Lecture 1

Space-time discretization Edit

We have so far analyzed time and space discretizations while respectively treating the complimentary dimension as continuous. Now we will consider how to discretize both space and time at once. In general, the methods we have outlined so far will work in conjunction. For example, in section **??**, the advection equation was discretized using second, fourth and sixth order differences. Since the motions are oscillatory we know that we need to use a time-stepping scheme that is [conditionally] stable for wave-like motions, such as leap-frog, Huen or Runge-Kutta.

Here, we will analyze particular combinations of space-time discretizations that yield new properties. You can try out various combinations of spatial and temporal discretization using the routines here.

1.1 Forward in Time, Upwind in Space Edit

The upwind scheme uses a side-difference in space biased in the upwind direction so that for positive flow ($c > 0$) the scheme is:

$$\frac{\theta_i^{n+1} - \theta_i^n}{\Delta t} + \frac{c}{\Delta x}\left(\theta_i^n - \theta_{i-1}^n\right) = 0 \tag{1.1}$$

An example here illustrates the evolution of a θ field with this approach. The final picture is a "Hovmöller diagram" showing the amplitude as a function of space and time.

To find the dispersion relation for the numerical solution we substitute in a wave solution of the form $e^{-(\lambda+i\omega)t+ikx}$:

$$e^{-(\lambda+i\omega)\Delta t} = 1 - C\left(1 - e^{-ik\Delta x}\right)$$

where $C = \frac{c\Delta t}{\Delta x}$ is the Courant number. Separating the imaginary and real components gives:

$$\begin{aligned} e^{-\lambda\Delta t}\sin\omega\Delta t &= C\sin k\Delta x \\ e^{-\lambda\Delta t}\cos\omega\Delta t &= 1 - C + C\cos k\Delta x \end{aligned}$$

Solving for $e^{\lambda\Delta t}$ and $\tan\omega\Delta t$ gives:

$$\begin{aligned} \tan\omega\Delta t &= \frac{C\sin k\Delta x}{1 - C(1-\cos k\Delta x)} \\ &= \frac{2C\sin\frac{k\Delta x}{2}\cos\frac{k\Delta x}{2}}{1 - 2C\sin^2\frac{k\Delta x}{2}} \end{aligned}$$

and

$$\begin{aligned} e^{-2\lambda\Delta t} &= (1 - C(1-\cos k\Delta x))^2 + (C\sin k\Delta x)^2 \\ &= 1 - 4C(1-C)\sin^2\frac{k\Delta x}{2} \end{aligned}$$

Since $\sin^2\frac{k\Delta x}{2}$ varies between 0 (for long waves) and 1 (for the grid-scale waves) stability depends on the sign of the quantity $4C(1-C)$; if either $C < 0$ or $C > 1$ then $4C(1-C) < 0$ and the solution grows with time. Therefore stability is conditional on:

$$0 \leq 4C(1-C) \leq 1$$

or simply $0 \leq C \leq 1$. The strongest damping occurs at $C = 1/2$ which maximizes $4C(1-C)$. Since $C < 0$ is unstable, we can infer that the downwind difference scheme is unconditionally unstable.

The waves are dispersive since $\tan\omega\Delta t$ depends on k. If $C < 1/2$ then the frequency is less than k so that the scheme is decelerating and if $C > 1/2$ then the frequency is either larger than k or changes sign. $C = 1/2$ is a special point because the denominator becomes $1 - \cos^2\frac{k\Delta x}{2}$ so that the whole expression becomes $2C\tan\frac{k\Delta x}{2}$ and the frequency is exactly correct.

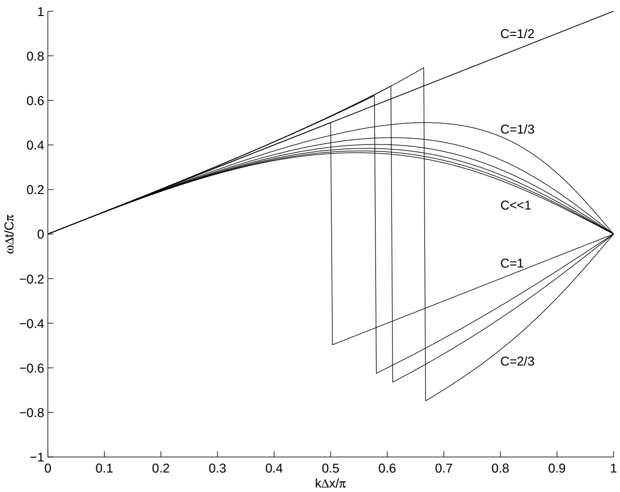

Figure 1.1: The frequency of the FTUS (or "upwind") scheme for various Courant number $C = \frac{\Delta t c}{\Delta x}$. $C = 1/2$ falls on $w = k$ which is the frequency of the continuum.

The upwind scheme also exhibits a special property of preserving extrema. This can be seen by re-arranging the difference equation for the future unknown value:

$$\theta_i^{n+1} = C\theta_{i-1}^n + (1-C)\theta_i^n$$

This is simply a linear interpolation between θ_{i-1}^n and θ_i^n with the C being the sliding parameter. Hence, θ_i^{n+1} must fall on or between θ_{i-1}^n and θ_i^n so long as $0 \leq C \leq 1$.

We know from section **??** that both the forward in time and side difference are both of first order accuracy. We also learnt in section **??** that the forward scheme is generally unstable when used for oscillatory motions. It is therefore some surprise the scheme is stable at all. One way of looking at why the scheme works is to express it as a FTCS (forward in time, centered in space) scheme with a specific amount of diffusion:

$$\theta^{n+1} - \theta^n = -C\left(\frac{\theta_{i+1}^n - \theta_{i-1}^n}{2} - \frac{|C|}{C}\frac{\theta_{i+1}^n - 2\theta_i^n + \theta_{i-1}^n}{2}\right) \quad (1.2)$$

$$= -C\delta_i\overline{\theta^n}^i + \frac{|C|}{2}\delta_{ii}\theta^n \quad (1.3)$$

In this form, the upwind scheme appears as a space centered derivative but with a diffusion term with diffusion coefficient $\frac{|C|}{2}\frac{\Delta x^2}{\Delta t}$ that is required and sufficient enough to make the scheme stable.

As a final note, advection schemes are often most useful when written in flux form (i.e. as a divergence of a flux). The last form allows us to write the scheme:

$$\frac{\theta^{n+1}-\theta^n}{\Delta t} = -\frac{1}{\Delta x}\delta_i F$$

where F is defined as

$$F = c\overline{\theta^n}^i - \frac{|c|}{2}\delta_i\theta^n$$

which a general way to write the upwind flux.

1.2 The Lax-Wendroff Method Edit

In a similar vain, the Lax-Wendroff method adds diffusion to the FTCS scheme:

$$\frac{\theta^{n+1}-\theta^n}{\Delta t} = -\frac{1}{\Delta x}\delta_i\left(c\overline{\theta^n}^i - \frac{cC}{2}\delta_i\theta^n\right) \qquad (1.4)$$

where the diffusion term has an implied diffusion coefficient equal to $c^2\Delta t/2$. That is, the last term is an approximation to:

$$\frac{c^2\Delta t}{2}\partial_{xx}\theta$$

Although the scheme is written as a forward difference in time, it is in fact second order accurate in time and space; the truncation error of the forward difference on the LHS is:

$$\frac{\Delta t}{2}\partial_{tt}\theta = \frac{\Delta t}{2}\partial_t\left(-c\partial_x\theta\right) = \frac{c^2\Delta t}{2}\partial_{xx}\theta$$

The diffusion term therefore cancels the leading truncation error from the forward time difference. This is an example of how treating time and space together leads to a substantially different scheme than would be obtained by discretizing the dimensions independently. The example here shows the relatively weak numerical diffusivity, but also some dispersion. Try halving both the time and space steps (doubling N).

An alternative way of accessing the second order nature of the Lax-Wendroff scheme is to break it down into a two stage scheme:

$$\theta^{*n+\frac{1}{2}} = \overline{\theta^n}^i - \frac{\Delta t}{2}\frac{c}{\Delta x}\delta_i\theta^n \tag{1.5}$$

$$\theta^{n+1} = \theta^n - \frac{\Delta t c}{\Delta x}\delta_i\theta^{*n+\frac{1}{2}} \tag{1.6}$$

Here, the time marching looks likes the mid-point second order Runge-Kutta method and the mid-point values are staggered in space.

The dispersion relation for the Lax-Wendroff scheme is:

$$\tan\omega\Delta t = \frac{2C\sin\frac{k\Delta x}{2}\cos\frac{k\Delta x}{2}}{1-2C^2\sin^2\frac{k\Delta x}{2}}$$

and amplification:

$$e^{-2\lambda\Delta t} = 1 - 4C^2(1-C^2)\sin^2\frac{k\Delta x}{2}$$

These expressions both take a similar form to those of the FTUS scheme except for the second order dependence on C. However, the Lax-Wendroff scheme does not conserve extrema like the FTUS method does.

1.3 Flux limiters Edit

The FTUS and Lax-Wendroff methods each have their advantages and disadvantages; the FTUS is only first order accurate and very diffusive but does conserve extrema while the Lax-Wendroff scheme doesn't conserve extrema but is second order accurate. Our objective in this section is to try to blend these two schemes, capturing the desired features of each.

First, we write the system in flux form:

$$\frac{1}{\Delta t}\left(\theta_i^{n+1} - \theta_i^n\right) = -\frac{1}{\Delta x}(F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}})$$

And now we cast the advective flux, F, as some unknown combination of the upwind flux, F^{US}, and Lax-Wendroff flux, F^{LW}:

$$F_{i+\frac{1}{2}} = \psi_{i+\frac{1}{2}}F^{LW} + (1-\psi_{i+\frac{1}{2}})F^{US}$$

$$F^{US}_{i+\frac{1}{2}} = c\theta_i$$

$$F^{LW}_{i+\frac{1}{2}} = c\theta_i + \frac{c(1-C)}{2}(\theta_{i+1} - \theta_i)$$

so that the advective flux is:

$$F_{i+\frac{1}{2}} = c\theta_i + \frac{c(1-C)}{2}\psi_{i+1/2}(\theta_{i+1} - \theta_i)$$

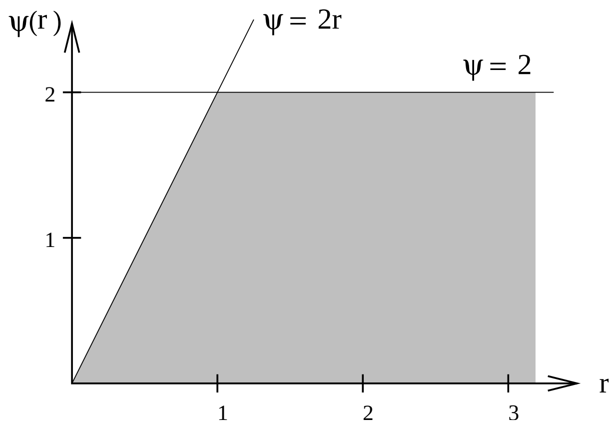

Figure 1.2: The two lines indicate the bounds on the limiter function, $\psi(r)$ given by the constraints for the scheme to be TVD (total variance diminishing).

The factor $\psi_{i+\frac{1}{2}}$ is some function, yet to be determined. In some texts, this form of the flux is justified by casting the Lax-Wendroff as above and arguing that the second term is a correction to the upwind flux and hence adjustable. Substituting into the prognostic equation gives:

$$\theta_i^{n+1} = \theta_i^n - \left(C - \frac{C(1-C)}{2}\psi_{i-\frac{1}{2}}\right)(\theta_i^n - \theta_{i-1}^n) - \frac{C(1-C)}{2}\psi_{i+\frac{1}{2}}(\theta_{i+1}^n - \theta_i^n)$$

The last term can be re-written as:

$$\frac{C(1-C)}{2}\psi_{i+\frac{1}{2}}(\theta_{i+1}^n - \theta_i^n) = \frac{C(1-C)}{2}\frac{\psi_{i+\frac{1}{2}}}{r_{i+\frac{1}{2}}}(\theta_i^n - \theta_{i-1}^n)$$

where $r_{i+\frac{1}{2}}$ is the slope ratio which is defined as:

$$r_{i+\frac{1}{2}} = \frac{(\theta_i^n - \theta_{i-1}^n)}{(\theta_{i+1}^n - \theta_i^n)}$$

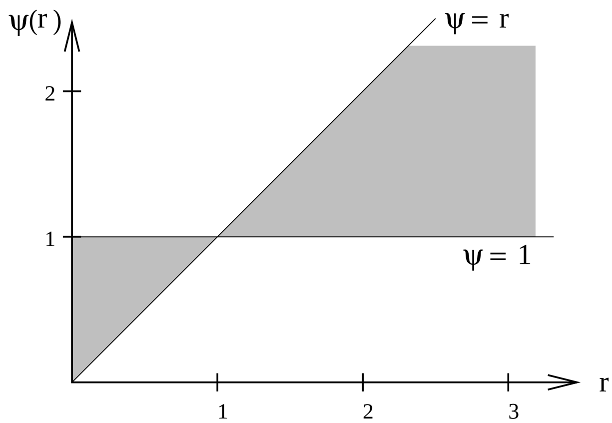

Figure 1.3: The two lines indicate the special limiter functions, $\psi(r) = 1$, which yields the Lax-Wendroff flux, and $\psi(r) = r$ which yields the Warming and Beam flux. Since both these schemes are of second order, any linear combination of these schemes is also of second order. The shaded region between them is thus the space where $\psi(r)$ will yield a second order scheme.

Now we will re-write the prognostic equation using the last expression:

$$\theta_i^{n+1} = \theta_i^n - C\left[1 - \frac{(1-C)}{2}\psi_{i-\frac{1}{2}} + \frac{(1-C)}{2}\frac{\psi_{i+\frac{1}{2}}}{r_{i+\frac{1}{2}}}\right](\theta_i^n - \theta_{i-1}^n)$$

which looks likes an FTUS (upwind) scheme but with a modified Courant number. The upwind scheme is both monotone and stable if the "effective" Courant number is both non-negative and less than one:

$$0 \leq C\left[1 - \frac{(1-C)}{2}\psi_{i-\frac{1}{2}} + \frac{(1-C)}{2}\frac{\psi_{i+\frac{1}{2}}}{r_{i+\frac{1}{2}}}\right] \leq 1$$

or

$$\frac{-2}{1-C} \leq \frac{\psi_{i+\frac{1}{2}}}{r_{i+\frac{1}{2}}} - \psi_{i-\frac{1}{2}} \leq \frac{2}{C}$$

Since $C \geq 0$ then the above is satisfied if the following stronger constraint is satisfied:

$$\left|\frac{\psi_{i+\frac{1}{2}}}{r_{i+\frac{1}{2}}} - \psi_{i-\frac{1}{2}}\right| \leq 2$$

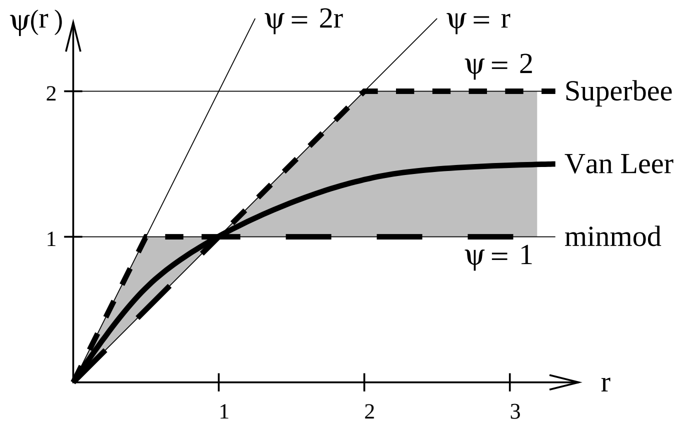

Figure 1.4: The shaded region is the intersection between the TVD region and region of second order accuracy. The short dashed line is the Superbee limiter, the long dash line is the minmod limiter and the solid curve is the Van Leer limiter.

Now we will allow the "limiter function" ψ be a function of the slope ratio:

$$\psi = \psi(r)$$

If $r < 0$, the slope must have changed sign and indicates a local extrema. In this instance we should limit the advective flux to take the form of the upwind flux since it is the only (linear) scheme capable of conserving extrema. Therefore, for $r < 0$ we set $\psi(r) = 0$.

If $r > 0$, the above inequality is satisfied when

$$0 \leq \frac{\psi(r)}{r} \leq 2 \quad \text{and} \quad 0 \leq \psi(r) \leq 2$$

and we must find functions that meet these criteria if the new scheme is to conserve extrema. The limits on $\psi(r)$ are indicated by the shaded region in Fig. 1.2. This region is said to be TVD (total variance diminishing) which means that the norm of gradients of a field can not be increased by the scheme. It happens that the FTUS scheme is TVD. Although we haven't directly use the TVD definition or cast the constraints as associated with TVD behaviour we have nevertheless derived constraints on a non-linear flux limiter that will make it TVD.

A further constraint on the limiter function is the preferred order of accuracy of the resulting scheme. We don't show the proof here but if the scheme can be shown to be an interpolation of the Lax-Wendroff method ($\psi(r) = 1$) and Warming and Beam method (which corresponds to $\psi(r) = r$) then the scheme will be of second order accuracy. $\psi(r)$ must therefore fall in the region indicated in Fig. 1.3 to be second order accurate. Finally, if $\psi(r)$ is chosen to fall in the intersection of these two regions, second order and TVD, then the resulting scheme will be both TVD and second order accurate.

There are many possible limiters but the most widely used are:

- **Superbee**: $\psi(r) = \max(0, \min(1, 2r), \min(2, r))$ due to Roe, 1985,
- **minmod**: $\psi(r) = \max(0, \min(1, r))$ due to who?,
- **Van Leer**: $\psi(r) = \frac{r+|r|}{1+|r|}$ due to Van Leer, 1974.

each of which is shown in Fig. 1.4.The example here shows the Superbee limiter; you can edit the file to look at the others.

A comparison of many different advection schemes is illustrated in Figs. 1.5 and 1.6. We have grouped the schemes in the following way: a) linear schemes with upwind bias (odd order), b) linear space centered schemes, c) second-order flux limited schemes (as discussed above) and d) higher order flux limited schemes. The reason for the grouping is the common behaviour within each group. The upwind biased schemes are smoother than the space-centered schemes. All the linear schemes, except for the upwind scheme, have false extrema. The flux limited schemes are monotone. The second order flux limited schemes tend to have reduced amplitude of extrema due to diffusion. The higher order flux limited methods can correct this.

1.4 Modified equations Edit

In section 1.2 we implicitly made use of a "modified equation". Modified equations provide a way of anticipating the behaviour of a numerical method by examing terms implied by a particular approximation. For a given difference equation, the corresponding modified equation is a continuous equation for which the same difference equation is a higher order approximation! For example, the difference equation

$$\frac{1}{\Delta t}(\theta_i^{n+1} - \theta_i^n) + \frac{c}{\Delta x}(\theta_i^n - \theta_{i-1}^n) = 0 \tag{1.7}$$

is the F.T.U.S. scheme (same as equation 1.1), which is a first order approximation in time and space to

$$\partial_t \theta + c\partial_x \theta = 0.$$

However, it is also a second order in time and space to the continuous equation

$$\partial_t \theta + \frac{\Delta t}{2}\partial_{tt}\theta + c\partial_x \theta - \frac{c\Delta x}{2}\partial_{xx}\theta = 0. \tag{1.8}$$

This is because the $O(\Delta t)$ trunction term arising from

$$\frac{1}{\Delta t}(\theta_i^{n+1} - \theta_i^n) = \partial_t \theta + \frac{\Delta t}{2!}\partial_{tt}\theta + \frac{\Delta t^2}{3!}\partial_{ttt}\theta + \ldots$$

exactly matches the second term in (1.8) leaving the $O(\Delta t^2)$ as the time-truncation error. Similarly, the $O(\Delta x)$ truncation term arising from the side difference

$$\frac{c}{\Delta x}(\theta_i^n - \theta_{i-1}^n) = c\partial_x \theta - c\frac{\Delta x}{2!}\partial_{xx}\theta + c\frac{\Delta x^2}{3!}\partial_{xxx}\theta - \ldots$$

exactly matches the last term in (1.8).

We can eliminate the second order time derivative from (1.8) by differentiating the governing equation:$\partial_{tt}\theta = c^2\partial_{xx}$. Thus, the first order approximation to the advection equation (1.7) is also a second order approximation to the modified equation

$$\partial_t \theta + c\partial_x \theta - \frac{c\Delta x}{2}(1 - \frac{c\Delta t}{\Delta x})\partial_{xx}\theta = 0.$$

This allows us to interpret the F.T.U.S. scheme; the modified equation differes from the governing equation by a diffusive term with diffusivity

$$\frac{c\Delta x}{2}(1 - \frac{c\Delta t}{\Delta x}).$$

This effective diffusivity is propotional to the flow c, it decreases with decreasing Δx (i.e. higher resolution is less diffusive) and also decreases with the Courant number $c\Delta t/\Delta x$. This is consistent with the special case of $c\Delta t/\Delta x = 1$ for which the F.T.U.S. scheme gives the exact answer.

Now we'll take the approach a step further and derive a third order modified equation. To get there we will use the following relations which are simply repeated differentiations of the *second-order modified equation* (1.8):

$$\begin{aligned}
\partial_{xt}\theta &= -c\partial_{xx}\theta + \frac{c\Delta x}{2}(1-C)\partial_{xxx}\theta \\
\partial_{xxt}\theta &= -c\partial_{xxx}\theta + \frac{c\Delta x}{2}(1-C)\partial_{xxxx}\theta = -c\partial_{xxx}\theta + O(\Delta x) \\
\partial_{tt}\theta &= -c\partial_{xt}\theta + \frac{c\Delta x}{2}(1-C)\partial_{xxt}\theta \\
&= -c\left(-c\partial_{xx}\theta + \frac{c\Delta x}{2}(1-C)\partial_{xxx}\theta\right) + \frac{c\Delta x}{2}(1-C)\left(-c\partial_{xxx}\theta\right) + O(\Delta x^2) \\
&= c^2\partial_{xx}\theta - c^2\Delta x(1-C)\partial_{xxx}\theta + O(\Delta x^2) \\
\partial_{xtt}\theta &= -c\partial_{xxt}\theta + O(\Delta x) = c^2\partial_{xxx} + O(\Delta x) \\
\partial_{ttt}\theta &= -c\partial_{xtt}\theta + O(\Delta x) = -c^3\partial_{xxx} + O(\Delta x)
\end{aligned}$$

where $C = \frac{c\Delta t}{\Delta x}$.

Now we substitute in for the first and second order Taylor series terms in the F.T.U.S scheme:

$$\begin{aligned}
&\frac{1}{\Delta t}(\theta_i^{n+1} - \theta_i^n) + \frac{c}{\Delta x}(\theta_i^n - \theta_{i-1}^n) \\
&\quad= \partial_t\theta + c\partial_x\theta + \frac{\Delta t}{2!}\partial_{tt}\theta - c\frac{\Delta x}{2!}\partial_{xx}\theta + \frac{\Delta t^2}{3!}\partial_{ttt}\theta + c\frac{\Delta x^2}{3!}\partial_{xxx}\theta + \ldots \\
&\quad= \partial_t\theta + c\partial_x\theta - \frac{c\Delta x}{2}(1-C)\partial_{xx}\theta + \frac{c\Delta x^2}{6}(1-C)(1-2C)\partial_{xxx}\theta \qquad (1.9)
\end{aligned}$$

keeping only $O(\Delta x^2)$ terms. This is the modified equation to which (1.7) is an $O(\Delta x^3)$ approximation. The second term is diffusive as we saw before with the second order modified equation. The last term in (1.9) causes waves to be dispersive (it leads to a $-ik^3$ term in the dispersion relation).

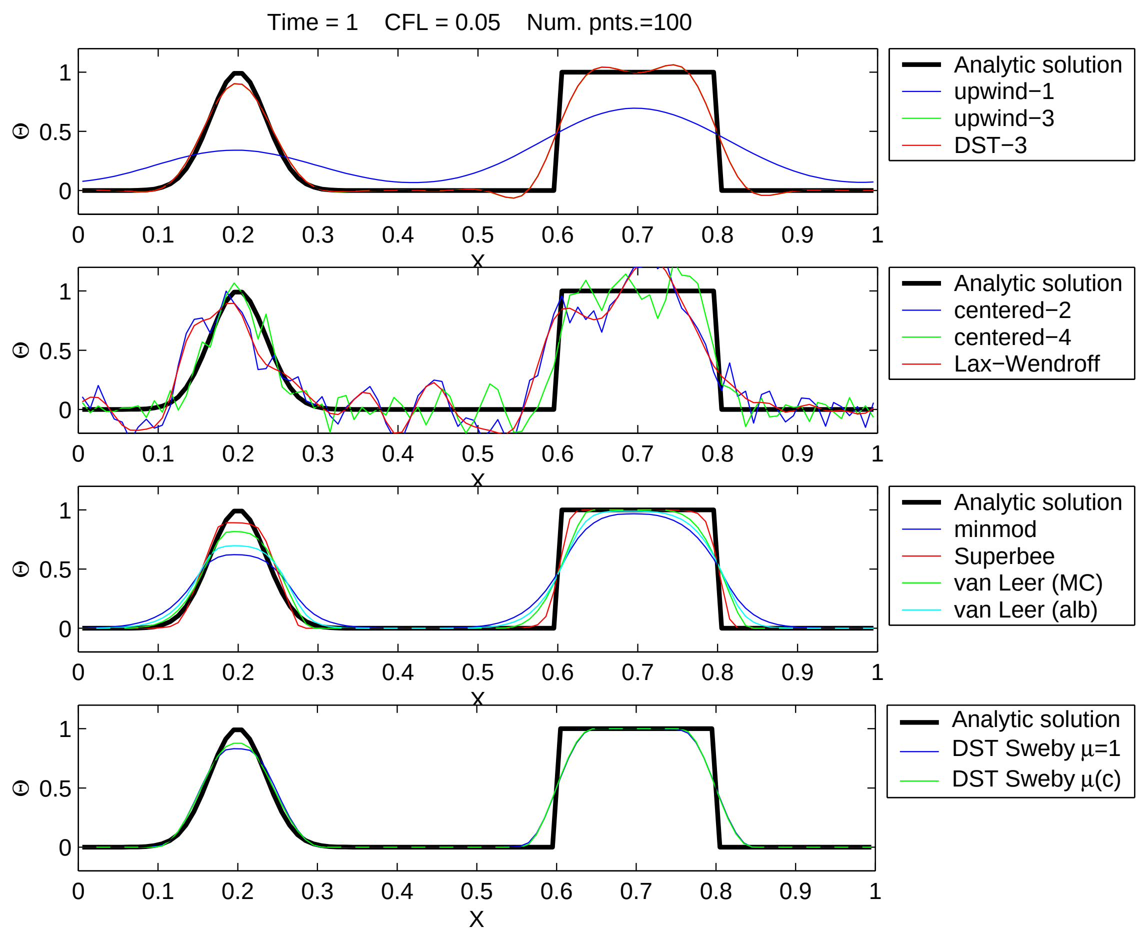

Figure 1.5: Solution obtained with a small Courant number. The analytic solution is the thick solid line in each panel. The upwind scheme does not exhibit false extrema but is clearly very diffusive. The third order upwind is much better at preserving the shapes. The even order methods (second panel) have multiple false extrema but the Lax-Wendroff method is at least smoother. The third panel shows the second order limited solutions, all of which conserve extrema. The fourth panel shows some third order flux limited solutions which preserve amplitude better.

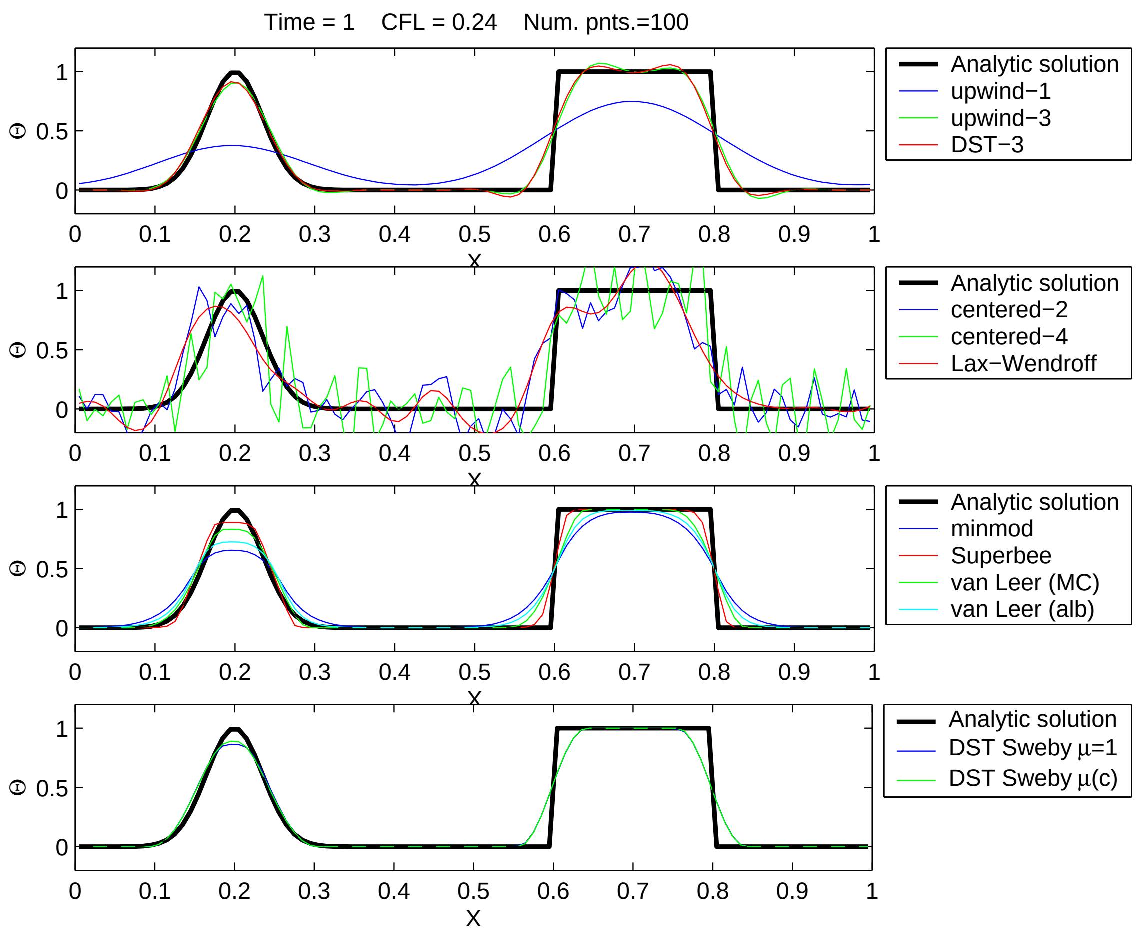

Figure 1.6: Solution obtained with a large Courant number. As for Fig. 1.5 but notice that the noise levels in the unlimited schemes is worse.