1 Finite-Differences

Under finite-difference discretizations, the function f is defined at spatial points

$$x_j = j\Delta x. \tag{1}$$

Recall the following three definitions of the derivative

$$\frac{df}{dx}(x_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}, \tag{2}$$

$$\frac{df}{dx}(x_0) = \lim_{\Delta x \to 0} \frac{f(x_0) - f(x_0 - \Delta x)}{\Delta x}, \tag{3}$$

$$\frac{df}{dx}(x_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x}. \tag{4}$$

Discrete approximations to the derivative at x_0 using the above definitions as motivation are then

$$\frac{df}{dx}(x_0) \approx \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{f(x_1) - f(x_0)}{\Delta x}, \tag{5}$$

$$\frac{df}{dx}(x_0) \approx \frac{f(x_0) - f(x_0 - \Delta x)}{\Delta x} = \frac{f(x_0) - f(x_{-1})}{\Delta x}, \tag{6}$$

$$\frac{df}{dx}(x_0) \approx \frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{\Delta x} = \frac{f(x_1) - f(x_{-1})}{2\Delta x}. \tag{7}$$

These three approximations differ in terms of accuracy. To measure the accuracy of these approximations, we look to the *truncation error*, obtained from substituting the exact solution into the discretization.

Recall that if $f(x)$ is sufficiently smooth, we can use a *Taylor series* to write

$$f(x_0 + \Delta x) = f(x_0) + \Delta x \frac{df}{dx}(x_0) + \frac{(\Delta x)^2}{2}\frac{d^2 f}{dx^2}(x_0) + \frac{(\Delta x)^3}{6}\frac{d^3 f}{dx^3}(x_0) + \cdots. \tag{8}$$

Substituting this into (5) then leads to

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} - \frac{df}{dx}(x_0) = \frac{\Delta x}{2}\frac{d^2 f}{dx^2}(x_0) + \mathcal{O}(\Delta x^2). \tag{9}$$

The expression on the right-hand-side is the *truncation error*. We say that this discretization is *first-order accurate* since the smallest right-hand-side term is $\mathcal{O}(\Delta x)$.

First-order accuracy means that halving the grid spacing Δx leads to halving of the error associated with this term.

If we perform a similar analysis on the centered difference scheme (7), we obtain

$$\frac{f(x_0 + \Delta x) - f(x_0 - \Delta x)}{2\Delta x} - \frac{df}{dx}(x_0) = \frac{(\Delta x)^2}{6}\frac{d^3 f}{dx^3}(x_0) + O(\Delta x^4), \tag{10}$$

implying that this approximation is second-order accurate. This means that halving Δx leads to a four-times improvement in the error of the scheme.

Higher-order discretizations can also be derived using polynomial interpolation, for example.

2 A First Look at the Advection Equation

The 1D advection equation is defined as

$$\boxed{\frac{\partial \psi}{\partial t} + c\frac{\partial \psi}{\partial x} = 0, \quad c > 0,} \tag{11}$$

where ψ is some scalar field.

One can quickly verify that for sufficiently smooth initial data ψ_0, this equation has solutions of the form

$$\psi(x,t) = \psi_0(x - ct). \tag{12}$$

If we substitute (12) into the LHS of (11) we obtain

$$-c\psi_0' + c\psi_0' = 0, \tag{13}$$

and so the equation is satisfied. As a consequence, we say that solutions of the advection equation are *advected* by the flow.

In 1D we seek the solution of this equation at points x_j and at time t^n, defined via

$$x_j = j\Delta x, \qquad t^n = n\Delta t, \tag{14}$$

where Δx is the *spatial step* and Δt is the *time step*. We approximate the solution at each point via

$$\psi(x_j, t^n) \approx \psi_j^n. \tag{15}$$

We can discretize the advection equation using the FD operators from the previous section,

$$\frac{\psi_j^{n+1} - \psi_j^n}{\Delta t} + c\frac{\psi_j^n - \psi_{j-1}^n}{\Delta x} = 0, \tag{16}$$

which can be rearranged to

$$\boxed{\psi_j^{n+1} = \psi_j^n - \frac{c\Delta t}{\Delta x}\left(\psi_j^n - \psi_{j-1}^n\right).} \tag{17}$$

This approximation is known as the *upwind scheme* for the advection equation. The quantity

$$\nu = \frac{c\Delta t}{\Delta x}, \tag{18}$$

frequently appears in these approximations, and is given the name the *Courant-Friedrichs-Lewy (CFL) number*.

We can again apply truncation error analysis to (16), obtaining

$$\left[\frac{\psi_j^{n+1} - \psi_j^n}{\Delta t} + c\frac{\psi_j^n - \psi_{j-1}^n}{\Delta x}\right] - \left[\frac{\partial \psi}{\partial t} + c\frac{\partial \psi}{\partial x}\right] = \frac{\Delta t}{2}\frac{\partial^2 \psi}{\partial t^2} - c\frac{\Delta x}{2}\frac{\partial^2 \psi}{\partial x^2} + \cdots. \tag{19}$$

The order of accuracy is determined by the lowest powers of Δt and Δx appearing on the right-hand-side. Hence, we say that this scheme is *first-order accurate* in time and *first-order accurate* in space.

If the truncation error of the finite-difference scheme approaches zero as $\Delta x \to 0$ and $\Delta t \to 0$, we say that a scheme is *consistent*. Clearly, the above scheme is consistent.

3 Stability

The notion of *stability* is closely related to the concept of energy in a system of equations. If the energy is bounded as $t \to \infty$ then we say that the system of equations is stable.

Generally, any positive-definite quantity can be used to represent energy. Some choices are more intuitive than others.

For the advection equation, one notion of energy is $E = \frac{1}{2}\psi^2$. If we multiply the advection equation (11) through by ψ, we obtain

$$\psi\frac{\partial\psi}{\partial t} + c\psi\frac{\partial\psi}{\partial x} = 0, \quad \iff \quad \frac{\partial\left(\frac{1}{2}\psi^2\right)}{\partial t} + c\frac{\partial\left(\frac{1}{2}\psi^2\right)}{\partial x} = 0, \quad \iff \quad \frac{\partial E}{\partial t} + c\frac{\partial E}{\partial x} = 0, \tag{20}$$

which implies that E also satisfies the advection equation. If we consider some closed-system, such as a periodic domain $[0, 1]$, then the total energy is defined as

$$E_{tot} = \int_0^1 E(x)dx. \tag{21}$$

If we now integrate the energy equation, we obtain

$$\frac{d}{dt}\int_0^1 E(x)dx + c\int_0^1 \frac{dE}{dx}dx = 0. \tag{22}$$

This equation can now be expanded using the fundamental theorem of calculus,

$$\frac{dE_{tot}}{dt} + c(E(1) - E(0)) = 0. \tag{23}$$

But since the domain is periodic, we have $E(1) = E(0)$ and so

$$\frac{dE_{tot}}{dt} = 0. \tag{24}$$

Hence, in the continuous equation, total energy is exactly conserved.

In the discrete case we do not have such guarantees. In fact, we can sometimes come up with discretizations that lead to a system where energy increases without bound. Such discretizations are referred to as "unstable." On the other hand, if energy is bounded for all time t, we refer to those discretizations as "stable."

Directly proving that a method is stable using energy is often very complicated, so in practice it is common to instead use von Neumann analysis.

4 Von Neumann Analysis

Von Neumann analysis is a very powerful method for determining the stability properties of a numerical method. However, it can only be applied to problems on periodic domains and to linear discretizations.

The basic idea behind von Neumann analysis is to transform a solution into the sum of individual waves. For a problem on a periodic domain $[0, 2\pi]$, we again define a numerical grid $x_j = j\Delta x$ for $j = 0, \ldots, N-1$ and $\Delta x = 2\pi/N$. First, observe that we can write any solution as a finite combination of Fourier modes

$$\psi_j = \frac{1}{N} \sum_{k=0}^{N-1} a_k \exp(ikj\Delta x), \tag{25}$$

It is a well-known fact (and it can be easily proven) that

$$\sum_{j=0}^{N-1} (\psi_j)^2 = \frac{1}{N} \sum_{k=0}^{N-1} (a_k)^2, \tag{26}$$

which is known as Parseval's theorem. Observe that the left-hand-side of this equation is closely related to the discrete analogue of the total energy at time t^n,

$$2E_{tot} = \sum_{j=0}^{N-1} (\psi_j)^2. \tag{27}$$

Hence, if we can show that all Fourier components a_k^n are bounded, it must follow that the total energy is also bounded.

The advantage of looking at Fourier components is that each Fourier mode is an eigenvalue of any linear discrete equation, and further, each mode is orthogonal to any other mode. These two results imply that each mode can be analyzed in isolation.

To determine the behavior of one Fourier mode in isolation, we utilize a trial wave solution

$$\psi_j^n = (a_k)^n \exp(ij\theta), \tag{28}$$

where $\theta = k\Delta x$ is the dimensionless wave-number (here $0 \leq \theta \leq \pi$). If we substitute (28) into the upwind scheme (17), we obtain

$$(a_k)^{n+1} \exp(ij\theta) = (a_k)^n \exp(ij\theta) - \nu \left((a_k)^n \exp(ij\theta) - (a_k)^n \exp(i(j-1)\theta) \right). \tag{29}$$

Simplifying gives

$$S_k = \frac{(a_k)^{n+1}}{(a_k)^n} = 1 - \nu(1 - \exp(-i\theta)). \tag{30}$$

The quantity on the left-hand-side is the so-called *Fourier symbol* of the upwind scheme for the k^{th} Fourier mode (a complex quantity). A related quantity is the *amplification factor*, which is defined by $A_k = |S_k|$, and determines the absolute amplification of this Fourier mode. Hence, in order for this scheme to be stable, we must have $|A|_k \leq 1$, or equivalently, $-1 \leq S_k \leq 1$.

Using (30), the amplification factor for this scheme can be written as

$$(A_k)^2 = (1 - \nu(1 - \cos\theta))^2 + \nu^2 \sin^2\theta = 1 - 2\nu(1-\nu)(1 - \cos\theta). \tag{31}$$

Stability of this method will be dependent on the CFL number ν. It can be quickly verified that in order for us to have $(A_k)^2 \leq 1$, we must have

$$0 \leq \nu \leq 1. \tag{32}$$

This stability restriction is known as the *CFL condition*.

5 The Dispersion Relation

As we have seen in (31), different Fourier wave modes will have potentially different behavior. Let's look into this a bit further.

We assume wave-like solutions of the form

$$\psi(x,t) = \psi_0 \exp(i(kx - \omega t)). \tag{33}$$

Substituting this solution into the advection equation (11) gives

$$-i\omega\psi_0 \exp(i(kx - \omega t)) + cik\psi_0 \exp(i(kx - \omega t)) = 0. \tag{34}$$

Simplifying implies that a wave-like solution (33) is only a solution of the advection equation if

$$\boxed{\omega(k) = ck}. \tag{35}$$

This equation relates each wave frequency to its wavenumber k, and is known as the *dispersion relation*.

There two specific quantities which are closely related to the dispersion relation. First, the *phase velocity*, which is given by

$$c_p = \frac{\omega}{k}, \tag{36}$$

defines the velocity at which a single wave moves. A related quantity is the *group velocity*, which is given by

$$c_g = \frac{d\omega}{dk}, \tag{37}$$

defines the velocity at which a packet of wave moves within the system. The group velocity is also the velocity at which energy propagates within the system. For the advection equation, both of these quantities are equal to c; however, this is not true for general equations.

The discrete case analogue of

$$\psi_j^n = \psi_0 \exp(i(j\theta - \omega n\Delta t)). \tag{38}$$

Substituting this solution into the upwind discretization (17) and simplifying, we obtain

$$\exp(-i\omega\Delta t) = 1 - \nu(1 - \cos(k\Delta x) + i\sin(k\Delta x)). \tag{39}$$

Observe that the right-hand-side of this expression is simply the Fourier symbol S_k (30). This equation can be inverted to obtain ω, giving the *numerical dispersion relation*,

$$\omega = -\frac{1}{\Delta t}\arctan\left(\frac{\nu\sin(k\Delta x)}{1 - \nu(1 - \cos(k\Delta x))}\right). \tag{40}$$

Observe that if $\nu = 0$ or $\nu = 1$ this dispersion relation exactly matches (35), but for any other value of ν the wave does not behave the same as the physical dispersion relation.

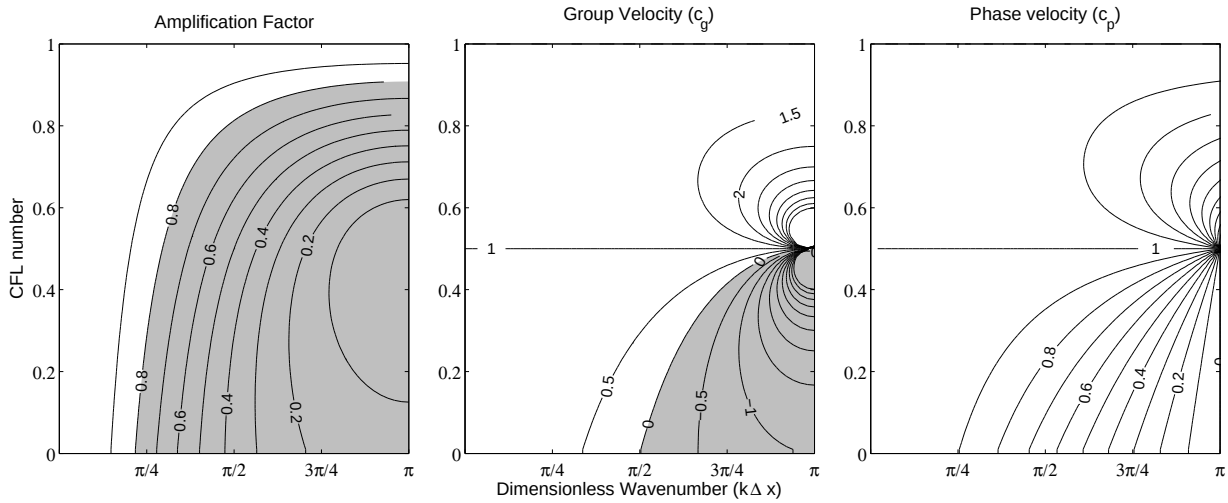

Figure 1: Diffusion and dispersion properties of the upwind scheme.

6 Timestepping Schemes

Upon applying a spatial discretization to a PDE, we obtain a semi-discretization of the form

$$\frac{\partial \mathbf{q}}{\partial t} = \mathcal{F}(\mathbf{q}). \tag{41}$$

Your favourite timestepping scheme from numerical ODEs can then be applied to this form of the equation.

Forward Euler (first-order accurate, usually unstable)

$$\mathbf{q}^{n+1} = \mathbf{q}^n + \Delta t \mathcal{F}(\mathbf{q}^n). \tag{42}$$

Leapfrog (second-order accurate, requires filtering)

$$\mathbf{q}^{n+1} = \mathbf{q}^{n-1} + 2\Delta t \mathcal{F}(\mathbf{q}^n). \tag{43}$$

Strong-Stability Preserving RK3 Scheme (third-order accurate, commonly used with Galerkin methods)

$$\begin{aligned}
\mathbf{q}^{(1)} &= \mathbf{q}^n + \Delta t \mathcal{F}(\mathbf{q}^n), && (44)\\
\mathbf{q}^{(2)} &= \tfrac{3}{4}\mathbf{q}^n + \tfrac{1}{4}\mathbf{q}^{(1)} + \tfrac{1}{4}\Delta t \mathcal{F}(\mathbf{q}^{(1)}), && (45)\\
\mathbf{q}^{n+1} &= \tfrac{1}{3}\mathbf{q}^n + \tfrac{2}{3}\mathbf{q}^{(2)} + \tfrac{2}{3}\Delta t \mathcal{F}(\mathbf{q}^{(2)}). && (46)
\end{aligned}$$

RK4 Scheme (fourth-order accurate)

$$\begin{aligned}
\mathbf{q}^{(1)} &= \mathbf{q}^n + \tfrac{1}{2}\Delta t \mathcal{F}(\mathbf{q}^n), && (47)\\
\mathbf{q}^{(2)} &= \mathbf{q}^n + \tfrac{1}{2}\Delta t \mathcal{F}(\mathbf{q}^{(1)}), && (48)\\
\mathbf{q}^{(3)} &= \mathbf{q}^n + \Delta t \mathcal{F}(\mathbf{q}^{(2)}), && (49)\\
\mathbf{q}^{n+1} &= -\tfrac{1}{3}\mathbf{q}^n + \tfrac{1}{3}\mathbf{q}^{(1)} + \tfrac{2}{3}\mathbf{q}^{(2)} + \tfrac{1}{3}\mathbf{q}^{(3)} + \tfrac{1}{6}\Delta t \mathcal{F}(\mathbf{q}^{(3)}). && (50)
\end{aligned}$$

7 The Shallow-Water Equations

The shallow-water equations are perhaps the simplest fluid equations that possess many of the properties also present in atmospheric flow. In 1D, these equations are

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} = 0, \tag{51}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + g\frac{\partial h}{\partial x} = 0, \tag{52}$$

where h is the fluid height, u is the fluid velocity and g is the gravitational constant $g = 9.80616$ m / s^2.

To study the linear behavior of these equations (which is generally dominant for atmospheric flows), we substitute

$$h = H + h', \qquad u = u', \tag{53}$$

and remove products of primed terms. This leads to the *linearized shallow-water equations*,

$$\frac{\partial h'}{\partial t} + H\frac{\partial u'}{\partial x} = 0, \tag{54}$$

$$\frac{\partial u'}{\partial t} + g\frac{\partial h'}{\partial x} = 0. \tag{55}$$

We wish to study wave-like solutions of these equations, given by

$$h' = h_0 \exp(i(kx - \omega t)), \qquad u' = u_0 \exp(i(kx - \omega t)). \tag{56}$$

Substituting these relations into (54)-(55) gives

$$[-i\omega h_0 \exp(i(kx - \omega t))] + H\,[iku_0 \exp(i(kx - \omega t))] = 0, \tag{57}$$

$$[-i\omega u_0 \exp(i(kx - \omega t))] + g\,[ikh_0 \exp(i(kx - \omega t))] = 0, \tag{58}$$

or more simply

$$\underbrace{\begin{bmatrix} -\omega & Hk \\ gk & -\omega \end{bmatrix}}_{\mathcal{M}} \begin{bmatrix} h_0 \\ u_0 \end{bmatrix} = 0. \tag{59}$$

For wave-like solutions to exist, we must have $\det(\mathcal{M}) = 0$, or equivalently,

$$\omega^2 - gHk^2 = 0, \tag{60}$$

which implies a dispersion relation of the form

$$\omega = \pm\sqrt{gH}k. \tag{61}$$

Observe that this system supports either left-going or right-going waves, both of which travel with wave speed $\sqrt{gH}$.

8 The Shallow Water Equations in 2D

The 2D shallow water equations, with Coriolis terms, can be written in conservative form as

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) + \frac{\partial}{\partial y}(hv) = 0, \tag{62}$$

$$\frac{\partial (hu)}{\partial t} + \frac{\partial}{\partial x}\left(hu^2 + \tfrac{1}{2}gh^2\right) + \frac{\partial}{\partial y}\left(huv\right) = fhv, \tag{63}$$

$$\frac{\partial (hv)}{\partial t} + \frac{\partial}{\partial y}\left(huv\right) + \frac{\partial}{\partial y}\left(hv^2 + \tfrac{1}{2}gh^2\right) = -fhu. \tag{64}$$

To linearize these equations, we write

$$h = H + h', \qquad u = u', \quad \text{and} \quad v = v', \tag{65}$$

and drop all products of primed terms. This procedure yields

$$\frac{\partial h'}{\partial t} + H\frac{\partial u'}{\partial x} + H\frac{\partial v'}{\partial y} = 0, \tag{66}$$

$$\frac{\partial u'}{\partial t} + g\frac{\partial h'}{\partial x} - fv = 0, \tag{67}$$

$$\frac{\partial v'}{\partial t} + g\frac{\partial h'}{\partial y} + fu = 0. \tag{68}$$

We again study wave-like solutions of the form

$$\mathbf{q}' = \mathbf{q}_0 \exp(i(kx + \ell y - \omega t)), \tag{69}$$

where $\mathbf{q}' = (h', u', v')$ and $\mathbf{q}_0 = (h_0, u_0, v_0)$. Substituting (69) into the linearized shallow-water equations yields

$$\underbrace{\begin{pmatrix} -i\omega & iHk & iH\ell \\ igk & -i\omega & -f \\ ig\ell & f & -i\omega \end{pmatrix}}_{\mathcal{M}} \begin{pmatrix} h_0 \\ u_0 \\ v_0 \end{pmatrix} = 0. \tag{70}$$

For wave-like solutions to exist, the matrix must have zero determinant, which implies

$$\det(\mathcal{M}) = i\omega(\omega^2 - f^2 - gH(k^2 + \ell^2)) = 0, \tag{71}$$

and so the possible solutions are

$$\omega = 0, \tag{72}$$

which is the so-called "geostrophic mode," and

$$\omega = \pm\sqrt{f^2 + gH(k^2 + \ell^2)}, \tag{73}$$

the so-called "inertia-gravity wave modes." The geostrophic mode is stationary, with zero phase and group velocity, and so corresponds to a steady-state. The inertia-gravity wave modes are non-stationary. Notice that for $\ell = 0$ and $f \neq 0$ the intertia-gravity wave modes are dispersive; namely $c_p \neq c_g$.

9 Finite-Volume Methods 1

Conservation is an important part of climate models; above all else, climate models are at least expected to conserve mass. Without conservation of mass, models often tend to slowly lose their atmosphere slowly which can lead to a drastically different climate.

Conservation laws are written in the form

$$\frac{\partial q}{\partial t} + \nabla \cdot \mathbf{F}(q) = 0, \tag{74}$$

where $F(q)$ is known as a *flux function*. If we integrate this conservation law over some region $\mathcal{Z}$, we obtain

$$\frac{d}{dt}\int_{\mathcal{Z}} q dV + \int \nabla \cdot \mathbf{F}(q) dV = 0, \tag{75}$$

which can be simplified by Gauss' divergence theorem to

$$\frac{d}{dt}\int_{\mathcal{Z}} q dV + \oint_{\partial \mathcal{Z}} \mathbf{F}(q) \cdot \mathbf{n} dV = 0, \tag{76}$$

where $\partial\mathcal{Z}$ denotes the boundary of $\mathcal{Z}$. The first term is equal to the change of the total mass of quantity q and the second term represents the flow through the boundaries of $\mathcal{Z}$. Intuitively, this means that

$$\langle \text{Change in total mass of } q \rangle = \langle \text{Flux of } \mathbf{q} \text{ through the boundary of } \mathcal{Z} \rangle. \tag{77}$$

If we take $\mathcal{Z}$ to be the entire atmosphere, then we obtain

$$\langle \text{Change in total mass of } q \rangle = 0, \tag{78}$$

implying that the mass of q does not change within the atmosphere.

Finite-volume methods are built on the idea of preserving conservation properties, by utilizing a conservation law for those conserved quantities. In a finite-volume model, instead of modeling point values we instead look at averaged values of a quantity over some grid element. In 1D we can write

$$\overline{q}_j = \frac{1}{\Delta x}\int_{\mathcal{Z}_j} q dx, \tag{79}$$

where $\mathcal{Z}_j$ represents the j^{th} element, which spans the interval $[(j-1/2)\Delta x, (j+1/2)\Delta x]$.

The conservation law (74) can also be integrated, as before, to give

$$\boxed{\frac{d\overline{q}_j}{dt} + \frac{F(q_{j+1/2}) - F(q_{j-1/2})}{\Delta x} = 0,} \tag{80}$$

where the flux function F is evalued at edges (denoted by half-indices). Using this approach, we obtain a conservation property similar to in the continuous case, since (80) implies

$$\underbrace{\frac{d}{dt}\sum_{j=a}^{b} \overline{q}_j \Delta x}_{\text{Change in total mass of } q} + \underbrace{(F(q_{b+1/2}) - F(q_{a-1/2}))}_{\text{Flux through boundary}} = 0. \tag{81}$$

In practice, we must define a numerical flux function $F^*_{j+1/2}$ in place of $F(q_{j+1/2})$. The majority of finite-volume methods only differ in the definition of $F^*_{j+1/2}$ and the timestepping scheme.

10 Finite-Volume Methods 2

There are two immediately obvious choices of F^* for the advection equation:

(A) The Upwind Flux

$$F^*_{j+1/2} = q_j$$

(B) The Central Flux

$$F^*_{j+1/2} = \frac{q_j + q_{j+1}}{2}$$

If combined with a Forward Euler timestep, the upwind flux leads to an evolution equation of the form

$$q_j^{n+1} = q_j^n - \frac{c\Delta t}{\Delta x}\left(q_j^n - q_{j-1}^n\right), \tag{82}$$

which is exactly the upwind scheme! This result also implies that the upwind method is conservative.

The central flux instead leads to

$$q_j^{n+1} = q_j^n - \frac{c\Delta t}{2\Delta x}\left(q_{j+1}^n - q_{j-1}^n\right), \tag{83}$$

which is the forward-in-time central-in-space (FTCS) scheme. This method is first-order-accurate in time, second-order-accurate in space, but is unconditionally unstable for any choice of CFL number $\nu = c\Delta t/\Delta x$.

11 Achieving second-order accuracy

First-order accurate schemes are generally far too diffusive to use in atmospheric models; all sharp profiles are dramatically "smeared" by the model. Hence, we are interested in schemes which obtain at least second-order accuracy overall.

It has been shown (see Exercises) that the forward-in-time central in space scheme is unconditionally unstable. This scheme takes the form

$$\frac{\psi_j^{n+1} - \psi_j^n}{\Delta t} + c\frac{\psi_{j+1}^n - \psi_{j-1}^n}{2\Delta x} = 0. \tag{84}$$

Observe that we may rewrite this scheme in flux-form (80) by choosing

$$F^*_{j+1/2} = \frac{c}{2}\left(\psi_{j+1}^n + \psi_j^n\right). \tag{85}$$

In order to "stabilize" this scheme, we typically add a diffusive term to the flux. The diffusive term has the responsibility of smoothing out instabilities.

In this section we present three common choices of the flux that incorporate a diffusion term.

Lax-Wendroff scheme:

$$F^*_{j+1/2} = \frac{c}{2}\left(\psi_{j+1}^n + \psi_j^n\right) - \underbrace{\frac{\nu^2\Delta x}{2\Delta t}\left(\psi_{j+1}^n - \psi_j^n\right)}_{\text{Diffusive term}} \tag{86}$$

To show that these terms are actually diffusive, we substitute this flux into (80) to obtain

$$\psi_j^{n+1} = \psi_j^n - c\Delta t\underbrace{\left(\frac{\psi_{j+1}^n - \psi_{j-1}^n}{2\Delta x}\right)}_{\text{Advective term}} + \underbrace{\frac{\nu^2\Delta x^2}{2}\left(\frac{\psi_{j+1}^n - 2\psi_j^n + \psi_{j-1}^n}{\Delta x^2}\right)}_{\text{Diffusive term}}. \tag{87}$$

Observe that the advective term is the same as in the centered difference scheme, implying that it is handled to second-order accuracy. The diffusive term is a second-order contribution to the error, and is proportional to $d\psi^2/dx^2$, which represents diffusion.

There are two more commonly used schemes of second-order accuracy, which again can be defined using appropriately chosen numerical fluxes.

Beam-Warming scheme:

$$F^*_{j+1/2} = \frac{c}{2}\left(3\psi_j^n - \psi_{j-1}^n\right) - \underbrace{\frac{\nu^2\Delta x}{2\Delta t}\left(\psi_j^n - \psi_{j-1}^n\right)}_{\text{Diffusive term}} \tag{88}$$

Fromm scheme:

$$F^*_{j+1/2} = \frac{1}{2}\left(F^*_{\text{Lax-Wendroff},j+1/2} + F^*_{\text{Beam-Warming},j+1/2}\right). \tag{89}$$

Unfortunately, all of these schemes suffer from dispersive errors, which can lead to overshoots and undershoots. We will return to this issue later.

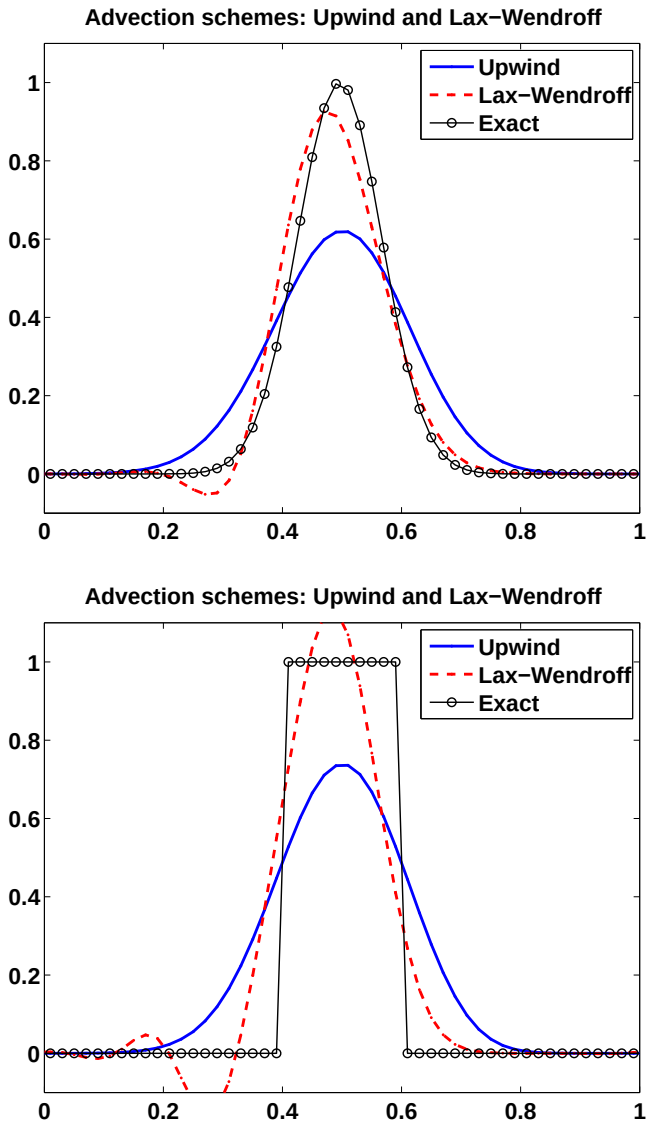

Figure 2: Advection equation with the upwind and Lax-Wendroff scheme after 4 rotations. The domain $[0, 1]$ is discretized with 50 elements. Observe that the Lax-Wendroff scheme produces a much better result than the Upwind scheme, but also leads to overshoots and undershoots near discontinuities. These overshoots and undershoots could lead to negative tracer concentrations, which has no physical meaning.

12 Conservative Semi-Lagrangian Schemes

Conservative semi-Lagrangian schemes achieve high-order accuracy by utilizing a so-called *sub-grid-scale spatial reconstruction*. The *sub-grid-scale reconstruction* is designed to describe the continuous behavior of the field ψ, and usually takes looks like a sum of polynomials, here given in 1D as

$$\psi_j(x) = \psi_j(0) + \left(\frac{\partial \psi_j}{\partial x}\right)_0 (x - x_{j,0}) + \left(\frac{\partial^2 \psi_j}{\partial x^2}\right)_0 \frac{(x - x_{j,0})^2}{2} + \cdots, \tag{90}$$

where $x_{j,0}$ is the center of element x_j. The number of terms used in this approximation determines the spatial order-of-accuracy of the underlying scheme.

The basic method used by the semi-Lagrangian scheme in 2D is depicted in Figure 3. For each Eulerian element A_k we trace velocities backward to obtain a source element a_k. Because we are working in a Lagrangian framework, the total mass within element a_k is exactly the total mass within element A_k. To obtain the mass within a_k, we simply integrate over the reconstruction within each element.

If we consider the constant reconstruction, the sub-grid-scale reconstruction takes the form

$$\psi_j(x) = \psi_j(0). \tag{91}$$

A depiction of this semi-Lagrangian method is given in Figure 4. Observe that the grid shift is exactly $c\Delta t$. If we perform the integration, we obtain a discrete evolution equation of the form

$$\psi_j^{n+1} = \nu \psi_{j-1}^n + (1 - \nu)\psi_j^n. \tag{92}$$

But this is exactly the upwind scheme (17)!

We now consider a linear reconstruction of the form

$$\psi_j(x) = \overline{\psi}_j + \left(\frac{\partial \psi_j}{\partial x}\right)_0 (x - x_0). \tag{93}$$

If we use the approximation

$$\left(\frac{\partial \psi_j}{\partial x}\right)_j = \frac{\overline{\psi}_{j+1} - \overline{\psi}_{j-1}}{2\Delta x}, \tag{94}$$

this leads to the Fromm scheme (exercise). Similarly, the one-sided approximations

$$\left(\frac{\partial \psi_j}{\partial x}\right)_j = \frac{\overline{\psi}_j - \overline{\psi}_{j-1}}{\Delta x}, \quad \text{and} \quad \left(\frac{\partial \psi_j}{\partial x}\right)_j = \frac{\overline{\psi}_{j+1} - \overline{\psi}_j}{\Delta x} \tag{95}$$

lead to the Beam-Warming and Lax-Wendroff schemes, respectively (exercise).

High-order reconstructions can be obtained by using more terms of the polynomial reconstruction. In this case, a third order reconstruction can be obtained by choosing

$$\psi_j(0) = -\tfrac{1}{24}\overline{\psi}_{j-1} + \tfrac{13}{12}\overline{\psi}_j - \tfrac{1}{24}\overline{\psi}_{j+1}, \qquad \left(\frac{\partial^2 \psi_j}{\partial x^2}\right)_k = \frac{\overline{\psi}_{j-1} - 2\overline{\psi}_j + \overline{\psi}_{j+1}}{\Delta x^2}, \tag{96}$$

and (94). The advection equation then takes the form (exercise)

$$\psi_j^{n+1} = \psi_j^n - \nu\left[-\frac{1}{6}(\nu^2 - 1)\psi_{j-2}^n + \frac{1}{2}(\nu + 1)(\nu - 2)\psi_{j-1}^n - \frac{1}{2}(\nu^2 - 2\nu - 1)\psi_j^n + \frac{1}{6}(\nu - 1)(\nu - 2)\psi_{j+1}^n\right]. \tag{97}$$

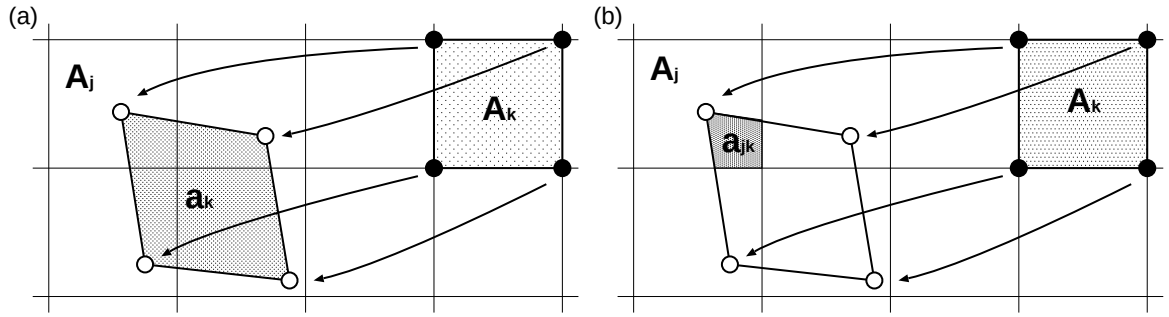

Figure 3: A depiction of the semi-Lagrangian scheme in 2D. The target element A_k is mapped backwards in time to a source element a_k. Reconstructions are constructed in all surrounding elements A_j and the total mass is remapped to a_k.

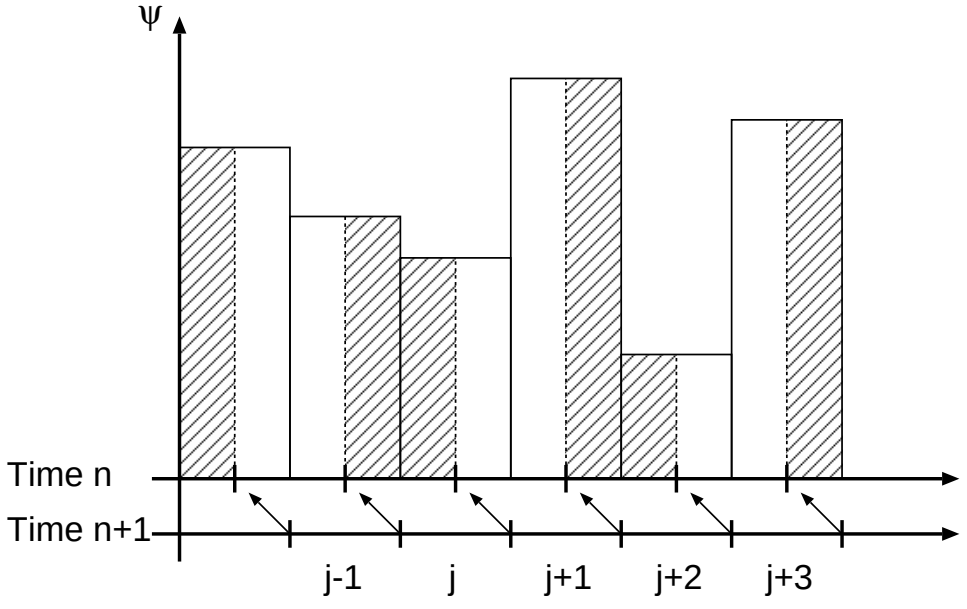

Figure 4: A depiction of the semi-Lagrangian scheme in 1D using piecewise constant reconstructions. Elements $k-2,\ldots,k+2$ are mapped to source elements at time n. We then integrate over the source elements.

13 Achieving monotonicity - Slope limiting

A well known result in computational fluid dynamics, known as Lax's theorem, states the following:

> *Linear numerical schemes for solving partial differential equations (PDE's), having the property of not generating new extrema (monotone scheme), can be at most first-order accurate.*

This unfortunate result suggests that it is impossible to construct a linear scheme (such as all the schemes we have seen so far) that does not generate unphysical minima or maxima. The most common strategy for dealing with this is instead to use a *nonlinear scheme*.

Recall that a linear sub-grid-scale reconstruction can be written as

$$\psi_j(x) = \psi_j(0) + \left(\frac{\partial \psi_j}{\partial x}\right)_0 (x - x_{j,0}). \tag{98}$$

The reason why schemes based on this reconstruction can generate unphysical minima and maxima can be attributed to the choice of first derivative $(\partial\psi_j/\partial x)$. Near a discontinuity in the solution, the first derivative may be estimated incorrectly (see Figure 5). A typical method to deal with this problem is to introduce a *slope limiter* which "limits" the choice of slope in regions where a discontinuity is detected.

To begin, we define the *ratio of consecutive slopes*,

$$\theta = \frac{\psi_j - \psi_{j-1}}{\psi_{j+1} - \psi_j}. \tag{99}$$

Note that for a smooth profile, the value of θ is very close to 1. Near sharp changes in the solution profile however, $|\theta|$ can become very large. As a consequence, we are motivated to introduce a linear sub-grid-scale reconstruction of the form

$$\psi_j(x) = \psi_j(0) + \phi(\theta)\frac{(\psi_{j+1} - \psi_j)}{\Delta x}(x - x_{j,0}), \tag{100}$$

where $\phi(\theta)$ is some nonlinear function of θ.

Some typical choices of $\phi(\theta)$ include the *MINMOD limiter*,

$$\phi_{mm}(\theta) = \max[0, \min(1, \theta)], \tag{101}$$

the *superbee limiter*,

$$\phi_{sb}(\theta) = \max[0, \min(2\theta, 1), \min(\theta, 2)], \tag{102}$$

and the *van Leer limiter*,

$$\phi_{vl}(\theta) = (\theta + |\theta|)/(1 + |\theta|). \tag{103}$$

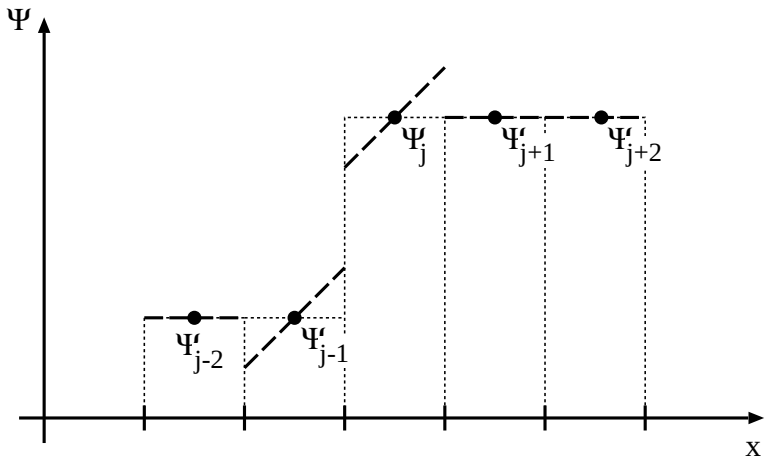

Figure 5: An illustration of a non-monotone reconstruction near a discontinuity in the solution.

14 The piecewise parabolic method

The piecewise parabolic method (PPM) was originally described in a paper by Colella and Woodward (1984). This method is incredibly popular, and is still used today in many climate and weather models. The basic idea behind this method is to enforce continuity of the sub-grid-scale reconstruction at cell boundaries. Here we use the formulation of Colella and Sekora (2009).

Edge values are defined at the left and right edges of cell j by $\psi_{L,j}$ (left side) and $\psi_{R,j}$ (right side). These are first approximated by

$$\psi_{R,j} = \psi_{L,j+1} = -\frac{1}{12}\left(\overline{\psi}_{j-1} + \overline{\psi}_{j+2}\right) + \frac{7}{12}\left(\overline{\psi}_{j} + \overline{\psi}_{j+1}\right). \tag{104}$$

The values are then scaled so as not to exceed the range $[\overline{\psi}_j, \overline{\psi}_{j+1}]$. We say

$$\psi_{R,j} = \min\left[\psi_{R,j}, \max\left(\overline{\psi}_{j+1}, \overline{\psi}_{j}\right)\right], \tag{105}$$
$$\psi_{R,j} = \max\left[\psi_{R,j}, \min\left(\overline{\psi}_{j+1}, \overline{\psi}_{j}\right)\right], \tag{106}$$

and similarly for left edge values $\psi_{L,j}$,

$$\psi_{L,j} = \min\left[\psi_{L,j}, \max\left(\overline{\psi}_{j-1}, \overline{\psi}_{j}\right)\right], \tag{107}$$
$$\psi_{L,j} = \max\left[\psi_{L,j}, \min\left(\overline{\psi}_{j-1}, \overline{\psi}_{j}\right)\right], \tag{108}$$

Note that at this point $\psi_{R,j}$ and $\psi_{L,j+1}$ should still be identical. Next the parabolic profile is modified so as not to have maxima within the element. If $(a_{j,R} - \overline{\psi}_j^n)(a_{j,L} - \overline{\psi}_j^n) > 0$, then we set

$$\psi_{j,R} = \psi_{j,L} = \overline{\psi}_j. \tag{109}$$

If $|\psi_{j,R} - \overline{\psi}_j| \geq 2|\psi_{j,L} - \overline{\psi}_j|$ then set

$$\psi_{j,R} = \overline{\psi}_j - 2(\psi_{j,L} - \overline{\psi}_j). \tag{110}$$

Similarly, if $|\psi_{j,L} - \overline{\psi}_j| \geq 2|\psi_{j,R} - \overline{\psi}_j|$ then set

$$\psi_{j,L} = \overline{\psi}_j - 2(\psi_{j,R} - \overline{\psi}_j). \tag{111}$$

Given left and right edge values, we define

$$\Delta\psi_j = \psi_{R,j} - \psi_{L,j}, \qquad \psi_{6,j} = 6\left(\overline{\psi}_j - \tfrac{1}{2}\left(\psi_{L,j} + \psi_{R,j}\right)\right). \tag{112}$$

The interpolation profile is then defined as

$$\psi(x) = \psi_{L,j} + \frac{(x - x_{j-1/2})}{\Delta x}\left(\Delta\psi + \psi_{6,j}\left(1 - \frac{(x - x_{j-1/2})}{\Delta x}\right)\right). \tag{113}$$

Integrating this profile backwards then gives an expression for the flux,

$$F^*_{j+1/2} = \nu\Delta x\left[\psi_{R,j} - \frac{\nu}{2}\left(\Delta\psi_j - \left(1 - \frac{2}{3}\nu\right)\psi_{6,j}\right)\right]. \tag{114}$$

15 Numerical methods for the shallow-water equations

Recall the 1D shallow-water equations

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} = 0, \tag{115}$$

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\tfrac{1}{2}u^2\right) + g\frac{\partial h}{\partial x} = 0, \tag{116}$$

where h is the fluid height, u is the fluid velocity and g is the gravitational constant $g = 9.80616$ m / s^2. One of the simplest discretization of these equations, which places height and velocity points at the same spot, is

$$h_j^{n+1} = h_j^n - \frac{\Delta t}{\Delta x}\left(h_{j+1}^n u_{j+1}^n - h_{j-1}^n u_{j-1}^n\right), \tag{117}$$

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{2\Delta x}\left((u_{j+1}^n)^2 - (u_{j-1}^n)^2\right) - \frac{\Delta t}{2\Delta x}\left(g h_{j+1}^n - g h_{j-1}^n\right). \tag{118}$$

This method is conservative in h (it is easy to see that the height equation satisfies a discrete conservation law), but is not conservative in u. The unstaggered grid can be viewed as the one-dimensional version of the Arakawa A-grid.

Instead of placing h and u nodes at the same point, it is popular in atmospheric models to instead *stagger* these variables. Essentially this means that the height variable exists in a finite-volume cell-averaged sense, whereas velocity is not conserved and is instead stored at cell edges. Again making use of the definition $x_j = j\Delta x$, we choose to store velocities at half indices, *i.e.* $j - 1/2$, $j + 1/2$ and $j + 3/2$. The discrete equations then take the form

$$h_j^{n+1} = h_j^n - \frac{\Delta t}{2\Delta x}\left(u_{j+1/2}^n\left(h_{j+1}^n + h_j^n\right) - u_{j-1/2}^n\left(h_j^n + h_{j-1}^n\right)\right), \tag{119}$$

$$u_{j+1/2}^{n+1} = u_{j+1/2}^n - \frac{\Delta t}{2\Delta x}\left((u_{j+3/2}^n)^2 - (u_{j-1/2}^n)^2\right) - \frac{\Delta t}{\Delta x}\left(h_{j+1}^n - h_j^n\right). \tag{120}$$

The staggered grid can be viewed as the one-dimensional version of the Arakawa C-grid.