Chapter 6

Transport and Fate

ABSTRACT
In this Chapter, we augment the diffusion equation of the preceding chapter to include the effects of advection (transport by the moving fluid) and fate (diffusion, plus possible source, and decay along the way). The numerical section begins with the design of schemes for advection in a fixed (Eulerian) framework and then extends those to include the discretization of diffusion and source/decay terms. Most of the developments are presented in one dimension before generalization to multiple dimensions.

6.1 COMBINATION OF ADVECTION AND DIFFUSION

When considering the heat (3.23), salt (3.14), humidity (3.15), or density (4.8) equations of Geophysical Fluid Dynamics, we note that they each include three types of terms. The first, a time derivative, tells how the variable is changing over time. The second is a group of three terms with velocity components and spatial derivatives, sometimes hidden in the material derivative $\mathrm{d}/\mathrm{d}t$. They represent the transport of the substance with the flow and are collectively called *advection*. Finally, the last group of terms, on the right-hand side, includes an assortment of diffusivities and second-order spatial derivatives. In the light of Chapter 5, we identify these with *diffusion*. They represent the spreading of the substance along and across the flow. Using a generic formulation, we are brought to study an equation of the type

$$\frac{\partial c}{\partial t}+u\frac{\partial c}{\partial x}+v\frac{\partial c}{\partial y}+w\frac{\partial c}{\partial z}=\frac{\partial}{\partial x}\left(\mathcal{A}\frac{\partial c}{\partial x}\right)+\frac{\partial}{\partial y}\left(\mathcal{A}\frac{\partial c}{\partial y}\right)+\frac{\partial}{\partial z}\left(\kappa\frac{\partial c}{\partial z}\right), \qquad (6.1)$$

where the variable c may stand for any of the aforementioned variables or represent a substance imbedded in the fluid, such as a pollutant in the atmosphere or in the sea. Note the anisotropy between the horizontal and vertical directions ($\mathcal{A}$ generally $\gg\kappa$).

The examples in the following figures illustrate the combined effects of advection and diffusion. Figure 6.1 shows the fate of the Rhône River waters as

Introduction to Geophysical Fluid Dynamics, Volume 101, ISSN 0074-6142, DOI: 10.1016/B978-0-12-088759-0.00006-7
Copyright © 2011, Elsevier Inc. All rights reserved.

FIGURE 6.1 Rhône River plume discharging in the Gulf of Lions (circa 43°N) and carrying sediments into the Mediterranean Sea. This satellite picture was taken on 26 February 1999. (*Satellite image provided by the SeaWiFS Project, NASA/Goddard Space Flight Center*). A color version can be found online at http://booksite.academicpress.com/9780120887590/

they enter the Mediterranean Sea. Advection pulls the plume offshore, whereas diffusion dilutes it. Figure 6.2 is a remarkable satellite picture, showing wind advection of sand from the Sahara Desert westward from Africa to Cape Verde (white band across the lower part of the picture) at the same time as, and independently from, marine transport of suspended matter southwestward from the Cape Verde islands (von Kármán vortices in left of middle of the picture). Although sand is being blown quickly and without much diffusion in the air, the sediments follow convoluted paths in the water, pointing to a disparity between the relative effects of advection and diffusion in the atmosphere and ocean.

Often, the substance being carried by the fluid is not simply moved and diffused by the flow. It may also be created or lost along the way. Such is the case of particle matter, which tends to settle at the bottom. Chemical species can be produced by reaction between parent chemicals and be lost by participating in other reactions. An example of this is sulfuric acid (H_2SO_4) in the atmosphere: It is produced by reaction of sulfur dioxide (SO_2) from combustion and lost by precipitation (acid rain or snow). Tritium, a naturally radioactive form of hydrogen enters the ocean by contact with air at the surface and disintegrates along its oceanic journey to become Helium. Dissolved oxygen in the sea is consumed by biological activity and is replenished by contact with air at the surface.

FIGURE 6.2 Sahara dust blown by the wind from the African continent over the ocean toward Cape Verde islands (15–17°N), while suspended matter in the water is being transported southwestward by a series of von Kármán vortices in the wake of the islands. Note in passing how these vortices in the water relate to the overlying cloud patterns. (*Jacques Descloitres, MODIS Land Science Team*)

To incorporate these processes, we augment the advection–diffusion equation (6.1) by adding source and sink terms in the right-hand side:

$$\frac{\partial c}{\partial t}+u\frac{\partial c}{\partial x}+v\frac{\partial c}{\partial y}+w\frac{\partial c}{\partial z} = \frac{\partial}{\partial x}\left(\mathcal{A}\frac{\partial c}{\partial x}\right)+\frac{\partial}{\partial y}\left(\mathcal{A}\frac{\partial c}{\partial y}\right)+\frac{\partial}{\partial z}\left(\kappa\frac{\partial c}{\partial z}\right)+S-Kc, \tag{6.2}$$

where the term S stands for the source, the formulation of which depends on the particular process of formation of the substance, and K is a coefficient of decay, which affects how quickly (large K) or slowly (small K) the substance is being lost.

At one dimension, say in the x-direction, and with constant diffusivity $\mathcal{A}$, the equation reduces to:

$$\frac{\partial c}{\partial t}+u\frac{\partial c}{\partial x}=\mathcal{A}\frac{\partial^2 c}{\partial x^2}+S-Kc. \tag{6.3}$$

Several properties of the advection–diffusion equation are worth noting because they bear on the numerical procedures that follow: In the absence of source and sink, the total amount of the substance is conserved, and, in the further absence of diffusion, the variance of the concentration distribution, too, is conserved over time.

When we integrate Eq. (6.2) over the domain volume $\mathcal{V}$, we can readily integrate the diffusion terms and, if the flux is zero at all boundaries, these vanish,

and we obtain the following:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{V}} c\,\mathrm{d}\mathcal{V} = -\int_{\mathcal{V}}\left(u\frac{\partial c}{\partial x}+v\frac{\partial c}{\partial y}+w\frac{\partial c}{\partial z}\right)\mathrm{d}\mathcal{V}+\int_{\mathcal{V}} S\mathrm{d}\mathcal{V}-\int_{\mathcal{V}} Kc\,\mathrm{d}\mathcal{V}.$$

After an integration by parts, the first set of terms on the right can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{V}} c\,\mathrm{d}\mathcal{V} = +\int_{\mathcal{V}} c\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)\mathrm{d}\mathcal{V}+\int_{\mathcal{V}} S\mathrm{d}\mathcal{V}-\int_{\mathcal{V}} Kc\,\mathrm{d}\mathcal{V},$$

as long as there is no flux or no advection at all boundaries. Invoking the continuity equation (4.21d) reduces the first term on the right to zero, and we obtain simply:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{V}} c\,\mathrm{d}\mathcal{V} = \int_{\mathcal{V}} S\mathrm{d}\mathcal{V}-\int_{\mathcal{V}} Kc\,\mathrm{d}\mathcal{V}. \tag{6.4}$$

Since the concentration c represents the amount of the substance per volume, its integral over the volume is its total amount. Equation (6.4) simply states that this amount remains constant over time when there is no source ($S=0$) or sink ($K=0$). Put another way, the substance is moved around but conserved.

Now, if we multiply Eq. (6.2) by c and then integrate over the domain, we can integrate the diffusion terms by parts and, if the flux is again zero at all boundaries, we have the following:

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{V}} c^2\mathrm{d}\mathcal{V} = &-\int_{\mathcal{V}}\left[\mathcal{A}\left(\frac{\partial c}{\partial x}\right)^2+\mathcal{A}\left(\frac{\partial c}{\partial y}\right)^2+\kappa\left(\frac{\partial c}{\partial z}\right)^2\right]\mathrm{d}\mathcal{V}\\ &+\int_{\mathcal{V}} Sc\,\mathrm{d}\mathcal{V}-\int_{\mathcal{V}} Kc^2\,\mathrm{d}\mathcal{V}.\end{aligned} \tag{6.5}$$

With no diffusion, source, or sink, the right-hand side is zero, and *variance* is conserved in time. Diffusion and decay tend to reduce variance, whereas a (positive) source tends to increase it.

This conservation property can be extended, still in the absence of diffusion, source, and sink, to any power c^p of c, by multiplying the equation by c^{p-1} before integration. The conservation property even holds for any function $F(c)$. It goes without saying that numerical methods cannot conserve all these quantities, but it is highly desirable that they conserve at least the first two (total amount and variance).

There is one more property worth mentioning, which we will state without demonstration but justify in a few words. Because diffusion acts to smooth the distribution of c, it removes the substance from the areas of higher concentration and brings it into regions of lower concentration. Hence, due to diffusion alone, the maximum of c can only diminish and its minimum can only increase. Advection, by contrast, redistributes existing values, thus not changing either

minimum and maximum. In the absence of source and sink, therefore, no future value of c can fall outside the initial range of values. This is called the *max-min property*. Exceptions are the presence of a source or sink, and the import through one of the boundaries of a concentration outside the initial range.

We call a numerical scheme that maintains the max-min property a *monotonic scheme* or *monotonicity preserving* scheme[1]. Alternatively, the property of *boundedness* is often used to describe a physical solution that does not generate new extrema. If c is initially positive everywhere, as it should be, the absence of new extrema keeps the variable positive at all future times, another property called *positiveness*. A monotonic scheme is thus positive but the reverse is not necessarily true.

6.2 RELATIVE IMPORTANCE OF ADVECTION: THE PECLET NUMBER

Since the preceding equations combine the effects of both advection and diffusion, it is important to compare the relative importance of one with the other. In a specific situation, could it be that one dominates over the other or that both impact concentration values to the same extent? To answer this question, we turn once again to scales. Introducing a length scale L, velocity scale U, diffusivity scale D, and a scale Δc to measure concentration differences, we note that advection scales like $U\Delta c/L$ and diffusion like $D\Delta c/L^2$. We can then compare the two processes by forming the ratio of their scales:

$$\frac{\text{advection}}{\text{diffusion}} = \frac{U\Delta c/L}{D\Delta c/L^2} = \frac{UL}{D}.$$

This ratio is by construction dimensionless. It bears the name of Peclet number[2] and is denoted by Pe:

$$Pe = \frac{UL}{D}, \tag{6.6}$$

where the scales U, L, and D may stand for the scales of either horizontal (u, v, x, y, and $\mathcal{A}$) or vertical (w, z, and κ) variables but not a mix of them. The Peclet number leads to an immediate criterion, as follows.

If $Pe \ll 1$ (in practice, if $Pe < 0.1$), the advection term is significantly smaller than the diffusion term. Physically, diffusion dominates, and advection is negligible. Diffusive spreading occurs almost symmetrically despite the directional bias of the weak flow. If we wish to simplify the problem, we may drop the

[1] Some computational fluid dynamicists do make a difference between these two labels, but this minor point lies beyond our present text.

[2] In honor of Jean Claude Eugène Péclet (1793–1857), French physicist who wrote a treatise on heat transfer.

advection term [$u\partial c/\partial x$ in (6.3)], as if u were zero. The relative error committed in the solution is expected to be on the order of the Peclet number, and the smaller Pe, the smaller the error. The methods developed in the preceding chapter were based on such simplification and thus apply whenever $Pe \ll 1$.

If $Pe \gg 1$ (in practice, if $Pe > 10$), it is the reverse: the advection term is now significantly larger than the diffusion term. Physically, advection dominates, and diffusion is negligible. Spreading is weak, and the patch of substance is mostly moved along, and possibly distorted by, the flow. If we wish to simplify the problem, we may drop the diffusion term [$\mathcal{A}\partial^2 c/\partial x^2$ in (6.3)], as if $\mathcal{A}$ were zero. The relative error committed in the solution by doing so is expected to be on the order of the inverse of the Peclet number ($1/Pe$), and the larger Pe, the smaller the error.

6.3 HIGHLY ADVECTIVE SITUATIONS

When a system is highly advective in one direction (high Pe number based on scales U, L, and D corresponding to that direction), diffusion is negligible *in that same direction*. This is not to say that it is also negligible in the other directions. For example, high advection in the horizontal does not preclude vertical diffusion, as this is often the case in the lower atmosphere. In such a case, the governing equation is

$$\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} + v\frac{\partial c}{\partial y} = \frac{\partial}{\partial z}\left(\kappa\frac{\partial c}{\partial z}\right) + S - Kc. \tag{6.7}$$

Because the diffusion terms are of higher-order (second derivatives) than those of advection (first derivatives), the neglect of a diffusion term reduces the order of the equation and, therefore, also reduces the need of boundary conditions by one in the respective direction. The boundary condition at the downstream end of the domain must be dropped: The concentration and flux there are whatever the flow brings to that point. A problem occurs when the situation is highly advective, but the small diffusion term is not dropped. In that case, because the order of the equation is not reduced, a boundary condition is enforced at the downstream end, and a locally high gradient of concentration may occur.

To see this, consider the one-dimensional, steady situation with no source and sink, with constant velocity and diffusivity in the x-direction. The equation is

$$u\frac{\mathrm{d}c}{\mathrm{d}x} = \mathcal{A}\frac{\mathrm{d}^2 c}{\mathrm{d}x^2}, \tag{6.8}$$

and its most general solution is

$$c(x) = C_0 + C_1 \mathrm{e}^{ux/\mathcal{A}}. \tag{6.9}$$

For $u > 0$, the downstream end is to the right of the domain, and the solution increases exponentially towards the right boundary. Rather, it could be said that the solution decays away from this boundary as x decreases away from it. In other words, a boundary layer exists at the downstream end. The e-folding length of this boundary layer is $\mathcal{A}/u$, and it can be very short in a highly advective situation (large u and small $\mathcal{A}$). Put another way, the Peclet number is the ratio of the domain length to this boundary-layer thickness, and the larger the Peclet number, the smaller the fraction of the domain occupied by the boundary layer. Why do we need to worry about this? Because in a numerical model it may happen that the boundary-layer thickness falls below the grid size. It is therefore important to check the Peclet number in relation to the spatial resolution. Should the ratio of the grid size to the length scale of the system be comparable with, or larger than, the inverse of the Peclet number, diffusion must be neglected in that direction, or, if it must be retained for some reason, special care must be taken at the downstream boundary.

6.4 CENTERED AND UPWIND ADVECTION SCHEMES

In GFD, advection is generally dominant compared with diffusion, and we therefore begin with the case of pure advection of a tracer concentration $c(x, t)$ along the x-direction. The aim is to solve numerically the following equation:

$$\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} = 0. \tag{6.10}$$

For simplicity, we further take the velocity u as constant and positive so that advection carries c in the positive $x-$direction. The exact solution of this equation is

$$c(x, t) = c_0(x - ut), \tag{6.11}$$

where $c_0(x)$ is the initial concentration distribution (at $t = 0$).

A spatial integration from $x_{i-1/2}$ to $x_{i+1/2}$ across a grid cell (Fig. 6.3) leads to the following budget

$$\frac{\mathrm{d}\bar{c}_i}{\mathrm{d}t} + \frac{q_{i+1/2} - q_{i-1/2}}{\Delta x} = 0, \quad q_{i-1/2} = uc|_{i-1/2}, \tag{6.12}$$

which is the basis for the finite-volume technique, as in (3.33). To close the system, we need to relate the local fluxes q to the cell-average concentrations $\bar{c}$. To do so, we must introduce an approximation, because we do not know the actual value of c at the interfaces between cells, but only the average value in the cell on each side of it. It appears reasonable to use the following, consistent, numerical interpolation for the flux:

$$\tilde{q}_{i-1/2} = u\left(\frac{\bar{c}_i + \bar{c}_{i-1}}{2}\right), \tag{6.13}$$

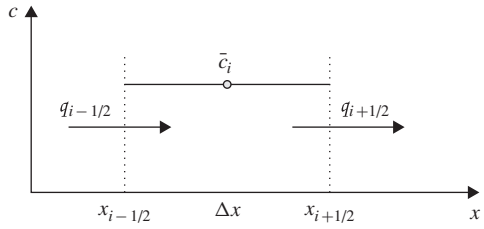


FIGURE 6.3 One-dimensional finite-volume approach with fluxes at the interfaces between grid cells for a straightforward budget calculation.

which is tantamount to assume that the local tracer concentration at the interface is equal to the mean of the surrounding cell averages. Before proceeding with time discretization, we can show that this centered approximation conserves not only the total amount of substance but also its variance, $\sum_i \bar{c}_i$ and $\sum_i (\bar{c}_i)^2$. Substitution of the flux approximation into (6.12) leads to the following semidiscrete equation for cell-averaged concentrations:

$$\frac{\mathrm{d}\bar{c}_i}{\mathrm{d}t} = -u\frac{\bar{c}_{i+1} - \bar{c}_{i-1}}{2\Delta x}. \tag{6.14}$$

Sum over index i leads to cancellation of terms by pairs on the right, leaving only the first and last $\bar{c}$ values. Then, multiplication of the same equation by c_i followed by the sum over the domain provides the time-evolution equation of the discretized variance:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i (\bar{c}_i)^2\right) = -\frac{u}{\Delta x}\sum_i \bar{c}_i\bar{c}_{i+1} + \frac{u}{\Delta x}\sum_i \bar{c}_i\bar{c}_{i-1},$$

where the sum covers all grid cells. By shifting the index of the last term from i to $i+1$, we note again cancellation of terms by pairs, leaving only contributions from the first and last grid points. Thus, except for possible contributions from the boundaries, the numerical scheme conserves both total amount and variance as the original equation does.

However, the conservation of global variance only holds for the semidiscrete equations. When time discretization is introduced as it must eventually be, conservation properties are often lost. In the literature, it is not always clearly stated whether conservation properties hold for the semidiscrete or fully discretized equations. The distinction, however, is important: The centered-space differencing conserves the variance of the semidiscrete solution, but a simple explicit time discretization renders the scheme unconditionally unstable and certainly does not conserve the variance. On the contrary, the latter quantity rapidly increases. Only a scheme that is both stable and consistent leads in the limit of vanishing time step to a solution that satisfies (6.12) and ensures conservation of the variance.

We might wonder why place emphasis on such conservation properties of semidiscrete equations, since by the time the algorithm is keyed into the

computer it will always rely on fully discretized numerical approximations in both space and time. A reason to look at semidiscrete conservation properties is that some special time discretizations maintain the property in the fully discretized case. We now show that the trapezoidal time discretization conserves variance.

Consider the more general linear equation

$$\frac{\mathrm{d}\tilde{c}_i}{\mathrm{d}t} + \mathcal{L}(\tilde{c}_i) = 0, \tag{6.15}$$

where $\mathcal{L}$ stands for a linear discretization operator applied to the discrete field $\tilde{c}_i$. For our centered advection, the operator is $\mathcal{L}(\tilde{c}_i) = u(\tilde{c}_{i+1} - \tilde{c}_{i-1})/(2\Delta x)$. Suppose that the operator is designed to satisfy conservation of variance, which demands that at any moment t and for any discrete field $\tilde{c}_i$ the following relation holds

$$\sum_i \tilde{c}_i \mathcal{L}(\tilde{c}_i) = 0, \tag{6.16}$$

because only then does $\sum_i \tilde{c}_i \, \mathrm{d}\tilde{c}_i/\mathrm{d}t$ vanish according to (6.15) and (6.16). The trapezoidal time discretization applied to (6.15) is

$$\frac{\tilde{c}_i^{n+1} - \tilde{c}_i^n}{\Delta t} = -\frac{\mathcal{L}\left(\tilde{c}_i^{n+1}\right) + \mathcal{L}\left(\tilde{c}_i^n\right)}{2} = -\frac{1}{2}\mathcal{L}\left(\tilde{c}_i^{n+1} + \tilde{c}_i^n\right), \tag{6.17}$$

where the last equality follows from the linearity of operator $\mathcal{L}$. Multiplying this equation by $\left(\tilde{c}_i^{n+1} + \tilde{c}_i^n\right)$ and summing over the domain then yields

$$\sum_i \frac{\left(\tilde{c}_i^{n+1}\right)^2 - \left(\tilde{c}_i^n\right)^2}{\Delta t} = -\frac{1}{2}\sum_i \left(\tilde{c}_i^{n+1} + \tilde{c}_i^n\right)\mathcal{L}\left(\tilde{c}_i^{n+1} + \tilde{c}_i^n\right). \tag{6.18}$$

The term on the right is zero by virtue of (6.16). Therefore, any spatial discretization scheme that conserves variance continues to conserve variance if the trapezoidal scheme is used for the time discretization. As an additional benefit, the resulting scheme is also unconditionally stable. This does not mean, however, that the scheme is satisfactory, as Numerical Exercise 6.9 shows for the advection of the top-hat signal. Furthermore, there is a price to pay for stability because a system of simultaneous linear equations needs to be solved at each time step if the operator $\mathcal{L}$ uses several neighbors of the local grid point i.

To avoid solving simultaneous equations, alternative methods must be sought for time differencing. Let us explore the leapfrog scheme on the finite-volume approach. Time integration of (6.12) from t^{n-1} to t^{n+1} yields

$$\bar{c}_i^{n+1} = \bar{c}_i^{n-1} - 2\frac{\Delta t}{\Delta x}\left(\hat{q}_{i+1/2} - \hat{q}_{i-1/2}\right), \tag{6.19}$$

where $\hat{q}_{i-1/2}$ is the time-average advective flux uc across the cell interfaces between cells $i-1$ and i during the time interval from t^{n-1} to t^{n+1}. Using centered operators, this flux can be estimated as

$$\hat{q}_{i-1/2} = \frac{1}{2\Delta t}\int_{t^{n-1}}^{t^{n+1}} uc|_{i-1/2}\,\mathrm{d}t \quad \rightarrow \quad \tilde{q}_{i-1/2} = u\left(\frac{\tilde{c}_i^n + \tilde{c}_{i-1}^n}{2}\right), \tag{6.20}$$

so that the ultimate scheme is as follows:

$$\tilde{c}_i^{n+1} = \tilde{c}_i^{n-1} - \mathrm{C}\left(\tilde{c}_{i+1}^n - \tilde{c}_{i-1}^n\right), \tag{6.21}$$

where the coefficient C is defined as

$$\mathrm{C} = \frac{u\Delta t}{\Delta x}. \tag{6.22}$$

The same discretization could have been obtained by straightforward finite differencing of (6.10).

The parameter C is a dimensionless ratio central to the numerical discretization of advective problems called the *Courant number* or *CFL parameter* (Courant, Friedrichs & Lewy, 1928). It compares the displacement $u\Delta t$ made by the fluid during one time step to the grid size Δx. More generally, the Courant number for a process involving a propagation speed (such as a wave speed) is defined as the ratio of the distance of propagation during one time step to the grid spacing.

To use (6.21), two initial conditions are needed, one of which is physical and the other artificial. The latter must be consistent with the former. As usual, an explicit Euler step may be used to start from the single initial condition $\tilde{c}_i^0$:

$$\tilde{c}_i^1 = \tilde{c}_i^0 - \frac{\mathrm{C}}{2}\left(\tilde{c}_{i+1}^0 - \tilde{c}_{i-1}^0\right). \tag{6.23}$$

In considering boundary conditions, we first observe that the exact solution of (6.10) obeys the simple law

$$c(x-ut) = \text{constant}. \tag{6.24}$$

By virtue of this property, a specified value of c somewhere along the line $x-ut=a$, called a *characteristic*, determines the value of c everywhere along that line. It is then easily seen (Fig. 6.4) that, in order to obtain a uniquely defined solution within the domain, a boundary condition must be provided at the upstream boundary, but no boundary condition is required at the outflow boundary. The centered space differencing, however, needs a value of $\tilde{c}$ given at each boundary. When discussing artificial boundary conditions (Section 4.7), we argued that these are acceptable as long as they are consistent with the mathematically correct boundary condition. But then, what requirement should the artificial boundary condition at the outflow obey with, since there is no physical

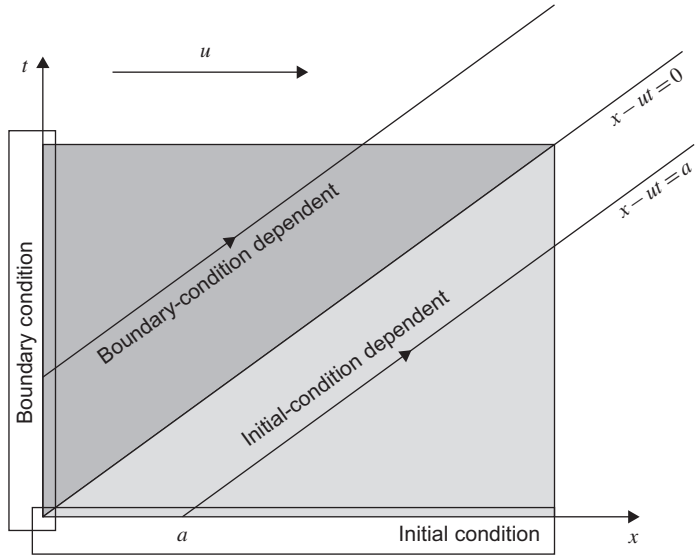


FIGURE 6.4 The characteristic line $x-ut=a$ propagates information from the initial condition or boundary condition into the domain. If the boundary is located at $x=0$, and the initial condition given at $t=0$, the line $x=ut$ divides the space-time frame into two distinct regions: For $x \leq ut$ the boundary condition defines the solution, whereas for $x \geq ut$ the initial condition defines the solution.

boundary condition for it to be consistent with? In practice, a one-sided space differencing is used at the outflow for the last calculation point $i=m$, so that its value is consistent with the local evolution equation:

$$\tilde{c}_m^{n+1} = \tilde{c}_m^n - \mathrm{C}\left(\tilde{c}_m^n - \tilde{c}_{m-1}^n\right). \tag{6.25}$$

This provides the necessary value at the last grid cell.

For the inflow condition, the physical boundary condition is imposed, and algorithm (6.21) can be used starting from $n=1$ and marching in time over all points $i=2, ..., m-1$. Numerically, we thus have sufficient information to calculate a solution that will be second-order accurate in both space and time, except near the initial condition and at the outflow boundary. In order to avoid any bad surprise when implementing the method, a stability analysis is advised.

For convenience, we use the von Neumann method written in Fourier-mode formalism (5.31)

$$\tilde{c}_i^n = A\mathrm{e}^{\mathrm{i}(k_x i\Delta x - \tilde{\omega} n\Delta t)}, \tag{6.26}$$

where the frequency $\tilde{\omega}$ may be complex. Substitution in the difference equation (6.21) provides the numerical dispersion relation

$$\sin(\tilde{\omega}\Delta t) = \mathrm{C}\sin(k_x\Delta x). \tag{6.27}$$

If $|C| > 1$, this equation admits complex solutions $\tilde{\omega} = \tilde{\omega}_r + i\tilde{\omega}_i$ for the $4\Delta x$ wave with $\tilde{\omega}_r \Delta t = \pi/2$ and $\tilde{\omega}_i$ such that

$$\sin(\tilde{\omega}_r \Delta t + i\tilde{\omega}_i \Delta t) = \cosh(\tilde{\omega}_i \Delta t) = C, \tag{6.28}$$

which admits two real solutions $\tilde{\omega}_i$ of opposite signs. One of the two solutions, therefore, corresponds to a growing amplitude, and the scheme is unstable.

For $|C| \leq 1$, dispersion relation (6.27) has two real solutions $\tilde{\omega}$, and the scheme is stable. Therefore, numerical stability requires the condition $|C| \leq 1$.

In the stable case, the numerical frequency $\tilde{\omega}$ may be compared with the exact value written in terms of discrete parameters

$$\omega = uk_x \quad \rightarrow \quad \omega \Delta t = Ck_x \Delta x. \tag{6.29}$$

Obviously, for $k_x \Delta x \rightarrow 0$ and $\Delta t \rightarrow 0$, the numerical relation (6.27) coincides with the exact relation (6.29). However, when $\tilde{\omega}$ is solution of (6.27) so is also $\pi/\Delta t - \tilde{\omega}$. The numerical solution thus consists of the superposition of the physical mode $\exp[i(k_x\, i\Delta x - \tilde{\omega}\, n\Delta t)]$, and a numerical mode that can be expressed as

$$\tilde{c}_i^n = Ae^{i(k_x i\Delta x + \tilde{\omega} n\Delta t)} e^{in\pi} \tag{6.30}$$

which, by virtue of $e^{in\pi} = (-1)^n$, flip-flops in time, irrespectively of how small the time step is or how well the spatial scale is resolved. This second component of the numerical solution, called *spurious mode* or *computational mode*, is traveling upstream, as indicated by the change of sign in front of the frequency. For the linear case discussed here, this spurious mode can be controlled by careful initialization (see Numerical Exercise 6.2), but for nonlinear equations, the mode may still be generated despite careful initialization and boundary conditions. In this case, it might be necessary to use time-filtering (see Section 10.6) to eliminate unwanted signals even if the spurious mode is stable for $|C| \leq 1$.

The leapfrog scheme is thus conditionally stable. The stability condition $|C| \leq 1$ was given a clear physical interpretation by Courant, Friedrichs and Lewy in their seminal 1928 paper. It is based on the fact that algorithm (6.21) defines a domain of dependence: Calculation of the value at point i and moment n (at the top of the gray pyramid in Fig. 6.5) implicates neighbor points $i \pm 1$ at time $n-1$ and the cell value i at time $n-2$. Those values in turn depend on their two neighboring and past values so that a network of points can be constructed that affect the value at grid point i and moment n. This network is the domain of numerical dependence. Physically, however, the solution at point i and time n depends only on the value along the characteristic $x - ut = x_i - ut^n$ according to (6.24). It is clear that, if this line does not fall inside the domain of dependence, there is trouble, for an attempt is made to determine a value from an irrelevant set of other values. Numerical instability is the symptom of this unsound approach. It is therefore necessary that the characteristic line passing through (i, n) be included in the domain of numerical dependence.

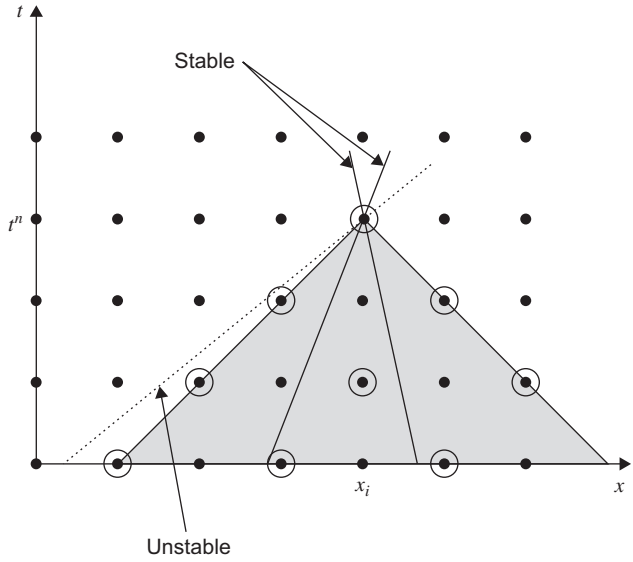


FIGURE 6.5 Domain of numerical dependence of the leapfrog scheme (in gray) covered by the points (circled dots) that influence the calculation at point i, n. This network of points is constructed recursively by identifying the grid points involved in prior calculations. The physical solution depends only on values along the characteristic. If the characteristic lies within the domain of numerical dependence (one of the solid lines, for example), this value can be captured by the calculations. On the contrary, when the physical characteristic lies outside the domain of numerical dependence (dashed line, for example), the numerical scheme relies on information that is physically unrelated to the advection process, and this flaw is manifested by numerical instability. Note also that the leapfrog scheme divides the grid points in two sets according to a checkerboard pattern (circled and noncircled dots). Unless some smoothing is performed, this risks to generate two numerically independent sets of values.

Except for the undesirable spurious mode, the leapfrog scheme has desirable features, because it is stable for $|C| \leq 1$, conserves variance for sufficiently small time steps, and leads to the correct dispersion relation for well-resolved spatial scales. But, is it sufficient to ensure a well-behaved solution? A standard test for advection schemes is the translation of a "top-hat" signal. In this case, the use of Eq. (6.21) leads to the result shown in Fig. 6.6, which is somewhat disappointing. The odd behavior can be explained: In terms of Fourier modes, the solution consists of a series of sine/cosine signals of different wavelengths, each of which by virtue of the numerical dispersion relation (6.27) travels at its own speed, thus unraveling the signal overtime. This also explains the unphysical appearance of both negative values and values in excess of the initial maximum. The scheme does not possess the monotonicity property but creates new extrema.

The cause of the poor performance of the leapfrog scheme is evident: The actual integration should be performed using upstream information exclusively,

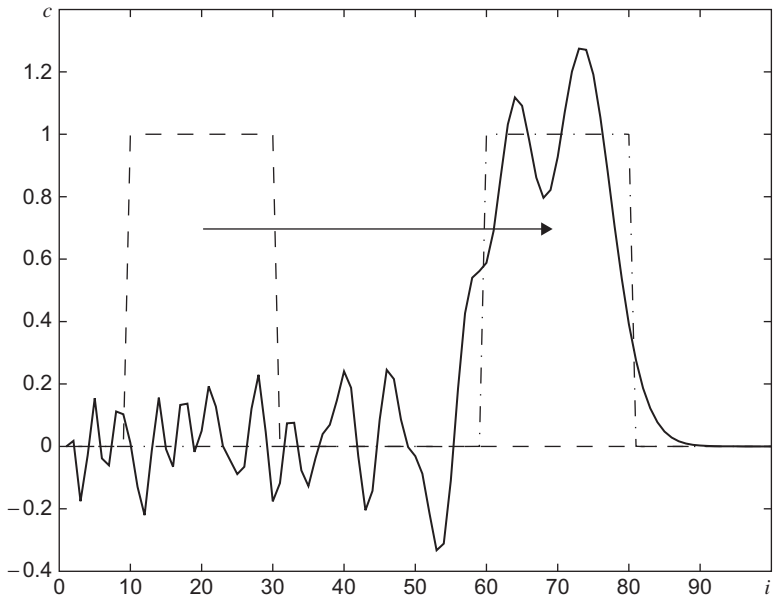


FIGURE 6.6 Leapfrog scheme applied to the advection of a "top-hat" signal with C = 0.5 for 100 times steps. The exact solution is a mere translation from the initial position (dashed curve on the left) by 50 grid points downstream (dash-dotted curve on the right). The numerical method generates a solution that is roughly similar to the exact solution, with the solution varying around the correct value.

whereas the scheme uses a central average that disregards the origin of the information. In other words, it ignores the physical bias of advection.

To remedy the situation, we now try to take into account the directional information of advection and introduce the so-called *upwind* or *donor cell* scheme. A simple Euler scheme over a single time step Δt is chosen, and fluxes are integrated over this time interval. The essence of this scheme is to calculate the inflow based solely on the average value across the grid cell from where the flow arrives (the donor cell). For positive velocity and a time integration from t^n to t^{n+1}, we obtain

$$\bar{c}_i^{n+1} = \bar{c}_i^n - \frac{\Delta t}{\Delta x}\left(\hat{q}_{i+1/2} - \hat{q}_{i-1/2}\right) \tag{6.31}$$

with

$$\hat{q}_{i-1/2} = \frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} q_{i-1/2}\mathrm{d}t \simeq u\tilde{c}_{i-1}^n, \tag{6.32}$$

so that the scheme is

$$\tilde{c}_i^{n+1} = \tilde{c}_i^n - \mathrm{C}\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right). \tag{6.33}$$

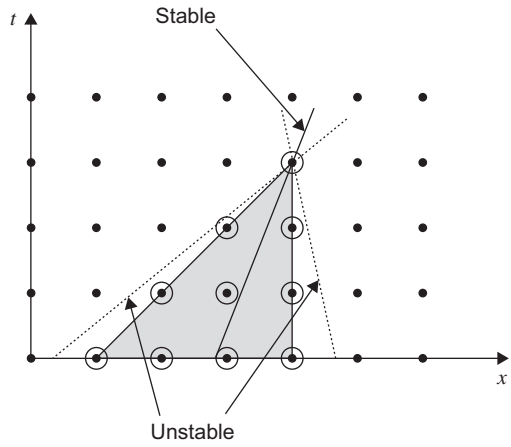


FIGURE 6.7 Domain of dependence of the upwind scheme. If the characteristic lies outside the numerical domain of dependence (dashed lines), unphysical behavior will be manifested as numerical instability. The necessary CFL stability condition therefore requires $0 \leq C \leq 1$ so that the characteristic lies within the numerical domain of dependence (solid lines). One initial condition and one upstream boundary condition are sufficient to determine the numerical solution.

Interestingly enough, the scheme can be used without need of artificial boundary conditions or special initialization, as we can see from algorithm (6.33) or the domain of numerical dependence (Fig. 6.7). The CFL condition $0 \leq C \leq 1$ provides the necessary condition for stability.

The stability of the scheme could be analyzed with the von Neumann method, but the simplicity of the scheme permits another approach, the so-called *energy method*. The energy method considers the sum of squares of $\tilde{c}$ and determines whether it remains bounded over time, providing a sufficient condition for stability. We start with (6.33), square it, and sum over the domain:

$$\sum_i \left(\tilde{c}_i^{n+1}\right)^2 = \sum_i (1-C)^2 \left(\tilde{c}_i^n\right)^2 + \sum_i 2C(1-C)\tilde{c}_i^n \tilde{c}_{i-1}^n + \sum_i C^2 \left(\tilde{c}_{i-1}^n\right)^2. \tag{6.34}$$

The first and last terms on the right can be grouped by shifting the index i in the last sum and invoking cyclic boundary conditions so that

$$\sum_i \left(\tilde{c}_i^{n+1}\right)^2 = \sum_i [(1-C)^2 + C^2] \left(\tilde{c}_i^n\right)^2 + \sum_i 2C(1-C)\tilde{c}_i^n \tilde{c}_{i-1}^n. \tag{6.35}$$

We can find an upper bound for the last term by using the following inequality:

$$0 \leq \sum_i \left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right)^2 = 2\sum_i \left(\tilde{c}_i^n\right)^2 - 2\sum_i \tilde{c}_i^n \tilde{c}_{i-1}^n, \tag{6.36}$$

which can be proved by using again the cyclic condition. If $C(1-C) > 0$, the last term in (6.35) may be replaced by the upper bound of (6.36) so that

$$\sum_i \left(\tilde{c}_i^{n+1}\right)^2 \leq \sum_i \left(\tilde{c}_i^n\right)^2, \tag{6.37}$$

and the scheme is stable because the norm of the solution does not increase in time. Although it is not related to a physical energy, the method derives its name from its reliance on a quadratic form that bears resemblance with kinetic energy. Methods which prove that a quadratic form is conserved or bounded over time are similar to energy-budget methods used to prove that the energy of a physical system is conserved.

The energy method provides only a sufficient stability condition because the upper bounds used in the demonstration do not need to be reached. But, since in the present case the sufficient stability condition was found to be identical to the necessary CFL condition, the condition $0 \leq C \leq 1$ is both necessary and sufficient to guarantee the stability of the upwind scheme.

Testing the upwind scheme on the "top-hat" problem (Fig. 6.8), we observe that, unlike leapfrog, the scheme does not create new minima or maxima, but somehow diffuses the distribution by reducing its gradients. The fact that the scheme is monotonic is readily understood by examining (6.33): The new value at point i is a linear interpolation of previous values found at i and $i-1$ so that no new value can ever fall outside the range of these previous values as long as the condition $0 \leq C \leq 1$ is satisfied.

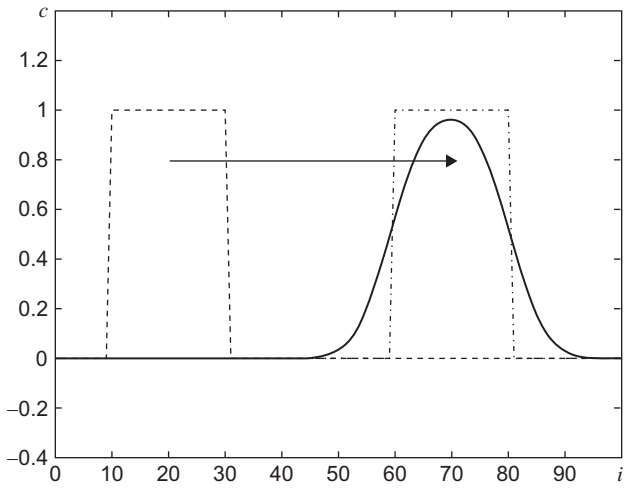


FIGURE 6.8 Upwind scheme with $C = 0.5$ applied to the advection of a "top-hat" signal after 100 times steps. Ideally the signal should be translated without change in shape by 50 grid points, but the solution is characterized by a certain diffusion and a reduction in gradient.

The diffusive behavior can be explained by analyzing the modified equation associated with (6.33). A Taylor expansion of the discrete solution around point (i, n) in (6.33) provides the equation that the numerical scheme actually solves the following:

$$\frac{\partial \tilde{c}}{\partial t} + \frac{\Delta t}{2}\frac{\partial^2 \tilde{c}}{\partial t^2} + \mathcal{O}\left(\Delta t^2\right) + u\left(\frac{\partial \tilde{c}}{\partial x} - \frac{\Delta x}{2}\frac{\partial^2 \tilde{c}}{\partial x^2} + \mathcal{O}\left(\Delta x^2\right)\right) = 0. \tag{6.38}$$

The scheme is only of first order as can be expected from the use of a one-sided finite difference. To give a physical interpretation to the equation, the second time derivative should be replaced by a spatial derivative. Taking the derivative of the modified equation with respect to t provides an equation for the second time derivative, which we would like to eliminate, but it involves a cross derivative[3]. This cross derivative can be obtained by differentiating the modified equation with respect to x. Some algebra ultimately provides

$$\frac{\partial^2 \tilde{c}}{\partial t^2} = u^2 \frac{\partial^2 \tilde{c}}{\partial x^2} + \mathcal{O}\left(\Delta t, \Delta x^2\right),$$

which can finally be introduced into (6.38) to yield the following equation

$$\frac{\partial \tilde{c}}{\partial t} + u\frac{\partial \tilde{c}}{\partial x} = \frac{u\Delta x}{2}(1 - \mathrm{C})\frac{\partial^2 \tilde{c}}{\partial x^2} + \mathcal{O}\left(\Delta t^2, \Delta x^2\right). \tag{6.39}$$

This is the equation that the upwind scheme actually solves.

Up to $\mathcal{O}\left(\Delta t^2, \Delta x^2\right)$, therefore, the numerical scheme solves an advection–diffusion equation instead of the pure advection equation, with diffusivity equal to $(1 - \mathrm{C})\, u\Delta x/2$. For obvious reasons, this is called an artificial diffusion or *numerical diffusion*. The effect is readily seen in Fig. 6.8. To decide whether this level of artificial diffusion is acceptable or not, we must compare its size with that of physical diffusion. For a diffusivity coefficient $\mathcal{A}$, the ratio of numerical to physical diffusion is $(1 - \mathrm{C})\, u\Delta x/(2\mathcal{A})$. Since it would be an aberration to have numerical diffusion equal or exceed physical diffusion (recall the error analysis of Section 4.8: Discretization errors should not be larger than modeling errors), the *grid Peclet number* $U\Delta x/\mathcal{A}$ may not exceed a value of $\mathcal{O}(1)$ for the upwind scheme to be valid.

When no physical diffusion is present, we must require that the *numerical* diffusion term be small compared with the *physical* advection term, a condition

[3] Note that using the *original* equations, the physical solution satisfies $\partial^2 c/\partial t^2 = u^2 \partial^2 c/\partial x^2$, which is sometimes used as a shortcut to eliminate the second time derivative from the modified equation. This is, however, incorrect because $\tilde{c}$ does not solve the original equation. In practice, this kind of shortcuts can lead to correct leading truncation errors, but without being sure that no essential term is overlooked.

that can be associated with another grid Peclet number:

$$\tilde{Pe} = \frac{UL}{U\Delta x\,(1-\mathrm{C})\,/2} \sim \frac{L}{\Delta x} \gg 1, \tag{6.40}$$

where L stands for the length scale of any solution component worth resolving. Even for well-resolved signals in GFD flows, this Peclet number associated with numerical diffusion is often insufficiently large, and numerical diffusion is a problem that plagues the upwind scheme.

The observation that the scheme introduces artificial diffusion is interesting and annoying, and the question now is to identify its origin in order to reduce it. Compared with the centered scheme, which is symmetric and of second order, the upwind scheme uses exclusive information from the upstream side, the donor cell, and is only of first order. Numerical diffusion must, therefore, be associated with the asymmetry in the flux calculation, and to reduce numerical diffusion, we must somehow take into account values of $\tilde{c}$ on both sides of the interface to calculate the flux and thereby seek a scheme that is second-order accurate.

This can be accomplished with the *Lax-Wendroff scheme*, which estimates the flux at the cell interface by assuming that the function is not constant within the cell but varies linearly across it:

$$\begin{aligned} \tilde{q}_{i-1/2} &= u\left[\frac{\tilde{c}_i^n + \tilde{c}_{i-1}^n}{2} - \frac{\mathrm{C}}{2}\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right)\right] \\ &= u\tilde{c}_{i-1}^n + \underbrace{(1-\mathrm{C})\frac{u\Delta x}{2}\,\frac{\tilde{c}_i^n - \tilde{c}_{i-1}^n}{\Delta x}}_{\simeq(1-\mathrm{C})\frac{u\Delta x}{2}\,\frac{\partial \tilde{c}}{\partial x}}. \end{aligned} \tag{6.41}$$

The last term, in addition to the upwind flux $u\tilde{c}_{i-1}^n$, is designed to oppose numerical diffusion. Substitution of this flux into the finite-volume scheme leads to the following scheme:

$$\tilde{c}_i^{n+1} = \tilde{c}_i^n - \mathrm{C}\left(\tilde{c}_i^n - \tilde{c}_{i-1}^n\right) - \frac{\Delta t}{\Delta x^2}(1-\mathrm{C})\frac{u\Delta x}{2}\left(\tilde{c}_{i+1}^n - 2\tilde{c}_i^n + \tilde{c}_{i-1}^n\right) \tag{6.42}$$

which, compared with the upwind scheme, includes an additional antidiffusion term with coefficient constructed to negate the numerical diffusion of the upwind scheme. The effect of this higher-order approach on the solution of our test case is a reduced overall error but the appearance of dispersion (Fig. 6.9). This is due to the fact that we eliminated the truncation error proportional to the second spatial derivative (an even derivative associated with dissipation) and now have a truncation error proportional to the third spatial derivative (an odd derivative associated with dispersion).

The same dispersive behavior is observed with the *Beam-Warming scheme*, in which the antidiffusion term is shifted upstream so as to anticipate the

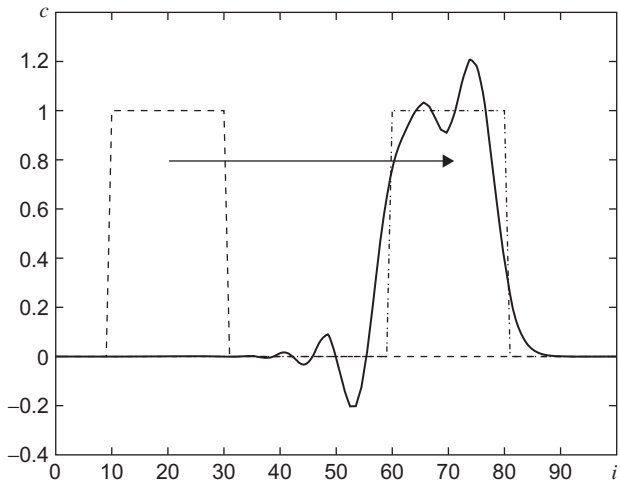


FIGURE 6.9 Second-order Lax-Wendroff scheme applied to the advection of a "top-hat" signal with C = 0.5 after 100 times steps. Dispersion and nonmonotonic behavior are noted.

gradient that will arrive later at the interface:

$$\tilde{q}_{i-1/2} = u\tilde{c}^n_{i-1} + (1-\mathrm{C})\,\frac{u}{2}\,\left(\tilde{c}^n_{i-1} - \tilde{c}^n_{i-2}\right). \tag{6.43}$$

This scheme is still of second order, since the correction term is only shifted upstream by Δx. The effect of anticipating the incoming gradients enhances stability but does not reduce dispersion (see Numerical Exercise 6.8).

Other methods spanning more grid points can be constructed to obtain higher-order integration of fluxes, implicit methods to increase stability, predictor-corrector methods, or combinations of all these schemes. Here, we only outline some of the approaches and refer the reader to more specialized literature for details (e.g., Chung, 2002; Durran, 1999).

A popular *predictor-corrector method* is the second-order MacCormack scheme: The predictor uses a forward spatial difference (antidiffusion),

$$\tilde{c}^\star_i = \tilde{c}^n_i - \mathrm{C}\left(\tilde{c}^n_{i+1} - \tilde{c}^n_i\right) \tag{6.44}$$

and the corrector a backward spatial difference on the predicted field (diffusion):

$$\tilde{c}^{n+1}_i = \tilde{c}^{n+1/2}_i - \frac{\mathrm{C}}{2}\left(\tilde{c}^\star_i - \tilde{c}^\star_{i-1}\right) \quad \text{with} \quad \tilde{c}^{n+1/2}_i = \frac{\tilde{c}^\star_i + \tilde{c}^n_i}{2}. \tag{6.45}$$

The elimination of the intermediate value $\tilde{c}^{n+1/2}_i$ from which starts the corrector step provides the expanded corrector step:

$$\tilde{c}^{n+1}_i = \frac{1}{2}\left[\tilde{c}^n_i + \tilde{c}^\star_i - \mathrm{C}\left(\tilde{c}^\star_i - \tilde{c}^\star_{i-1}\right)\right], \tag{6.46}$$

assuming $u > 0$ as usual. Substitution of the predictor step into the corrector step shows that the MacCormack scheme is identical to the Lax-Wendroff scheme in the linear case, but differences may arise in nonlinear problems.

An *implicit scheme* can handle centered space differencing and approximates the flux as

$$\tilde{q}_{i-1/2} = \alpha u \, \frac{\tilde{c}_i^{n+1} + \tilde{c}_{i-1}^{n+1}}{2} + (1-\alpha) u \, \frac{\tilde{c}_i^{n} + \tilde{c}_{i-1}^{n}}{2}. \tag{6.47}$$

For $\alpha = 1$, the scheme is fully implicit, whereas for $\alpha = 1/2$ it becomes a semi-implicit or trapezoidal scheme (also called *Crank-Nicholson scheme*). The latter has already been shown to be unconditionaly stable [see variance conservation and the trapezoidal scheme (6.17)]. The price to pay for this stability is the need to solve a linear algebraic system at every step. As for the diffusion problem, the system is tridiagonal in the 1D case and more complicated in higher dimensions. The advantage of the implicit approach is a robust scheme when C occasionally happens to exceed unity in a known dimension[4]. It should be noted, however, that for too large a Courant number accuracy degrades.

All of the previous schemes can be mixed in a *linear combination*, as long as the sum of the weights attributed to each scheme is unity for the sake of consistency. An example of combining two schemes consists in averaging the flux calculated with a lower-order scheme $\tilde{q}_{i-1/2}^{L}$ with that of a higher-order scheme $\tilde{q}_{i-1/2}^{H}$:

$$\tilde{q}_{i-1/2} = (1-\Phi)\, \tilde{q}_{i-1/2}^{L} + \Phi \, \tilde{q}_{i-1/2}^{H},$$

in which the weight Φ ($0 \leq \Phi \leq 1$) acts as a trade-off between the undesirable numerical diffusion of the lower-order scheme and numerical dispersion and loss of monotonicity of the higher-order scheme.

All these methods lead to sufficiently accurate solutions, but none except the upwind scheme ensures monotonic behavior. The reason for this disappointing fact can be found in the frustrating theorem by Godunov (1959) regarding the discretized advection equation:

A consistent linear numerical scheme that is monotonic can at most be first-order accurate.

Therefore, the upwind scheme is the inevitable choice if no overshoot or undershoot is permitted. To circumvent the Godunov theorem, state-of-the-art advection schemes relax the linear nature of the discretization and adjust the parameter Φ locally, depending on the behavior of the solution. The function

[4] Typically the vertical Courant number may be so variable that it becomes difficult to ensure that the local vertical C value remains below one. In particular, it is prudent to use an implicit scheme in the vertical when the model has nonuniform grid spacing, and when the vertical velocity is weak except on rare occasions.

that defines the way Φ is adapted locally is called a *limiter*. Such an approach is able to capture large gradients (fronts). Because of its advanced nature, we delay its presentation until Section 15.7. An example of a nonlinear scheme called TVD, however, is already included in the computer codes provided for the analysis of advection schemes in several dimensions.

6.5 ADVECTION–DIFFUSION WITH SOURCES AND SINKS

Having considered separately advection schemes (this chapter), diffusion schemes (Chapter 5) and time discretizations with arbitrary forcing terms (Chapter 2), we can now combine them to tackle the general advection–diffusion equation with sources and sinks. For a linear sink, the 1D equation to be discretized is

$$\frac{\partial c}{\partial t}+u\,\frac{\partial c}{\partial x}=-K\,c+\frac{\partial}{\partial x}\left(\mathcal{A}\frac{\partial c}{\partial x}\right). \tag{6.48}$$

Since we already have a series of discretization possibilities for each individual process, the combination of these provides an even greater number of possible schemes which we cannot describe exhaustively here. We simply show one example to illustrate two important facts that should not be forgotten when combining schemes: The properties of the combined scheme are neither simply the sum of the properties of the individual schemes, nor is its stability condition the most stringent condition of the separate schemes.

To prove the first statement, we consider (6.48) without diffusion ($\mathcal{A}=0$) and solve by applying the second-order Lax-Wendroff advection scheme with the second-order trapezoidal scheme applied to the decay term. The discretization, after some rearrangement of terms, is as follows:

$$\tilde{c}_i^{n+1}=\tilde{c}_i^n-\frac{\mathrm{B}}{2}\left(\tilde{c}_i^n+\tilde{c}_i^{n+1}\right)-\frac{\mathrm{C}}{2}\left(\tilde{c}_{i+1}^n-\tilde{c}_{i-1}^n\right)+\frac{\mathrm{C}^2}{2}\left(\tilde{c}_{i+1}^n-2\tilde{c}_i^n+\tilde{c}_{i-1}^n\right), \tag{6.49}$$

where $\mathrm{B}=K\Delta t$ and $\mathrm{C}=u\Delta t/\Delta x$. This scheme actually solves the following modified equation:

$$\begin{aligned}\frac{\partial\tilde{c}}{\partial t}+u\,\frac{\partial\tilde{c}}{\partial x}+K\,\tilde{c}&=-\frac{\Delta t}{2}\,\frac{\partial^2\tilde{c}}{\partial t^2}-K\frac{\Delta t}{2}\,\frac{\partial\tilde{c}}{\partial t}+\frac{u^2\Delta t}{2}\,\frac{\partial^2\tilde{c}}{\partial x^2}+\mathcal{O}(\Delta t^2,\Delta x^2)\\&=-\frac{uK\Delta t}{2}\,\frac{\partial\tilde{c}}{\partial x}+\mathcal{O}(\Delta t^2,\Delta x^2)\end{aligned} \tag{6.50}$$

where the last equality was obtained by a similar procedure as the one used to find the modified equation (6.39). It is not possible to cancel the leading term on the right unless $K=0$ or $u=0$, in which case we recover the second-order Lax-Wendroff or the second-order trapezoidal scheme. Thus, in the combined advection-decay case, what was expected to be a second-order scheme degenerates into a first-order scheme.

For the purpose of illustrating the second statement on stability, we combine the second-order Lax-Wendroff advection scheme (stability condition $|\mathrm{C}| \leq 1$) with the explicit Euler diffusion scheme (stability condition $0 \leq \mathrm{D} \leq 1/2$) and an explicit scheme for the sink term with rate K (stability condition $\mathrm{B} \leq 2$). The discretization, after some rearrangement of the terms, is as follows:

$$\tilde{c}_i^{n+1} = \tilde{c}_i^n - \mathrm{B}\,\tilde{c}_i^n - \frac{\mathrm{C}}{2}\left(\tilde{c}_{i+1}^n - \tilde{c}_{i-1}^n\right) + \left(\mathrm{D} + \frac{\mathrm{C}^2}{2}\right)\left(\tilde{c}_{i+1}^n - 2\tilde{c}_i^n + \tilde{c}_{i-1}^n\right), \quad (6.51)$$

where $\mathrm{D} = \mathcal{A}\Delta t/\Delta x^2$. Application of the von Neumann stability analysis yields the following amplification factor

$$\varrho = 1 - \mathrm{B} - 4\left(\mathrm{D} + \frac{\mathrm{C}^2}{2}\right)\sin^2\theta - \mathrm{i}\ 2\mathrm{C}\sin\theta\cos\theta, \quad (6.52)$$

where $\theta = k_x\Delta x/2$, so that

$$|\varrho|^2 = \left[1 - \mathrm{B} - 4\left(\mathrm{D} + \frac{\mathrm{C}^2}{2}\right)\xi\right]^2 + 4\mathrm{C}^2\xi(1-\xi), \quad (6.53)$$

where $0 \leq \xi = \sin^2\theta \leq 1$. For $\xi \simeq 0$ (long waves), we obtain the necessary stability condition $\mathrm{B} \leq 2$, corresponding to the stability condition of the sink term alone. For $\xi \simeq 1$ (short waves), we find the more demanding necessary stability condition

$$\mathrm{B} + 2\mathrm{C}^2 + 4\mathrm{D} \leq 2. \quad (6.54)$$

We can show that the latter condition is also sufficient (Numerical Exercise 6.13), which proves that the stability condition of the combined schemes is more stringent than each stability condition taken individually. Only when two processes are negligible does the stability condition revert to the stability condition of the single remaining process. This seems evident but is not always the case. In some situations, adding even an infinitesimally-small stable process can require a discontinuous reduction in time step (e.g., Beckers & Deleersnijder, 1993).

In other situations, adding a process can stabilize an otherwise unconditionally unstable scheme (Numerical Exercise 6.14). Therefore, in theory, it is not enough to consider the stability of each piece of the scheme separately, but the stability of the full scheme must be investigated. In practice, however, if a complete scheme is too difficult to analyze, subschemes (i.e., including only a few processes) are isolated with the hope that the full scheme does not demand a drastically shorter time step than the one required by the most stringent stability condition of each elementary scheme taken separately.

Stability is an important property of any scheme as is, at least for tracers, monotonic behavior. If we assume the scheme to be explicit, linear and covering a stencil spanning p grid points upstream and q downstream (for a total of

$p+q+1$ points), it can be written in the general form:

$$\tilde{c}_i^{n+1} = a_{-p}\tilde{c}_{i-p}^n + \cdots + a_{-1}\tilde{c}_{i-1}^n + a_0\tilde{c}_i^n + a_1\tilde{c}_{i+1}^n + \cdots + a_q\tilde{c}_{i+q}^n. \tag{6.55}$$

To be consistent with (6.48), we need at least to ensure $a_{-p}+\cdots+a_{-1}+a_0+a_1+\cdots+a_q = 1-\mathrm{B}$, otherwise, not even a spatially uniform field would be represented correctly.

If there is a negative coefficient a_k, the scheme will not be monotonic, for indeed, if the function is positive at point $i+k$ but zero everywhere else, it will take on a negative value $\tilde{c}_i^{n+1}$. However, if all coefficients are positive, the sum of the total weights is obviously positive but less than one because it is equal to $(1-\mathrm{B})$. The scheme thus interpolates while damping, in agreement with physical decay. And, since damping does not create new extrema, we conclude that positive coefficients ensure a monotonic behavior in all situations. For our example (6.51), this demands $\mathrm{B}+\mathrm{C}^2+2\mathrm{D} \le 1$ and $\mathrm{C} \le \mathrm{C}^2+2\mathrm{D}$. The former condition is a slightly more constraining version of the stability condition (6.54), whereas the latter condition imposes a constraint on the grid Peclet number:

$$Pe_{\Delta x} = \frac{u\Delta x}{\mathcal{A}} = \frac{\mathrm{C}}{\mathrm{D}} \le \frac{u^2\Delta t}{\mathcal{A}} + 2. \tag{6.56}$$

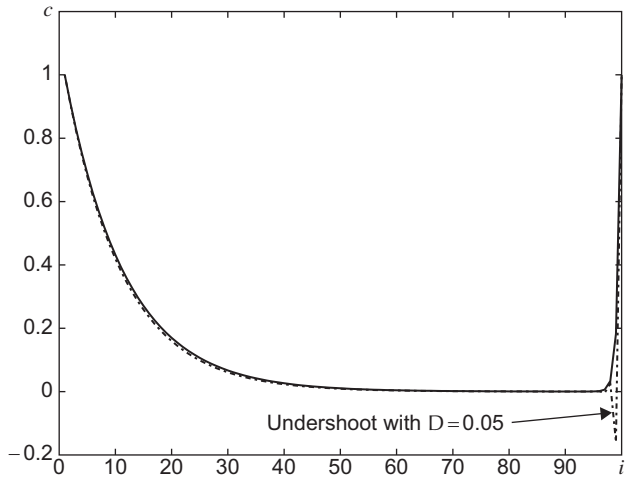


FIGURE 6.10 Simulation using (6.51) with $\mathrm{B}=0.05$, $\mathrm{C}=0.5$, and $\mathrm{D}=0.25$, after convergence to a stationary solution (solid line). With decreasing diffusion, the scheme eventually fails to resolve adequately the outflow boundary layer, and undershoot appears ($\mathrm{D}=0.05$, dot-dash line). This corresponds to a situation in which one of the coefficients in the numerical scheme has become negative. The program `advdiffsource.m` can be used to test other combinations of the parameter values.

For short time steps, this imposes a maximum value of 2 on the grid Peclet number. It is not a stability condition but a necessary condition for monotonic behavior.

The scheme is now tested on a physical problem. Because of the second derivative, we impose boundary conditions at both upstream *and* downstream and for simplicity hold $\tilde{c} = 1$ steady at these locations. We then iterate from a zero initial condition until the scheme converges to a stationary solution. This solution (Fig. 6.10) exhibits a boundary layer at the downstream end because of weak diffusion, in agreement with the remark made in Section 6.3. For weak diffusion, the grid Peclet number $Pe_{\Delta x}$ is too large and violates (6.56). Undershooting appears, although the solution remains stable. In conclusion, besides the parameters B, C, and D that control stability, the grid Peclet number C/D appears as a parameter controlling monotonic behavior.

6.6 MULTIDIMENSIONAL APPROACH

In addition to the various combinations already encountered in the 1D case, generalization to more dimensions allows further choices and different methods. Here, we concentrate on the 2D advection case because generalizations to 3D do not generally cause more fundamental complications.

The finite-volume approach can be easily extended to a 2D grid cell with the fluxes perpendicular to the interfaces (Fig. 6.11):

$$\frac{\tilde{c}_{i,j}^{n+1} - \tilde{c}_{i,j}^{n}}{\Delta t} + \frac{\tilde{q}_{x,i+1/2,j} - \tilde{q}_{x,i-1/2,j}}{\Delta x} + \frac{\tilde{q}_{y,i,j+1/2} - \tilde{q}_{y,i,j-1/2}}{\Delta y} = 0, \qquad (6.57)$$

where $\tilde{q}_{x,i\pm1/2,j}$ and $\tilde{q}_{y,i,j\pm1/2}$ are approximations of the actual fluxes uc and vc, respectively.

For any flux calculation, the least we require is that it be able to represent correctly a uniform tracer field $\mathcal{C}$. All of our 1D flux calculations do so and should do so. When (6.57) is applied to the case of a uniform concentration

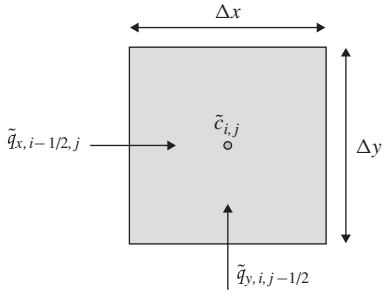


FIGURE 6.11 Finite volume in 2D with fluxes at the interfaces. The budget involves the total balance of inflowing and outflowing fluxes during one time step.

distribution $\tilde{c} = \mathcal{C}$, we obtain

$$\frac{\tilde{c}_{i,j}^{n+1} - \mathcal{C}}{\Delta t} + \frac{\tilde{u}_{i+1/2,j} - \tilde{u}_{i-1/2,j}}{\Delta x}\,\mathcal{C} + \frac{\tilde{v}_{i,j+1/2} - \tilde{v}_{i,j-1/2}}{\Delta y}\,\mathcal{C} = 0.$$

This can only lead to $\tilde{c}^{n+1} = \mathcal{C}$ at the next time step if the discrete velocity field satisfies the condition

$$\frac{\tilde{u}_{i+1/2,j} - \tilde{u}_{i-1/2,j}}{\Delta x} + \frac{\tilde{v}_{i,j+1/2} - \tilde{v}_{i,j-1/2}}{\Delta y} = 0. \tag{6.58}$$

Since this requirement is an obvious discretization of $\partial u/\partial x + \partial v/\partial y = 0$, the 2D form of the continuity equation (4.21d), it follows that a prerequisite to solve the concentration equation by the finite-volume approach is a nondivergent flow field *in its discretized form*. Ensuring that (6.58) holds is the role of the discretization of the dynamical equations, those governing velocity and pressure.

Here, in order to test numerical advection schemes, we take the flow field as known and obeying (6.58). We can easily generate such a discrete velocity distribution by invoking a discretized *streamfunction* ψ:

$$\tilde{u}_{i-1/2,j} = -\frac{\psi_{i-1/2,j+1/2} - \psi_{i-1/2,j-1/2}}{\Delta y} \tag{6.59}$$

$$\tilde{v}_{i,j-1/2} = \frac{\psi_{i+1/2,j-1/2} - \psi_{i-1/2,j-1/2}}{\Delta x}. \tag{6.60}$$

It is straighforward to show that these $\tilde{u}$ and $\tilde{v}$ values satisfy (6.58) for any set of streamfunction values.

Note that if we had discretized directly the continuous equation

$$\frac{\partial c}{\partial t} + u\,\frac{\partial c}{\partial x} + v\,\frac{\partial c}{\partial y} = 0, \tag{6.61}$$

we would have obtained

$$\frac{\partial \tilde{c}_{i,j}}{\partial t} + u_{i,j}\,\left.\frac{\partial \tilde{c}}{\partial x}\right|_{i,j} + v_{i,j}\,\left.\frac{\partial \tilde{c}}{\partial y}\right|_{i,j} = 0,$$

which guarantees that an initially uniform tracer distribution remains uniform at all later times regardless of the structure of the discretized velocity distribution as long as the discretized form of the spatial derivatives return zeros for a uniform distribution (a mere requirement of consistency). Such a scheme could appear to offer a distinct advantage, but it is easy to show that it has a major drawback. It looses important conservation properties, including conservation of the quantity of tracer (heat for temperature, salt for salinity, etc.).

Assuming the discrete velocity field to be divergence-free in the sense of (6.58), the first method that comes to mind is to calculate the flux components using the discretizations developed in 1D along each coordinate line separately (Fig. 6.12). The upwind scheme is then easily generalized as follows:

$$\tilde{q}_{x,i-1/2,j} = \tilde{u}_{i-1/2,j}\,\tilde{c}^{n}_{i-1,j} \quad \text{if} \quad \tilde{u}_{i-1/2,j} > 0, \quad \tilde{u}_{i-1/2,j}\,\tilde{c}^{n}_{i,j} \quad \text{otherwise} \tag{6.62a}$$

$$\tilde{q}_{y,i,j-1/2} = \tilde{v}_{i,j-1/2}\,\tilde{c}^{n}_{i,j-1} \quad \text{if} \quad \tilde{v}_{i,j-1/2} > 0, \quad \tilde{v}_{i,j-1/2}\,\tilde{c}^{n}_{i,j} \quad \text{otherwise.} \tag{6.62b}$$

The other 1D schemes can be generalized similarly. Applying such schemes to the advection of an initially cone-shaped distribution (single peak with same linear drop in all radial directions) embedded in a uniform flow field crossing the domain at 45°, we observe that the upwing scheme is plagued by a very strong numerical diffusion (left panel of Fig. 6.13). Using the TVD scheme keeps the signal to a higher amplitude but strongly distorts the distribution (right panel of Fig. 6.13).

This distortion is readily understood in terms of the advection process: The information should be carried by the oblique flow, but the flux calculation relies on information strictly along the x or y axes. In the case of a flow oriented at 45° from the $x-$axis, this ignores that grid point (i,j) is primarily influenced by point $(i-1,j-1)$ whereas points $(i-1,j)$ and $(i,j-1)$ are used in the flux calculations. In conclusion, the double 1D approach is unsatisfactory and rarely used.

The *Corner Transport Upstream* (CTU) scheme (e.g., Colella, 1990) takes into account the different contributions of the four grid cells involved in the displacement (Fig. 6.14). Assuming uniform positive velocities to illustrate the approach, the following discretization ensures that a diagonal flow brings to

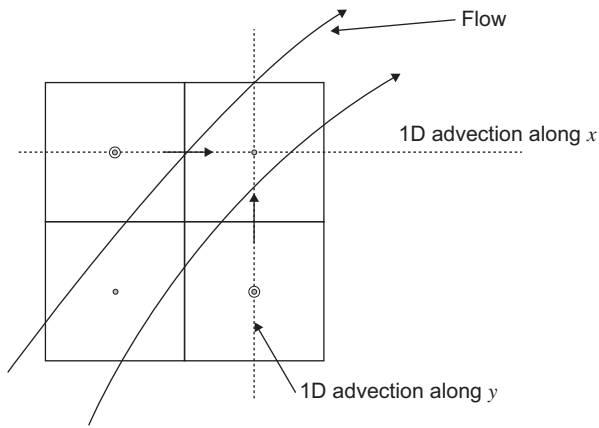


FIGURE 6.12 Naive 2D generalization using 1D methods along each coordinate line to approximate the advection operator as the sum of $\partial(uc)/\partial x$ and $\partial(vc)/\partial y$.

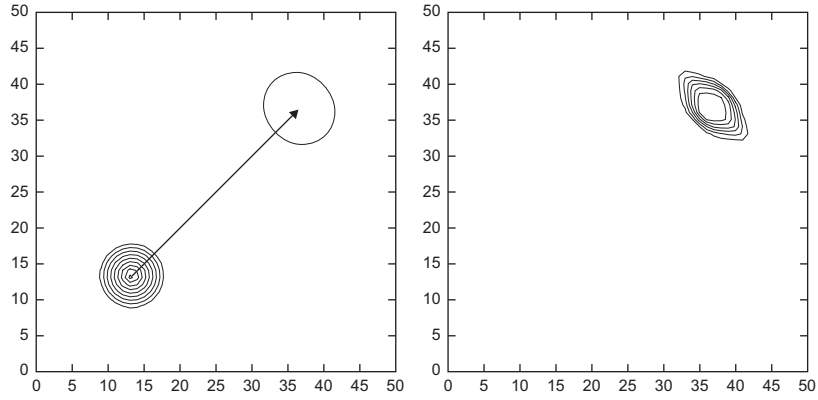


FIGURE 6.13 Oblique advection of a cone-shaped distribution using the upwind scheme generalized to 2D (left panel) and a TVD scheme (right panel) with $C_x = C_y = 0.12$. The upwind scheme severely dampens the signal, to less then 20% of its initial amplitude, whereas the TVD scheme used as a double 1D problem greatly distorts the solution.

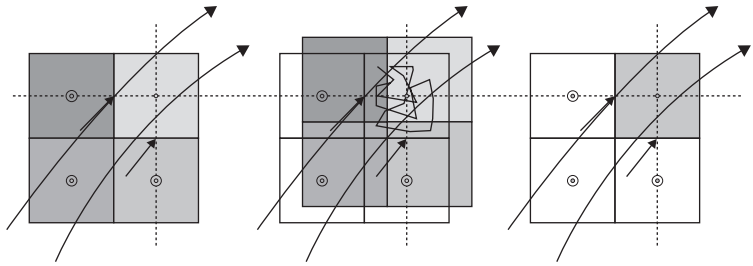


FIGURE 6.14 2D generalization designed to advect the field obliquely along streamlines. The associated numerical diffusion can then be interpreted as the necessary grid averaging (i.e., mixing) used in the finite-volume technique after displacement of the donor cells. The flux calculations (thick arrows) need to integrate the inflow of c along the flow instead of along the grid lines.

the interface a correct mixing of two donor cells:

$$\tilde{q}_{x,i-1/2,j} = \left(1 - \frac{C_y}{2}\right) \tilde{u}\, \tilde{c}^n_{i-1,j} + \frac{C_y}{2}\, \tilde{u}\, \tilde{c}^n_{i-1,j-1} \tag{6.63a}$$

$$\tilde{q}_{y,i,j-1/2} = \left(1 - \frac{C_x}{2}\right) \tilde{v}\, \tilde{c}^n_{i,j-1} + \frac{C_x}{2}\, \tilde{v}\, \tilde{c}^n_{i-1,j-1}, \tag{6.63b}$$

leading to the expanded scheme

$$\begin{aligned} \tilde{c}^{n+1}_{i,j} = \tilde{c}^n_{i,j} &- C_x\left(\tilde{c}^n_{i,j} - \tilde{c}^n_{i-1,j}\right) - C_y\left(\tilde{c}^n_{i,j} - \tilde{c}^n_{i,j-1}\right) \\ &+ C_x C_y\left(\tilde{c}^n_{i,j} - \tilde{c}^n_{i-1,j} - \tilde{c}^n_{i,j-1} + \tilde{c}^n_{i-1,j-1}\right), \end{aligned} \tag{6.64}$$

where the last term is an additional term compared with the double 1D approach. Two distinct Courant numbers arise, one for each direction:

$$\mathrm{C}_x = \frac{u\Delta t}{\Delta x}, \qquad \mathrm{C}_y = \frac{v\Delta t}{\Delta y}. \tag{6.65}$$

For $\mathrm{C}_x = \mathrm{C}_y = 1$, the scheme provides $\tilde{c}_{i,j}^{n+1} = \tilde{c}_{i-1,j-1}^{n}$, with obvious physical interpretation. The scheme may also be written as

$$\begin{aligned}\tilde{c}_{i,j}^{n+1} &= (1-\mathrm{C}_x)(1-\mathrm{C}_y)\,\tilde{c}_{i,j}^{n} \\ &+ (1-\mathrm{C}_y)\mathrm{C}_x\,\tilde{c}_{i-1,j}^{n} + (1-\mathrm{C}_x)\mathrm{C}_y\,\tilde{c}_{i,j-1}^{n} + \mathrm{C}_x\mathrm{C}_y\,\tilde{c}_{i-1,j-1}^{n}\end{aligned} \tag{6.66}$$

highlighting the relative weights attached to the four grid points involved in the calculation (Fig. 6.14). This expression proves that the method is monotonic for Courant numbers smaller than one (ensuring that all coefficients on the right-hand side are positive). The method is only of first order according to the Godunov theorem, but it causes less distortion of the solution (Fig. 6.15) than the previous approach. It still dampens excessively, however.

Other generalizations of the various 1D schemes to integrate along the current directions are possible but become rapidly complicated. We will therefore introduce a method that is almost as simple as solving a 1D problem but yet takes into account the multidimensional essence of the problem.

The method shown here is a special case of so-called *operator splitting* methods or *fractional steps*. To show the approach, we start from the semidiscrete equation

$$\frac{\mathrm{d}\tilde{c}_i}{\mathrm{d}t} + \mathcal{L}_1(\tilde{c}_i) + \mathcal{L}_2(\tilde{c}_i) = 0, \tag{6.67}$$

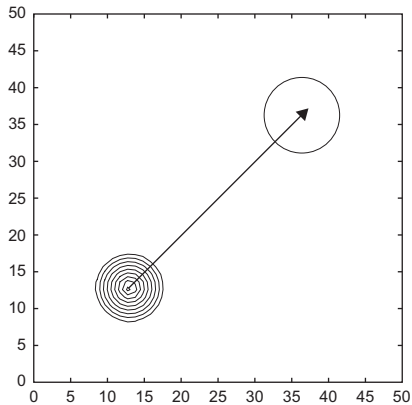


FIGURE 6.15 2D oblique advection using the CTU scheme (6.66). The solution's asymmetric deformation is reduced, but numerical diffusion still reduces the amplitude significantly.

where $\mathcal{L}_1$ and $\mathcal{L}_2$ are two discrete operators, which in the present case are advection operators along x and y. Temporal discretization by time splitting executes the following:

$$\frac{\tilde{c}_i^\star - \tilde{c}_i^n}{\Delta t} + \mathcal{L}_1(\tilde{c}_i^n) = 0 \tag{6.68a}$$

$$\frac{\tilde{c}_i^{n+1} - \tilde{c}_i^\star}{\Delta t} + \mathcal{L}_2(\tilde{c}_i^\star) = 0, \tag{6.68b}$$

where the second operator is marched forward with a value already updated by the first operator.

In this manner, we solve two sequential one-dimensional problems, which is not more complicated than what we just did and is a major improvement compared with the naive double 1D approach used in (6.62): The initial (predictor) step creates a field that is already advected in the direction of the $\mathcal{L}_1$ operator, and the second (corrector) step advects in the remaining direction the partially displaced field. In this way, point (i, j) is influenced by the upstream value, point $(i-1, j-1)$ in the case of positive velocities (Fig. 6.16).

To verify this, we can see how the splitting works with the 1D upwind scheme for positive and uniform velocities:

$$\tilde{c}_{i,j}^\star = \tilde{c}_{i,j}^n - \mathrm{C}_x\left(\tilde{c}_{i,j}^n - \tilde{c}_{i-1,j}^n\right) \tag{6.69a}$$

$$\tilde{c}_{i,j}^{n+1} = \tilde{c}_{i,j}^\star - \mathrm{C}_y\left(\tilde{c}_{i,j}^\star - \tilde{c}_{i,j-1}^\star\right). \tag{6.69b}$$

Substitution of the intermediate values $\tilde{c}_{i,j}^\star$ into the final step then proves that the scheme is identical to the CTU scheme (6.64) for uniform velocities. Such elimination, however, is not done in practice, and the sequence (6.69) is used. This is particularly convenient because, in the computer program, $\tilde{c}^\star$ may be stored during the first step in the future place of $\tilde{c}^{n+1}$ and then moved to the place of $\tilde{c}^n$ for the second step; $\tilde{c}^{n+1}$ can then be calculated and stored without need of additional storage.

Using always the same operator with current values and the other with "predicted" values is unsatisfactory because it breaks the symmetry between the two spatial dimensions. Hence, it is recommended to alternate the order in which operators are applied, depending on whether the time step is even or odd. Following a time step using (6.68) we then switch the order of operators by performing

$$\frac{\tilde{c}_i^\star - \tilde{c}_i^{n+1}}{\Delta t} + \mathcal{L}_2(\tilde{c}_i^{n+1}) = 0 \tag{6.70a}$$

$$\frac{\tilde{c}_i^{n+2} - \tilde{c}_i^\star}{\Delta t} + \mathcal{L}_1(\tilde{c}_i^\star) = 0. \tag{6.70b}$$

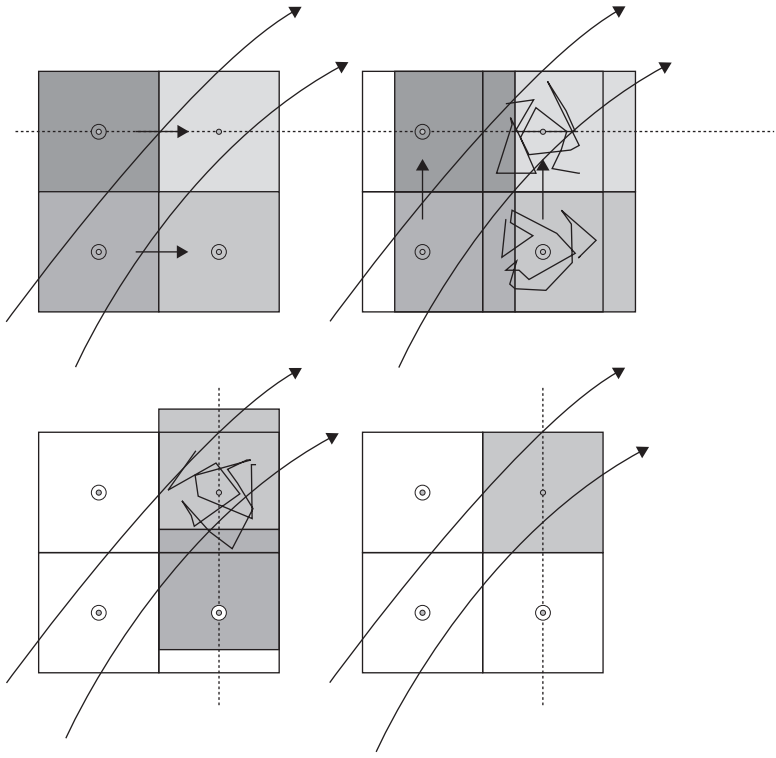


FIGURE 6.16 The splitting method uses two sequential 1D advection schemes. First, the signal is transported along the x-direction, and then the intermediate solution is advected along the y-direction. In this way, the information at upstream point $(i-1, j-1)$ is involved in the evolution of the value at (i, j) (case of positive velocity components).

This approach, alternating the order of the directional splitting, is a special case of the more general *Strang splitting* method designed to maintain second-order time accuracy when using time splitting (Strang, 1968).

The splitting approach thus seems attractive. It is not more complicated than applying two successive one-dimensional schemes. In the general case, however, attention must be given to the direction of the local flow so that "upwinding" consistently draws the information from upstream, whatever that direction may be. There is no other complication, and we now proceed with a test of the method with the TVD scheme. As Fig. 6.17 reveals, the result is a significant improvement with no increase of computational burden.

A more complicated case of advection can now be tried. For this, we choose an initially square distribution of tracer and place it in a narrow sheared flow (Fig. 6.18). We expect that the distribution will be distorted by the shear flow. To assess the quality of the advection scheme, we could try to obtain an analytical

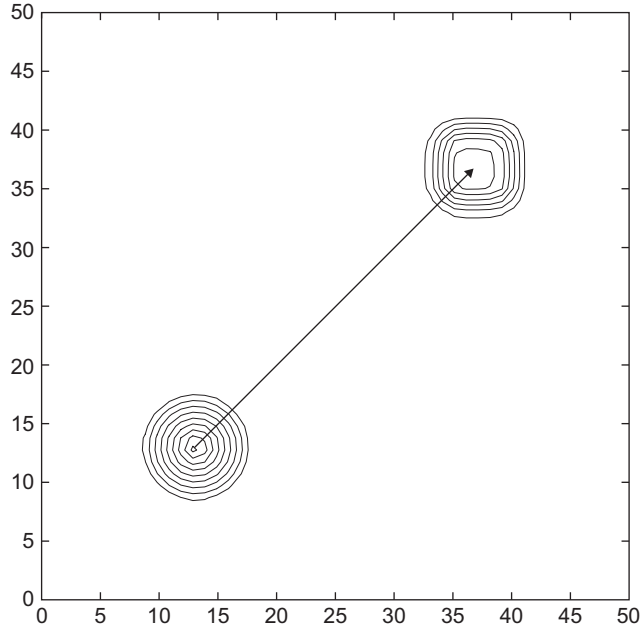


FIGURE 6.17 Advection at 45° of an initially conical distribution using the splitting method and the TVD scheme.

solution by calculating trajectories from the known velocity field, but a much simpler approach is to flip the sign of the velocity field after some time and continue the integration for an equal amount time. If the scheme were perfect, the patch would return to its original position and shape (without diffusion, the system is reversible and trajectories integrated forward and then backward should bring all particles back to their original position), but this won't be the case, and the difference between initial and final states is a measure of the error. Because some of the error generated during the flow in one direction may be negated during the return flow, we also need to consider the result at the moment of current reversal, i.e., the moment of farthest displacement.

For the method developed up to now, some degradation occurs, and bizarre results happen, even in regions of almost uniform flow (see Numerical Exercise 6.15). To discern the cause of this degradation, we first have to realize that the oblique advection test case is special in the sense that during a 1D step, the corresponding velocity is uniform. In the present case, the velocity during a substep is no longer uniform, and application of the first substep on a uniform field $\tilde{c} = \mathcal{C}$ yields

$$\frac{\tilde{c}^{\star}_{i,j} - \mathcal{C}}{\Delta t} + \frac{\tilde{u}_{i+1/2,j} - \tilde{u}_{i-1/2,j}}{\Delta x}\,\mathcal{C} = 0,$$

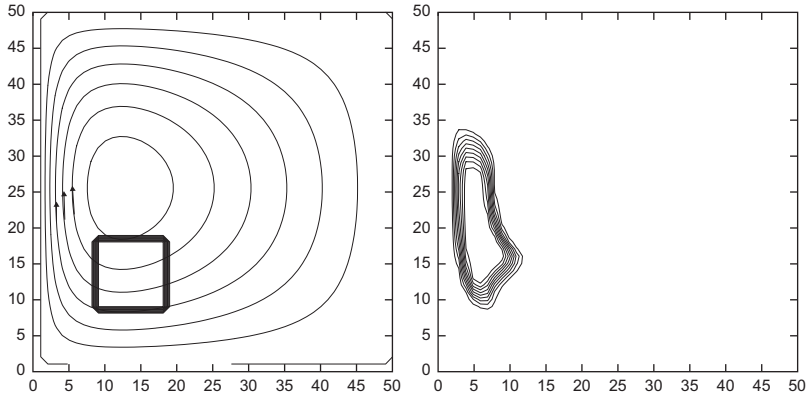


FIGURE 6.18 Advection of a square signal along a sheared boundary-layer current. *Left panel*: Initial distribution and streamlines. *Right panel*: After some advection. The distortion of the distribution is mostly due to the sheared current, which causes cross-stream squeezing and downstream stretching as the tracer enters the boundary layer.

which provides a value of $\tilde{c}^{\star}_{i,j}$ different from the constant $\mathcal{C}$. The next step is unable to correct this by returning the distribution back to a constant. The problem arises because the substep is not characterized by zero divergence of the 1D velocity field, and conservation of the tracer is not met. Conservation in the presence of a converging/diverging velocity in 1D is, however, encountered in another, physical problem: compressible flow. Mimicking this problem, we introduce a *pseudo-compressible approach*, which introduces a density-like variable ρ, to calculate the pseudo-mass conservation written as

$$\frac{\partial}{\partial t}(\rho)+\frac{\partial}{\partial x}(\rho u)+\frac{\partial}{\partial y}(\rho v)=0 \tag{6.71}$$

and the tracer budget as

$$\frac{\partial}{\partial t}(\rho c)+\frac{\partial}{\partial x}(\rho uc)+\frac{\partial}{\partial y}(\rho vc)=0. \tag{6.72}$$

The splitting method starts with a constant ρ during the first substep and yields

$$\frac{\rho^{\star}-\rho}{\Delta t}+\frac{\tilde{u}_{i+1/2,j}-\tilde{u}_{i-1/2,j}}{\Delta x}\,\rho=0$$

for the pseudo-mass equation and

$$\frac{\rho^{\star}\tilde{c}^{\star}_{i}-\rho\tilde{c}^{n}_{i}}{\Delta t}+\rho\mathcal{L}_1(\tilde{c}^{n}_{i})=0. \tag{6.73}$$

for the tracer concentration. In each calculation, the constant ρ is a multiplicative constant, which can be taken out of the advection operator $\mathcal{L}_1$. The second

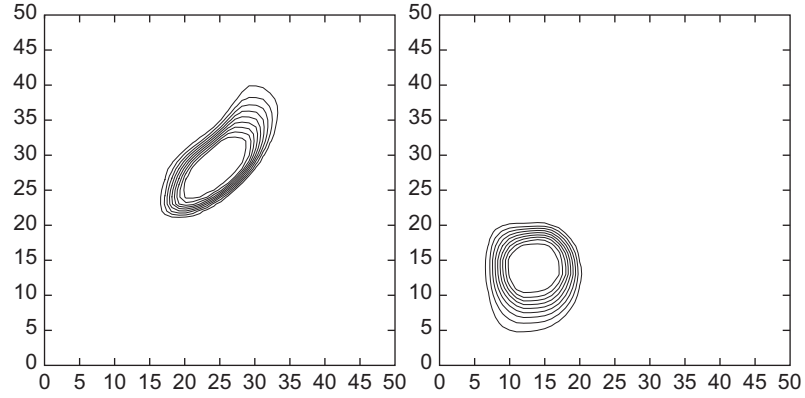


FIGURE 6.19 Advection with TVD scheme, Strang splitting, and pseudo-compressibility. The initial condition is as shown in the left panel of Fig. 6.18. *Left panel*: The patch of tracer at its furthest distance from the point of release, at the time of flow reversal. Its deformation is mostly physical and should ideally be undone during the return travel. *Right panel*: End state after return travel. The patch has nearly returned to its original location and shape, indicative of the scheme's good level of performance.

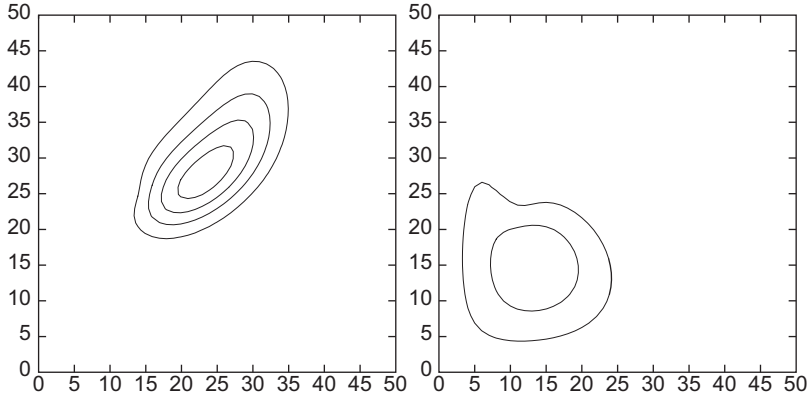


FIGURE 6.20 Same as Fig. 6.19 but with advection by the upwind scheme, using Strang splitting and including pseudo-compressibility. The situation at time of flow reversal (left panel) and after return (right panel) shows that numerical diffusion is clearly stronger than with the TVD scheme. The final distribution is hardly identifiable with the initial condition. The contour values are the same as in Fig. 6.19.

substep similarly follows with

$$\frac{\rho^{n+1}-\rho^{\star}}{\Delta t}+\frac{\tilde{v}_{i,j+1/2}-\tilde{v}_{i,j-1/2}}{\Delta y}\,\rho=0,$$
$$\frac{\rho^{n+1}\tilde{c}_i^{n+1}-\rho^{\star}\tilde{c}_i^{\star}}{\Delta t}+\rho\mathcal{L}_2(\tilde{c}_i^{\star})=0, \tag{6.74}$$

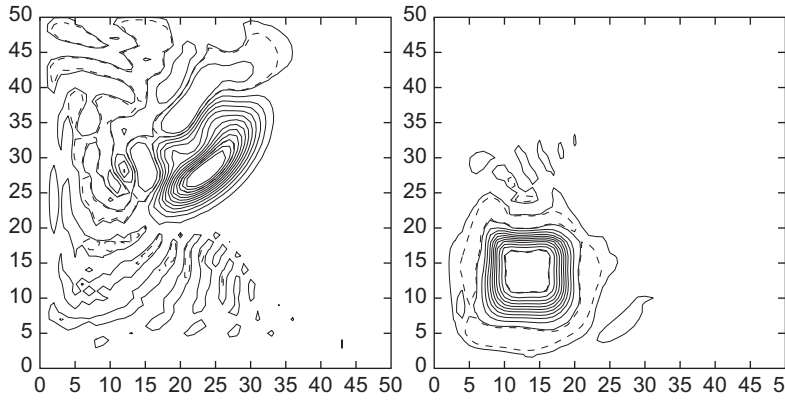


FIGURE 6.21 Same as Fig. 6.19 but with advection by the Lax-Wendroff scheme, using Strang splitting and including pseudo-compressibility. The situation at flow reversal (left panel) shows much dispersion, which is partly undone during the return travel (right panel). At the end, the distribution has been fairly well reconstructed but there is some undershooting around the edges and overshooting in the center. Dashed lines indicate values outside the initial range.

with the same constant ρ used again in the spatial operator. Setting the pseudo-density ρ^{n+1} equal to its previous value ρ, so that it disappears from the equations after a full time step, leads to the constraint that velocity is divergence-free in the sense of (6.58). If this is the case, when $\tilde{c}^n = \mathcal{C}$, it also guarantees $\tilde{c}^{n+1} = \mathcal{C}$.

Other splitting techniques have been devised (e.g., Pietrzak, 1998), but the approach remains essentially the same: pseudo-compression in one direction during the first substep, followed by a compensating amount of decompression during the second substep in the other direction.

With pseudo-compressibility, the sheared flow advection simulated with a flux limiter indicates that the scheme is quite accurate (Fig. 6.19), being both less diffusive than the upwind scheme (Fig. 6.20) and less dispersive than the Lax-Wendroff method (Fig. 6.21). The Matlab code `tvdadv2D.m` allows the reader to experiment with various strategies, by turning pseudo-compressibility on or off, enabling and disabling time splitting, and using different limiters in sheared and unsheared flow fields (Numerical Exercise 6.15).

ANALYTICAL PROBLEMS

6.1. Show that

$$c(x, y, t) = \frac{M}{4\pi \mathcal{A} t} \, \mathrm{e}^{-[(x-ut)^2+(y-vt)^2]/(4\mathcal{A}t)} \tag{6.75}$$

is the solution of the two-dimensional advection-diffusion equation with uniform velocity components u and v. Plot the solution for decreasing values of t and infer the type of physical problem the initial condition

is supposed to represent. Provide an interpretation of M. (*Hint*: Integrate over the infinite domain.)

6.2. Extend solution (6.75) to a radioactive tracer with decay constant K. (*Hint*: Look for a solution of similar structure but with one more exponential factor.)

6.3. Assuming a highly advective situation (high Peclet number), construct the 2D solution corresponding to the continuous release of a substance (S, in mass per time) from a punctual source (located at $x=y=0$) in the presence of velocity u in the x-direction and diffusion $\mathcal{A}$ in the y-direction.

6.4. An unreported ship accident results in an instantaneous release of a conservative pollutant. This substance floats along the sea surface and disperses for some time until it is eventually detected and measured. The maximum concentration, then equal to $c=0.1\ \mu g/m^2$, is found just West of the Azores at 38°30′N 30°00′W. A month later, the maximum concentration has decreased to $0.05\ \mu g/m^2$ and is located 200 km further South. Assuming a fixed diffusivity $\mathcal{A}=1000\ m^2/s$ and uniform steady flow, can you infer the amount of substance that was released from the ship, and the time and location of the accident? Finally, how long will it be before the concentration no longer exceeds $0.01\ \mu g/m^2$ anywhere?

6.5. Study the dispersion relation of the equation

$$\frac{\partial c}{\partial t}=\kappa\ \frac{\partial^p c}{\partial x^p} \tag{6.76}$$

where p is a positive integer. Distinguish between even and odd values of p. What should be the sign of the coefficient κ for the solution to be well behaved? Then, compare the cases $p=2$ (standard diffusion) and $p=4$ (biharmonic diffusion). Show that the latter generates a more scale-selective damping behavior than the former.

6.6. Explain the behavior found in Fig. 6.10 by finding the analytical solution of the corresponding physical problem (6.48).

6.7. In the interior of the Pacific Ocean, a slow upwelling compensating the deep convection of the high latitudes creates an average upward motion of about 5 m/year between depths of 4 km to 1 km. The average background turbulent diffusion in this region is estimated to be on the order of $10^{-4}\ m^2/s$. From the deep region, Radium ^{226}Ra found in the sediments, is brought up, while Tritium ^{3}H of atmospheric origin diffuses downward from the surface. Radium has a half-life (time for 50% decay) of 1620 years, and Tritium has a half-life of 12.43 years. Determine the steady-state solution using a one-dimensional vertical advection-diffusion model, assuming fixed and unit value of Tritium at the surface and zero at 4 km depth. For Radium, assume a unit value at depth of 4 km and

zero value at the surface. Compare solutions with and without advection. Which tracer is most influenced by advection? Analyze the relative importance of advective and diffusive fluxes for each tracer at 4 km depth and 1 km depth.

6.8. If you intend to use a numerical scheme with an upwind advection to solve the preceding problem for Carbon-14 ^{14}C (half-life of 5730 years), what vertical resolution would be needed so that the numerical calculation does not introduce an excessively large numerical diffusion?

NUMERICAL EXERCISES

6.1. Prove the assertion that a forward-in-time, central-in-space approximation to the advection equation is unconditionaly unstable.

6.2. Use `advleap.m` with different initialization techniques for the first time step of the leapfrog scheme. What happens if an inconsistent approach is used (e.g., zero values)? Can you eliminate the spurious mode totally by a clever initialization of the auxiliary initial condition c^1 when a pure sinusoidal signal is being advected?

6.3. Use the stability analysis under the form (5.31) using an amplification factor. Verify that the stability condition is $|C| \leq 1$.

6.4. Verify numerically that the leapfrog scheme conserves variance of the concentration distribution when $\Delta t \rightarrow 0$. Compare with the Lax-Wendroff scheme behavior for the same time steps.

6.5. Analyze the numerical phase speed of the upwind scheme. What happens for $C = 1/2$? Which particular behavior is observed when $C = 1$?

6.6. Design a fourth-order spatial difference and explicit time stepping for the 1D advection problem. What is the CFL condition of this scheme? Compare with the von Neumann stability condition. Simulate the standard advection test case.

6.7. Design a higher-order finite-volume approach by using higher-order polynomials to calculate the flux integrals. Instead of a linear interpolation as in the Lax-Wendroff scheme, use a parabolic interpolation.

6.8. Show that the von Neumann stability condition of the Beam-Warming scheme is $0 \leq C \leq 2$.

6.9. Implement the trapezoidal scheme with centered space difference using the tridiagonal algorithm `thomas.m`. Apply it to the standard problem of the top-hat signal advection and verify that you find the result shown in Fig. 6.22. Provide an interpretation of the result in terms of the numerical

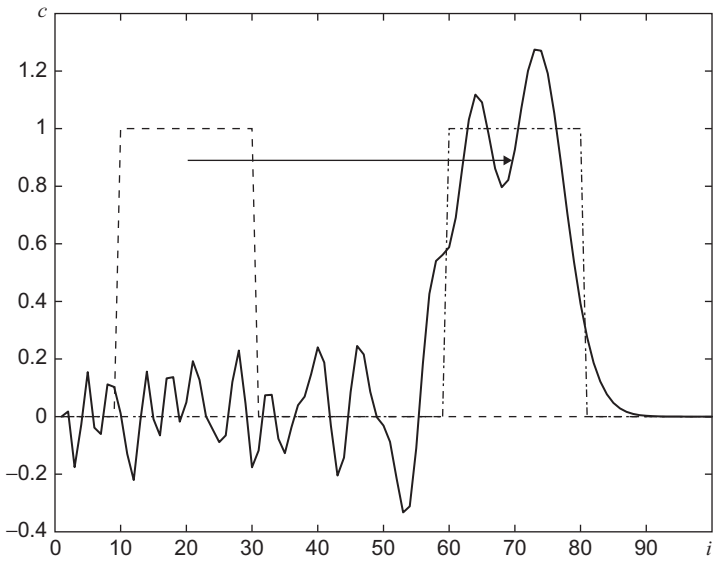


FIGURE 6.22 Standard test case with the trapezoidal scheme and centered spatial derivatives.

dispersion relation. Verify numerically that the variance is conserved exactly.

6.10. Show that the higher-order method for flux calculation at an interface using a linear interpolation on a nonuniform grid with spacing Δx_i between interfaces of cell i leads to the following flux, irrespective of the sign of the velocity

$$\tilde{q}_{i-1/2} = u\,\frac{(\Delta x_{i-1} - u\Delta t)\tilde{c}_i^n + (\Delta x_i + u\Delta t)\tilde{c}_{i-1}^n}{\Delta x_i + \Delta x_{i-1}}. \tag{6.77}$$

6.11. Use a leapfrog centered scheme for advection with diffusion. Apply it to the standard top-hat for different values of the diffusion parameter and interpret your results.

6.12. Find an explanation for why the $2\Delta x$ mode is stationary in all discretizations of advection. (*Hint*: Use a sinusoidal signal of wavelength $2\Delta x$ and zero phase, then sample it. Change the phase (corresponding to a displacement) by different values less than π and resample. What do you observe?)

6.13. Prove that (6.54) is the sufficient stability condition of scheme (6.51). (*Hint*: Rewrite $|\varrho|^2$ as $(\phi - 2C^2\xi)^2 + 4C^2\xi(1-\xi)$ and observe that as a

function of ϕ the amplification factor reaches its maxima at the locations of the extrema of ϕ, itself constrained by (6.54).)

6.14. Consider the one-dimensional advection-diffusion equation with Euler time discretization. For advection, use a centered difference with implicit factor α, and for diffusion the standard second-order difference with implicit factor β. Show that numerical stability requires $(1-2\alpha)C^2 \leq 2D$ and $(1-2\beta)D \leq 1/2$. Verify that, without diffusion, the explicit centered advection scheme is unstable.

6.15. Use `tvdadv2D.m` with different parameters (splitting or not, pseudo-mass conservation or not) under different conditions (sheared velocity field or solid rotation) and initial conditions (smooth field or strong gradients) with different flux limiters (upwind, Lax-Wendroff, TVD, etc.) to get a feeling of the range of different numerical solutions an advection scheme can provide compared with the analytical solution.

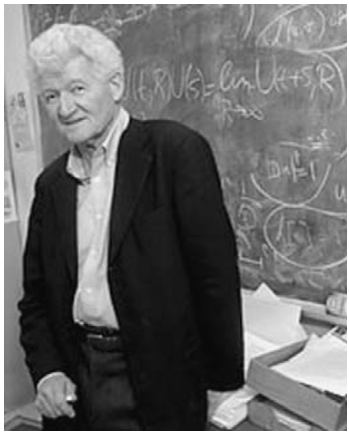
Richard Courant
1888–1972

Born in Upper Silesia, now in Poland but then part of Germany, Richard Courant was a precocious child and, because of economic difficulties at home, started to support himself by tutoring at an early age. His talents in mathematics led him to study in Göttingen, a magnet of mathematicians at the time, and Courant studied under David Hilbert with whom he eventually published in 1924 a famous treatise on methods of mathematical physics. In the foreword, Courant insists on the need for mathematics to be related to physical problems and warns against the trend of that time to loosen that link.

In 1928, well before the invention of computers, Richard Courant published with Kurt Friedrichs and Hans Lewy a most famous paper on the solution of partial difference equations, in which the now-called CFL stability condition was derived for the first time.

Courant left Germany for the United States, where he was offered a position at New York University. The Courant Institute of Mathematical Sciences at that institution is named after him. (*Photo from the MacTutor History of Mathematics archive at the University of St. Andrews*)

Peter David Lax
1926–



Born in Budapest (Hungary), Peter Lax quickly attracted attention for his mathematical prowess. His parents and he had barely moved to the United States, in December 1941 with the last ship from Lisbon during the war, and Peter was still in high school when he was visited in his home by John von Neumann (see biography at end of Chapter 5), who had heard about this outstanding Hungarian mathematician. After working on the top-secret Manhattan atomic bomb project in 1945–1946, he completed his first university degree in 1947 and obtained his doctorate in 1949, both at New York University (NYU). In his own words, "these were years of explosive growth in computing." Lax quickly gained a reputation for his work in numerical analysis.

Lax served as director of Courant Institute at NYU from 1972–1980 and was instrumental in getting the US Government to provide supercomputers for scientific research. (*Photo from the MacTutor History of Mathematics archive at the University of St. Andrews*)