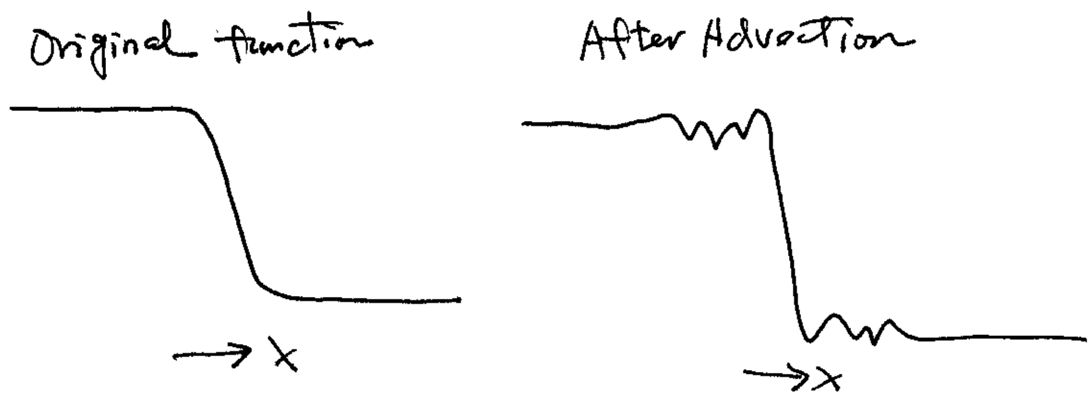
3.4. Monotonicity of Advection Schemes

3.4.1. Concept of Monotonicity

When numerical schemes are used to advect a monotonic function, e.g., a monotonically decreasing function of x, the numerical solutions do not necessarily preserve the mononotic property – in fact, most of the time they do not, and the errors tend to be large near sharp gradient. This is illustrated in the following:



A few example solutions are given below:

Sample Solutions to the Inviscid Burgers' Equation

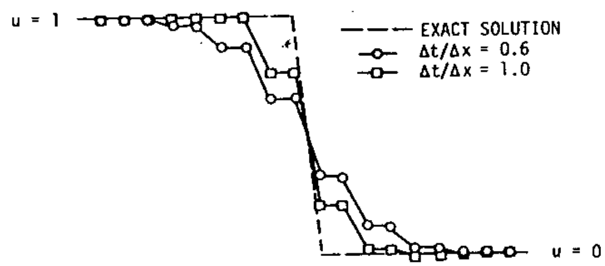


Figure 4-27 Numerical solution of Burgers' equation using Lax method.

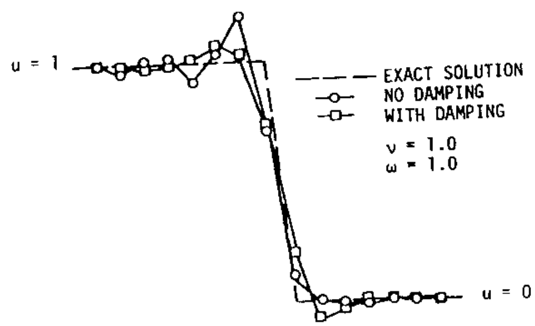
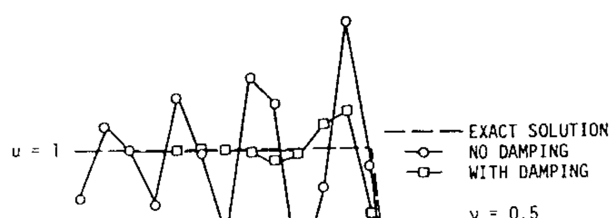


Figure 4-36 Solution for right-moving discontinuity time-centered implicit method, delta form.



Monotonic numerical schemes are ones which, given an initial distribution which is monotonic before advection, produce a monotonic distribution after advection.

A consequence of this property is that monotonic schemes neither create new extrema in the solution nor amplify existing extrema.

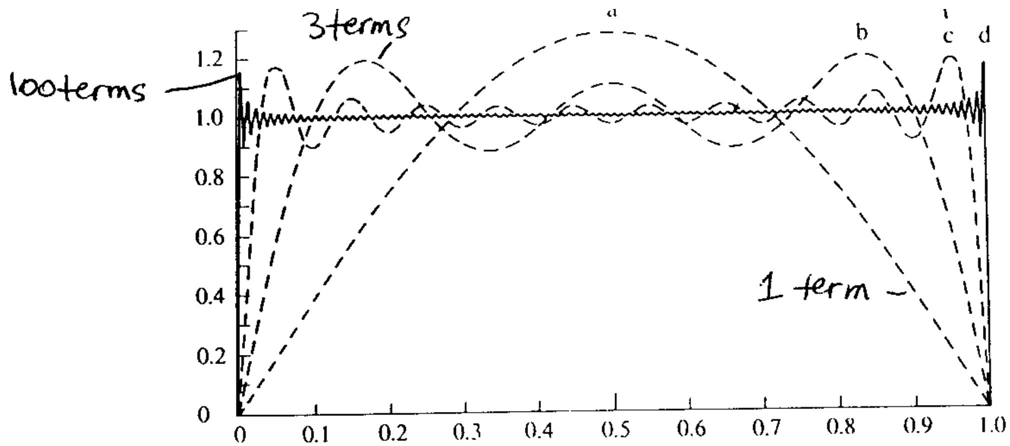
S.K. Godunov (1959) showed that no schemes having greater than first-order accuracy in space can be monotonic by construction (i.e., without using some artificial modification to ensure monotonicity). The highly dissipative upstream scheme is the classic example of a monotonic scheme.

Monotonic schemes are widely used in computational fluid dynamics because they do not allow the Gibbs Phenomenon to occur. This phenomenon results from attempting to represent a sharp gradient or discontinuity by a truncated number of waves, and always produces "undershoots and overshoots" relative to the amplitude of the initial distribution.

- These oscillations typically appear in the "wake" of a traveling wave which exhibits a sharp gradient, but do not necessarily grow in time.
- They are short waves that become noises in the solution – the damping of them results in smoothing of numerical solution.
- The oscillations can cause positive-definite fields, such as mass and water, to turn negative.

The Gibbs phenomenon is illustrated in the figure below, which shows how a square wave is represented by various numbers of waves in a Fourier expansion. Even if 100 terms are retained in the expansion, small over- and under-shoots remain. Monotonic schemes do not allow such oscillations to occur, i.e., one can think of the oscillations being removed by very selective damping.

Spectral methods use truncated spectral series to represent variable fields – they are particularly suspect to the Gibbs errors.



Monotonic schemes are often constructed by examining local features of the advected field, and adjust the advective fluxes of certain high-order schemes explicitly so that no new extrema is created in the solution.

3.4.2. Two basic classes of monotonic schemes

One is called the Flux-corrected transport (FCT) scheme, original proposed by Boris and Book (1973) and extended to multiple dimensions by Zalesak (1979).

With this scheme, the advective fluxes are essentially a weighted average of a lower-order monotonic scheme (usually 1st-order upwind) and a higher-order non-monotonic scheme. The idea is to use the high-order scheme as much as one can without violating the monotonicity condition. Details can also be found in Section 5.4 of Durran's book. In the ARPS, the FCT scheme is available as an option for scalar advection – it is three to four times as expensive as a regular 1st or second advection, however.

The other class is the so-called flux limiter method. With this method, the advective fluxes of a high-order scheme is directly modified (limited by a limiter) and the goal is that the total variation of the solution does not increase in time and this property is usually referred to as total variation diminishing (TVD).

The total variation of a function ϕ is defined as

$$TV(\phi) = \sum_{j-1}^{N-1} | \phi_{j+1} - \phi_j |$$

A TVD scheme ensures that $TV(\phi^{n+1}) \leq TV(\phi^n)$.

Sweby (1984) presented a systematic derivation of the flux limiter for this class (see also Durran Section 5.5.1).

With both methods, the flux correction or limiting is done grid point by grid point – in effect, the coefficients of the finite difference schemes are solution dependent therefore they are often called non-linear schemes.

Recommended Reading: Sections 5.2.1, 5.2.2., 5.3-5.5 of Durran.

Summarizing comments:

By now, you should have realized that no scheme is perfect, although some is better than the others. When we design or choose a scheme, we need to look at a number of properties, including accuracy (in terms of amplitude and phase), stability (implicit schemes tends to be more stable), computational complexity (implicit schemes cost more to solve per step), monotonicity (can we tolerate negative water generation?), and conservation properties etc. You need also consider the problem at hand – e.g., does it contain sharp gradient that is important to your solution? What is your target computer?
The computational and storage requirement are other factors to consider.

3.5. Multi-Dimensional Advection

Reading: Durran section 3.2.1. Smolarkievicz (1982 MWR).

Similar to the diffusion or heat transfer equations, there are three general approaches for solving multi-dimensional advection equations, namely:

1) Fully multi-dimensional methods
2) Direct extensions of 1-D schemes
3) Directional splitting methods

We will look at each in the following.

3.5.1. Direct Extension

Many 1-D advection schemes can be directly extended to multiple dimensions.

Multi-dimensional extension of 1-D explicit schemes often has a more restrictive stability condition.

We will look at the 2-D leapfrog centered scheme first.

For equation

$$\frac{\partial u}{\partial t} + c_x \frac{\partial u}{\partial x} + c_y \frac{\partial u}{\partial y} = 0 , \qquad (31)$$

the leapfrog centered discretization is

$$u_{m,j}^{n+1} - u_{m,j}^{n-1} = -\frac{c_x \Delta t}{\Delta x}(u_{m+1,j}^{n} - u_{m-1,j}^{n}) - \frac{c_y \Delta t}{\Delta y}(u_{m,j+1}^{n} - u_{m,j-1}^{n}) \tag{32}$$

$$\tau = O(\Delta x^2, \Delta y^2, \Delta t^2)$$

Let the individual wave component be

$$u_{m,j}^{n} = \lambda^n \exp[i(km\Delta x + lj\Delta y] \tag{33}$$

where k and l are wave number in x and y directions, respectively.

Substituting (33) into (32) and solve for λ, you can obtain (do it yourself):

$$\lambda_{\pm} = -i\left[\frac{c_x \Delta t}{\Delta x}\sin(k\Delta x) + \frac{c_y \Delta t}{\Delta y}\sin(l\Delta y)\right] \pm \left\{1 - \left[\frac{c_x \Delta t}{\Delta x}\sin(k\Delta x) + \frac{c_y \Delta t}{\Delta y}\sin(l\Delta y)\right]^2\right\}^{1/2} \tag{34}$$

Similar to the 1-D case, if

$$1 - \left[\frac{c_x \Delta t}{\Delta x}\sin(k\Delta x) + \frac{c_y \Delta t}{\Delta y}\sin(l\Delta y)\right]^2 \geq 0, \tag{35}$$

then $|\lambda_{\pm}| \equiv 1$, the scheme is stable (and has no amplitude error).

Inequality (35) is satisfied when

$$\left|\frac{c_x\Delta t}{\Delta x}\sin(k\Delta x)+\frac{c_y\Delta t}{\Delta y}\sin(l\Delta y)\right|\le 1. \qquad (36)$$

Let's consider the simpler case of $\Delta x = \Delta y = d$, and rewrite

$c_x = u_s \cos(\theta)$, $c_y = u_s \sin(\theta)$,

where u_s is the flow speed, (36) then becomes

$$\frac{u_s\Delta t}{d}\left|\cos(\theta)\sin(k\Delta x)+\sin(\theta)\sin(l\Delta y)\right|\le 1. \qquad (37)$$

Since we want (37) to be satisfied for <u>all possible</u> waves, we choose the most stringent case of sin ($k\Delta x$) = 1 and sin($l\Delta y$) = 1, (37) the becomes

$$\frac{u_s\Delta t}{d}\left|\cos(\theta)+\sin(\theta)\right|\le 1.$$

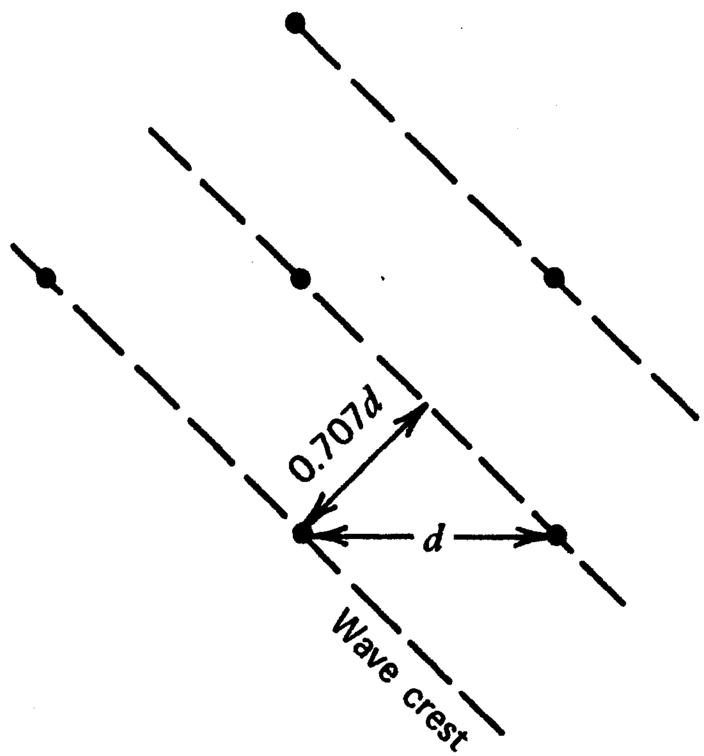
The maximum value of $\cos(\theta)+\sin(\theta)$ is $\sqrt{2}$ which occurs when $\theta = \pi/4$, the result is the stability condition for 2-D advection equation in the case of $\Delta x = \Delta y$:

$$\frac{u_s\Delta t}{d}\sqrt{2}\le 1 \quad \text{or} \quad \frac{u_s\Delta t}{d}\le 0.707 \qquad (38)$$

i.e., the Courant number has to be less than 0.707, instead of 1 as we get for 1-D case.

P.S. To find the max value of $\cos(\theta)+\sin(\theta)$, set $d[\cos(\theta)+\sin(\theta)]/d\theta = -\sin\theta + \cos\theta = 0$ → $\sin\theta = \cos\theta$ → $\tan\theta = 1$ → $\theta = \frac{\pi}{4}$ and $\frac{5\pi}{4}$. The maximum value occurs when $\theta = \frac{\pi}{4}$.

The reason that Δt has to be about 30% smaller is explained by the following diagram:



As seen from the figure, for a wave propagating from SW to NE, the effective distance between two grid points is d/$\sqrt{2}$ instead of d. A wave signal cannot propagate more than one (effective) grid interval with this explicit second-order leapfrog-centered scheme for stability.

Similar reduction of time step size occurs for most other explicit schemes, including the upwind scheme.

Note: For the three dimensional case, with the same FTCS scheme, the stability requirement is

$$\left| \frac{c_x \Delta t}{\Delta x} \sin(k\Delta x) + \frac{c_y \Delta t}{\Delta y} \sin(l\Delta y) + \frac{c_z \Delta t}{\Delta z} \sin(m\Delta z) \right| \le 1.$$

For the $\Delta x = \Delta y = \Delta z = d$ case, let

$$c_x = u_s \cos(\phi)\cos(\theta), \;\; c_x = u_s \cos(\phi)\sin(\theta), \;\; c_z = u_s \sin(\phi)$$

where θ and ϕ are, respectively, the azimuth and elevation angles of the velocity vector having speed u_s.

$$\frac{u_s \Delta t}{d} \left| \cos(\phi)\cos(\theta)\sin(k\Delta x) + \cos(\phi)\sin(\theta)\sin(l\Delta y) + \sin(\phi)\sin(m\Delta z) \right| \le 1.$$

Since we want the above satisfied for all possible waves, we choose the most stringent case of sin ($k\Delta x$) = 1 and sin($l\Delta y$) = 1 and sin($m\Delta z$) = 1, then the above condition becomes

$$\frac{u_s \Delta t}{d} \left[\cos(\phi)\cos(\theta) + \cos(\phi)\sin(\theta) + \sin(\phi) \right] \le 1.$$

The maximum possible value of $\cos\phi\cos\theta + \cos\phi\sin\theta + \sin\phi$ is $\sqrt{3}$ which occurs when

when $\theta = \pi/4$ and $\phi=\pi/6$.

P.S. To find the max value of $\cos\phi\cos\theta + \cos\phi\sin\theta + \sin\phi$, set

$$d[\cos\phi\cos\theta + \cos\phi\sin\theta + \sin\phi]/d\theta = \cos\phi[-\sin\theta + \cos\theta] = 0$$
$$d[\cos\phi\cos\theta + \cos\phi\sin\theta + \sin\phi]/d\phi = -\sin\phi[\cos\theta + \sin\theta] + \cos\phi = 0$$

or

$$\cos\phi[-\sin\theta + \cos\theta] = 0$$
$$\tan\phi[\cos\theta + \sin\theta] = 1$$

They are satisfied when $\theta = \dfrac{\pi}{4}$ and $\tan\phi = \sqrt{2}/2$ and the maximum value is

$$\cos\phi[\cos\theta + \sin\theta] + \sin\phi = \sqrt{\frac{1}{1+\tan^2\phi}}[\cos\theta + \sin\theta] + \sqrt{1 - \frac{1}{1+\tan^2\phi}}.$$
$$= \sqrt{\frac{1}{1+1/2}}\sqrt{2} + \sqrt{1 - \frac{1}{1+1/2}} = \sqrt{\frac{2}{3}}\sqrt{2} + \sqrt{\frac{1}{3}} = \frac{2+1}{\sqrt{3}} = \sqrt{3}$$

Such a situation occurs when the wind vector points from near lower left corner of a cube to the far upper right corner of the cube. The effective grid spacing is the distance from the near lower left corner at (0, 0, 0) to a plane cutting through (1, 0, 0), (0, 1, 0) and (0, 0, 1).

3.5.2. Fully Multi-Dimensional Method

Not all direct extensions of 1-D schemes are stable, unfortunately.

Consider the Lax-Wendroff (also called Crowley) scheme we derived earlier using both second-order interpolation method (section 2.2.3 of Chapter 2):

1-D Lax-Wendroff or Crowley scheme:

$$\frac{u_i^{n+1}-u_i^n}{\Delta t}=-c\frac{u_{i+1}^n-u_{i-1}^n}{2\Delta x}+\frac{c^2\Delta t}{2}\frac{u_{i+1}^n-2u_i^n+u_{i-1}^n}{(\Delta x)^2} \qquad (39)$$

The scheme is stable when $|\mu|\le 1$.

Using the notation of finite-difference operators, (39) becomes

$$u_i^{n+1}=u_i^n-c\Delta t\delta_{2x}u^n+\frac{(c\Delta t)^2}{2}\delta_{xx}u^n. \qquad (40)$$

Direct extension of (40) into 2-D is:

$$u_i^{n+1}=u_i^n-\Delta t(c_x\delta_{2x}u^n+c_y\delta_{2y}u^n)+\frac{(c_x\Delta t)^2}{2}\delta_{xx}u^n+\frac{(c_y\Delta t)^2}{2}\delta_{yy}u^n. \qquad (41)$$

It turns out that (41) is <u>absolutely unstable</u>. This is because the cross-derivative terms are neglected!

To see it, we need to go back to original derivation of the Lax-Wendroff scheme:

$$u^{n+1} = u^n + \Delta t u_t + \frac{1}{2}(\Delta t)^2 u_{tt} + O(\Delta t^3) \qquad (42)$$

Use $u_t = -c_x u_x - c_y u_y$

and $u_{tt} = -c_x u_{tx} - c_y u_{ty} = c_x^2 u_{xx} + c_y^2 u_{yy} + 2c_x c_y u_{xy}$,

and replace the spatial derivatives with the corresponding finite differences, (42) becomes

$$u_i^{n+1} = u_i^n - \Delta t(c_x \delta_{2x} u^n + c_y \delta_{2y} u^n) + \frac{(c_x \Delta t)^2}{2} \delta_{xx} u^n + \frac{(c_y \Delta t)^2}{2} \delta_{yy} u^n + (c_x c_y \Delta t^2) \delta_{xy} u^n \qquad (43)$$

Clearly, the last term on the RHS is additional, compared to (41).

Note that we can also obtain (43) using the characteristics method plus quadratic interpolation, as long as all terms in the second-order 2-D polynomial are retained.

Equation (43) is an example of fully multidimensional scheme, which is different from the direct extension of 1-D counterpart.

Smolarkiewicz (1982 MWR) discuss the MD Crowley scheme in details (handout).

3.5.3. Directional Splitting

It turned out that by using directional splitting method (i.e., applying 1-D scheme in one direction at a time), the effect of cross-derivative terms can also be retained and a stable scheme result, with the Lax-Wendroff scheme.

The algorithms is

$$(u_{i,j}^{n+1})^* = u_{i,j}^n - c_x \Delta t \delta_{2x} u_{i,j}^n + \frac{(c_x \Delta t)^2}{2} \delta_{xx} u_{i,j}^n \quad (44a)$$

$$u_{i,j}^{n+1} = (u_{i,j}^{n+1})^* - c_y \Delta t \delta_{2y} (u_{i,j}^{n+1})^* + \frac{(c_y \Delta t)^2}{2} \delta_{yy} (u_{i,j}^{n+1})^* \quad (44b)$$

In this case, we preserve the stability of each step and $\lambda = \lambda_x \lambda_y$.

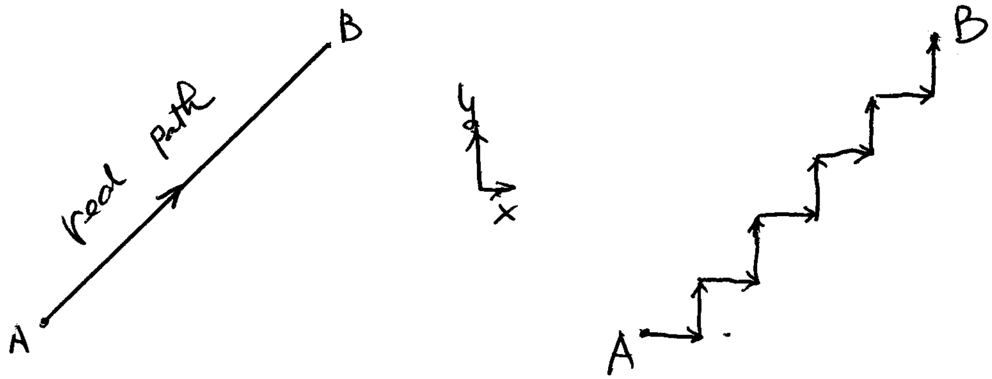
With the above scheme, we have

Advantages:

1. 1-D advection is straightforward – properties of schemes are well understood.
2. The time step constraint is not as severe as for true multi-dimensional problems.

Disadvantages:

1. We implicitly assume that features move obliquely to the grid may be represented as a series of orthogonal steps in the coordinate directions:



In an implicit scheme, where the time step can be large, these errors can be substantial.