n+1/2}$  as in step (1).

$$f_{j+1/2,n+1/2}^{*(L)} = \begin{cases} f_{j,n+1/2} & (\Delta_{j+1/2,n+1/2} \Delta_{j-1/2,n+1/2} \leq 0) \\ f_{j,n+1/2} + \frac{\Delta_{j-1/2,n+1/2}}{2} \min \left( 1, \left| \frac{\Delta_{j+1/2,n+1/2}}{\Delta_{j-1/2,n+1/2}} \right| \right) & (\text{otherwise}) \end{cases}, \quad (1.160)$$

$$f_{j+1/2,n+1/2}^{*(R)} = \begin{cases} f_{j+1,n+1/2} & (\Delta_{j+1/2,n+1/2} \Delta_{j+3/2,n+1/2} \leq 0) \\ f_{j+1,n+1/2} - \frac{\Delta_{j+3/2,n+1/2}}{2} \min \left( 1, \left| \frac{\Delta_{j+1/2,n+1/2}}{\Delta_{j+3/2,n+1/2}} \right| \right) & (\text{otherwise}) \end{cases}, \quad (1.161)$$

$$\Delta_{j-1/2,n+1/2} = f_{j,n+1/2} - f_{j-1,n+1/2}, \quad (1.162)$$

$$\Delta_{j+1/2,n+1/2} = f_{j+1,n+1/2} - f_{j,n+1/2}, \quad (1.163)$$

$$\Delta_{j+3/2,n+1/2} = f_{j+2,n+1/2} - f_{j+1,n+1/2}. \quad (1.164)$$

- (5) Obtain the numerical flux from  $f_{j+1/2,n+1}^{(L)}$  and  $f_{j+1/2,n+1/2}^{(R)}$ .<sup>11</sup>

$$F_{j+1/2,n+1/2}^* = \frac{1}{2} \left[ \frac{\left(f_{j+1/2,n+1/2}^{*(L)}\right)^2}{2} + \frac{\left(f_{j+1/2,n+1/2}^{*(R)}\right)^2}{2} \right] - \frac{|\lambda|}{2} \left( f_{j+1/2,n+1/2}^{*(L)} - f_{j+1/2,n+1/2}^{*(R)} \right), \quad (1.165)$$

$$|\lambda| = \begin{cases} \frac{|f_{j+1/2,n+1/2}^{*(L)} + f_{j+1/2,n+1/2}^{*(R)}|}{2} & \left( \frac{|f_{j+1/2,n+1/2}^{*(L)} + f_{j+1/2,n+1/2}^{*(R)}|}{2} \geq \varepsilon \right) \\ \varepsilon & (otherwise) \end{cases}, \quad (1.166)$$

$$\varepsilon = \max \left( 0, \frac{f_{j+1/2,n+1/2}^{*(L)} - f_{j+1/2,n+1/2}^{*(R)}}{2} \right), \quad (1.167)$$

- (6) Obtain  $f_{j,n}$  using the numerical flux obtained in (5).

$$f_{j,n} = f_{j,n} + \frac{\Delta t}{\Delta x} \left( F_{j+1/2,n+1/2}^* - F_{j-1/2,n+1/2}^* \right). \quad (1.168)$$

The above procedure is called MUSCL and can be applied to the hydrodynamical equations. When MUSCL is applied to the hydrodynamical equations, the question arises as to which variables should be extrapolated. For beginners, we recommend extrapolation of the density, velocity, and temperature ( $T \equiv P/\rho$ ). Extrapolation of the pressure and extrapolation of the state variable  $\mathbf{U}$  may cause difficulty near the shock front. Extrapolation of the Riemann invariants succeed in solving the magnetohydrodynamical equations (see, Fukuda and Hanawa 1999, ApJ, 517, p. 226). Extrapolation of the Riemann invariants also enables solution of the hydrodynamical equations. However, the extrapolation of the Riemann solver increases the computation cost and the complexity of the computation code. Thus, we do not recommend the extrapolation of the Riemann invariants for beginners.

## 1.7 Extension to Multi-dimensional Problems

In the previous sections, we restricted ourselves to the consideration of one-dimensional problems. In this section, we introduce numerical methods for solving the two-dimensional and three-dimensional hydrodynamical equations. We use Cartesian coordinates in this section. The use of cylindrical coordinates and spherical coordinates is discussed in Section 1.9.

The hydrodynamical equations for a multi-dimensional problem are expressed as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (1.169)$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} + \frac{1}{\rho} \nabla P = 0, \quad (1.170)$$

$$\frac{\partial}{\partial t} (\rho E) + \nabla \cdot (\rho \mathbf{v} H) = 0. \quad (1.171)$$

---

<sup>11</sup>Although it is omitted here for simplicity, we should take the entropy condition into account, .

As in the case for the one-dimensional problem, they can be rewritten in conservation form as follows:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}_x}{\partial x} + \frac{\partial \mathbf{G}_y}{\partial y} + \frac{\partial \mathbf{H}_z}{\partial z} = 0, \quad (1.172)$$

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ \rho E \end{pmatrix}, \quad (1.173)$$

$$\mathbf{F} = \begin{pmatrix} \rho u \\ \rho u^2 \\ \rho uv \\ \rho uw \\ \rho H u \end{pmatrix}, \quad (1.174)$$

$$\mathbf{G} = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 \\ \rho vw \\ \rho H v \end{pmatrix}, \quad (1.175)$$

$$\mathbf{H} = \begin{pmatrix} \rho w \\ \rho wu \\ \rho wv \\ \rho w^2 \\ \rho H w \end{pmatrix}. \quad (1.176)$$

Here,  $u$ ,  $v$ , and  $w$  denote the  $x$ -,  $y$ -, and  $z$  components of the velocity, respectively. The density is evaluated as

$$\rho_{i,j,k,n} \equiv \rho(x = i \Delta x, y = j \Delta y, z = k \Delta z, t = n \Delta t) \quad (1.177)$$

at the center of the small rectangular box having a volume of  $\Delta x \times \Delta y \times \Delta z$ . Similarly, each component of the velocity,  $E$  and  $H$ , is evaluated at the center of the rectangular box. The last subscript indicates the time, as in the case of the one-dimensional flow. For later convenience, we use the following notation:

$$\mathbf{U}_{i,j,k,n} \equiv \mathbf{U}(x = i \Delta x, y = j \Delta y, z = k \Delta z, t = n \Delta t). \quad (1.178)$$

Then the difference equation for a multi-dimensional flow is expressed as

$$\begin{aligned} & \frac{\mathbf{U}_{i,j,k,n+1} - \mathbf{U}_{i,j,k,n}}{\Delta t} + \frac{\mathbf{F}_{i+1/2,j,k,n}^* - \mathbf{F}_{i-1/2,j,k,n+1}^*}{\Delta x} \\ & + \frac{\mathbf{G}_{i,j+1/2,k,n}^* - \mathbf{G}_{i,j-1/2,k,n}^*}{\Delta y} \\ & + \frac{\mathbf{H}_{i,j,k+1/2,n}^* - \mathbf{H}_{i,j,k-1/2,n}^*}{\Delta z} = 0, \end{aligned} \quad (1.179)$$

where  $\mathbf{F}_{i+1/2,j,k,n}^*$ ,  $\mathbf{G}_{i,j+1/2,k,n}^*$ , and  $\mathbf{H}_{i,j,k+1/2,n}^*$  denote the numerical fluxes in the  $x$ -,  $y$ -, and  $z$ -directions, respectively. The numerical fluxes should of course be upwind fluxes.

First, we obtain the numerical flux in the  $x$ -direction,  $\mathbf{F}_{i+1/2,j,k,n}$ . Since the state vector  $\mathbf{U}$  and the flux vector  $\mathbf{F}$  have three components in the one-dimensional problem, the velocity matrix  $\mathbf{A}$  has three rows and three columns. Consequently, there exist three eigenvalues, three right eigenvectors, and three left eigenvectors. The state and flux vectors ( $\mathbf{U}$ ,  $\mathbf{F}$ ,  $\mathbf{G}$ , and  $\mathbf{H}$ ) have five components in the three-dimensional hydrodynamical equations. Thus, we have five characteristic speeds, five right eigenvectors, and five left eigenvectors, which are derived as follows. First, we express each component of the flux vector as a function of the state:

$$\mathbf{F} = \left\{ \begin{array}{c} (\gamma - 1) \left[ U_5 - \frac{(U_2)^2 + (U_3)^2 + (U_4)^2}{U_1} \right] + \frac{(U_2)^2}{U_1} \\ \frac{U_2 U_3}{U_1} \\ \frac{U_2 U_4}{U_1} \\ \frac{U_2 U_5}{U_1} + (\gamma - 1) \left[ U_5 - \frac{(U_2)^2 + (U_3)^2 + (U_4)^2}{U_1} \right] \frac{U_2}{U_1} \end{array} \right\}. \quad (1.180)$$

The velocity matrix ( $\mathbf{A} \equiv \partial \mathbf{F} / \partial \mathbf{U}$ ) is then expressed as

$$\mathbf{A} = \left[ \begin{array}{ccc} 0 & 1 & \\ \frac{(\gamma - 1)}{2} q - u^2 & (3 - \gamma) u & \\ -uv & v & \\ -uw & w & \\ -\gamma u E + (\gamma - 1) u q & \gamma E - \frac{(\gamma - 1)}{2} (2u^2 + q) & \\ 0 & 0 & 0 \\ (1 - \gamma) v & (1 - \gamma) w & (\gamma - 1) \\ u & 0 & 0 \\ 0 & u & 0 \\ (1 - \gamma) uv & (1 - \gamma) uw & \gamma u \end{array} \right], \quad (1.181)$$

$$q = u^2 + v^2 + w^2. \quad (1.182)$$

This velocity matrix has five eigenvalues,

$$\lambda_1 = u - a, \quad (1.183)$$

$$\lambda_2 = u, \quad (1.184)$$

$$\lambda_3 = u, \quad (1.185)$$

$$\lambda_4 = u, \quad (1.186)$$

$$\lambda_5 = u + a, \quad (1.187)$$

where  $a$  denotes the speed of sound as in the case of the case of a one-dimensional flow. The

corresponding right eigenvectors are expressed as follows:

$$\mathbf{r}_1 = \begin{pmatrix} 1 \\ u - a \\ v \\ w \\ H - au \end{pmatrix}, \quad (1.188)$$

$$\mathbf{r}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -a \\ -wa \end{pmatrix}, \quad (1.189)$$

$$\mathbf{r}_3 = \begin{pmatrix} 0 \\ 0 \\ a \\ 0 \\ va \end{pmatrix}, \quad (1.190)$$

$$\mathbf{r}_4 = \begin{pmatrix} 1 \\ u \\ v \\ w \\ \frac{q}{2} \end{pmatrix}, \quad (1.191)$$

$$\mathbf{r}_5 = \begin{pmatrix} 1 \\ u + a \\ v \\ w \\ H + au \end{pmatrix}. \quad (1.192)$$

Similarly, the left eigenvectors are expressed as follows:

$$\ell_1 = \left[ \frac{1}{2} \left( b_1 + \frac{u}{a} \right), -\frac{1}{2} \left( \frac{1}{a} + b_2 u \right), -\frac{b_2 v}{2}, -\frac{b_2 w}{2}, \frac{b_2}{2} \right], \quad (1.193)$$

$$\ell_2 = \left( \frac{w}{a}, 0, 0, -\frac{1}{c}, 0 \right), \quad (1.194)$$

$$\ell_3 = \left( -\frac{v}{a}, 0, \frac{1}{c}, 0, 0 \right), \quad (1.195)$$

$$\ell_2 = \left( \frac{w}{a}, 0, 0, -\frac{1}{c}, 0 \right), \quad (1.196)$$

$$\ell_4 = (1 - b_1, b_2 u, b_2 v, b_2 w, -b_2), \quad (1.197)$$

$$\ell_5 = \left[ \frac{1}{2} \left( b_1 - \frac{u}{a} \right), \frac{1}{2} \left( \frac{1}{a} - b_2 u \right), -\frac{b_2 v}{2}, -\frac{b_2 w}{2}, \frac{b_2}{2} \right], \quad (1.198)$$

where

$$b_1 = \frac{q}{2} \frac{\gamma - 1}{a^2}, \quad (1.199)$$

$$b_2 = \frac{\gamma - 1}{a^2}. \quad (1.200)$$

The right and left eigenvectors are normalized so that

$$\boldsymbol{\ell}_m \cdot \mathbf{r}_n = \delta_{m,n}. \quad (1.201)$$

These eigenvalues and eigenvectors are, of course, similar to those for the one-dimensional hydrodynamical equations. The two eigenvectors, the eigenvalues of which are  $v \pm a$ , are modified slightly to include the  $y$ - and  $z$ -components of the velocity. The eigenvalue  $u$  is degenerated three times to obtain two additional eigenvectors. The eigenvector  $\mathbf{r}_4$  denotes the entropy flow, as in the one-dimensional hydrodynamical equations. The eigenvectors  $\mathbf{r}_2$  and  $\mathbf{r}_3$  denote the shear flows in which the  $z$ - and  $y$ -components of the velocity, respectively, change in the  $x$ -direction.

If we evaluate the eigenvalues and eigenvectors by the Roe average defined as

$$\bar{u} = \frac{\sqrt{\rho_{i+1,j,k}} u_{i+1,j,k} + \sqrt{\rho_{i,j,k}} u_i}{\sqrt{\rho_{i+1,j,k}} + \sqrt{\rho_{i,j,k}}}, \quad (1.202)$$

$$\bar{v} = \frac{\sqrt{\rho_{i+1,j,k}} v_{i+1,j,k} + \sqrt{\rho_{i,j,k}} v_{i,j,k}}{\sqrt{\rho_{i+1,j,k}} + \sqrt{\rho_{i,j,k}}}, \quad (1.203)$$

$$\bar{w} = \frac{\sqrt{\rho_{i+1,j,k}} w_{i+1,j,k} + \sqrt{\rho_{i,j,k}} w_{i,j,k}}{\sqrt{\rho_{i+1,j,k}} + \sqrt{\rho_{i,j,k}}}, \quad (1.204)$$

$$\bar{q} = \bar{u}^2 + \bar{v}^2 + \bar{w}^2, \quad (1.205)$$

$$\bar{H} = \frac{\sqrt{\rho_{i+1,j,k}} H_{i+1,j,k} + \sqrt{\rho_{i,j,k}} H_{i,j,k}}{\sqrt{\rho_{i+1,j,k}} + \sqrt{\rho_{i,j,k}}}, \quad (1.206)$$

$$\bar{a} = (\gamma - 1) \left( \bar{H} - \frac{\bar{q}}{2} \right), \quad (1.207)$$

then the velocity matrix satisfies the U property. Thus, the numerical flux is given by

$$\mathbf{F}_{i+1/2,j,k}^* = \frac{1}{2} (\mathbf{F}_{i+1,j,k} + \mathbf{F}_{i,j,k}) - \frac{1}{2} \mathbf{R} |\mathbf{\Lambda}| \mathbf{L} (\mathbf{U}_{i+1,j,k} - \mathbf{U}_{i,j,k}). \quad (1.208)$$

Equation (1.208) is rewritten as

$$\mathbf{F}_{j+1/2,n}^* = \frac{1}{2} \left[ \mathbf{F}_{j+1,n} + \mathbf{F}_{j,n} - \sum_{k=1}^5 |\lambda_k| w_k \mathbf{r}_k \right], \quad (1.209)$$

where

$$w_1 = \frac{1}{2\bar{a}} \left[ \frac{P_{j+1} - P_j}{\bar{a}} - \bar{\rho} (v_{j+1} - v_j) \right], \quad (1.210)$$

$$w_2 = -\frac{\bar{\rho}}{\bar{a}} (w_{j+1} - w_j), \quad (1.211)$$

$$w_3 = -\frac{\bar{\rho}}{\bar{a}} (v_{j+1} - v_j), \quad (1.212)$$

$$w_4 = \rho_{j+1} - \rho_j - \frac{P_{j+1} - P_j}{\bar{a}^2}, \quad (1.213)$$

$$w_5 = \frac{1}{2\bar{a}} \left[ \frac{P_{j+1} - P_j}{\bar{a}} + \bar{\rho} (v_{j+1} - v_j) \right], \quad (1.214)$$

because

$$\mathbf{U}_{j+1} - \mathbf{U}_{j+1} = \sum_{k=1}^5 w_k \mathbf{r}_k, \quad (1.215)$$

for any  $U_j$  and  $U_{j+1}$ .

The remaining numerical fluxes,  $G_{i,j+1/2,k,n}$  and  $H_{i,j,k+1/2,n}$ , are obtained in a similar manner. However, we can obtain the corresponding formula by replacing  $u$ ,  $v$ , and  $w$  in a cyclic manner. The details are omitted here in order to save space.

We strongly recommend using the same function (or subroutine) to compute  $F_{i+1/2,j,n}$ ,  $G_{i,j+1/2,k,n}$ , and  $H_{i,j,k+1/2,n}$  when coding the program in C or in Fortran. If the arguments are changed appropriately, the function (or subroutine) can compute all of the numerical fluxes. Two or more functions that are essentially the same are not easy to handle.

By substituting the numerical fluxes into Equation (1.172), we obtain a new state vector,  $U_{i,j,k,n+1}$ , which is of the first-order in space and time.

Next, we consider the Courant condition. The time step should be smaller in the three-dimensional flow than in the one-dimensional flow. If the time step is smaller than

$$\frac{1}{\Delta t} \geq \frac{\lambda_{x,max}}{\Delta x} + \frac{\lambda_{y,max}}{\Delta y} + \frac{\lambda_{z,max}}{\Delta z}, \quad (1.216)$$

$$\lambda_{x,max} = |v_x| + c_s, \quad (1.217)$$

$$\lambda_{y,max} = |v_y| + c_s, \quad (1.218)$$

$$\lambda_{z,max} = |v_z| + c_s, \quad (1.219)$$

we can obtain the stable solution safely.<sup>12</sup>

The solution of second-order accuracy in time is obtained by the following two-stage procedure:

1. Compute the numerical fluxes  $F_{i+1/2,j,k,n}^*$ ,  $G_{i,j+1/2,k,n}^*$ , and  $H_{i,j,k+1/2,n}^*$  from the initial state  $U_{i,j,k,n}$ .
2. Obtain the intermediate state  $U_{i,j,k,n+1/2}$ , at  $t = t_0 + \Delta t/2$  using the numerical flux.
3. Obtain the numerical fluxes  $F_{i+1/2,j,k,n+1/2}^*$ ,  $G_{i,j+1/2,k,n+1/2}^*$ , and  $H_{i,j,k+1/2,n+1/2}^*$  from the intermediate state  $U_{i,j,k,n+1/2}$  at  $t = t_0 + \Delta t/2$  so that the numerical fluxes are of first-order accuracy in space.
4. Obtain the new state  $U_{i,j,k,n+1}$  at  $t = t_0 + \Delta t$  using the numerical flux obtained in the previous stage.

We have treated the three numerical fluxes separately, which is referred to as directional splitting. Directional splitting is widely used because it is easy to implement. However, directional splitting does have some disadvantages, as discussed below.

When the gradient of the density (or that of the velocity) is parallel to the coordinate, the flow is essentially one-dimensional and directional splitting is not problematic. However, when the gradient is inclined with respect to the coordinate, the directional splitting produces a spurious feature.

The weak points are relaxed if we take nearby cells into account when we interpolate and extrapolate the density and velocity. As shown in Section 1.6, the physical variables, such as the density and velocity, are extrapolated in order to achieve higher-order accuracy. The variables are extrapolated along the coordinate in directional splitting. If the surrounding cells are taken into account, the gradient will be improved.

The numerical fluxes for the two-dimensional hydrodynamical equations are derived easily from the numerical fluxes for the three-dimensional hydrodynamical equations.

---

<sup>12</sup>The time step can be taken to be slightly longer than that shown here. Here, a very safe criterion is given.

## 1.8 Inclusion of Gravity, Heating and Cooling

Thus far, we have ignored gravity in the hydrodynamical equations. However, gravity plays an important role in astrophysics. Gravity appears as a source term in the hydrodynamical equations.

If we take gravity,  $\mathbf{g}$ , into account, the equation of motion and the equation for energy conservation are rewritten as

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} + \frac{1}{\rho} \nabla P = \rho \mathbf{g}, \quad (1.220)$$

and

$$\frac{\partial}{\partial t} (\rho E) + \nabla (\rho \mathbf{v} H) = \rho \mathbf{v} \cdot \mathbf{g}, \quad (1.221)$$

respectively. The right-hand side of Equations (1.220) and (1.221) are the source terms due to gravity. The gravitational acceleration,  $\mathbf{g}$ , is given by an external field (i.e., by an explicit function of the coordinates) or by the solution of the Poisson equation for a given density distribution. The Poisson equation is solved by iteration. In this textbook, we assume that the gravitational acceleration has already been obtained.

The right-hand sides of Equations (1.220) and (1.221) are added separately after solving the left-hand sides. In other words, we obtain the solution of the first-order in space and time as

$$\frac{U_{j,n+1} - U_{j,n}}{\Delta t} = \frac{F_{j+1/2,n}^* - F_{j-1/2,n}^*}{\Delta x} + S_{j,n} \quad (1.222)$$

$$\mathbf{S} = \begin{pmatrix} 0 \\ \rho \mathbf{g} \\ \rho \mathbf{g} \cdot \mathbf{v} \end{pmatrix}. \quad (1.223)$$

The solution of the second-order accuracy in time is obtained by adding the source term ( $\mathbf{S}$ ) at each stage.

Inclusion of the gravity is easy, as shown above. However, a certain degree of care must be taken. First, the grid spacing,  $\Delta x$ , should be small, so that

$$|\mathbf{g}| \Delta x \ll a^2. \quad (1.224)$$

The left-hand side of Equation (1.224) denotes the potential energy difference over the grid spacing, which is equal to the potential energy change when the gas flows from a cell to its neighbor. The right-hand side of Eq. (1.224) denotes the square of the speed of sound, which is equal to the specific thermal energy. Thus, Equation (1.224) requires that the spatial change in the potential energy should be much smaller than the thermal energy. This is equivalent to the requirement that the grid spacing be much smaller than the pressure scale height. If Equation (1.224) is divided by  $|\mathbf{g}|$ , the right-hand denotes the pressure scale height. When the grid spacing is smaller than one-tenth the pressure scale height, gravity is taken into account fairly accurately. If the grid spacing is larger than half the pressure scale height, the accuracy of the solution is very limited. In addition, unnatural features may be observed due to the large gravity.

Inclusion of the gravity appears to place another constraint on the time step,  $\Delta t$ . Suppose that the initial velocity vanishes. The gas element moves by  $g \Delta t^2 / 2$  in a time step. The movement should be smaller than the grid spacing,  $\Delta x$ . This gives the following condition:

$$\frac{g \Delta t^2}{2} \leq \Delta x, \quad (1.225)$$

which is equivalent to

$$\Delta t \leq \sqrt{\frac{2}{g \Delta x}} \Delta x. \quad (1.226)$$

This constraint is automatically satisfied if the Courant condition and Equation (1.224) are satisfied simultaneously. In practice, for safety, we recommend lowering the CFL number slightly when gravity is included in the computation.

Heating and cooling by radiation and nuclear and thermal reactions appear as source terms in the hydrodynamical equations:

$$\frac{\partial}{\partial t}(\rho E) + \nabla(\rho \mathbf{v} H) = \rho \mathbf{v} \cdot \mathbf{g} + \Gamma - \Lambda, \quad (1.227)$$

where  $\Gamma$  and  $\Lambda$  denote the heating and cooling rates per unit volume, respectively. When expressed as functions of density and temperature, they can be included easily. Heating by nuclear and chemical reactions can be expressed as a function of density and temperature. In addition, heating and cooling by radiation can be approximated as a function of density and temperature, when the system is optically thin. When the system is optically thick, the radiative transfer equations should be solved simultaneously. However, this is beyond the scope of this textbook.

## 1.9 Extension to the Cylindrical and Spherical Coordinates

Next, in this section, we introduce a numerical method by which to solve the hydrodynamical equations using cylindrical or spherical coordinates. When the hydrodynamical equations are expressed in cylindrical or spherical coordinates, we need to modify them to apply the upwind scheme.

Suppose that the density and velocity are independent of  $\varphi$  in cylindrical coordinates,  $(r, \varphi, z)$ , for simplicity. The mass conservation is then expressed as

$$\frac{\partial \rho}{\partial t} + \frac{1}{r} \frac{\partial}{\partial r}(r \rho v_r) + \frac{\partial}{\partial z}(r \rho v_z) = 0. \quad (1.228)$$

The second term diverges from the axis ( $r = 0$ ) and is difficult to treat in this form. To avoid the divergence, we rewrite the mass conservation as

$$\frac{\partial}{\partial t}(r \rho) + \frac{\partial}{\partial r}(r \rho v_r) + \frac{\partial}{\partial z}(r \rho v_z) = 0, \quad (1.229)$$

by multiplying by  $r$ . We can rewrite this differential equation in integral form as

$$\int_V \frac{\partial \rho}{\partial t} dV + \int_S \rho \mathbf{v} \cdot d\mathbf{S} = 0, \quad (1.230)$$

by integrating in the  $r$ - and  $z$ -directions. Here,  $dV$  denotes the volume integral and is equal to  $r dr d\varphi dz$ . The integral form (Gauss's law) denotes that the temporal change in mass within a volume is equal to the sum of the mass flux flowing into and out from the volume thorough the surface. Equation (1.229) is better than Equation (1.228) because the mass conservation is more clearly shown in this expression.

If we multiply the hydrodynamical equations by  $r$ , they are expressed as follows:

$$\frac{\partial}{\partial t}(r \mathbf{U}) + \frac{\partial}{\partial r}(r \mathbf{F}_r) + \frac{\partial}{\partial z}(r \mathbf{F}_z) = \mathbf{S}, \quad (1.231)$$

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho v_r \\ \rho v_\varphi \\ \rho v_z \\ \rho E \end{pmatrix}, \quad (1.232)$$

$$\mathbf{F}_r = \begin{pmatrix} \rho v_r \\ \rho v_r^2 + P \\ \rho v_r v_\varphi \\ \rho v_r v_z \\ \rho H v_r \end{pmatrix}, \quad (1.233)$$

$$\mathbf{F}_z = \begin{pmatrix} \rho v_z \\ \rho v_r v_z \\ \rho v_\varphi v_z \\ \rho v_z^2 + P \\ \rho H v_z \end{pmatrix}, \quad (1.234)$$

$$\mathbf{S} = \begin{pmatrix} 0 \\ \rho v_\varphi^2 + P \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (1.235)$$

$$E = \frac{v_r^2 + v_\varphi^2 + v_z^2}{2} + \frac{1}{\gamma - 1} \frac{P}{\rho}, \quad (1.236)$$

$$H = \frac{v_r^2 + v_\varphi^2 + v_z^2}{2} + \frac{\gamma}{\gamma - 1} \frac{P}{\rho}. \quad (1.237)$$

Note that the extra source terms appear in this expression. The first source term denotes the centrifugal force, and the second source term is related to the pressure gradient because

$$\frac{\partial P}{\partial r} = \frac{1}{r} \frac{\partial}{\partial r} (rP) - \frac{P}{r}. \quad (1.238)$$

These source terms, as well as the gravity, should be taken into account in the cylindrical coordinates.

The state vector ( $\mathbf{U}$ ) and the flux vectors ( $\mathbf{F}_r$  and  $\mathbf{F}_z$ ) are also multiplied by  $r$ . The radius,  $r$ , is evaluated at the cell surface when the numerical flux is computed. When we compute the numerical flux between the cells centered at  $r = r_j$  and at  $r = r_{j+1}$ , the radius should be evaluated simply as  $r = r_{j+1/2} \equiv (r_j + r_{j+1})/2$ . Note that the value of  $r$  need not be evaluated at the upwind side. Thus, the numerical flux in the  $r$ -direction ( $\mathbf{F}_r$ ) is evaluated by applying Roe's formula as given in Section 1.5 after replacing  $(v_x, v_y, v_z)$  with  $(v_r, v_\varphi, v_z)$ .

Care must be taken in evaluating the rotation velocity  $v_\varphi$  near the axis, because it is small and proportional to the radius near the  $z$ -axis. However, the obtained value should be accurate because the centrifugal force is  $v_\varphi^2/r$ . Otherwise, the error is amplified by a factor of  $1/r$ . The best way to evaluate the rotation velocity is to assume that the angular velocity  $\Omega \equiv v_\varphi/r$  is approximately constant and changes smoothly near the axis. When the flow is symmetric around the axis ( $\partial/\partial\varphi = 0$ ), this is a fairly good approximation. Thus, the boundary condition

near the axis is given by

$$\frac{\partial \rho}{\partial r} = 0, \quad (1.239)$$

$$v_r = 0, \quad (1.240)$$

$$\frac{\partial \Omega}{\partial r} = 0, \quad (1.241)$$

$$\frac{\partial v_z}{\partial r} = 0. \quad (1.242)$$

This boundary condition is equivalent to assuming that  $\rho$ ,  $\Omega$ , and  $v_z$  are even with respect to the reversal of  $r$ , while  $v_r$  is odd.

Note that the third component of Equation (1.231) has no source term and hence the total angular momentum around the  $z$ -axis is conserved in this finite difference method. This is an advantage of the conservation form. The total angular momentum in the computation volume is conserved within the round-off error as long as the boundary conditions are appropriate.

It is possible to include the dependence on  $\varphi$  in the cylindrical coordinates. However, in practice, this is not possible near the  $z$ -axis because many cells are adjacent to each other on the axis of  $r = 0$ . Special care should be taken if the three-dimensional hydrodynamical equations are computed in cylindrical coordinates.

When spherical coordinates are applied, the mass conservation should be rewritten as

$$\frac{\partial}{\partial t} (r^2 \sin \theta \rho) + \frac{\partial}{\partial r} (r^2 \sin \theta \rho v_r) + \frac{\partial}{\partial \theta} (r^2 \sin \theta \rho v_\theta) + \frac{\partial}{\partial \varphi} (r^2 \sin \theta \rho v_\varphi) = 0. \quad (1.243)$$

Similarly, the other components of the hydrodynamical equations should be multiplied by a factor  $r^2 \sin \theta$ . Then, the source terms appear in the conservation form because of the centrifugal force and pressure. They can be treated similarly to the case of the cylindrical coordinates. Details are omitted to save space.

## 1.10 Boundary Conditions

The boundary condition, as well as the initial condition, is necessary for specifying a solution for a given partial differential equation. In this section, we introduce the discretization of the boundary condition.

The Dirichlet boundary condition and the Neumann boundary conditions are famous as mathematically well-defined boundary conditions. The former specifies the value of the variable on the boundary, while the latter specifies the gradient of the variable. In numerical simulations, we employ conventional boundary conditions other than well-defined boundary conditions in order to maintain the computation volume finite. First, we consider the Dirichlet and Neumann boundary conditions.

Equations (1.239) through (1.242), which are given on the boundary of  $r = 0$ , are equivalent to either the Dirichlet boundary condition ( $v_r$ ) or the Neumann boundary condition ( $\rho$ ,  $\Omega$ , and  $v_z$ ). For convenience, we assume that the radius can be negative. Then, Equations (1.239) through (1.242) are equivalent to

$$\rho(-r) = \rho(r), \quad (1.244)$$

$$v_r(-r) = -v_r(r), \quad (1.245)$$

$$\Omega(-r) = \Omega(r), \quad (1.246)$$

$$P(-r) = P_r(r). \quad (1.247)$$

Suppose that the grids are placed on

$$r_j = \left(j - \frac{1}{2}\right) \Delta r, \quad (1.248)$$

in the  $r$ -direction. Then, we have the following numerical boundary conditions:

$$\rho(r_0) = \rho(r_1), \quad (1.249)$$

$$v_r(r_0) = -v_r(r_1), \quad (1.250)$$

$$\Omega(r_0) = \Omega(r_1), \quad (1.251)$$

$$P(r_0) = P_r(r_1). \quad (1.252)$$

and

$$\rho(r_{-1}) = \rho(r_2), \quad (1.253)$$

$$v_r(r_{-1}) = -v_r(r_2), \quad (1.254)$$

$$\Omega(r_{-1}) = \Omega(r_2), \quad (1.255)$$

$$P(r_{-1}) = P_r(r_2). \quad (1.256)$$

The density, velocity and pressure in the region of negative  $r$  are obtained from those in the region of positive  $r$  by the symmetry with respect to  $r = 0$ . In other words, there is no boundary at  $r = 0$  in a practical sense. The same procedure works for any symmetric boundary or reflection boundary.

As stated earlier, we want to limit the computation box to be a finite volume around stars or galaxies in numerical simulations. The surface of the computation box condition is artificial and is not limited by any physical laws. However, we want to reduce the artificial effects due to the boundaries as much as possible. For this purpose, several passive boundary conditions have been proposed.

The simplest (least expensive) condition is to place the reflection boundary at a very large distance from the center of the computation box. Although the waves reflected at the boundary are artificial and unnatural, the effects are rather limited if the density is extremely low near the boundary. If the grid spacing is larger near the boundary, then the artificial effects are further reduced because the boundary extends further. We can also add artificial damping in the narrow zone near the boundaries to reduce the reflected waves.

A more sophisticated boundary condition admits outgoing waves and inhibits incoming waves. This boundary condition is called the radiation boundary condition. Variations in density, velocity and pressure can be decomposed into waves, as described in previous sections. Thus, it is possible to exclude only incoming waves. However, it is not possible to perfectly cancel out a specific component of waves because the waves are nonlinear and the components are coupled with each other.

## 1.11 Extension to MHD Equations

The magnetohydrodynamical (MHD) equations can also be solved by the upwind scheme. The magnetohydrodynamical equations take the magnetic force and the induction of the magnetic

field into account and are expressed as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (1.257)$$

$$\rho \left\{ \left( \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} \right) \right\} + \nabla P + \frac{(\nabla \times \mathbf{B}) \times \mathbf{B}}{4\pi} = 0, \quad (1.258)$$

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{v} \times \mathbf{B}) = 0, \quad (1.259)$$

$$\frac{\partial}{\partial t} (\rho E) + \nabla (\rho H) = 0, \quad (1.260)$$

$$E = \frac{|\mathbf{v}|^2}{2} + \frac{1}{\gamma - 1} \frac{P}{\rho} + \frac{|\mathbf{B}|^2}{8\pi\rho}, \quad (1.261)$$

$$H = \frac{|\mathbf{v}|^2}{2} + \frac{\gamma}{\gamma - 1} \frac{P}{\rho} + \frac{|\mathbf{B}|^2}{4\pi\rho}, \quad (1.262)$$

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0. \quad (1.263)$$

The MHD equations are more complicated than the hydrodynamical equations. First of all, they have eight components, and we need to solve  $\mathbf{B}$  as well as  $\rho$ ,  $P$ , and  $\mathbf{v}$ . Second, they are associated with the constraint  $\nabla \cdot \mathbf{B} = 0$  and only seven of these components are independent. Moreover, the MHD waves can be degenerate, i.e., the phase velocity of an MHD wave may coincide with that of another. Even when the waves are degenerate, the eigenvectors should be independent of one other. Fortunately, all of these technical complexities have been resolved. We can solve the one-dimensional MHD equation as follows.

First, we rewrite the one-dimensional MHD equation into conservation form as follows:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0, \quad (1.264)$$

where

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho v_x \\ \rho v_y \\ \rho v_z \\ B_y \\ B_z \\ \rho E \end{pmatrix}, \quad (1.265)$$

$$\mathbf{F} = \begin{pmatrix} \rho v_x^2 + P + \frac{\rho v_x}{B_y^2 + B_z^2 - B_x^2} \\ \rho v_x v_y - \frac{B_x 8\pi}{4\pi} \\ \rho v_x v_z - \frac{B_x 4\pi}{4\pi} \\ v_x B_y - v_y B_x \\ v_x B_z - v_z B_x \\ \rho H v_x - \frac{B_x (B_x v_x + B_y v_y + B_z v_z)}{4\pi} \end{pmatrix}. \quad (1.266)$$

The numerical flux is expressed as

$$\mathbf{F}_{j+1/2}^* = \frac{1}{2} (\mathbf{F}_{j+1} + \mathbf{F}_j) - \sum_k \delta w_k |\lambda_k| \mathbf{r}_k, \quad (1.267)$$

where  $\delta w_k$ ,  $\lambda_k$ , and  $\mathbf{r}_k$  denote the amplitude, the phase velocity, and the eigenvector of the  $k$ -th eigenmode. For later convenience, we use the suffix 0 for the entropy wave, the suffix  $A_{\pm}$  for the Alfvén waves, the suffix  $f_{\pm}$  for the fast waves, and the suffix  $s_{\pm}$  for the slow waves.

The eigenvalues are denoted as follows:

$$\lambda_0 = \bar{v}_x, \quad (1.268)$$

$$\lambda_{A+} = \bar{v}_x + b_x, \quad (1.269)$$

$$\lambda_{A-} = \bar{v}_x - b_x, \quad (1.270)$$

$$\lambda_{f+} = \bar{v}_x + c_f, \quad (1.271)$$

$$\lambda_{f-} = \bar{v}_x - c_f, \quad (1.272)$$

$$\lambda_{s+} = \bar{v}_x + c_s, \quad (1.273)$$

$$\lambda_{s-} = \bar{v}_x - c_s, \quad (1.274)$$

where

$$\bar{\rho} = \sqrt{\rho_{j+1}\rho_j}, \quad (1.275)$$

$$\bar{v}_x = \frac{\sqrt{\rho_{j+1}} v_{x,j+1} + \sqrt{\rho_j} v_{x,j}}{\sqrt{\rho_{j+1}} + \sqrt{\rho_j}}, \quad (1.276)$$

$$\bar{v}_y = \frac{\sqrt{\rho_{j+1}} v_{y,j+1} + \sqrt{\rho_j} v_{y,j}}{\sqrt{\rho_{j+1}} + \sqrt{\rho_j}}, \quad (1.277)$$

$$\bar{v}_z = \frac{\sqrt{\rho_{j+1}} v_{z,j+1} + \sqrt{\rho_j} v_{z,j}}{\sqrt{\rho_{j+1}} + \sqrt{\rho_j}}, \quad (1.278)$$

$$\bar{B}_y = \frac{\sqrt{\rho_{j+1}} B_{y,j} + \sqrt{\rho_j} B_{y,j+1}}{\sqrt{\rho_{j+1}} + \sqrt{\rho_j}}, \quad (1.279)$$

$$\bar{B}_z = \frac{\sqrt{\rho_{j+1}} B_{z,j} + \sqrt{\rho_j} B_{z,j+1}}{\sqrt{\rho_{j+1}} + \sqrt{\rho_j}}, \quad (1.280)$$

$$\bar{H} = \frac{\sqrt{\rho_{j+1}} H_{j+1} + \sqrt{\rho_j} H_j}{\sqrt{\rho_{j+1}} + \sqrt{\rho_j}}, \quad (1.281)$$

$$\bar{P} = \frac{\sqrt{\rho_{j+1}} P_j + \sqrt{\rho_j} P_{j+1}}{\sqrt{\rho_{j+1}} + \sqrt{\rho_j}}, \quad (1.282)$$

The eigenvectors are expressed as follows:

$$\delta b^2 = \frac{\gamma - 1}{\gamma - 2} \frac{(B_{y,j+1} - B_{y,j})^2 + (B_{z,j+1} - B_{z,j})^2}{8\pi (\sqrt{\rho_{j+1}} + \sqrt{\rho_j})^2}, \quad (1.283)$$

$$a^2 = (\gamma - 1) \left( \bar{H} - \frac{\bar{v}_x^2 + \bar{v}_y^2 + \bar{v}_z^2}{2} + \frac{B_x^2 + \bar{B}_y^2 + \bar{B}_z^2}{4\pi\rho} - \delta b^2 \right), \quad (1.284)$$

$$a_*^2 = (\gamma - 1) \left( \bar{H} - \frac{\bar{v}_x^2 + \bar{v}_y^2 + \bar{v}_z^2}{2} - \delta b^2 \right), \quad (1.285)$$

$$= (\gamma - 2) \left( \frac{B_x^2 + \bar{B}_y^2 + \bar{B}_z^2}{4\pi\rho} \right), \quad (1.286)$$

$$b_x = \frac{|B_x|}{\sqrt{4\pi\bar{\rho}}} \quad (1.287)$$

$$c_{f,s}^2 = \frac{a_*^2 \pm \sqrt{a_*^4 - 4a^2 b_x^2}}{2}, \quad (1.288)$$

$$\mathbf{r}_0 = \begin{pmatrix} 1 \\ \bar{v}_x \\ \bar{v}_y \\ \bar{v}_z \\ 0 \\ 0 \\ \frac{(\bar{\mathbf{v}})^2}{2} + \delta b^2 \end{pmatrix}, \quad (1.289)$$

$$\mathbf{r}_{A+} = \begin{pmatrix} 0 \\ 0 \\ -\beta_z \operatorname{sgn}(B_x) \\ \beta_y \operatorname{sgn}(B_x) \\ \beta_z \sqrt{4\pi/\bar{\rho}} \\ -\beta_y \sqrt{4\pi/\bar{\rho}} \\ -(\beta_z \bar{v}_y - \beta_y \bar{v}_z) \operatorname{sgn}(B_x) \end{pmatrix}, \quad (1.290)$$

$$\mathbf{r}_{A-} = \begin{pmatrix} 0 \\ 0 \\ \beta_z \operatorname{sgn}(B_x) \\ -\beta_y \operatorname{sgn}(B_x) \\ \beta_z \sqrt{4\pi/\bar{\rho}} \\ -\beta_y \sqrt{4\pi/\bar{\rho}} \\ (\beta_z \bar{v}_y - \beta_y \bar{v}_z) \operatorname{sgn}(B_x) \end{pmatrix}, \quad (1.291)$$

and

$$\mathbf{r}_{f+} = \begin{pmatrix} \alpha_f \\ \alpha_f (\bar{v}_x + c_f) \\ \alpha_f \bar{v}_y - \alpha_s \beta_y b_x \operatorname{sgn}(B_x) \\ \alpha_f \bar{v}_z - \alpha_s \beta_z b_x \operatorname{sgn}(B_x) \\ \alpha_s \beta_y c_f \sqrt{4\pi/\bar{\rho}} \\ \alpha_s \beta_z c_f \sqrt{4\pi/\bar{\rho}} \\ \alpha_f \left\{ \frac{(\bar{\mathbf{v}})^2}{2} + \delta b^2 + c_f \bar{v}_x + \frac{c_f^2}{\gamma - 1} \right. \\ \left. + \frac{\gamma - 2}{\gamma - 1} (c_f^2 - a^2) \right\} - \alpha_s b_x \operatorname{sgn}(B_x) (\beta_y \bar{v}_y + \beta_z \bar{v}_z) \end{pmatrix}, \quad (1.292)$$

$$\mathbf{r}_{f-} = \begin{pmatrix} \alpha_f \\ \alpha_f (\bar{v}_x - c_f) \\ \alpha_f \bar{v}_y + \alpha_s \beta_y b_x \operatorname{sgn}(B_x) \\ \alpha_f \bar{v}_z + \alpha_s \beta_z b_x \operatorname{sgn}(B_x) \\ \alpha_s \beta_y c_f \sqrt{4\pi/\bar{\rho}} \\ \alpha_s \beta_z c_f \sqrt{4\pi/\bar{\rho}} \\ \alpha_f \left\{ \frac{(\bar{\mathbf{v}})^2}{2} + \delta b^2 - c_f \bar{v}_x + \frac{c_f^2}{\gamma_t - 1} \right. \\ \left. + \frac{\gamma - 2}{\gamma - 1} (c_f^2 - a^2) \right\} + \alpha_s b_x \operatorname{sgn}(B_x) (\beta_y \bar{v}_y + \beta_z \bar{v}_z) \end{pmatrix}, \quad (1.293)$$

$$\mathbf{r}_{s+} = \begin{pmatrix} \alpha_s \\ \alpha_s (\bar{v}_x + c_s) \\ \alpha_s \bar{v}_y + \alpha_f \beta_y a \operatorname{sgn}(B_x) \\ \alpha_s \bar{v}_z + \alpha_f \beta_z a \operatorname{sgn}(B_x) \\ - \frac{\alpha_f \beta_y a^2 \sqrt{4\pi}}{c_f \sqrt{\bar{\rho}}} \\ - \frac{\alpha_f \beta_z a^2 \sqrt{4\pi}}{c_f \sqrt{\bar{\rho}}} \\ \alpha_s \beta_z c_f \sqrt{4\pi/\bar{\rho}} \\ \alpha_s \left\{ \frac{(\bar{\mathbf{v}})^2}{2} + \delta b^2 + c_s \bar{v}_x + \frac{c_s^2}{\gamma - 1} \right. \\ \left. + \frac{\gamma - 2}{\gamma - 1} (c_s^2 - a^2) \right\} + \alpha_f a \operatorname{sgn}(B_x) (\beta_y \bar{v}_y + \beta_z \bar{v}_z) \end{pmatrix}, \quad (1.294)$$

$$\mathbf{r}_{s-} = \begin{pmatrix} \alpha_s \\ \alpha_s (\bar{v}_x - c_s) \\ \alpha_s \bar{v}_y - \alpha_f \beta_y a \operatorname{sgn}(B_x) \\ \alpha_s \bar{v}_z - \alpha_f \beta_z a \operatorname{sgn}(B_x) \\ - \frac{\alpha_f \beta_y a^2 \sqrt{4\pi}}{c_f \sqrt{\bar{\rho}}} \\ - \frac{\alpha_f \beta_z a^2 \sqrt{4\pi}}{c_f \sqrt{\bar{\rho}}} \\ \alpha_s \beta_z c_f \sqrt{4\pi/\bar{\rho}} \\ \alpha_s \left\{ \frac{(\bar{\mathbf{v}})^2}{2} + \delta b^2 - c_s \bar{v}_x + \frac{c_s^2}{\gamma - 1} \right. \\ \left. + \frac{\gamma - 2}{\gamma - 1} (c_s^2 - a^2) \right\} - \alpha_f a \operatorname{sgn}(B_x) (\beta_y \bar{v}_y + \beta_z \bar{v}_z) \end{pmatrix}, \quad (1.295)$$

where

$$\beta_y = \frac{\bar{B}_y}{\sqrt{\bar{B}_y^2 + \bar{B}_z^2}}, \quad (1.296)$$

$$\beta_z = \frac{\bar{B}_z}{\sqrt{\bar{B}_y^2 + \bar{B}_z^2}}, \quad (1.297)$$

$$\beta_y^2 + \beta_z^2 = 1, \quad (1.298)$$

$$\alpha_f^2 + \frac{b_x^2}{c_f^2} \alpha_s^2 = 1. \quad (1.299)$$

The wave amplitudes are given as follows:

$$\delta w_{A+} = \frac{1}{2} \left[ -\bar{\rho}(\beta_z \Delta v_y - \beta_y \Delta v_z) \text{sgn}(B_x) + \sqrt{\frac{\bar{\rho}}{4\pi}} (\beta_z \Delta B_y - \beta_y \Delta B_z) \right], \quad (1.300)$$

$$\delta w_{A-} = \frac{1}{2} \left[ \bar{\rho}(\beta_z \Delta v_y - \beta_y \Delta v_z) \text{sgn}(B_x) + \sqrt{\frac{\bar{\rho}}{4\pi}} (\beta_z \Delta B_y - \beta_y \Delta B_z) \right], \quad (1.301)$$

$$\delta w_{f+} + \delta w_{f-} = \frac{\alpha_f}{c_f^2} \left( \Delta P + \frac{\tilde{B}_y \Delta B_y + \tilde{B}_z \Delta B_z}{4\pi} \right) \quad (1.302)$$

$$+ \left\{ \frac{\alpha_s}{a^2 c_f} [(\gamma - 1)c_s^2 - (\gamma - 2a^2)] \sqrt{4\pi \bar{\rho}} \right. \quad (1.303)$$

$$\left. + (\gamma - 2) \sqrt{\bar{B}_y^2 + \bar{B}_z^2} \frac{\alpha_f}{c_f^2} \right\} \frac{\beta_y \Delta B_y + \beta_z \Delta B_z}{4\pi}, \quad (1.304)$$

$$\delta w_1 - \delta w_7 = \frac{\alpha_f}{c_f} \bar{\rho} \Delta v_x - \frac{\alpha_s c_s}{c_f a} \text{sgn}(B_x) \bar{\rho} (\beta_y \Delta v_y + \beta_z \Delta v_z), \quad (1.305)$$

$$\delta w_{s+} + \delta w_{s-} = \frac{\alpha_s}{a^2} \left( \Delta P + \frac{\tilde{B}_y \Delta B_y + \tilde{B}_z \Delta B_z}{4\pi} \right) \quad (1.306)$$

$$+ \left\{ \alpha_f \left[ \frac{\gamma - 2}{c_f} - (\gamma - 1) \frac{c_f}{a^2} \right] \sqrt{4\pi \bar{\rho}} \right. \quad (1.307)$$

$$\left. + (\gamma - 2) \sqrt{\bar{B}_y^2 + \bar{B}_z^2} \frac{\alpha_s}{a^2} \right\} \frac{\beta_y \Delta B_y + \beta_z \Delta B_z}{4\pi}, \quad (1.308)$$

$$\delta w_{s+} - \delta w_{s-} = \frac{\alpha_s b_x}{c_f a} \bar{\rho} \Delta v_x + \frac{\alpha_f}{a} \text{sgn}(B_x) \bar{\rho} (\beta_y \Delta v_y + \beta_z \Delta v_z), \quad (1.309)$$

$$\delta w_0 = \Delta \rho - \alpha_f (\delta w_1 + \delta w_7) - \alpha_s (\delta w_3 + \delta w_5) \quad (1.310)$$

These eigenvalues and eigenvectors satisfy the U property, as follows:

$$\mathbf{U}_{j+1} - \mathbf{U}_j = \sum_k \delta w_k \mathbf{r}_k, \quad (1.311)$$

$$\mathbf{F}_{j+1} - \mathbf{F}_j = \sum_k \delta w_k \lambda_k \mathbf{r}_k. \quad (1.312)$$

This type of numerical flux for the MHD equation was first given by Brio and Wu (1988). Parameters such as  $\alpha_f$  were introduced by Ryu and Jones (1995) to handle the degeneracy. The numerical flux was modified to satisfy the U property given in Cargo and Gallice (1997).

## 1.12 Some Other Numerical Schemes

Although in this chapter we described only Roe's method in detail, there are a number of other popular schemes. This section briefly describes these schemes, with emphasis on their advantages.

Before introducing specific methods, we shall discuss the qualities of a good numerical scheme. A perfect scheme can solve any problem stably with a high accuracy in a short time. Of course, perfect schemes do not exist. Perhaps perfect schemes have not yet been discovered because of the existence of contradicting requirements. For example, high-accuracy schemes tend to sacrifice short computation time. Robustness, which guarantees the ability to solve any problem, also requires greater computation cost in order to handle extreme conditions. We are often forced to compromise between two competing factors. Thus, a particular method recommended by a textbook or a paper should be regarded as being a good method for solving the particular problem in which the authors are interested and should not be considered to be good for all problems. As computer performance increases, the computation time for a specific problem is reduced. In the near future, some methods that are currently considered to be impractical may become practical. At present, however, a good method is only good for a particular type of problem.

Some good methods pursue higher-order accuracy, i.e., third-order or fourth-order accuracy. The Piecewise Parabolic Method (PPM) approximates the density, velocity, and temperature by piecewise parabolic functions in a cell and achieves third-order accuracy in space. Other methods employ the Runge-Kutta method, which is often used for solving ordinary differential equations, to achieve fourth-order accuracy in time. These schemes reduce truncation errors in the region where the density and velocity change smoothly.

Another method has been developed in order to capture the discontinuity more sharply. The Godunov method gives the exact solution for the initial density and velocity and regards the average over a cell as the solution. This method succeeds in treating a strong shock wave, in front of which the density and velocity change dramatically. The Cubic InterPolation (CIP) method proposed by Yabe and his collaborators is designed to capture the contact discontinuity more sharply. The CIP method succeeds in solving the problem of a solid body floating in liquid.

The reduction of computation time is another consideration when developing a new scheme. The HLL scheme reduces the time required to compute numerical flux by taking only the maximum and minimum phase speeds into account. This simplification reduces the computation time appreciably without reducing the quality of the solution greatly. If the computation time per cell is reduced, we can increase the number of cells and improve the resolution when the total computation time is fixed.

Generally, computation load is reduced when accuracy is reduced. Thus, it is not recommended to pursue excessive quality for unimportant regions. It is important to evaluate the specific requirements of the evaluation method before starting numerical simulations. Although robust code, which can manage any problem, provides good results, it consumes a great deal of computation time.

A number of computation codes for numerical simulations of astrophysical problems are available online. ZEUS, which was developed by J. Stone, is one such code. Another is CANS, which was developed by R. Matsumoto and his collaborators. The latter code is used for this winter school and can be downloaded from <http://www.astro.phys.s.chiba-u.ac.jp/netlab/pub/>. Both Japanese and English guide books are available for the CANS. The Japanese guide book was written by H. Hanayama, and the English guide book was translated from the Japanese by

S. Miyaji (Chapter 3).

For further study, we recommend “Riemann Solvers And Numerical Methods for Fluid Dynamics: A Practical Introduction” by E. Toro and “Numerical Computation of Internal and External Flows” by C. Hirsch.