Finite difference and finite volume methods for transport and conservation laws

Boualem Khouider

PIMS summer school on stochastic and probabilistic methods for atmosphere, ocean, and dynamics.
University of Victoria, July 14-18, 2008.

Contents

Foreword

The celebrated Chapman-Kolmogorov equation for a diffusion Markovian process reduces to the well known Fokker-Planck equation [Gardiner, 2004]

$$\frac{\partial p(\mathbf{z}, t/\mathbf{y}, t')}{\partial t} + \nabla_{\mathbf{z}} \cdot (\mathbf{A}(\mathbf{z}, t) p(\mathbf{z}, t/\mathbf{y}, t)) = \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial z_i \partial z_j} \left(B_{i,j}(\mathbf{z}, t) p(\mathbf{z}, t/\mathbf{y}, t) \right). \tag{1}$$

Here $p(\mathbf{z}, t/\mathbf{y}, t')$ is the probability density distribution of the random variable $\mathbf{z}$ at time t given $\mathbf{y}$ at time t' of the underlying Markovian process. $\nabla_{\mathbf{z}}$ is the gradient differential operator with respect to the variable $\mathbf{z}$, $\mathbf{A}(\mathbf{z}, t)$ is a vector function known as the drift, representing the deterministic dynamics of the process, and $B = [B_{i,j}]$ is the diffusion matrix describing the Gaussianity or randomness of the process. When $B = 0$ the Focker-Planck equation is also known as the Liouville equation, describing the evolution of the probability distribution of a random process undertaking deterministic dynamics. Taking the derivative with respect to the known variable y at time t', instead, yields the famous backward equation [Gardiner 2004]

$$\frac{\partial p(\mathbf{x}, t/\mathbf{y}, t')}{\partial t'} + \mathbf{A}(\mathbf{y}, t') \cdot \nabla_{\mathbf{y}} p(\mathbf{x}, t/\mathbf{y}, t') = -\frac{1}{2} \sum_{i,j} B_{i,j}(\mathbf{y}, t') \frac{\partial^2}{\partial y_i \partial y_j} \left(p(\mathbf{x}, t/\mathbf{y}, t') \right). \tag{2}$$

Note that while the two partial differential equations above are given in terms of the variables $\mathbf{z}$ and t or $\mathbf{y}$ and t', respectively, the remaining variables $(\mathbf{y}, t')$ for the first equation and $(\mathbf{x}, t)$ for the second) can be treated as parameters and therefore ignored when we are only concerned with numerical or analytic solution methodology for theses PDEs. This kind of equations are wide spread in the applied physical sciences. The term involving the vector $\mathbf{A}$ is also known as a transport process. In fluid mechanics, for example, it models the action of the flow field on the dynamical quantity under consideration, such as temperature, density, or momentum. It appears in either a conservative form

$$\partial_t q + \nabla \cdot (\mathbf{A} q) = 0 \tag{3}$$

as in the forward Focker-Plank equation (1) or in advective form

$$\partial_t q + \mathbf{A} \cdot \nabla q = 0 \tag{4}$$

as in the backward equation (2). Also under the obvious condition of ellipticity, the matrix B can be diagonalized and the associated diffusion operator is reduced to the more standard Laplace operator. Therefore, we are interested here in the numerical solution of the advection diffusion equation

$$\partial_t q + \mathbf{A} \cdot \nabla q = D \Delta q \tag{5}$$

or

$$\partial_t q + \nabla \cdot (\mathbf{A} q) = D \Delta q \tag{6}$$

which is a superposition of the transport equation of conservative or advective type and the diffusion equation

$$u_t = D \Delta u$$

also known as the heat equation. In this series of lectures we will discuss some standard numerical technics for these types of equations. Special emphasis will be given to finite difference and finite

volume methods for the advection and conservation equations in (4) and (3), respectively. We will treat in some details the case when the advection field $\mathbf{A}$ depends on the solution q that leads to shock formation and other types of singularities, which are important in gas dynamics and other fields of practical importance.

Unless otherwise stated, from now on, we consider only partial differential equations in the 2 variables (x,t), where $-\infty < x < +\infty$ represents the space variable and $t \geq 0$ is time. The solution is denoted by $u(x,t)$.

1 Introduction to finite differences: The heat equation

We introduce some basics of the finite difference methodology for partial differential equation through the simple case of the heat or diffusion equation in 1 dimension

$$u_t = Du_{xx},$$

where $D > 0$ is a constant heat conduction or diffusion coefficient.

The finite differences method applied to the heat equation above, starts by the approximation of the partial derivatives, u_t and u_{xx} by their corresponding finite difference quotients. For a smooth function $f(x)$ of the variable x, we have according to Taylor expansion

$$f(x_0+h) = f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)h^n + \frac{1}{(n+1)!}f^{(n+1)}(\xi)h^(n+1)$$

where h is a non zero increment or displacement along the real line, starting from a fixed point x_0 and ξ is between x_0 and x_0+h. Recall that, a function g is said a big O of h and we write $g = O(h^p)$ if

$$\lim_{h \longrightarrow 0} \frac{g(h)}{h^p} = \text{ Constant } .$$

Assuming $h > 0$ and using the Taylor approximation for $f(x_0+h)$ and $f(x_0-h)$, the forward and backward difference formulas follow immediately,

Forward Formula:

$$f'(x_0) = \frac{f(x_0+h)-f(x_0)}{h} - \frac{1}{2}f''(\xi)h = \frac{f(x_0+h)-f(x_0)}{h} + O(h) \approx \frac{f(x_0+h)-f(x_0)}{h}, \quad (7)$$

Backward Formula:

$$f'(x_0) = \frac{f(x_0)-f(x_0-h)}{h} + \frac{1}{2}f''(\xi)h = \frac{f(x_0)-f(x_0-h)}{h} + O(h) \approx \frac{f(x_0)-f(x_0-h)}{h}, \quad (8)$$

Furthermore, the 3rd order Taylor approximations of the difference $f(x_0+h)-f(x_0-h)$ yields the

Centered Formula:

$$\begin{aligned} f'(x_0) &= \frac{f(x_0+h)-f(x_0-h)}{2h} - \frac{1}{6}\frac{f'''(\xi_1)+f'''(\xi_2)}{2}h^2 \\ &= \frac{f(x_0+h)-f(x_0-h)}{2h} + O(h^2) \approx \frac{f(x_0+h)-f(x_0-h)}{2h}, \end{aligned} \quad (9)$$

where $x_0 - h \le \xi_1 \le x_0 \le \xi_2 \le x_0 + h$, whereas the 4th order Taylor approximation of the sum $f(x_0 + h) + f(x_0 - h)$ leads to an approximation for the second order derivative $f''(x_0)$.

Centered Formula for the second order derivative:

$$f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + \frac{1}{24}(\frac{(}{f}'''(\xi_1) + f'''(\xi_2))h^2$$
$$= \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2} + O(h^2) \approx \frac{f(x_0+h) - 2f(x_0) + f(x_0-h)}{h^2}, \qquad (10)$$

The formulas on the very-right hand sides of (7) to (10) are only some of the very basic examples of finite difference approximations for the first and second order derivatives to a first and second order accuracy, respectively. (The forward and backward finite difference approximations in (7) and (8) are first order accurate, therefore called first order approximations while those in (9) and (10) are second order accurate and are called second order approximations.) Finite difference formulas of higher order and for higher order derivative can be derived by using similar manipulations of Taylor approximations or polynomial approximations (e.g. interpolation). Also different combinations of points to the left or the right of the point x_0 can be considered separately.

A finite difference method for a given partial differential equation PDE consists of the approximation of the partial derivatives of its (unknown) solution u by a corresponding finite difference formula of a certain order.

1.1 Explicit scheme for the heat equation

Consider the heat equation

$$u_t = Du_{xx}$$

on a finite rod $x \in (0, L)$ with the initial condition $u(x, 0) = u_0(x)$ and boundary conditions $u(0, t) = \alpha(t)$ and $u(L, t) = \beta(t)$, $t \in [0, T]$. Consider a discretization of the rectangle $[0, L] \times [0, T]$ into a finite number of nodes $(x_j, t^n), j = 0, 1, \cdots, M + 1, n = 0, 1, \cdots, N$ such that $x_j = jh$ and $t_n = n\Delta t$ where $h \equiv \Delta x = L/(M + 1)$ and $\Delta t = T/N$. The set of all points (x_j, t^n) is called a grid or a mesh while h and Δt are respectively called the time step and spatial grid size. Let $u_j^n = u(x_j, t^n)$. Using a forward finite difference approximation for the time derivative combined with a centered formula for the second order spacial derivative, applied at each node (j, n), the heat equation can be rewritten as

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + O(\Delta t) = D\frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2} + O(h^2). \qquad (11)$$

Let w_j^n be the finite sequence of real numbers satisfying the following *difference scheme*

$$w_j^{n+1} = w_j^n + \frac{D\Delta t}{h^2}(w_{j+1}^n - 2w_j^n + w_{j-1}^n), j = 1, \cdots, M, n = 0, 1, \cdots N - 1, \qquad (12)$$

obtained from (11) by dropping the small error terms $O(\Delta t)$ and $O(h^2)$, known as the *truncation error*, and using the initial and boundary conditions

$$w_j^0 = u_0(x_j), w_0^n = \alpha(t_n), w_{M+1}^n = \beta(t_n).$$

This is the main philosophy behind finite differences of obtaining an approximate solution for the given PDE at the interior grid points

$$u(x_j, t_n) \approx w_j^n, \quad j = 1, \cdots, M, \quad n = 1, 2 \cdots, N.$$

As we shall see below, for the scheme (12) for the heat equation, we have

$$u(x_j, t^n) = w_j^n + O(\Delta t) + O(h^2),$$

i.e, the approximation is first order in time and second order in space, which is a statement of convergence as well, as $\Delta t, h \longrightarrow 0$, which is proved below after the introduction of the notions of consistency and accuracy. For a given finite discretization, an approximation for the solution $u(x,t)$ at a given interior point $(x,t) \in (0,L) \times (0,T)$–not on part of the grid can be obtained by 2d interpolation of the discrete solution w_j^n. Ideally, the order of interpolation would match that of the numerical approximation to obtain an optimal approximation in terms of efficiency and accuracy.

Note that the initial condition $u_0(x)$ and the boundary conditions $\alpha(t), \beta(t)$ provide the starting points $w_j^0, j = 0, 1, \cdots, M+1$ and the lateral grid point values w_0^n and $w_M^n, n = 1, 2, \cdots, N$ respectively while the numerical scheme is evolved from time step to time step using the formula (12) to provide the solution at time $t_n + \Delta t$, which is given explicitly in terms of the solution at time t_n. Such a scheme is called explicit as opposed to an implicit scheme where obtaining w_j^{n+1} from w_j^n involves the inversion of a linear or non-linear system of algebraic equations (see next subsection).

Definition 1 (Consistency):
A numerical scheme, $\mathcal{L}(w_j^n) = 0$, for a given PDE, $\mathcal{P}(u(x,t)) = 0$, is said consistent if the truncation error

$$\tau_h(h, \Delta t) \equiv \mathcal{P}(u(x,t)) - \mathcal{L}(u_j^n) \longrightarrow 0, \quad h, \Delta t \longrightarrow 0.$$

The scheme is said consistent of order (p,q) if

$$\tau_h(h, \Delta t) = O(h^p) + O(\Delta t^q).$$

For the forward explicit scheme (12) for the heat equation we have

$$u_t - D u_{xx} - \left(\frac{u_j^{n+1} - u_j^n}{\Delta t} - D \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{h^2} \right) = O(h^2) + O(\Delta t),$$

i.e, the forward scheme for the heat equation is consistent of order $(2,1)$.

Definition 2 (Stability):
A numerical scheme, $\mathcal{L}(w_j^n) = 0$ for an evolution partial differential equation on $[0,T]$ is said stable if the discrete solution satisfies

$$\max_{j=1,\cdots,M} |w_j^n| \leq C, \quad \forall n = 1, \cdots, N$$

where $C > 0$ is a constant independent on the grid size, $h, \Delta t$.

Theorem 1 (Convergence, Lax equivalence theorem)
If the original PDE problem is well posed then, the discrete solution of the numerical scheme converges to the solution of the PDE when $h, \Delta t \longrightarrow 0$ if and only if the scheme is consistent and stable.

Proof:
For simplicity in exposition, we assume that both the PDE and the numerical scheme are linear. Let $u(x,t)$ be the the solutions to the corresponding PDE and w_j^n the discrete solution of the numerical scheme. Let $u_j^n = u(x_j, t^n)$. The linear numerical scheme can be written as

$$w^{n+1} = L_{h,\Delta t} w^n$$

where $w^n = (w_j^n)$ is the $\mathbb{R}^n$ vector representing the discrete solution and $L_{h,\Delta t}$ is the linear operator (a matrix) associated with the numerical scheme, which depends on the grid size parameters h and Δt. For the forward scheme (12) for the heat equation, we have

$$(L_{h,\Delta t} w^n)_j = w_j^n + \frac{D\Delta t}{h^2}(w_{j+1}^n - 2w_j^n - w_{j-1}^n).$$

Note that the stability requirement implies that the operator L and all its powers stay bounded as $N, M \longrightarrow +\infty$ (or equivalently as $h, \Delta t \longrightarrow 0$), i.e,

$$\text{Stability} \implies ||L_{h,\Delta t}||^n \leq C, \ \forall N, M, n \geq 0, n \leq N.$$

To satisfy such condition if suffices to have

$$||L_{h,\Delta t}|| \leq 1.$$

However this later condition can be relaxed to

$$||L_{h,\Delta t}|| \leq 1 + O(\Delta t).$$

By the consistency requirement, we have

$$u^{n+1} = L_{h,\Delta t} u^n + O(\Delta t^2) + O(h^2)\Delta t.$$

Thus,

$$u^{n+1} - w^{n+1} = L_{h,\Delta t}(u^n - w^n) + O(\Delta t^2) + O(h^2)\Delta t.$$

Using basic linear algebra we have

$$||u^{n+1} - w^{n+1}|| \leq ||L_{h,\Delta t}||||u^n - w^n|| + O(\Delta t^2) + O(h^2)\Delta t,$$

and by induction on n, we arrive to

$$||u^{n+1} - w^{n+1}|| \leq ||L_{h,\Delta t}||^{n+1}||u^0 - w^0|| + (||L_{h,\Delta t}||^n + ||L_{h,\Delta t}||^{n-1} + \cdots + ||L_{h,\Delta t}|| + 1)(O(\Delta t^2) + O(h^2)\Delta t).$$

From the initial condition, we have $u^0 - w^0 = 0$. Thus,

$$||u^{n+1} - w^{n+1}|| \leq ((1 + O(\Delta t))^n + (1 + O(\Delta t))^{n-1} + \cdots + (1 + O(\Delta t)) + 1)(O(\Delta t^2) + O(h^2)\Delta t)$$

$$\leq \frac{(1+O(\Delta t))^{n+1}-1}{\Delta t}(O(\Delta t^2)+O(h^2)\Delta t) \leq (e^T-1)(O(\Delta t)+O(h^2)) \longrightarrow 0, \Delta t, h \longrightarrow 0,$$

and the rate of convergence is the same as the order of consistency, i.e, linear in time and quadratic in space.

Numerical Tests and stability of the forward scheme
Consider the PDE

$$\begin{aligned} u_t = \frac{1}{16}u_{xx}, \quad x \in (0,1), t \in (0,T) \\ u_(x,0) = \sin(2\pi x) \\ u(0,t) = u(1,t) = 0. \end{aligned} \tag{13}$$

The exact analytical solution for this PDE is given by

$$u(x,t) = e^{-\frac{1}{4}\pi^2 t}\sin(2\pi x).$$

The matlab code for solving this problem using the forward/explicit scheme (12) is given below in (14) and the results obtained with two different time-step sizes are plotted in Figure 13. The spacial discretization consists of 11 grid points and the time step is $\Delta t = 0.02$ for the top panel and $\Delta ?t = 0.2$ for the bottom panel. With $\Delta t = 0.02$, the numerical scheme (12) provides an accurate solution for this problem while the larger time step value $\Delta t = 0.2$ leads to a numerical solution that grows without bounds. Such behavior is known as a numerical instability. In fact, we show below that the forward scheme (12) is *conditionally stable* ; It is stable only for a relatively small Δt values; the largest eigenvalue of the linear operator associated with the forward scheme (12) is smaller or equal to one up to $O(\Delta t)$ provided $D\Delta t/h^2 \leq \frac{1}{2}$. Instead of going through the tedious task of computing the eigenvalues of the matrix, we use an alternate methodology for stability of difference schemes known as the *von Neumann analysis*.

```
%%%Forward scheme for the heat equation:
%%% INPUT
%%%%%Advection velocity:
mu = 1/16;
%%%%Grid size; Use periodic boundary conditions
X=1;M=10;Tend=4;
h=1/(M+1);
Dt=0.02;
x= 0:h:X; %%% x(1) =0, x(2) = h, , ..., x(M) = X -h; x(M+1) = X;
 wn=sin(2*pi*x);
 time=0;
 mu = mu*Dt/h^2;
 while(time<Tend)
     wn(2:M+1) = wn(2:M+1) +mu*(wn(3:M+2)-2*wn(2:M+1)+wn(1:M));
     time=time+Dt;
 end
 figure(2)
 xx=0:1/1000:1;
 plot(xx, exp(-time/4*pi^2)*sin(2*pi*xx))
```

```
hold on
 plot(x,wn,'x','linewidth',2)
```

Matlab code for the explicit scheme for the heat equation. (14)

1.2 Stability of the forward scheme: von Neumann analysis

Consider the forward in time centered in space numerical scheme for the heat equation

$$w_j^{n+1} = w_j^n + \frac{\Delta t D}{\Delta x^2}\left(w_{j+1}^n - 2w_j^n + w_{j-1}^n\right).$$

Consider simple solution for this difference equation on the form of

$$w_j^n = \rho^n e^{2\pi i j l h} \tag{15}$$

where $i = \sqrt{-1}$ and l is an integer, which can be thought of as a discrete version of the Fourier harmonics where the amplitude ρ is known as the *amplification factor.*

Theorem 2 (von Neumann)
A numerical scheme for an evolution equation is stable if and only if the associated largest amplification factor satisfies

$$|\rho| \leq 1 + O(\Delta t).$$

We skip the details of the proof here but in the nutshell it is due to the fact that in the linear case the largest eigenvalue of the difference scheme matches the amplification factor of von Neumann.

Inserting the expression of w_j^n in (15) into the forward scheme (12) yields

$$\rho = 1 + \frac{D\Delta t}{h^2}\left(e^{2\pi i l h} - 2 + e^{-2\pi i l h}\right) = 1 - 2\frac{D\Delta t}{h^2}\left(1 - \cos(2\pi l h)\right)$$

i.e

$$|\rho| \leq 1 \iff 2\frac{D\Delta t}{h^2}\sin^2(\pi h l) \leq 1 \iff \Delta t \leq \frac{h^2}{2D}.$$

Thus, the difference scheme (12) is conditionally stable. For the example (13) with $D = 1/16$ and $h = 1/11$ we have stability if and only if

$$\Delta t \leq \frac{16}{2 \times 11^2} \approx 0.06,$$

which explains the results in Figure 13. The condition of a time step being as small as h^2 is very bad news for this method especially if we want to integrate for a long period of time.

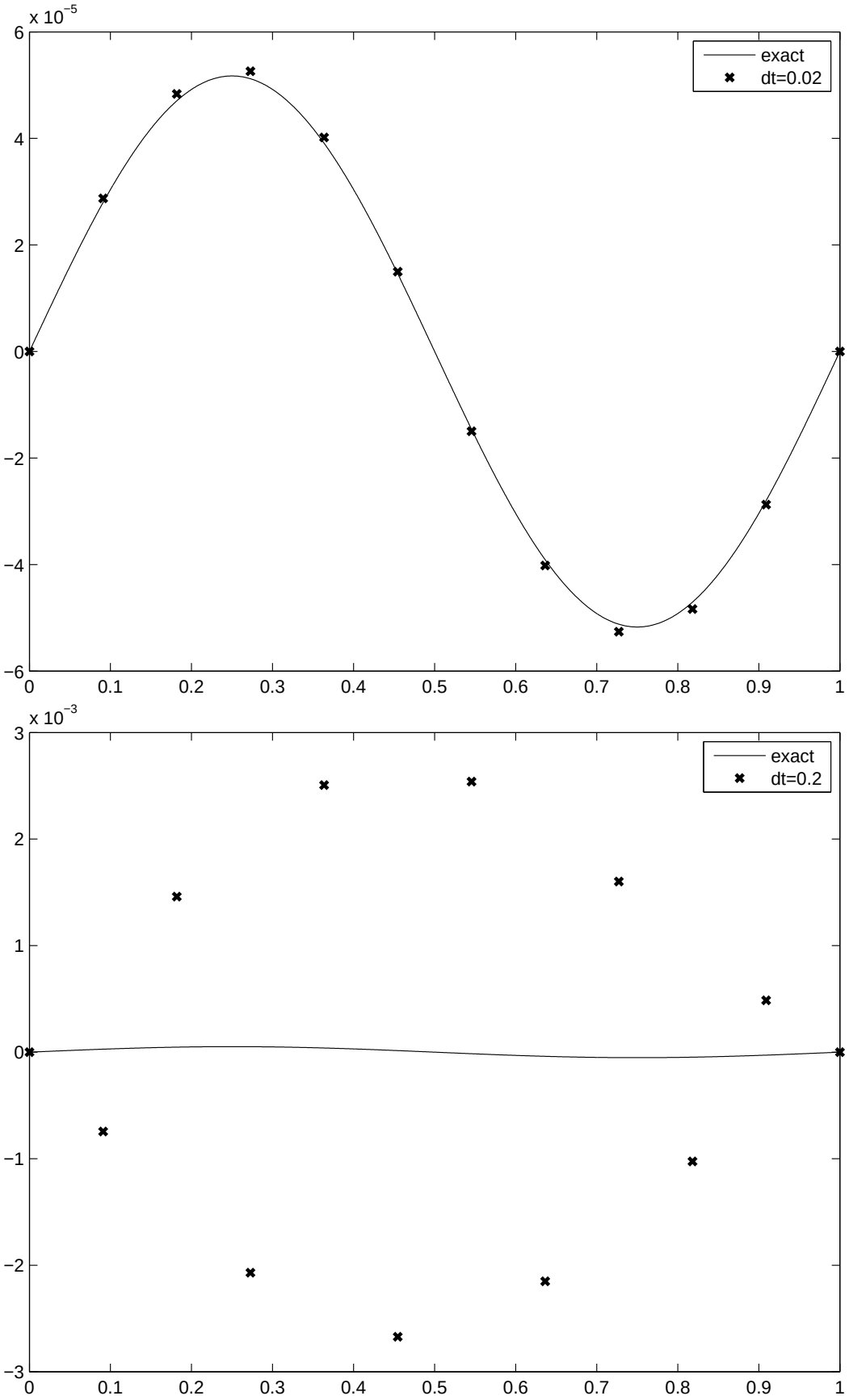

Figure 1: Explicit scheme for the heat equation. Numerical solution (crosses) compared to the exact solution (solid) for (13) at time t=4 with 11 spatial grid points. Top: $\Delta t = 0.02$, Bottom: $\Delta t = 0.2$

1.3 Implicit scheme for the heat equation

Instead of approximating the derivatives for the heat equation at time t_n using a forward finite deference in time let us instead consider an approximation at time t_{n+1} and use a backward difference in time. We arrive at the implicit/backward scheme for the heat equation

$$w_j^{n+1} = w_j^n + \frac{D\Delta t}{h^2}\left(w_{j+1}^{n+1} - 2w_j^{n+1} + w_{j-1}^{n+1}\right). \tag{16}$$

Note that because w_j^{n+1} is not given "explicitly" in terms of w_j^n, the time evolution of this scheme necessitates the inversion of a linear system of equations:

$$(I - \mu A)X^{n+1} = X^n + F^{n+1}$$

where

$$X = \begin{pmatrix} w_1^n \\ w_2^n \\ \vdots \\ w_M^n \end{pmatrix}, A = \begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix}, F^{n+1} = D\begin{pmatrix} \alpha(t_{n+1}) \\ 0 \\ \vdots \\ 0 \\ \beta(t_{n+1}) \end{pmatrix}$$

where $D = \frac{D\Delta t}{h^2}$.

It is clear from its derivation that this new scheme is consistent and is first order in time and second order in space.

Let us look at the stability properties of (16) using von Neumann's method. The amplification factor in this case satisfies

$$\rho = 1 - 4\frac{D\Delta t}{h^2}\rho\sin^2(\pi l h)$$

i.e,

$$0 \leq \rho = \frac{1}{1 + 4\frac{D\Delta t}{h^2}\rho\sin^2(\pi l h)} \leq 1$$

independently on the values of $D, \Delta t$ or h. The implicit scheme is *unconditionally stable*.

Because of this unconditional stability property the implicit scheme appears to be much superior than its explicit counterpart because in principle it can be run with an arbitrarily large time step and will still provide sensible results but it is much more expensive in terms of numerical operations per time step. Moreover, because it is only first order accurate in time, the time step required to achieve an accuracy on the order of h^2 for a given spatial grid size h is $\Delta t \approx h^2$, i.e, as small as the time step required to achieve a numerical stability with the explicit scheme...Ideally, we want a method which is both accurate and stable for large values of Δt at least as large as h. The Crank-Nicholson scheme described below is one of such methods.

1.4 The Crank-Nicholson scheme

Crank-Nicholson's scheme combines the forward/explicit and the backward/implicit schemes in (12) and (16) to provide a method which is both second order in time and space and unconditionally stable. A straight average of the two schemes (12) and (16), yields the Crank-Nicholson scheme for the heat equation

$$w_j^{n+1} = w_j^n + \frac{\mu}{2}\left(w_{j+1}^{n+1} - 2w_j^{n+1} + w_{j-1}^{n+1} + w_{j+1}^n - 2w_j^n + w_{j-1}^n\right). \tag{17}$$

Notice that the Crank-Nicholson scheme is implicit and as the backward-method it involves the solution of a linear system at each iteration.

$$(I - \frac{1}{2}\mu A)X^{n+1} = (I + \frac{1}{2}\mu A)X^n + \frac{1}{2}(F^{n+1} + F^n).$$

Exercise 1 *Show that the Crank-Nicholson scheme is consistent to the second order in both time and space and it is unconditionally stable. Hint: To show that it is second order accurate in both time and space, consider Taylor expansion in both time and space for the solution $u(x,t)$ about the cell-center $(x_j, t_n + \Delta t/2)$, highlighting the fact that the Crank-Nicholson method is centered in both time and space.*

2 Time splitting methods

Time splitting is an useful technique which consists on breaking down a complex PDE equation into a few simple parts for which numerical schemes are easily constructed and analyzed. For illustration we consider an advection diffusion equation in one space dimension, a reduced model for the Focker-Plank equation

$$u_t + a(x,t)u_x = Du_{xx}.$$

After discretization of the spatial derivatives, using the appropriate difference schemes or some other technique to approximate the spatial derivatives, we arrive to a linear system of differential equations with respect to time

$$\frac{d}{dt}w = Aw + Bw \tag{18}$$

where $-A$ is the discrete advection operator and B the discrete diffusion operator, and $w(t) = (w_j(t))_{1\le j\le M} \approx (u(x_j,t))_{1\le j\le M}$. Time splitting consists in dividing this linear systems onto two natural systems that are integrated separately and successively during each time step, each corresponding to the operators A and B, respectively. To integrate the discrete system (18) from t to $t + \Delta t$, we proceed as follows.

The time splitting algorithm:

1. Let $w_j^n = w_j(t_n)$ be given at time t_n

2. Solve $\frac{d}{dt}w_j^1 = Aw^1$ on $[t, t+\Delta t]$, with $w^1(t) = w(t)$

3. Solve $\frac{d}{dt}w_j^2 = Bw^2$ on $[t, t+\Delta t]$, with $w^2(t) = w^1(t+\Delta t)$

4. Set $w(t+\Delta t) = w^2(t+\Delta t)$, $t = t+\Delta t$ and proceed to step 2.

The main advantage of the time splitting method is that it permits to use numerical schemes that are known to converge and perhaps readily implemented for each one of the differential operators separately, e.g. the advection operator for which various numerical schemes will be designed below and the diffusion operator introduced above. However, the splitting methods introduce splitting errors which limit the overall order of accuracy to first order, as it is revealed by the consistency analysis performed next.

Let us analyze the consistency of the splitting methodology to see whether this is a sensitive method to use in practice. Assume A and B are two linear time independent operators, as in the case of the advection-diffusion problem. According to the theory of linear systems of differential equations, the solution to the total linear system (18) is given by

$$w(t+s) = e^{s(A+B)}w(t) = \left(I + s(A+B) + \frac{s^2}{2}(A+B)^2 + O(s^3)I\right) w(t)$$

$$= \left(I + s(A+B) + \frac{s^2}{2}(A^2+B^2+AB+BA) + O(s^3)I\right) w(t),$$

for s small (thinking $s = \Delta t$). One step of the splitting scheme yields

$$\tilde{w}(t+s) = e^{sB}w^2(t) = e^{sB}e^{sA}w(t) = \left(I + sB + \frac{s^2}{2}B^2 + O(s^3)I\right)\left(I + sA + \frac{1}{2}s^2A^2 + O(s^3)I\right) w(t)$$

$$= \left(I + s(A+B) + \frac{s^2}{2}(A^2+B^2+2BA) + O(s^3)I\right) w(t).$$

Thus $w(t+\Delta t) - \tilde{w}(t+\Delta t) = O(\Delta t^3)$ if $AB = BA$, i.e, if A and B commute with each other and $w(t+\Delta t) - \tilde{w}(t+\Delta t) = O(\Delta t^2)$ if A and B do not commute, which is generally the case in practice. Therefore the time splitting method is only first order accurate in time, but surprisingly it often provides good results. Nevertheless, a more elaborate version of the time splitting method which is second order accurate regardless the commutativity of the operators A and B is introduced by Strang and, therefore, known as the Strang-splitting method. Stang splitting method consists in symmetrizing the splitting operation by introducting an extra step to the time splitting algorithm, where one of the two operators is solved twice with one half time step. The Strang-splliting method is given next.

Strang-Splitting

1. Let $w_j^n = w_j(t_n)$ be given at time t_n

2. Solve $\frac{d}{dt}w_j^1 = Aw^1$ on $[t, t+\Delta t/2]$, with $w^1(t) = w(t)$

3. Solve $\frac{d}{dt}w_j^2 = Bw^2$ on $[t, t+\Delta t]$, with $w^2(t) = w^1(t+\Delta t/2)$

4. Solve $\frac{d}{dt}w_j^3 = Aw^3$ on $[t + \Delta/2, t + \Delta t]$, with $w^3(t + \Delta t/2) = w^2(t + \Delta t)$

5. Set $w(t + \Delta t) = w^3(t + \Delta t)$, $t = t + \Delta t$ and proceed to step 2.

Notice that in practice, when this algorithm is called successively in time, only the first and last time steps need to involve half time steps: performing step 5 followed by step 2 is equivalent to performing one full time step with the operator A. This may explain why the standard time splitting performs very well in practice.

Exercise 2 *Show that the Strang-splitting method is second order accurate.*

3 Introduction to quasi-linear equations and scalar conservation laws

A partial differential equation of the form

$$u_t + a(x, t, u)u_x + b(x, t, u) = 0, \tag{19}$$

is said *quasi-linear*. If in addition the coefficient a is independent of u and b is linear in u, then the equation is said linear. For practical purposes, more often than not, our 'computational domain' will be the finite interval $0 \leq x \leq 1$, for which a boundary condition is needed at least at one end of the domain. Also we assume that the equation in (19) is supplemented by an initial condition at $t = 0$:

$$u(x, 0) = u_0(x). \tag{20}$$

The couple pde + initial condition (19) and (20) is often referred to as a *the Cauchy problem*.

3.1 Prototype examples

Our main focus here is on the following two prototypes of equations: advection equations

$$u_t + a(x, t)u_x = 0 \tag{21}$$

and scalar conservation laws

$$u_t + (f(x, t, u))_x = 0, \tag{22}$$

both for being very important in many applications and as simple prototypes for more general and more complex models used in applications. For instance, in the general area of fluid mechanics, the advection equation is used as a model for the transport of *tracers* such as temperature, density, or the concentration of a certain chemical by the fluid flow, when the latter does not change with changes in the tracer u. It is widely used in biology and atmosphere-ocean sciences to model the concentration of pollutants and other substances in the presence of wind and or ocean currents. The term $a(x, t)$ represents the flow velocity, i.e, the wind or the stream of water–called as the

advection velocity or speed of propagation and $u(x,t)$ is some measure of the tracer–called the advected variable. More often advection models involve two to three space dimensions:

$$u_t + a_1(x,y,z,t)u_x + a_2(x,y,z,t)u_y + a_3(x,y,z,t)u_z = 0$$

$$\text{or } u_t + \mathbf{a}(x,y,z,t)\cdot\nabla u = 0$$

where $\mathbf{a} = (a_1, a_2, a_3)$ represents the three dimensional velocity field and ∇ is the gradient operator. Here we consider only the one space dimension case, extension to higher dimensions is mostly only technical. In some applications, the advection coefficient is stochastic, i,e, depends on a random variable or that the PDE/conservation law is forced by a stochastic forcing. Solving such a stochastic model numerical can be easily handled by the techniques developed here.

The conservation law equation on the other hand, models the evolution of a conservative quantity. That is a quantity whose variation inside a closed domain is equal to its flux across the boundaries, i.e, the amount which flows in minus the amount which flows out, of the closed domain. In one space dimension this can be stated as follows. Let $u(x,t)$ be the concentration density of such a conserved quantity within an interval $[x, x+dx]$. Let $f(x,t,u)$ be the flux of u at the extremity x and $f(x+dx,t,u)$ the flux at $x+dx$. This is illustrated in Figure 2 for both the 1D and 2D cases. By assuming the rate of change in the total quantity $\int_x^{x+dx} u(\xi,t,u)dy$ is equal to the total flux through x and $x+dx$, we can write

$$\frac{\partial}{\partial t}\int_x^{x+dx} u(x,t)dx = -f(x+dx,t,u(x+dx,t)) + f(x,t,u(x,t)). \tag{23}$$

Notice, the minus sign in front of $f(x+dx,t,u)$ guarantees that the flux on the right side is directed inward if f is positive and outward if f is negative. The opposite is true for the flux on the left side. The equation (23) is known as the integral form of the conservation law. Notice that no regularity with respect to x for u or f is required when we derived this equation. As we will see later this has some important consequences in designing numerical schemes for solving such equations.

Dividing by dx both sides of (23) and letting dx go to zero yields the differential form of the conservation law in (22), provided both u and f are smooth. However, in many textbooks and research papers, only the differential notation is used even when u or f are not smooth, in this case u is no longer a solution in the classical sense but only in a weak sense, which will be clarified below.

Another way to derive (22) from (23) is by noting that

$$-f(x+dx,t,u(x+dx,t)) + f(x,t,u(x,t)) = -\int_x^{x+dx} (f(y,t,u))_x dy,$$

which leads to the equality of the integrands because the length dx is arbitrary. This is actually the way the conservation law generalizes to 2 space dimensions and more. In fact, consider a closed domain Ω of $\mathbb{R}^2$ and let $\partial\Omega$ be its boundary. Then equating the total rate of change of u in Ω to the total flux of u across the boundary of Ω yields

$$\frac{\partial}{\partial t}\int_\Omega u(x,y,t)dxdy = \int_{\partial\Omega} F(x,y,t,u)\cdot\mathbf{n}d\Gamma$$

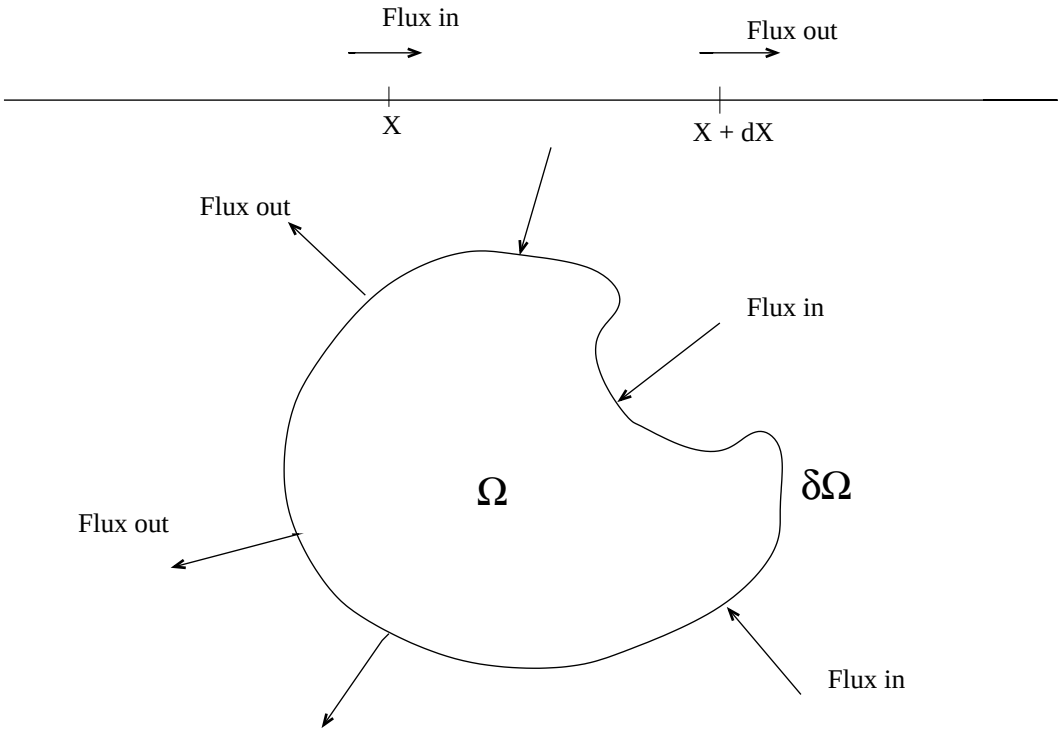

Figure 2: The integral form of the conservation law states that the total rate of change in u is compensated by the flux-in minus the flux-out.

where F is the flux vector and $\mathbf{n}$ is the unit normal vector to $\partial\Omega$ which is directed outside Ω. Invoking the divergence theorem yields

$$\frac{\partial}{\partial t}\int_{\Omega} u(x,y,t)dxdy = -\int_{\Omega} \nabla \cdot F(x,y,t,u)\, dxdy$$

for all bounded domain Ω. This yields the conservation law in differential form

$$u_t + \nabla \cdot F(x,y,t,u) = 0.$$

Links between the advection and the conservation law equations

The advection and the conservation law equations are intimately interconnected. In the case when $f = f(u)$ the conservation equation (22) can be rewritten in the advective form as

$$u_t + \frac{df(u)}{du}u_x = 0$$

and when a is constant with respect to x the advection equation can also be viewed as a conservation law. Moreover, note that in general every quasi-linear equation can be written as a conserved part plus a forcing

$$u_t + (f(t,x,u))_x + c(x,t,u) = 0.$$

Such equations are sometimes called balance laws and they are widely used in practice. The search for adequate–*numerically well balanced* schemes for this kind of equations has recently became a very active research area.

Finally, we note the advection equation in (21) is linear while the conservation equation in (22) can be non-linear if $\partial_u f \neq 0$. A very commonly used prototype example of a nonlinear conservation law is the celebrated Burger's equation:

$$u_t + \frac{1}{2}(u^2)_x = 0, \tag{24}$$

which becomes

$$u_t + u\, u_x = 0,$$

when written in advective form.

3.2 Solutions by the method of characteristics

Consider the quasi-linear equation in (19). Let $x = x(t)$ be a parametric curve in the (x, t) plane such that $\dot{x} = a(x, t, u(x(t), t))$ where $u(x(t), t) \equiv z(t)$ is the solution to (19) along this curve. Using the chain rule and plugging into the equation in (19) yields

$$\dot{z} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}\dot{x} = \frac{\partial u}{\partial t} + a(x, t, u)\frac{\partial u}{\partial x} = -b(x, t, u) = -b(x, t, z).$$

i.e, finding a solution for the quasi-linear equation reduces to solving the following system of first order ordinary differential equations.

$$\begin{aligned} \dot{x} &= a(x, t, z) \\ \dot{z} &= -b(x, t, z) \\ x(0) &= x_0, z_0 = u(x_0, 0) = u_0(x_0) \end{aligned} \tag{25}$$

Equations (25) are known as the characteristic equations and the resulting solution curves $x = x(t), x(0) = x_0$ are called characteristic curves.

Example 1: Solution to the advection equation
For simplicity we assume that the advection speed a is constant, in which case the advection equation reduces to since

$$u_t + a u_x = 0.$$

The characteristic equations for this simple example are

$$\begin{aligned} \dot{x} &= a \\ \dot{z} &= 0 \end{aligned}$$

whose solution is $x(t) = x_0 + at$, $z(t) \equiv u(x(t), t) = u_0(x_0)$. Two key important points should be noted here. i) The characteristic curves are straight lines. ii) The solution u is constant along the characteristic lines. The characteristic curves for the advection equation are sketched on Figure 3 for both $a > 0$ and $a < 0$. Note that when $a > 0$ the characteristics are directed to the right and when $a < 0$ they are directed to the left. In some sense the sign of a indicates the direction of propagation of information. In fact the advection equation is also called the one-way wave equation, where a is the speed of propagation of the wave.

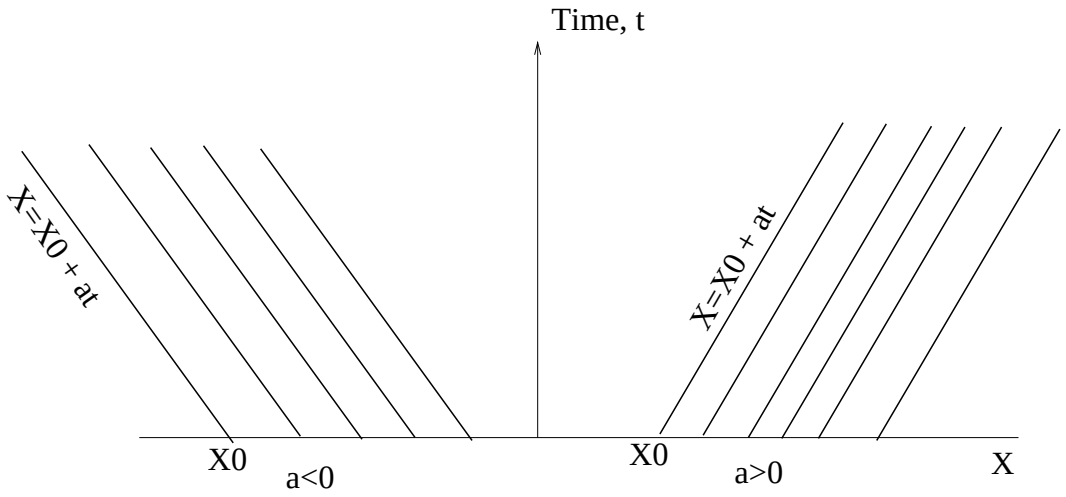

Figure 3: Characteristic lines for the advection equation. When $a > 0$ the characteristics are directed to the right and when $a < 0$ they are directed to the left.

To find the solution $u(x,t)$ at an arbitrary point (x,t) in the x-t plane, one needs to follow the characteristic line passing through (x,t) back to its original point at $t = 0$: $(x_0, 0)$ with $x = x_0 + at$. This leads to

$$u(x,t) = u_0(x_0) = u_0(x - at). \tag{26}$$

Example 2: Burger's equation
The Burger equation constitutes a somewhat more complex example a quasi-linear PDE. However, we still can, in principle, construct exact solutions using the method of characteristics.

The system of characteristic equations for Burger's equation is given by

$$\dot{x} = z$$
$$\dot{z} = 0.$$

The characteristic solution is thus given by

$$u(x(t),t) = u_0(x_0); \text{ where } x(t) = x_0 + u_0(x_0)t. \tag{27}$$

Again note that the characteristics are straight lines and the solution is constant along the characteristic lines, with one important difference, however; the characteristic curves are no longer parallel to each other. As we will see below this has rather "unpleasant" consequences. Provided the characteristics lines do not cross each other, which is guaranteed for at least a short period of time if the initial data u_0 is continuous, the solution to Burgers equation is given by the following implicit formula

$$u(x,t) = u_0(x - u_0(x_0)t), \quad x = x_0 + u_0(x_0)t.$$

The characteristic lines associated with Burger's equation are illustrated in Figure 4.

Exercise 3 *Use the method of characteristics to solve the following quasi-linear equations.*

$$u_t + xu_x = 0$$

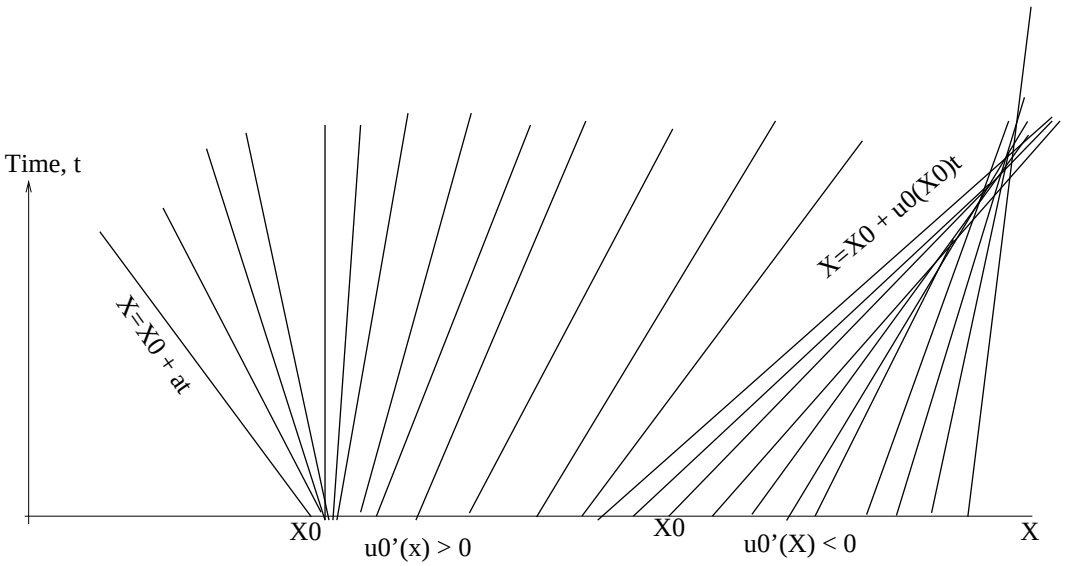

Figure 4: Characteristic lines for Burger's equation. When $u_0'(x) > 0$ the characteristics are divergent and when $u_0'(x) < 0$ they are converging toward each other.

and

$$u_t + u_x + x = 0.$$

Write down the solution $u(x,t)$ and draw the characteristic curves.

3.3 Notion of shocks and weak solutions

Note that because the slope of the characteristic curves $x = x_0 + u_0(x_0)t$ for Burger's equation (4) increases when $u_0(x_0)$ increases and decreases when $u_0(x_0)$ decreases, the characteristic curves will accordingly diverge or converge toward each other (see Figure 4). Two convergent characteristic lines will ultimately cross each other at some point in the x-t plane. Beyond such intersection point the characteristic solution is no-longer valid, because the value of $u(x,t)$ at such a point is not univalued–one can follow back either one of the two intersecting characteristic lines.

One way to correct for this flaw is by stopping the characteristic lines as soon as they cross each other. Let Σ be the set of such crossing points in the x-t plane. The solution can then be defined on both sides of Σ by following the corresponding characteristic line back to its origin. Below we will see that Σ is a parametric curve on the form $x = s(t)$ as show on Figure 5 and constitutes a curve of discontinuity for $u(x,t)$. Such a curve is called a *shock* by analogy to gas dynamics. One of the main difficulties in practice is to find the shock curve $x = s(t)$. For any given (x_1, t_1) one has to determine whether two characteristic lines cross each other prior to time t_1 along the curve $x = x_1$.

Exercise 4 *Show that a shock forms in the solution for Burger's equation if and only if the initial condition satisfies*

$$u_0'(x) < 0 \text{ for some } x$$

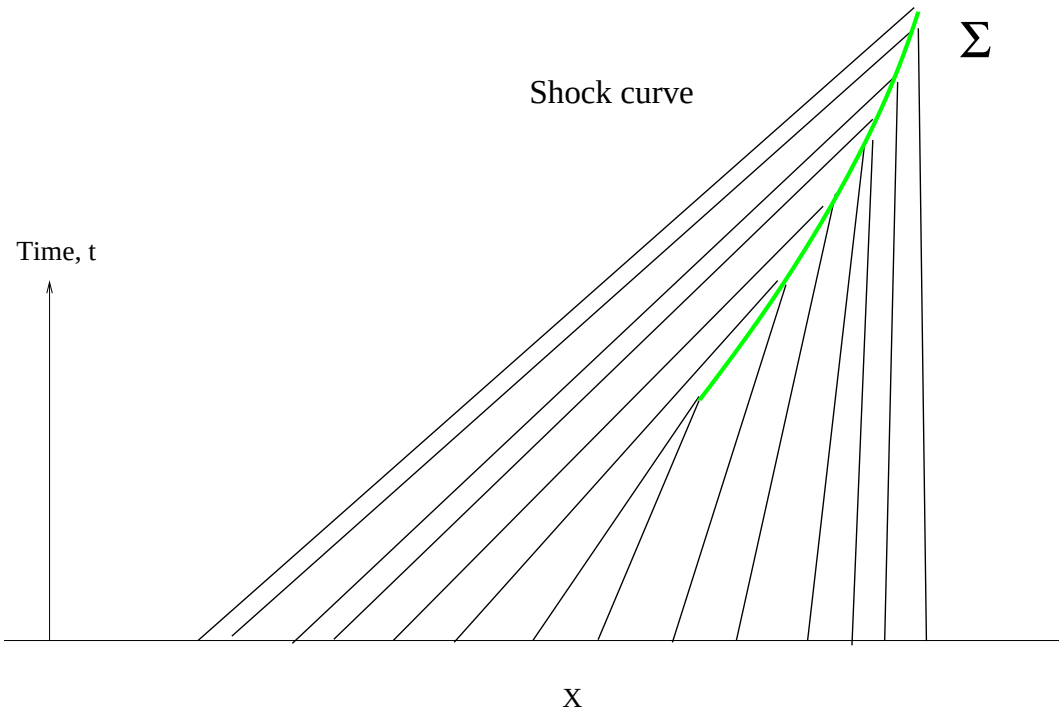

Figure 5: The shock curve Σ separates two regions of the x-t plane where the solution is smooth and is uniquely determined by the characteristics. The solution is discontinuous across the shock curve.

and that the first time a shock occurs is given by

$$T_* = -\frac{1}{\min_x u_0'(x)}.$$

After a shock is formed the solution $u(x,t)$ is no longer valid in the classical sense except for its restrictions on the sub-domains located on either side on the shock. Nevertheless, such solution can be defined in the *weak sense* on the whole x-t plane.

Definition 3 *A function $u(x,t)$ is said to be a weak solution for the conservation law*

$$u_t + (f(x,t,u))_x = 0$$

if for any test function $\phi(x,t)$ *sufficiently smooth (e.g. $\mathcal{C}^1$) with a compact support*[1] *in* $(-\infty,+\infty)\times(0,+\infty)$*, the solution $u(x,t)$ satisfies*

$$\int_0^{+\infty}\int_{-\infty}^{+\infty} u(x,t)\phi_t(x,t)\,dxdt + \int_0^{+\infty}\int_{-\infty}^{+\infty} f(x,t,u)\phi_x(x,t)\,dxdt = 0. \tag{28}$$

Remark:
Note that according to the definition of weak solutions, given above, a $\mathcal{C}^1$ function $u(x,t)$ is a

[1] i.e, there exist a bounded rectangle $[t_1,t_2]\times[a,b]\subset(-\infty,+\infty)\times(0,+\infty)$ such that $phi(x,t)=0$ out side this rectangle.

solution to the conservation law in the classical sense if and only if it is a solution in the weak sense. Therefore the notion of weak solutions is more general and the set of weak solutions contains discontinuous solutions as well as the classical $\mathcal{C}^1$ solutions as a special subset. However, in some situations weak solutions are not unique, in the sense that one initial value problem can have more than one weak solution. Selecting the physically relevant solution can be tricky. For Burger's equation for example, the physical solution coincides with the vanishing viscosity solution:

$$\bar{u}(x,t) = \lim_{\epsilon \longrightarrow 0} u_\epsilon(x,t)$$

where

$$\frac{\partial u_\epsilon}{\partial t} + u_\epsilon \frac{\partial u_\epsilon}{\partial x} = \epsilon \frac{\partial^2 u_\epsilon}{\partial x^2},$$

under the grounds that the unviscid Burger equation is a mathematical idealization of the viscous Burger equation when the viscosity is very small. However, given two weak solutions it is not easy to identify which one is the limiting viscosity solution and which one is not. The answer is provided by an extra condition known as the *entropy* condition satisfied by the physical solution. We will see this in detail below.

Note that the notion of weak solutions is very abstract and it is not obvious how one would handle this in practice. Nevertheless the theorem below provides the necessary ingredients both for constructing weak solutions and to gain physical insight.

Theorem 3 (Rankine-Hugoniot Condition) *Let Σ be a curve in $(-\infty,+\infty)\times(0,+\infty)$ parametrized by $x = s(t)$. Let $u(x,t)$ be a $\mathcal{C}^1$ function on both side of but possibly not defined and discontinuous across the curve Σ. Assume u is a solution to the conservation law*

$$u_t + (f(u))_x = 0$$

on all points (x,t) not on Σ. For each point $(x_1,t_1) \in \Sigma$ we set

$$u_\pm(x_1,t_1) = \lim_{\Omega_1 \cup \Omega_2 \ni (x,t) \longrightarrow (x_1,t_1)_\pm} u(x,t).$$

i.e. the limits from the right and from the left of Σ. Then $u(x,t)$ is a weak solution for the conservation law if and only if the shock speed $\dot{s}$ satisfies

$$\dot{s} \equiv \frac{ds}{dt} = \frac{f(u_+) - f(u_-)}{u_+ - u_-}. \tag{29}$$

The proof of this theorem is not terribly hard but it is quite technical and therefore left as an exercise for the interested student.

3.4 Discontinuous initial data and the Riemann problem

As it is pointed out above, the notion of weak solutions permits to define discontinuous solutions for conservation laws. Here we propose to construct such weak solutions with discontinuous initial

data. For simplicity, we consider the Burger equation with discontinuous initial data consisting of two constant states, a left and a right state:

$$\begin{aligned} &u_t + uu_x = 0 \\ &u_0(x) = \begin{cases} u_L & \text{if } x < 0 \\ u_R & \text{if } x > 0. \end{cases} \end{aligned} \tag{30}$$

The problem in (30) is known as the Riemann problem.

We propose to construct simple weak solutions for the Riemann problem associated with Burger's equation, using the Rankine-Huguniot condition (29).

Shock waves:
Consider a discontinuous function which consists of the two left and right constant states on both sides of a shock curve $\Sigma : x = s(t); s(0) = 0$

$$u(x,t) = u_R \text{ if } x > s(t) \tag{31}$$

$$u(x,t) = u_L \text{ if } x < s(t). \tag{32}$$

According to the Rankine-Hugoniot condition we have

$$\dot{s} = \frac{1}{2}\frac{u_R^2 - u_L^2}{u_R - u_L} = \frac{u_R + u_L}{2}.$$

Note that the shock speed in this case is constant and the curve Σ is a straight line. Also recall that the speed of the characteristic lines on each side of the shock is simply u_L and u_R, respectively. Therefore the speed of the shock is exactly halfway between the left and right characteristic speeds. This has physical sense and the associated weak solution is called a shock wave. A rough sketch of the shock wave solution is given in Figure 6 for both cases when $u_R < u_L$ and when $u_L > u_R$. Note that in the first case the characteristics run into the shock line and stop and therefore the solution on both sides of the shock is consistent with the characteristic solution while in the second case the characteristics diverge away and the region surrounding the shock is not reached by any of the characteristics.

Rarefaction waves:
Now assume that $u_L < u_R$ so the characteristics emanating from both side of the discontinuity are divergent from each other. In this case we can actually construct another weak solution to the Riemann problem associated with the Burger's equation. First note that for $t > 0$ the function $u(x,t) = x/t$, satisfies Burger's equation

$$u_t + uu_x = 0.$$

For $t > 0$ consider

$$u(x,t) = \begin{cases} u_L & \text{if } x < tu_L \\ x/t & \text{if } tu_L < x < tu_R \\ u_R & \text{if } tu_R < x. \end{cases} \tag{33}$$

First note that $u(x,t)$ is a solution to Burger's equation on each one of the designated parts of the domain and is continuous on the x-t plane. We can therefore show by simple integration by

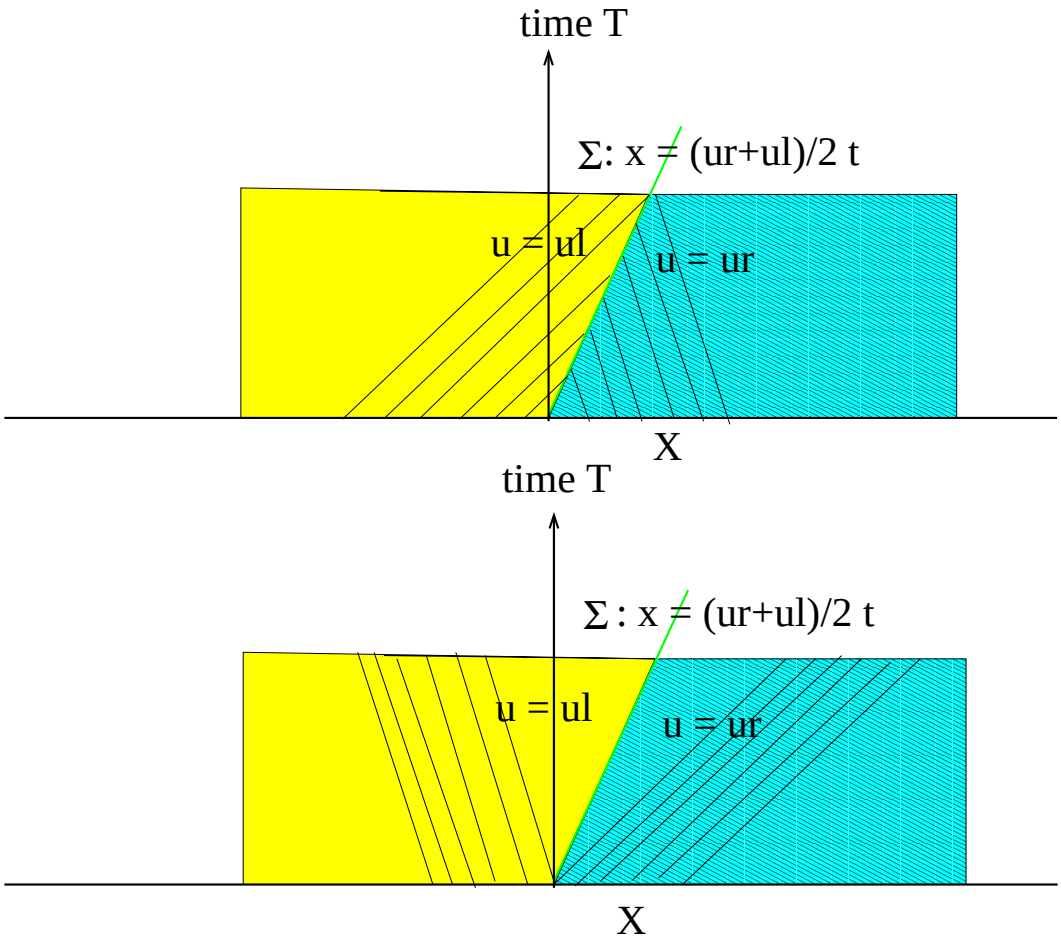

Figure 6: Shock wave solutions for Burger's equation. The two cases when $u_L > u_R$ (top) and when $u_L < u_R$ (bottom) are shown.

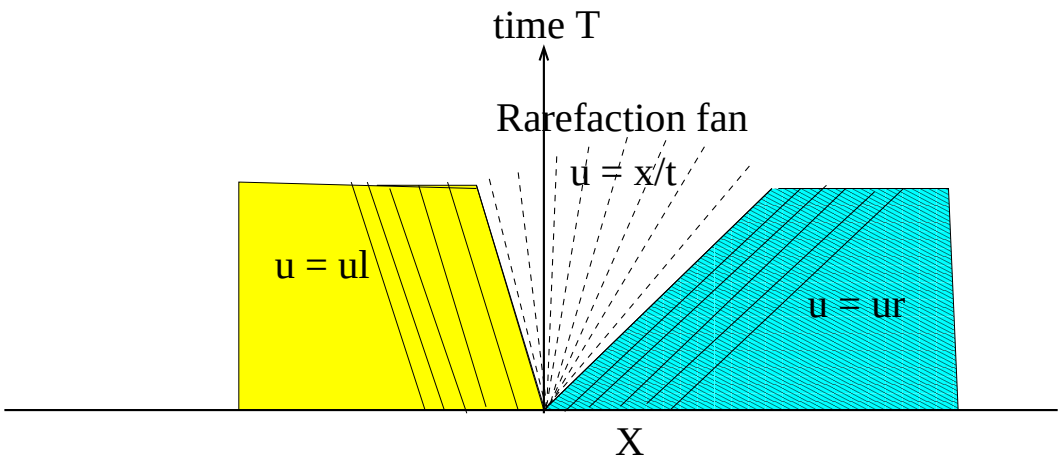

Figure 7: Rarefaction wave solution for Burger's equation. The rarefaction fan is shown.

parts that indeed $u(x,t)$ is a weak solution (see exercise 5 below). This type of solution is called *a rarefaction wave* by analogy to compressible gas dynamics and the solution $u = x/t$ in the middle is referred to as *a rarefaction fan*, see the illustration in Figure 7.

In summary, this shows that the Riemann problem for Burger's equation has at least two weak solutions when $u_L < u_R$, one is a shock wave and the other is a rarefaction wave. Therefore, weak solutions for conservation laws are in general non-unique.

Exercise 5 *Let Ω be a bounded open set in the xt-plane. Let Σ be a curve passing through Ω dividing it onto two disjoint open subsets $\Omega_{1,2}$ such that $\Omega = \Omega_1 \cup \Sigma \cup \Omega_2$. Let $\phi(x,t)$ be a smooth function (e.g. $\mathcal{C}^1$) supported in Ω, that is ϕ vanishes outside a compact set $K \subset \Omega$. Let*

$$u(x,t) = \begin{cases} u_1(x,t) & \text{if } (x,t) \in \Omega_1 \\ u_2(x,t) & \text{if } (x,t) \in \Omega_2, \end{cases}$$

where u_1, u_2 are two $\mathcal{C}^1$ functions satisfying the Burger equation in Ω_1, Ω_2, respectively. Use integration by parts to show that if in addition u is continuous across Σ, then

$$\int_\Omega u\phi_t + \frac{1}{2}u^2\phi_x \, dxdt = 0.$$

Deduce that (33) is a weak solution to Burger's equation.

3.5 Non-uniqueness of weak solutions and the entropy condition

As illustrated above with the example of a Riemann problem for Burger's equation, weak solutions are non-unique. However, common sense suggests that for any given Cauchy problem, only one solution is physically relevant. We need an additional constraint to choose this physically relevant solution among all the weak solutions. In fact, in reality some viscosity is always associated with a given conservation law, so instead we have

$$u^\epsilon_t + (f(u^\epsilon))_x = \epsilon \frac{\partial^2 u^\epsilon}{\partial x^2}, \tag{34}$$

and the zero viscosity limit, $\epsilon \longrightarrow 0$, is just a convenient mathematical idealization. It is true in many physical applications! Therefore, one universally accepted criterion states that the physically relevant weak solution for the conservation law $u_t + (f(u))_x = 0$ is the limit of the solution to the viscous equation (34) when $\epsilon \longrightarrow 0$. On the other hand it is easy to show that, when combined with appropriate initial conditions, the latter has a unique solution. In practice, however, it is not clear how to establish if a given weak solution is actually the vanishing viscosity limit or not. The answer to this question is provided by the concept of *entropic solutions*. In a nutshell, the entropy condition states that the physical solution satisfies a general principle of thermodynamics that the entropy always decreases. It remains to find which among the weak solutions for a given conservation law satisfies the so-called *entropy condition*. There are many versions of the entropy condition for a given conservation law

$$u_t + (f(u))_x = 0.$$

i) A somewhat abstract version of the entropy condition, but with a clear physical significance is as follows. Given a conservation law

$$u_t + f(u)_x = 0.$$

A convex function $\Phi(u)$ and a flux function $\Psi(u)$ are called an entropy/entropy flux pair if

$$\Psi'(u) = \Phi'(u) f'(u).$$

Given an entropy/entropy flux pair (Φ, Ψ), a solution u for the conservation law is said to be an entropic solution if it satisfies

$$\frac{\partial \Phi(u)}{\partial t} + \frac{\partial \Psi(u)}{\partial x} \leq 0, \tag{35}$$

in the weak sense. In many physical problem the entropy function $\Phi(u)$ is some measure of energy. It can be shown that for Burger's equation, an entropy/entropy flux pair is given by

$$\phi(u) = u^2; \Psi(u) = \frac{2}{3}u^3.$$

Here $\Phi(u) = u^2$ can be thought of as an energy density and $\Psi(u)$ is the energy flux. The most apparent merit of this formulation is that it generalizes 'easily' to systems of conservation laws.

ii) A more practical version of the entropy condition is the following:

$$u(x+z,t) - u(x,t) \leq C\left(1 + \frac{1}{t}\right) z, \;\; z, t > 0. \tag{36}$$

iii) Perhaps the most abstract of them is due to Kruzkov. A (weak) solution to the conservation law is said an entropy solution in the sense of Kruzkov if, in addition, it satisfies

$$\int_0^{\infty} \int_{-\infty}^{+\infty} sign(u-k)\left[(u-k)\phi_t + (f(u) - f(k))\phi_x\right] dxdt \geq 0 \tag{37}$$

for all real constants k and all test functions $\phi \geq 0$. This version of the entropic solution is useful for theory. It is mainly used by mathematicians to prove existence and uniqueness theorems.

iv) Finally, we give the *Lax entropy condition* for shocks: A shock solution for the Riemann problem with left and right states u_L, u_R is said an entropic shock if the shock speed, $\dot{s}$, satisfies

$$f'(u_R) < \dot{s} < f'(u_L). \tag{38}$$

It is easy to see that only one of the two shock solutions for the Riemann problem for Burger's equation is an entropic solution. Namely, the one associated with the case $u_R < u_L$ (by the Lax entropy condition iv) as well as condition ii)). Moreover, it can be shown that the unique entropic solution in the case $u_L < u_R$ is the rarefaction wave in (33).

4 Finite difference schemes for the advection equation

We start by discussing some basic simple finite difference schemes for the advection equation

$$u_t + au_x = 0,$$

where a is a positive constant. In principle these schemes are easily generalized to non-constant advection speeds with an arbitrary sign.

4.1 Some simple basic schemes

Throughout this paper we will assume the following discretization of space-time domain

$$x_j = j\Delta x; \quad t_n = n\Delta t,$$

where $\Delta x, \Delta t > 0$ are respectively the spatial and time step sizes. We denote by u_j^n the approximate/numerical solution to the solution $u(x_j, t_n)$.

Perhaps the most obvious scheme to attempt for the advection equation, which unfortunately turns out to be *unstable*, is obtained by taking a first order forward finite differencing in time and a centred differencing in space:

$$u_j^{n+1} = u_j^n - \frac{a\Delta t}{2\Delta x}(u_{j+1}^n - u_{j-1}^n). \tag{39}$$

This scheme is referred to below, simply, as the centred scheme. Other simple possibilities are to take a first order derivative in space either to the left or to the right, combined with the forward differencing in time, yielding the so-called first order *upwind* and *downwind* schemes, respectively:

$$u_j^{n+1} = u_j^n - \frac{a\Delta t}{\Delta x}(u_j^n - u_{j-1}^n) \tag{40}$$

and

$$u_j^{n+1} = u_j^n - \frac{a\Delta t}{\Delta x}(u_{j+1}^n - u_j^n). \tag{41}$$

Note that those two schemes are also known as the upstream and downstream schemes; depending on the application: hydrodynamics or gas dynamics (ocean v.s. atmosphere). The word upwind

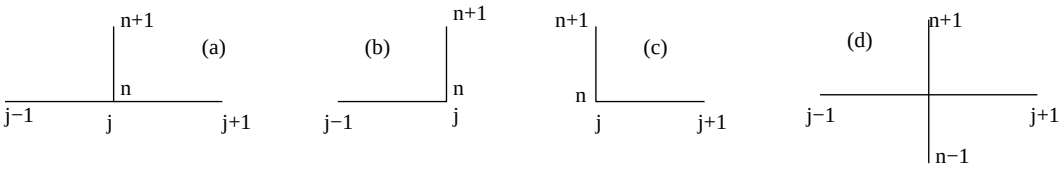

Figure 8: Simple finite difference stencils for the advection equation: (a) forward in time centred in space, (b) upwind, (c) downwind, (d) leap-frog.

refers to the fact that the finite differencing is performed in the direction opposite to the wind and downwind when the difference scheme follows the wind direction. Accordingly, when $a < 0$ the scheme in (41) becomes upwind and the scheme (40) becomes downwind.

Warning: As we will see below both the centred and the downwind schemes (39) and (41) are not recommended in practice because they are both unstable.

A slightly more sophisticated scheme is the *leap frog scheme*, which uses centred differences in both space and time:

$$u_j^{n+1} = u_j^{n-1} - \frac{a\Delta t}{\Delta x}(u_{j+1}^n - u_{j-1}^n). \tag{42}$$

The stencils for the four schemes listed above are given in Figure 8.

4.2 Accuracy and consistency

Definition 4 *Let $\mathcal{L}_h(u_h) = 0$ denote the numerical discretization using a certain method for a given partial differential equation denoted by $\mathcal{L}(u(x,t)) = 0$, with a time step, Δt, and grid spacing, Δx. The numerical scheme is said to be* consistent *if the truncation error:*

$$\tau_h = \mathcal{L}(u) - \mathcal{L}_h(u) \tag{43}$$

satisfies

$$\lim_{\Delta t, \Delta x \longrightarrow 0} \tau_h = 0.$$

The scheme is said consistent of order (p,q) accurate or simply of order (p,q) if

$$\tau_h = O((\Delta x)^p + (\Delta t)^q).$$

Examples:

Consider the advection equation

$$\mathcal{L}(u) \equiv u_t + a u_x = 0.$$

Using simple Taylor expansions, it is easy to see that the "centred scheme" (39) is consistent of order (2,1), namely

$$u_t + au_x - \frac{u(x,t+\Delta t) - u(x,t)}{\Delta t} - a\frac{u(x+\Delta x,t) - u(x-\Delta x,t)}{2\Delta x} = -\frac{\Delta t}{2}u_{tt}(x,\eta) - a\frac{(\Delta x)^2}{6}u_{xxx}(\xi,t)$$
$$= O(\Delta t + (\Delta x)^2).$$

i.e, this scheme is first order accurate in time and second order in space while the upwind and downwind schemes (40) and (41) are only first order in both space and time,

$$\tau_h(upwind/downwind) = O(\Delta t + \Delta x).$$

The leap-frog scheme (42), on the other hand is second order in both space and time

$$\tau_h(leapfrog) = O((\Delta t)^2 + (\Delta x)^2).$$

Exercise 6 *Show that the leap-frog scheme (42) is second order accurate in both time and space.*

4.3 Stability and convergence: the CFL condition and Lax-equivalence theorem

Definition 5 *A numerical scheme for an evolution equation on a finite interval* $[0,T]$*, on the form*

$$u_j^{n+1} = S_h(u_j^n)$$

is said stable if there exists a constant $C > 0$ *such that*

$$||u^n|| \leq C||u^0||$$

for a certain norm $||.||$ *in* R^N*, where* N *is the number of spatial grid points.*

Convergence of the numerical scheme
The convergence of a numerical solution u_h, obtained by a given numerical scheme, to the solution u of the original continuous equation is solved by the celebrated Lax-equivalence theorem.

Theorem 4 (Lax-equivalence theorem) *The approximate numerical solution to a well posed linear problem converges to the solution of the continuous equation if and only if the numerical scheme is linear, consistent, and stable. The rate of convergence and the order of accuracy of the numerical solution is equal to the truncation error of the numerical scheme.*

This elegant and powerful theorem is often summarized as follows

$$\text{consistency} \ + \ \text{stability} \ = \ \text{convergence} \ .$$

von Neumann stability analysis
The study of stability of a numerical scheme can be very tedious but the use of Fourier analysis when appropriate simplifies it a great deal. This idea was first used by von Neumann–apparently. For simplicity we assume that any discrete function, f_j, i.e, defined on the grid points $x_j = j\Delta x$ by its values $f_j = f(x_j)$ can be expanded in discrete Fourier modes

$$f_j = \sum_{l=0}^{N/2} \rho_l e^{ij\Delta x 2\pi l} + \text{ complex conjugate terms}$$

where $i = \sqrt{-1}$ and ρ_l are complex Fourier coefficients. N here is the number of spatial grid points and $N/2$ is known as the *Nyquist number.* It represents the largest wavenumber represented on an N-points grid.

To simplify the notation we set

$$\phi_l = 2\pi l \Delta x.$$

For the numerical solution u_j^n evolving in the discrete time, t_n, we have

$$u_j^n = \sum_{l=0}^{N/2} \rho_l^n e^{ij\phi_l}.$$

Note that the upper script n is an index not a power.

Theorem 5 (von Neumann Stability) *A numerical scheme for an evolution PDE is stable (in the sense of von Neumann) if and only if the ratio*

$$\rho_l = \frac{\rho_l^{n+1}}{\rho_l^n}$$

known as the amplification factor*, of discrete Fourier coefficients of the numerical solution, satisfies*

$$|\rho_l| \leq (1 + \Delta t), l = 0, \cdots, N/2.$$

Although the proof of this theorem is almost trivial, by Parseval's equality, it has a huge significance and a big impact on our way of studying numerical methods, because it is very easy to use, especially for linear problems. In fact, when the numerical scheme is linear, it is enough to consider solutions on the form of a single Fourier mode

$$u_j^n = \rho_l^n e^{ij\phi_l}. \tag{44}$$

Plugging in (44) into the

- the centred scheme (39) yields

$$\rho^{n+1} e^{ij\phi_l} = \rho^n e^{ij\phi_l} - \rho^n \frac{\mu}{2} e^{ij\phi_l} \left(e^{i\phi_l} - e^{-i\phi_l} \right)$$

where $\mu = a\Delta t/\Delta x$. Thus the amplification factor, $\rho \equiv \frac{\rho^{n+1}}{\rho^n}$, is given by

$$\rho = 1 - \mu i \sin(\phi_l).$$

Thus, $|\rho|^2 = 1 + \mu^2 \sin^2(\phi_l) > 1 + \Delta t$ for some values of ϕ_l, for all Δt sufficiently small, therefore the centred scheme (39) is unstable as suggested above.

- the backward (in space) first-order scheme (40) yields

$$\begin{aligned}\rho &= 1 - \mu(1 - e^{-i\phi_l}) = 1 - \mu(1 - \cos(\phi_l)) - i\mu \sin(\phi_l),\\ |\rho|^2 &= 1 + \mu^2(1 + \cos^2(\phi_l) - 2\cos(\phi_l)) - 2\mu(1 - \cos(\phi_l) + \mu^2 \sin^2(\phi_l)\\ &= 1 + 2\mu^2 + 2\mu(1 - \mu)\cos(\phi_l)) - 2\mu.\end{aligned}$$

Clearly if $\mu < 0$, i.e $a < 0$, $\rho > 1 + \Delta t$ for some values of ϕ_l, for all Δt sufficiently small, and when $\mu > 0$ ($a > 0$), $\rho \leq 1$ provided $0 \leq \mu \leq 1$. In other words, the upwind scheme is conditionally stable: $\Delta t \leq \Delta x/|a|$ or $|a|\Delta t \leq \Delta x$ while the downwind scheme is always unstable. Similarly we can show that the forward-scheme (41) is stable if $-1 \leq \mu \leq 0$ and unstable otherwise. Notice that the upwind scheme amounts to taking the derivative in the direction opposite to the advection speed, i.e, toward the direction where the information came from.

- the leap-frog scheme (42) yields

$$\rho^2 = 1 - 2\rho\mu i \sin(\phi_l),$$

which has 2 roots

$$\rho_\pm = -i\mu \sin(\phi_l) \pm \sqrt{1 - \mu^2 \sin^2(\phi_l)} \text{ if } |\mu| < 1 \tag{45}$$

$$|\rho_\pm|^2 = 1 \text{ if } |\mu| < 1.$$

When $|\mu| > 1$, one can always find a value for ϕ_l for which the two roots $\rho_\pm$ are both imaginary and then necessarily one of them has to be strictly larger than one for all values of Δt. Thus, as the upwind scheme, the leapfrog scheme is conditionally stable: $\Delta t \leq \Delta x/|a|$. Note however, that unlike the previous schemes, von Neumann analysis for the leap-frog method leads to two amplitude modes, $\rho_\pm$, for a given spatial mode $\exp(ij\phi_l)$. Nevertheless, as we will see below, only one of them is physical, namely, ρ_+, that is it represents an approximation (converges) to the exact solution, while the other one is an artifact of the numerical discretization. The latter is often called a computational or a parasitic mode.

CFL condition:
The CFL condition, named after its discoverers, Courant, Friedrichs, and Lewy, states that if a difference scheme for an evolution equation is stable then the *domain of dependence*, corresponding to one time step, of the numerical scheme contains the domain of the dependence of the original continuous equation, when both the time step and the spatial grid size $\Delta t, \Delta x \longrightarrow 0$.

Let's first clarify the notion of domain of dependence. Given the partial differential equation,

$$u_t = F(u, u_x)$$

Table 1: Some simple schemes for the advection equation and their properties.

Scheme	Order of accuracy	Stability	CFL condition
Centred	$O(\Delta t + (\Delta x)^2)$	Unstable	Satisfied if $\Delta t \leq \Delta x/\|a\|$
Upwind	$O(\Delta t + \Delta x)$	Stable if $\Delta t \leq \Delta x/\|a\|$	Satisfied if $\Delta t \leq \Delta x/\|a\|$
Downwind	$O(\Delta t + \Delta x)$	Unstable	Not satisfied
Leap-frog	$O((\Delta t)^2 + (\Delta x)^2)$	Stable if $\Delta t \leq \Delta x/\|a\|$	Satisfied if $\Delta t \leq \Delta x/\|a\|$

with and initial condition

$$u(x,t) = u_0(x),$$

the domain of dependence of this equation at some point (x,t) is the set of all points x_0 such that $u(x,t)$ depends on the value of u_0 at x_0). For instance according to the solution by the method of characteristics to the advection equation (26), the domain of dependence of the advection equation is reduced to one point

$$D_{(x,t)} = \{x_0 = x - at\}.$$

The domain of dependence corresponding to one time step is

$$D_{(x,t+\Delta t)} = \{(x - a\Delta t, t)\}.$$

This is illustrated in figure 9 for $a > 0$. The numerical domain of dependence of the numerical schemes considered so far are as follows

$$D_{centred}(x, t+\Delta t) = \{(x - \Delta x, t), (x + \Delta x, t)\}$$

$$D_{backward}(x, t+\Delta t) = \{(x - \Delta x, t), (x, t)\}$$

$$D_{forward}(x, t+\Delta t) = \{(x, t), ((x + \Delta x, t)\}$$

$$D_{leapfrog}(x, t+\Delta t) = \{(x - \Delta x, t), (x, t - \Delta t), (x + \Delta x, t)\}$$

Note that, as shown in Figure 9, the point $(x - a\Delta x, t)$ is within the range the domain of dependence of all the above schemes so that the CFL condition is satisfied, provided the stability condition $|\Delta t \leq \Delta x/|a|$ is satisfied, except for the forward scheme, corresponding to downwind differencing in this case, which explains its instability. Notice that this makes physical sense. As it is suggested by the method of characteristics, when $a > 0$ the solution to the advection equation uses information on the left to advance forward in time while the forward scheme uses information on the right. Note that the upwind and the leapfrog schemes do not have this problem. Also, the condition $|a|\Delta t \leq \Delta x$ can be interpreted as a requirement that the numerical speed of propagation of information being smaller than the advection speed of the continuous problem.

Interestingly note that the centred scheme (39) satisfies the CFL condition, if $\Delta t \leq \Delta x/|a|$, as for the leap-frog scheme, but is not stable. This is a good example illustrating the important fact that the CFL condition is a necessary condition for stability but it is not sufficient.

Table 1 summarizes the properties of the four simple schemes we have covered so far.

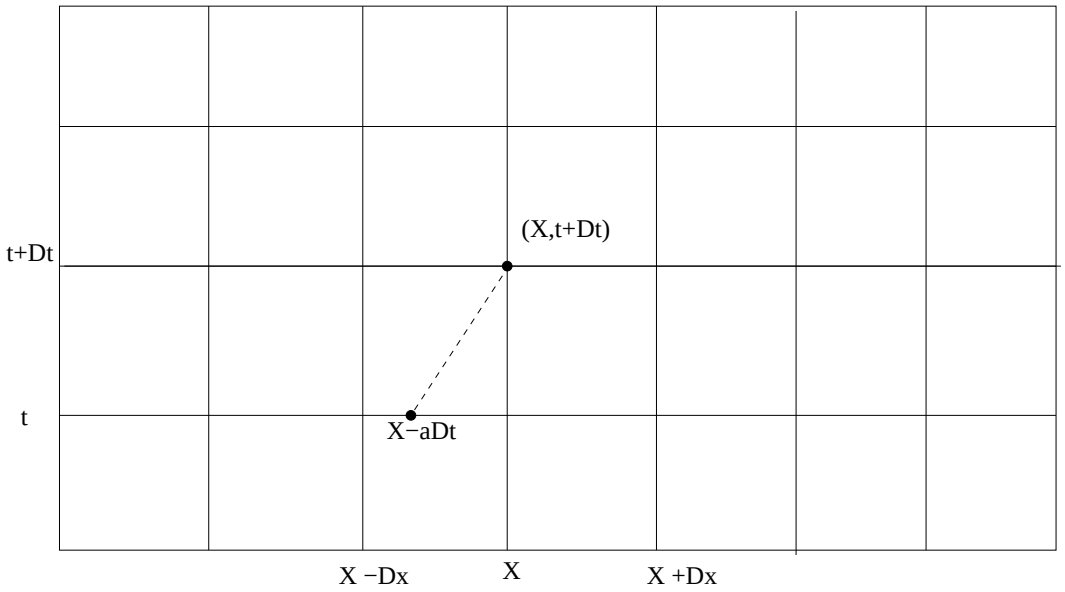

Figure 9: Domain of dependence of the advection equation for $a > 0$ and the CFL condition.

4.4 More on the leap-frog scheme: the parasitic mode and the Robert-Asselin filter

According to table 1, the best we have, among the simple schemes seen so far, is the leap-frog scheme; it is stable and second order accurate in both space and time. Notice also as such it is relatively cheap and easy to implement. This may explain in part why this scheme is so popular in the engineering and atmosphere/ocean communities. However, it has at least two drawbacks: 1) it is a multistep (2 steps or 3 levels) scheme which means it needs the knowledge of the solution at two successive time steps in order to advance to the next one and 2) it carries a parasitic mode which may ruin the numerical solution when used for long time integrations, if it is not filtered carefully. More details on the leap-frog scheme and its parasitic mode can be found in the literature (e.g. Durran). Here we briefly illustrate and demonstrate the behaviour of this parasitic mode and give a strategy for controlling it in practice.

Recall the von Neumann amplification factors associated with the leap-frog scheme

$$\rho_\pm = -i\sigma \pm \sqrt{1-\sigma^2}$$

where we set $\sigma = \mu \sin(\phi_l)$. Since $|\rho_\pm| = 1$ we have

$$\rho_\pm = e^{i\psi_\pm}, \quad \psi_\pm = \arctan\left(\frac{-\sigma}{\pm\sqrt{1-\sigma^2}}\right)$$

with $\psi_0 = \arcsin(\sigma)$ we have

$$\psi_+ = -\psi_0, \quad \psi_- = -\pi + \psi_0,$$
$$\rho_+ = e^{-i\psi_0}, \quad \rho_- = e^{i(-\pi+\psi_0)},$$

and the Fourier mode solution for the leap-frog scheme is given by

$$u_j^n = (c_1\rho_+^n + c_2\rho_-^n)e^{ji\phi_l} = c_1 e^{i(j\phi_l - n\psi_0)} + c_2(-1)^n e^{i(j\phi_l + n\psi_0)}$$
$$\approx c_1 e^{2\pi l i(x_j - at_n)} + c_2(-1)^n e^{2\pi l i(x_j + at_n)} \tag{46}$$

where we used the fact that $\mu = a\Delta t/\Delta x, \phi_l = 2\pi l\Delta x, n\Delta t = t_n; j\Delta x = x_j$ and the approximation

$$\arcsin(\sigma) \approx \sigma; \sin(\phi_l)/\phi_l \approx 1.$$

The expression in (46) clarifies that the term representing ρ_+ has the form $f(x-at)$ and therefore provides an approximation for the solution to the advection equation while the remaining term represents a wave moving in the opposite direction (to the left) and its amplitude oscillates between positive and negative values. Clearly, the latter is an artifact of the numerical discretization, called the computational or parasitic mode, and may lead to serious damage to the solution if it is not controlled in some way.

There are many ways how to control the computational mode. One of them is to make sure that the 'extra initial condition', i.e, solution at first time step, needed to advance the leap-frog method is chosen so that the coefficient c_2, of the parasitic mode, is zero fin the decomposition of u_0 into Fourier modes. One easy way to guarantee that initially the parasitic mode is zero is to set as second initial data at $t = \Delta t$, required to evolve the multi-step leap-frog method, to be

$$u_j^1 = \sum_l \rho_l^+ \rho_l^0 e^{ij\phi_l}$$

given at $t = 0$ we have

$$u_j^0 = \sum_l \rho_l^0 e^{ij\phi_l}.$$

However, we known that for long integration periods the parasitic mode can be excited and grow just from round-off errors.

A safe commonly used strategy for controlling the computational mode for the leap-frog scheme, known as the Robert-Asselin filter, is given next. The Robert-Asselin filter consists in averaging the solution at every time step by using the previous and the future solution, at $t - \Delta t$ and $t + \Delta t$, respectively, by introducing an extra-filtering step. Let $\bar{u}_j^n$ denote the solution filtered in such a way. The two step leap-frog plus Robert-Asselin filtering scheme is given by

$$u_j^{n+1} = \bar{u}_j^{n-1} - \mu(u_{j+1}^n - u_{j-1}^n)$$
$$\bar{u}_j^n = u_j^n + \gamma(u_j^{n+1} - 2u_j^n + \bar{u}_j^{n-1}), \tag{47}$$

where γ is a small filtering parameter usually taking to be $\gamma = 0.06$. Some atmospheric models used for cloud physics use values as large as $\gamma = 0.3$ (Durran). It is important to note that the filtering step destroys the second order accuracy in time and the resulting scheme is only first order.

4.5 The Lax-Friedrichs scheme

The Lax-Friedrichs scheme is a clever modification for the unstable centred scheme (39), which makes it stable. It consists of replacing the term u_j^n by the average from the neighbouring cells to

obtain

$$u_j^{n+1} = \frac{u_{j+1}^n + u_{j-1}^n}{2} - \frac{a\Delta t}{2\Delta x}\left(u_{j+1}^n - u_{j-1}^n\right). \tag{48}$$

This is the celebrated Lax-Friedrichs scheme. As the centred scheme the Lax-Friedrichs scheme (48) is second order accurate in space and first order in time. However, it is stable under the CFL condition $|\mu| \equiv |a|\Delta t/\Delta x < 1$. In fact, the associated von Neumann amplification factor is given by

$$\rho = \cos(\phi_l) - i\mu\sin(\phi_l)$$
$$\implies |\rho|^2 = \cos(\phi_l)^2 + \mu^2\sin(\phi_l)^2 \leq 1 \text{ if } |\mu| \leq 1.$$

Exercise 7 *Show that the Lax-Friedrich's scheme is first order in time and second order in space and that its amplification factor is given by*

$$\rho = \cos(\phi_l) - i\mu\sin(\phi_l).$$

The apparent advantages of the Lax-Friedrichs scheme, when compared to the three other stable schemes listed above is that it is second order in space and is a one step method, therefore easier to implement in practice compared to the leap-frog method. However, it is only first order accurate in time, which limits its use in practical applications and it is extremely *dissipative* as we will see below.

4.6 Second order schemes: the Lax-Wendroff scheme

Perhaps the simplest way to achieve a one-step (2 levels) scheme, which is second order in both time and space for the advection equation, is to resort to Taylor expansion in the time variable

$$u(x, t+\Delta t) = u(x,t) + \Delta t u_t(x,t) + \frac{(\Delta t)^2}{2}u_{tt}(x,t) + \cdots$$

using the advection equation $u_t = -au_x$ yields

$$u(x, t+\Delta t) = u(x,t) - a\Delta t u_x(x,t) + \frac{a^2(\Delta t)^2}{2}u_{xx}(x,t) + \cdots$$

Now, we use second order centred differencing to approximate the spatial derivatives u_x and u_{xx} to obtain the **Lax-Wendroff** scheme

$$u_j^{n+1} = u_j^n - \frac{\mu}{2}\left(u_{j+1}^n - u_{j-1}^n\right) + \frac{\mu^2}{2}\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right). \tag{49}$$

It is easy to show that the Lax-Wendroff scheme (49) is second order accurate in both space and time and it is stable under the CFL condition $a\Delta t/\Delta x \leq 1$. The amplification factor for this scheme is

$$\rho = 1 - i\mu\sin(\phi_l) - \mu^2(1 - \cos(\phi_l)).$$

Exercise 8 *Show that the Lax-Wendroff scheme is second order in both time and space and it is stable under the CFL condition* $a\Delta t/\Delta x \leq 1$.

If instead we use a second order approximation of the first order space derivative in the Taylor expansion using points that are only in the upwind direction, namely, x_{j-2}, x_{j-1}, x_j when $a > 0$, we obtain the following scheme, known as the **Beam-Warming's** scheme:

$$u_j^{n+1} = u_j^n - \frac{\mu}{2}\left(3u_j^n - 4u_{j-1}^n + u_{j-2}^n\right) + \frac{\mu^2}{2}\left(u_j^n - 2u_{j-1}^n + u_{j-2}^n\right). \tag{50}$$

This scheme can be derived by using polynomial interpolation. In fact, let

$$P_2(x) = u_{j-2} + \frac{u_{j-1} - u_{j-2}}{\Delta x}(x - x_{j-2}) + \frac{u_j - 2u_{j-1} + u_{j-2}}{2(\Delta x)^2}(x - x_{j-1})(x - x_{j-2})$$

be the 2nd degree polynomial interpolating u on the grid points x_j, x_{j-1}, x_{j-2}. Using this polynomial to approximate the derivatives u_x, u_{xx} at $x = x_j$ yields the Beam-Warming scheme. The details are left as an exercise for the student.

Note that because the stencil of the Beam-Warming scheme uses two grid cells on the left of x_j, the associated CFL condition becomes $0 \leq a\Delta t/\Delta x < 2$. In fact we can show that this scheme is stable under this CFL condition. Compared to Lax-Wendroff's scheme this new scheme allows larger time steps. However, if the advection speed is not too large the time step might be restricted to that of the Lax-Wendroff for accuracy reasons.

Exercise 9 *Show that the Beam-Warming scheme is second order in both time and space and it is stable under the CFL condition* $a\Delta t/\Delta x \leq 2$.

Exercise 10 *Derive the version of the Beam-Warming scheme when* $a < 0$.

4.7 Some numerical experiments

Here we assess the performance of each one of the (stable) methods listed above, for the advection equation, the interval $[0, 1]$ with periodic boundary conditions and two different initial data. We consider a smooth initial condition consistent of a hump-like shaped profile,

$$u_0(x) = \exp(-100(x - .5)^2),$$

and a non-smooth–piece-wise constant, initial condition, called here *a square wave*

$$u_0(x) = 0 \text{ if } |x - 0.5| > 0.25, \quad u_0(x) = 1 \text{ if } |x - 0.5| < 0.25.$$

The advection velocity is assumed constant and normalized to $a = 1$ and integrate to time $t = 1$ so that the initial profile is moved forward by a distance equal to the length of the interval $[0, 1]$ and by periodicity we have $u(x, t = 1) = u_0(x)$. We use a mesh size $\Delta x = 0.01$ and a time step $\Delta t = 0.008$ corresponding to a *Courant number* $\mu = \Delta t/(|a|\Delta x) = 0.8$.

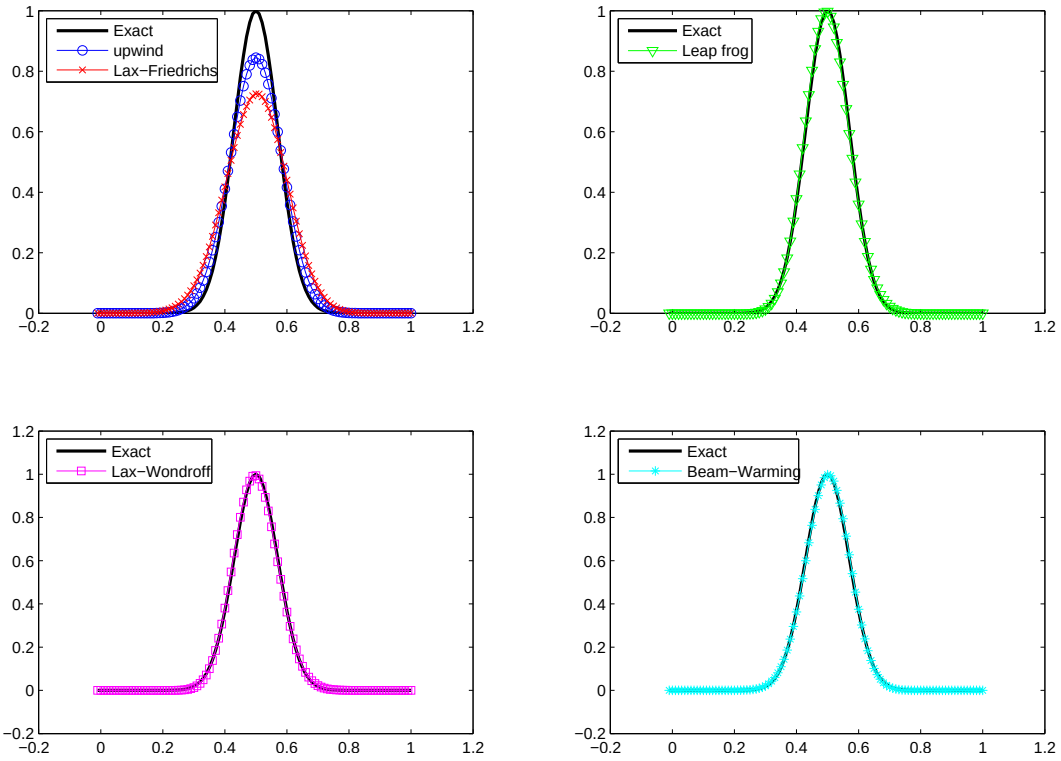

Figure 10: Solution to the advection equation using different finite difference schemes. Case of a smooth hump advected periodically through the interval $[0, 1]$

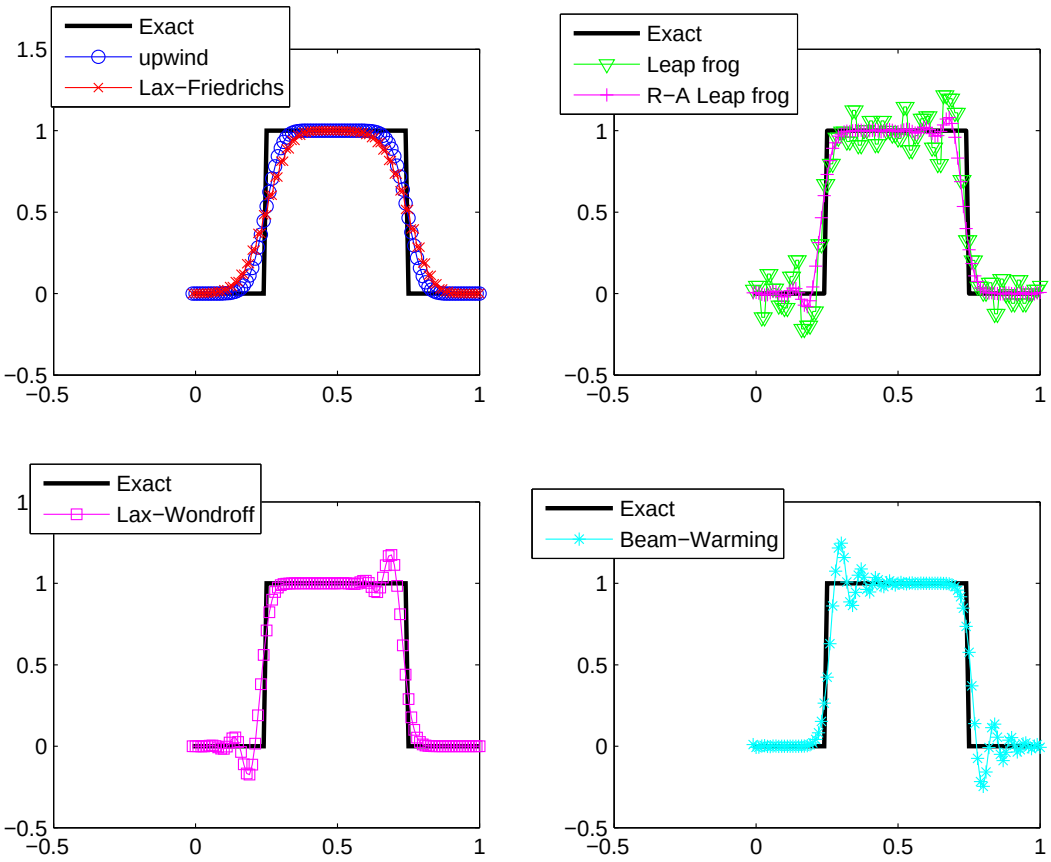

Figure 11: Same as Fig. 10, except for the non-smooth, piece-wise constant, square wave. Note that for this case results with both the plain leap-frog and leap-frog combined with the Robert-Asselin filter are shown.

In Figure 10, we plot the numerical solutions, obtained with each one of the 5 methods, against the exact solution for the advection equation. In Figure 11 we report the similar plots corresponding to the non-smooth square wave. Here we note a few important points.

- First, for the smooth solution in Figure 10, as expected the second order methods (leapfrog, Lax-Wendroff, and Beam-Warming) are highly accurate whereas the first order upwind and Lax-Friedrichs methods yield very unsatisfactory results. They both suffer from an excessive *dissipation*; that is the wave amplitude decays in time. Surprisingly, the upwind scheme seem to perform better than Lax-Friedrichs, although, the latter is second order accurate in space.
- Second, for the non-smooth case in Figure 11, on the other hand, the two first-order methods seem to perform better than the three second order methods, although they tend to smooth out the discontinuity. The second order schemes exhibit high oscillations near the discontinuities. This is typical for high order methods, they are exhibit an oscillatory behaviour near shocks. Note also that Lax-Wendroff is somewhat better than Beam-Warming and that Lax-Wendroff produces oscillations behind the shock while the oscillations in the Beam-Warming are located in front of the discontinuity.

  Below, we will see how to design *high resolution or non-oscillatory* scheme which in principle are second order accurate in regions where the solution is smooth and only first order–non oscillatory near the shocks.
- Third, note that the leap-frog scheme seem to the exhibit the worst oscillatory behaviour for the non-smooth solution. In fact, most of this is due to the presence of the computational mode which tend to amplify the oscillations. Things look a lot better when the the computational mode is controlled by the Robert-Asselin filter. Notice that the filtered leap-frog is somewhat between the Lax-Wendroff and a first order method for it exhibits some oscillations behind the shock and smooths out the discontinuity at the same time.

The *matlab* programs used to generate those results can be obtained upon request from the author.

Below, we analyze the different schemes in detail to understand better their performances.

4.8 Numerical diffusion, dispersion, and the modified equation

Consider the upwind scheme (40) for the advection equation. With some simple manipulations this scheme can be rewritten as

$$\frac{1}{2\Delta t}(u_j^{n+1}-u_j^{n-1})+\frac{\Delta t}{2}\frac{u_j^{n+1}-2u_j^n+u_j^{n-1}}{(\Delta t)^2}=-a\frac{u_{j+1}^n-u_{j-1}^n}{2\Delta x}+\frac{a\Delta x}{2}\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{(\Delta x)^2}$$

or

$$\frac{u_j^{n+1}-u_j^{n-1}}{2\Delta t}+a\frac{u_{j+1}^n-u_{j-1}^n}{2\Delta x}=\frac{a\Delta x}{2}\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{(\Delta x)^2}+\frac{\Delta t}{2}\frac{u_j^{n+1}-2u_j^n+u_j^{n-1}}{(\Delta t)^2}. \qquad (51)$$

For all practical purposes this scheme can be thought of as an approximation for the *diffusive* advection equation

$$u_t+au_x=\frac{a\Delta x}{2}u_{xx}+\frac{\Delta t}{2}u_{tt},$$

with a second order truncation error, $O((\Delta x)^2 + (\Delta x)^2)$. Differentiating the advection equation once with respect to t, yields $u_{tt} = au_{xt} = (au_t)_x = a^2 u_{xx}$, i.e, the diffusive equation above becomes

$$u_t + au_x = \frac{a\Delta x}{2}\left(1 - \frac{a\Delta t}{\Delta x}\right) u_{xx}.$$

This equation is sometimes called the modified equation approximated by the upwind scheme to the second order. In essence, the upwind scheme approximates the viscosity solution introduced in (34) corresponding to the advection equation with the viscosity $\epsilon = \frac{a\Delta x}{2}(1-\mu)$ where $\mu = a\Delta t/\Delta x$ is the Courant number. On the one hand this can make us believe that in fact the upwind scheme computes the right physical solution, by approximating the vanishing viscosity solution, and on the other hand it provides an explanation on the poor performance of this scheme, especially in the case of the smooth solution in Figure 10, of under-predicting the solution.

Note the viscosity coefficient, introduced by the upwind scheme, is zero if the Courant number if choosing to be perfectly one. Such choice is avoided in practice because it may lead to numerical instabilities because of round-off errors. Further as noted above a little viscosity is necessary to guarantee convergence to the physical solution.

This phenomenon is known as *numerical dissipation or diffusion.* It is due to the term $\frac{a\Delta x}{2}\frac{u^n_{j+1}-2u^n_j+u^n_{j-1}}{(\Delta x)^2} + \frac{\Delta t}{2}\frac{u^{n+1}_j-2u^n_j+u^{n-1}_j}{(\Delta t)^2}$ on the right hand side of (51). In fact, multiplying the viscous advection equation by u and integrate on $[0,1]$ yields

$$\frac{d}{dt}\int_0^1 \frac{u^2}{2}dx + \frac{a}{2}\int_0^1 (u^2)_x dx = \epsilon \int_0^1 uu_{xx}dx. \tag{52}$$

Where we used $uu_t = (u^2)_t/2$ and $uu_x = (u^2)_x/2$. By integration by parts we have

$$\int_0^1 uu_{xx}dx = uu_x|_0^1 - \int_0^1 (u_x)^2 dx = -\int_0^1 (u_x)^2 dx$$

provided the boundary terms vanish, which happens if we use periodic or homogeneous Dirichlet or Neumann boundary conditions. The second term on the right of the integral equation (52) vanishes for the same reason and yields the following energy dissipation principle

$$\frac{d}{dt}\int_0^1 \frac{u^2}{2}dx = -\epsilon\int_0^1 (u_x)^2 dx < 0 \text{ if } u_x \neq 0.$$

Thus if integrated to infinity the upwind scheme will smooth out all the gradients in the solution and will ultimately converge to a constant-flat solution.

As we can surmise from Figures 10 and 11, the Lax-Friedrichs scheme is more dissipative than the upwind scheme. Indeed, the Lax-Friedrichs scheme is equivalent to

$$\frac{u^{n+1}_j - u^n_j}{\Delta t} + a\frac{u^n_{j+1} - u^n_{j-1}}{2\Delta x} = \frac{(\Delta x)^2}{2\Delta t}\frac{u^n_{j+1} - 2u^n_j + u^n_{j-1}}{(\Delta x)^2} - \frac{\Delta t}{2}\frac{u^{n+1}_j - 2u^n_j + u^{n-1}_j}{(\Delta t)^2}. \tag{53}$$

which yields a dissipative modified equation, which is the second order approximation of the Lax-Friedrichs scheme, with a viscosity $\epsilon = \frac{(\Delta x)^2}{2\Delta t} - a^2\frac{\Delta t}{2} = (\Delta x)^2(1-\mu^2)/(2\Delta t)$ which is typically larger than that of the upwind scheme.

This dissipation or diffusivity problem is typical to first order schemes for hyperbolic systems. Let us now consider the second order schemes and derive the modified equations associated with those schemes. Since they are already second order accurate, we should consider approximations with an order of accuracy higher than two. Below, we take a slightly different route than it is done above, to derive the modified equation. Namely, we use Taylor expansion as to compute the truncation error but instead we keep the first neglected term and add it to the advection equation to form the modified equation.

For the leap-frog method we have

$$\frac{u_j^{n+1}-u_j^{n-1}}{2\Delta x}+a\frac{u_{j+1}^n-u_{j-1}^n}{2\Delta x}=u_t+\frac{(\Delta t)^2}{6}u_{ttt}(x,t)+O((\Delta t)^4)+au_x+a\frac{(\Delta x)^2}{6}u_{xxx}(x,t)+O((\Delta x)^4)$$

(because all even terms in the Taylor series cancel out). Again using the advection equation we have: $u_{ttt}=-a^3u_{xxx}$ and therefore the modified equation, which is approximated to the fourth order by the leap-frog scheme, is

$$u_t+au_x=-\frac{a(\Delta x)^2}{6}\left(1-\mu^2\right)u_{xxx}. \tag{54}$$

This equation is known as *the dispersive wave equation*. First, it is easy to show that this equation conserves energy (see exercise 11 below). This is consistent with the fact that the von Neumann amplification factors for the leap-frog method satisfy $|\rho_\pm|=1$ when $\mu<1$.

Second, dispersion refers to a physical system where waves of different wavelengths propagate at different wave speeds. Let's look at wave like solution on the form

$$u=e^{i(kx-\omega t)}$$

for the dispersive wave equation. Here $k=2\pi l$ is the wavenumber, ω is called the phase or frequency, ω/k defines the phase speed, the speed at which the wave propagates. Plugging this ansatz into (54) yields, the *dispersion relation*

$$\omega=ak-\frac{a(\Delta x)^2}{6}\left(1-\mu^2\right)k^3,$$

i.e the phase speed, $c_k=a-\frac{a(\Delta x)^2}{6}\left(1-\mu^2\right)k^2$, depends highly on the wavenumber, and higher the wavenumber faster the wave disturbance moves relative to the advective motion. Short waves move relatively faster. This is very unlike the plain-original advection equation were all waves move at the same speed–the advection speed a.

Now, we reconsider the von Neumann amplification factors for the leap-frog method in (45). Recall that we have two solutions (one is physical and the other is a numerical artifact), which we denote here by:

$$u_+=e^{i(kj\Delta x-n\psi_0)};\quad u_-=(-1)^ne^{i(kj\Delta x+n\psi_0)}).$$

Let us consider only the physical mode, u_+. From the discussion above, using $n=t_n/\Delta t$, the associated phase speed is given by

$$c_+=\frac{\psi_0}{k\Delta t}=\frac{\arcsin(\mu\sin(\phi_l))}{k\Delta t}$$

Using Taylor expansion to expand both the sine and the arcsine functions, for small ϕ_l values (resolved modes), yields

$$\psi_0 \approx \mu\phi_l - \frac{\mu}{6}(1-\mu^2)\phi_l^3 = a\Delta t k - \frac{(\Delta x)^2}{6}a\Delta t(1-\mu^2)k^3$$

which implies that the phase speed of the physical mode for the leapfrog scheme satisfies

$$c_+ \approx a - a\frac{(\Delta x)^2}{6}(1-\mu^2)k^2.$$

This matches exactly the expression for c_k found above, using the dispersive wave equation.

Now, we can explain why in Figure 11, the leapfrog scheme exhibits oscillations. They result from the fact that the small truncation errors located near the discontinuity, which also accumulate with time, are viewed as wave disturbances of much shorter wavelengths which then propagate at their own speeds different from that of the actual solution.

Recall that the amplification factors for the previous two first order schemes, namely the upwind and the Lax-Friedrichs, are respectively given by

$$\rho_{uw} = 1 - \mu(1-\cos(\phi_l)) - i\mu\sin(\phi_l) \text{ and } \rho_{LF} = \cos(\phi_l) - i\mu\sin(\phi_l).$$

In the light of the analysis done in the previous paragraphs for the leap-frog scheme, we set $\rho^n \equiv e^{-in\Delta t\omega}$ where ω is a generalized phase so that $\Re(\omega)/k$ yields the phase speed of the numerical wave solution and $\Im(\omega)$ yields the exponential growth or damping rate of the wave amplitude. We have for the upwind and Lax-Friedrichs schemes, respectively,

$$e^{-i\omega_{uw}\Delta t} = 1 - \mu(1-\cos(\phi_l)) - i\mu\sin(\phi_l) \quad , e^{-i\omega_{LF}\Delta t} = \cos(\phi_l) - i\mu\sin(\phi_l).$$

For small $\omega\Delta t$, we have

$$e^{-i\omega\Delta t} \approx 1 - i\omega\Delta t.$$

Hence

$$\omega_{uw}\Delta t \approx \mu\sin(\phi_l) - i\mu(1-\cos(\phi_l)) \text{ and } \omega_{LF}\Delta t \approx \mu\sin(\phi_l) - i(1-\cos(\phi_l)).$$

First note that both schemes have quite strong damping rates,

$$\Im(\omega_{uw}) = \frac{\mu}{\Delta t}(1-\cos(k\Delta x)) \approx -\frac{1}{2}k^2\Delta x, \Im(\omega_{LF}) = -\frac{1-\cos(\phi_l)}{\Delta t} \approx -\frac{a}{2\mu}\Delta x k^2,$$

respectively, and consistently with the previous results, the LF scheme being more damped–by a factor of $1/\mu$.

With $\phi_l = k\Delta x$, the phase speeds are equal and are given by

$$\frac{\Re(\omega_{up})}{k} = \frac{\Re(\omega_{LF})}{k} = \frac{a}{k\Delta x}\sin(k\Delta x) \approx a\left(1 - \frac{k^2(\Delta x)^2}{6}\right).$$

This clearly shows that both schemes are dispersive, similarly to the leap-frog scheme. Nevertheless, small wave disturbances typically occurring at the grid scale decrease in magnitude at a faster rate than the domain-scale physical wave.

Now we turn to the Lax-Wendroff and Beam-Warming schemes. We have

$$\rho_{LW} = 1 - i\mu \sin(\phi_l) - \mu^2(1 - \cos(\phi_l))$$

yielding

$$\omega_{LW} = \frac{\mu}{\Delta t} \sin(\phi_l) - i\frac{\mu^2}{\Delta t}(1 - \cos(\phi_l)), \tag{55}$$

i.e, for the phase speed, we have the same dispersive behaviour as for the previous schemes but a much smaller damping rate

$$\Im(w_{LW}) = -\frac{\mu^2}{\Delta t}(1 - \cos(\phi_l)) \approx -\frac{\mu^2}{2\Delta t} k^2 (\Delta x)^2,$$

of the same order as the phase speed deviation. For grid scale disturbances, the wavenumber, k, scales as $1/\Delta x$, therefore the damping rates for those small scale waves is on the order $O(\frac{\mu^2}{2\Delta t})$ which seems to be small, explaining the propagation of the small oscillations, in Figure 10, away from the discontinuity before they get damped.

More importantly note the propagation of the oscillations for the Lax-Wendroff scheme, in Figure 10, to the left of the discontinuities, is associated with the fact that according to the dispersion relation (55) smaller wavelength disturbances move slower than those with larger wavelengths. We have a similar behaviour for the Lax-Wendroff scheme. But the Beam-Warming scheme exhibits somehow larger amplitude oscillations moving to the right, which seems to anticipate that the Beam-Warming scheme has both a weaker damping rate and dispersive phase speeds characterized by smaller wavelengths propagating faster than larger ones. The details are left as an exercise.

Exercise 11 *Show that the dispersive wave equation*

$$u_t + a u_x = \gamma u_{xxx}$$

conserves energy, that is

$$\frac{d}{dt} \int_0^1 \frac{u^2}{2} dx = 0.$$

Exercise 12 *Determine the exponential damping rate and the phase speed for the Beam-Warming scheme. Deduce that this scheme is dispersive and that wave disturbances with larger wavenumbers (smaller wavelengths) propagate faster than those with smaller wavenumbers.*

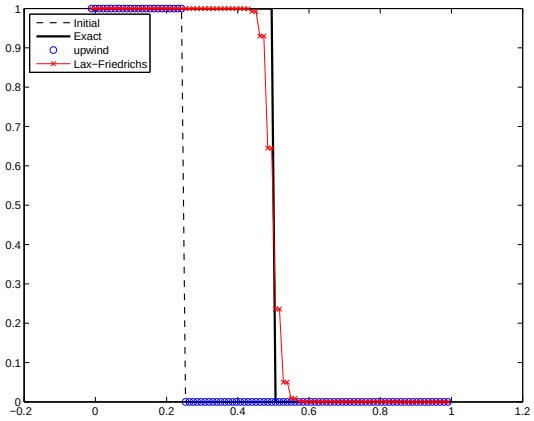

Figure 12: Riemann Problem for Burger's equation solved with both upwind and Lax-Friedrichs schemes, showing that the upwind schemes predicts a wrong shock speed.

5 Finite volume methods for scalar conservation laws

5.1 Wrong shock speed and importance of conservative form

Consider Burger's equation with the discontinuous initial condition

$$u_t + \frac{1}{2}(u^2)_x = 0$$
$$u_0(x) = \begin{cases} 1 \text{ if } x < 0.25 \\ 0 \text{ if } x > 0.25. \end{cases} \tag{56}$$

We view Burger's equation as an advection equation with a non-linear advection speed $a(u) = u$ and attempt to solve this problem with both the upwind and Lax-Friedrichs methods. The upwind scheme is generalized to Burger's equation as follows.

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}\left(\max(u_j^n, 0)(u_j^n - u_{j-1}^n) + \min(u_j^n, 0)(u_{j+1}^n - u_j^n)\right). \tag{57}$$

Whereas the formula for Lax-Friedrichs' scheme amounts to approximating the flux derivative using centred differences.

$$u_j^{n+1} = \frac{1}{2}(u_{j+1}^n + u_{j-1}^n) - \frac{\Delta t}{2\Delta x}\left((u_{j+1}^n)^2/2 - (u_{j-1}^n)^2/2\right). \tag{58}$$

The results are shown in Figure 12 where the two numerical solutions are compared to the exact solution to the Riemann problem. Note that the Lax-Friedrichs method predicts well the propagation of the shock wave, with a significant smearing of the discontinuity, as excepted, due its high viscosity, while the upwind scheme predicts a steady solution which doesn't change with time. In fact, this latter result can be easily recovered analytically from the upwind scheme. The numerical solution in this case remains zero on the right of the shock because the advection velocity is zero and remains one on the left because the backward finite difference derivative is zero.

The main reason for why the Lax-Friedrichs scheme out performs the upwind scheme, in this case, is because the former has a conservative form. A numerical scheme for a conservation law

$$u_t + (f(u))_x = 0$$

is said **conservative** if it can be written on the form

$$u_j^{n+1} = u_j^n + \left[F_{j+1/2} - F_{j-1/2}\right]. \tag{59}$$

Where $F_{j+1/2}$, called a *numerical flux* which is in some sense an approximation of the flux $f(u)$ at the cell interface, $x_{j+1/2} = x_j + \Delta x/2$.

Here we show that indeed the Lax-Friedrichs scheme has a conservative form. Let

$$F_{j+1/2}^{LF} = \frac{1}{2}(u_{j+1}^n - u_j^n) - \frac{\mu}{2}(f(u_{j+1}^n) + f(u_j^n).$$

Then the Lax-Friedrichs schemes can be written as

$$u_j^{n+1} = u_j^n + F_{j+1/2}^{LF} - F_{j-1/2}^{LF}.$$

Thus it is conservative. We will see below that, the Lax-Wendroff and the Beam-Warming methods are also conservative.

Remark:
It is worthwhile noting that the formula for the Lax-Friedrichs flux, F^{LF}, at the interface $j+1/2$, is the average of the left and right fluxes, $f(u_j), f(u_{j+1})$, plus a centered difference, $u_{j+1} - u_j$, suggesting a *diffusive* term. Thus, the Lax-Friedrichs scheme can be viewed as a discrete version of the diffusive equation

$$u_t + (f(u) - \epsilon u_x)_x = 0$$

where $\epsilon = \frac{(\Delta x)^2}{2\Delta t}$ is the viscosity coefficient.

Convergence of Conservative Schemes:
It is shown in the literature that **if a numerical scheme for a conservation law** $u_t + (f(u))_x = 0$ **is consistent, stable, and has a conservative form, then the resulting numerical solution converges to a weak solution to the conservation law.**

This statement can be viewed as an extension of the Lax-equivalence theorem to non-linear conservation laws. Notice however, since weak solutions are not unique, it is not guaranteed that the numerical solution obtained by a consistent, stable, and conservative scheme converges to the physical entropic solution. For that some extra-properties of the numerical scheme are needed to enforce the entropy condition.

5.2 Godonuv's first order scheme

Let $x_j = j\Delta x, \quad j = \cdots, -2, -1, 0, 1, 2, \cdots$ and $t_n = n\Delta t, \ n = 0, 1, 2, \cdots$ be a discretization of the space-time domain (x, t). We divide the real line into sub-intervals $[x_{j-1/2}, x_{j+1/2}]$, called *grid cell*, where $x_{j+1/2} = x_j + \Delta x/2$. We define the *cell average* of the solution $u(x, t)$ at each grid cell as

$$\bar{u}_j^n = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t_n)dx. \tag{60}$$

Consider the conservation law

$$u_t + (f(u))_x = 0.$$

Next, we integrate this equation on the rectangle $[t_n + t_{n+1}] \times [x_{j-1/2}, x_{j+1/2}]$, often called *the control or finite volume*, hence the name *finite volume method*. We have

$$\int_{t_n}^{t_{n+1}} \int_{x_{j-1/2}}^{x_{j+1/2}} u_t(x,t)dxdt + \int_{t_n}^{t_{n+1}} \int_{x_{j-1/2}}^{x_{j+1/2}} (f(u(x,t)))_x dxdt = 0$$

or

$$\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,t_{n+1})dx - \int_{x_{j-1/2}}^{x_{j+1/2}} u(x,t_n)dx + \int_{t_n}^{t_{n+1}} (f(u(x_{j+1/2},t)))dt - \int_{t_n}^{t_{n+1}} (f(u(x_{j-1/2},t)))dt = 0.$$

Dividing by the mesh size Δx and introducing the averages $\bar{u}_j^n$ above, yields

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{\Delta t}{\Delta x}\left[F_{j+1/2}^n - F_{j-1}^n\right], \tag{61}$$

where

$$F_{j+1/2}^n = \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} (f(u(x_{j+1/2},t)))dt,$$

is the average flux of u through the interface $x = x_{j+1/2}, t_n \le t \le t_{n+1}$. Provided we find an adequate numerical approximation to the time integral, the formula (61) provides a numerical scheme for the conservation law, in conservative form.

Piecewise constant approximations and the Riemann problem at the cell interfaces: Assume that at time $t = t_n$, the cell averages $\bar{u}_j^n$ are known. Consider the approximation:

$$u(x,t_n) \approx \bar{u}_j^n, \text{ if } x_{j-1/2} \le x \le x_{j+1/2}.$$

Then finding an approximation for the flux $F_{j+1/2}$, amounts to solving the Riemann problem

$$u_t + (f(u))_x = 0, \quad t \in [t_n, t_{n+1}] \tag{62}$$
$$u(x,t_n) = \begin{cases} u_L & \text{if } x < x_{j+1/2} \\ u_R & \text{if } x > x_{j+1/2} \end{cases}$$

where the left and right states are given, respectively, by $u_L = \bar{u}_j$ and $u_R = \bar{u}_{j+1}^n$.

To compute the flux $F_{j+1/2}$ we need to know the solution u along the interface $x = x_{j+1/2}, t_n \le t \le t_n + \Delta t$. We introduce the shock speed at the cell interface,

$$\dot{s} = \frac{f(u_R) - f(u_L)}{u_R - u_L},$$

according to the Rankine-Hugoniot jump condition. Under the condition that $f(u)$ is convex, using the method of characteristics, introduced above, we have, for $t_n \le t \le t_n + \Delta t$,

$$u(x_{j+1/2},t) = \begin{cases} u_L & \text{if } f'(u_L) > 0 \ \& \ f'(u_R) > 0 \\ u_R & \text{if } f'(u_L) < 0 \ \& \ f'(u_R) < 0 \\ u_L & \text{if } f'(u_L) \ge 0 \ \& \ f'(u_R) \le 0 \ \& \ \dot{s} > 0 \\ u_R & \text{if } f'(u_L) \ge 0 \ \& \ f'(u_R) \le 0 \ \& \ \dot{s} < 0 \\ u_* & \text{if } f'(u_L) \le 0 \ \& \ f'(u_R) \ge 0, \end{cases} \tag{63}$$

where in (63), u_*, *called a sonic point*, is defined such that $f'(u_*) = 0$ and corresponds to a rarefaction wave solution. Notice that the convexity of f guarantees that u_* exists and is unique. Note also the fist four cases correspond either to an entropic shock solution where, according to Lax's criterion,

$$f'(u_L) \geq \dot{s} \geq f'(u_R)$$

so that $u_{j+1/2} = u_L$ if $\dot{s} > 0$ and $u_{j+1/2} = u_R$ if $\dot{s} < 0$ or a rarefaction waves where the rarefaction fan is either completely to the left or completely to right of the interface according to whether $f'(u_{R,L}) < 0$ or $f'(u_{R,L}) > 0$, respectively. The last case, however, corresponds to rarefaction wave where the rarefaction fan contains characteristics going both to the left and to the right. It is called *a transonic rarefaction,* by analogy to gas dynamics, where such a rarefaction happens when the fluid on one side of the wave moves at a speed smaller than the speed of sound (subsonic) and on the other side it moves with a speed larger than the speed of sound (supersonic).

Here we show that the sonic point u_* is in fact the root of the equation $f'(u_*) = 0$. Assume $f'(u_L) < 0 < f'(u_R)$, then in this case there is a rarefaction wave connecting the left and right states. Inspired by the solution to Burger's equation, a rarefaction wave solution is sought on the form $u(x,t) = v(x/t)$. Plugging this ansatz into the conservation law $u_t + (f(u))_x = 0$, yields

$$-\frac{x}{t^2}v'(x/t) + f'(v(x/t))v'(x/t)\frac{1}{t} = 0.$$

Hence

$$f'(v(x/t)) = \frac{x}{t}.$$

Along the interface $x = 0$ we have $u_* = u(0,t) = v(0)$ and $f'(v(0)) = 0$.

The celebrated Godunov's method is now obtained by simply using the solution to the Riemann problem in (63) at each interface $x_{j+1/2}$ to computes the numerical fluxes $F_{j-1/2}, F_{j+1/2}$ in (61), yielding

$$F_{j+1/2} = \begin{cases} f(u_L) & \text{if } f'(u_L) > 0 \ \& \ f'(u_R) > 0 \\ f(u_R) & \text{if } f'(u_L) < 0 \ \& \ f'(u_R) < 0 \\ f(u_L) & \text{if } f'(u_L) \geq 0 \ \& \ f'(u_R) \leq 0 \ \& \ \dot{s} > 0 \\ f(u_R) & \text{if } f'(u_L) \geq 0 \ \& \ f'(u_R) \leq 0 \ \& \ \dot{s} < 0 \\ f(u_*) & \text{if } f'(u_L) \leq 0 \ \& \ f'(u_R) \geq 0. \end{cases} \tag{64}$$

Notice that the solution u_* in (62) can be replaced by the shock solution, i.e, u_R if $\dot{s} < 0$ and u_L if $\dot{s} > 0$, even when $f'(u_L) < 0 < f'(u_R)$, and will still provide a weak solution satisfying the Rankine-Hugoniot jump condition, but as we already know this will lead to an unphysical weak solution which violates the entropy condition. The resulting numerical solution is referred to as *an all shock solution* and provides an approximation to the weak solution to the conservation law, which is not necessarily the physically relevant–entropic solution.

Case of the advection equation and Stability of Godunov's method

If we replace the conservation law by the simple advection equation with a positive advection speed, $a > 0$, then the solution to the Riemann problem at the cell interface, $x_{j+1/2}$, reduces to

$$u(x_{j+1/2}, t) = u_j,$$

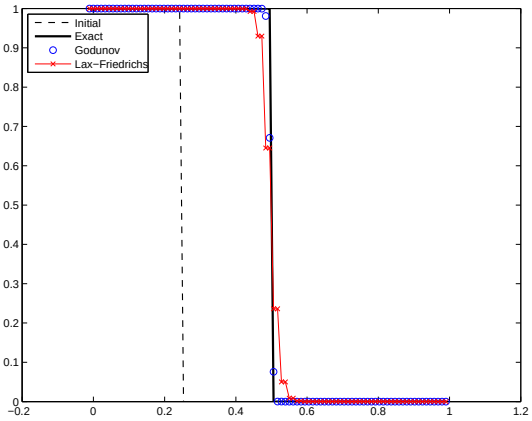

Figure 13: Same as Figure 12 but with Godunov's method instead of the upwind scheme.

and Godunov's method reduces to the upwind scheme

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{a\Delta t}{\Delta x}(\bar{u}_j^n - \bar{u}_{j-1}^n). \tag{65}$$

Recall that the upwind method is stable under the CFL condition. Therefore, provided the CFL condition

$$\max(f'(u_L), f'(u_R))\Delta t \leq \Delta x \tag{66}$$

is satisfied, Godunov's method is stable.

Consistency and convergence
Moreover, given the way the fluxes were computed, i.e, using the exact solution of the Riemann problem using first order approximation of the initial data, namely piecewise constant cell averages, we can show that Godunov's method is consistent of order 1 in both space and time, just like the upwind scheme. Moreover, since it is also conservative by construction, Godunov's method is guaranteed to convergence to the entropic weak solution, provided we always choose the entropic solution for the Riemann problem. The numerical test of Figure 12, is repeated with the Godunov method and the results are shown in Figure 13, where Godunov's method is compared to the Lax-Friedrichs method. As we expect Godunov's method predicts the right shock speed and has much less dissipation than the Lax-Friedrichs method. However, one big disadvantage of Godunov's method is that it requires the solution of a Riemann problem at each cell interface and every time step. This can be very costly especially for non-linear systems of conservation laws.

An other serious shortcoming of Godunov's method is that it is only first order in both time and space. This limits its use in practice but it constitutes an important cornerstone for more sophisticated so called *high resolution methods developed below.*

5.3 High resolution, TVD, and MUSCL schemes

As noted above one of the major shortcomings of Godunov's method is that it is only first accurate in both space and time. As noted above a crucial step in deriving this method is the piecewise

constant approximation step, when the solution $u(x,t)$ at time $t = t_n$ is approximated by the cell average

$$u(x,t_n) \approx \bar{u}_j^n, \quad x_{j-1/2} \le x \le x_{j+1/2},$$

to advance to the next time step, t_{n+1}, which is only a first order approximate/interpolation. Since Godunov's method computes the cell averages at each new time step, this piecewise approximation appears to be very convenient, indeed. Given the cell averages $\bar{u}_j^n$, we can use piecewise linear interpolation to yield a second order reconstruction

$$u(x,t_n) \approx \tilde{u}(x,t_n) \equiv \bar{u}_j^n + \sigma_j(x - x_j), \quad x_{j-1/2} < x < x_{j+1/2},$$

which can be used instead to advance the solution to next step (see Figure 14). The slopes σ_j can be computed in many different ways; using many different finite differencing formulas as we will see below.

To illustrate let us consider the advection equation

$$u_t + au_x = 0, \quad a > 0.$$

Given the piecewise linear reconstruction values, $\tilde{u}(x,t_n)$, at time $t = t_n$, the "exact" solution for this equation at $t = t_n + \Delta t$ is given by

$$u(x,t_n + \Delta t) = \tilde{u}(x - a\Delta t, t_n).$$

Provided $a\Delta t/\Delta x < 1$, we have

$$u(x,t_n+\Delta t) = \begin{cases} \bar{u}_{j-1}^n + \sigma_{j-1}(x - a\Delta t - x_{j-1}) & \text{if } x_{j-1/2} < x < x_{j-1/2} + a\Delta t \\ \bar{u}_j^n + \sigma_j(x - a\Delta t - x_j) & \text{if } x_{j-1/2} + a\Delta t < x < x_{j+1/2}. \end{cases} \tag{67}$$

Averaging $\tilde{\tilde{u}}(x,t_{n+1})$ on the intervals $(x_{j-1/2}, x_{j+1/2})$,

$$\begin{aligned} \bar{u}_j^{n+1} =& \frac{a\Delta t}{\Delta x}\bar{u}_{j-1}^n + \frac{\sigma_{j-1}}{\Delta x}\int_{x_{j-1/2}}^{x_{j-1/2}+a\Delta t}(x - a\Delta t - x_{j-1})dx + \\ & \frac{\Delta x - a\Delta t}{\Delta x}\bar{u}_j^n + \frac{\sigma_j}{\Delta x}\int_{x_{j-1/2}+a\Delta t}^{x_{j+1/2}}(x - a\Delta t - x_j)dx, \end{aligned}$$

yields the cell averages at $t_n + \Delta t$, given by

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{a\Delta t}{\Delta x}(\bar{u}_j^n - \bar{u}_{j-1}^n) - \frac{a\Delta t}{2\Delta x}(\Delta x - a\Delta t)(\sigma_j - \sigma_{j-1}). \tag{68}$$

If we set the slopes to zeros, $\sigma_j = 0$, then we recover the first order Godunov (upwind) method in (65). A second order scheme is obtained, if the slope is choosing to be a first order finite differencing formula which approximates the derivatives $u_x(x_j,t_n)$ in the j'th grid cell, using neighbouring cell average values. The following three choices of slopes yield three popular second order schemes

Centered:	$\sigma_j = \dfrac{\bar{u}_{j+1} - \bar{u}_{j-1}}{2\Delta x}$ (Fromm)
Upwind:	$\sigma_j = \dfrac{\bar{u}_j - \bar{u}_{j-1}}{\Delta x}$ (Beam-Warming)
Downwind:	$\sigma_j = \dfrac{\bar{u}_{j+1} - \bar{u}_j}{\Delta x}$ (Lax-Wendroff).

Exercise 13 *Show that when $a < 0$ the second order scheme for the advection equation using piecewise linear reconstruction is given by*

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{a\Delta t}{\Delta x}(\bar{u}_{j+1}^n - \bar{u}_j^n) + \frac{a\Delta t}{2\Delta x}\left(\Delta x + a\Delta t\right)(\sigma_{j+1} - \sigma_j). \tag{69}$$

Reconstruct, Solve, Average:
The algorithm followed above to derive Godunov's method as well as its second order version for the advection equation involves three main steps.

1) Reconstruct: Given the cell averages $\bar{u}_j^n$, we (re)construct piecewise constant or piecewise linear approximations

$$\tilde{u}(x, t_n) \approx \bar{u}_n^j \text{ or } \tilde{u}(x, t_n) \approx \bar{u}_n^j + \sigma_j(x - x_j), \quad x_{j-1} < x < x_{j+1/2}$$

2) Solve: Solve the Riemann problem

$$u_t + (f(u))_x = 0, t_n \leq t \leq t_{n+1}, \quad u(x, t_n) = \tilde{u}(x, t_n)$$

3) Average: Average the solution at t_{n+1}

$$\bar{u}_j^{n+1} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t_{n+1})\, dx.$$

A few remarks on the Reconstruct-Solve-Average algorithm:
Note the first step is crucial, it sets the order of accuracy of the method. Piecewise constants yield a first order method while a piecewise linear approximation yields a second order scheme. In fact methods using parabolic approximations exist are often used in practice, they are known are **PPM** (for piecewise parabolic methods). Piecewise parabolic reconstruction yields a third order method and so on. For, the second step it is desirable to be able to solve the conservation law exactly. This is possible in the case of the advection equation and for the conservation law in general when piecewise constants are used in step one, yielding Godunov's method. But in general, we need to resort to quadrature formulas to approximate the flux integrals along the time interval $[t_n, t_n + \Delta t]$.

Finally, when a finite volume method, such as Godunov's, is applied to a conservation law, it yields the average values $\bar{u}_j^{n+1}$ directly, which takes care of step three.

High order finite volume method for conservation laws
Recall the finite volume derived in (61) for Godunov's method

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{\Delta t}{\Delta x}\left[F_{j+1/2}^n - F_{j-1}^n\right],$$

where

$$F_{j+1/2}^n = \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} (f(u(x_{j+1/2}, t)))dt.$$

To reconcile with the case when $u(x,t)$ is not piece-wise constant at time $t = t_n$, we approximate the time integral by a rectangular rule yielding an Euler step

$$\bar{u}_{j+1} = \bar{u}_j^n - \frac{\Delta t}{\Delta x}\left(\tilde{F}_{j+1/2}^n - \tilde{F}_{j-1/2}^n\right) \tag{70}$$

where

$$\tilde{F}_{j+1/2}^n = f(u^*(x_{j+1/2}, t_n^+)).$$

Here

$$u^*(x_{j+1/2}, t_n^+) = \lim_{t \longrightarrow t_n} u(x_{j+1/2}, t)$$

is obtained by solving the Riemann problem with left and right states

$$u_L = \tilde{u}(x_{j+1/2,-}, t_n), u_R = \tilde{u}(x_{j+1/2,+}, t_n)$$

where again

$$\tilde{u}(x_{j+1/2,-}, t_n) = \lim_{x \longrightarrow x_{j+1/2}, x < x_{j+1/2}} \tilde{u}(x, t_n) \text{ and } \tilde{u}(x_{j+1/2,+}, t_n) = \lim_{x \longrightarrow x_{j+1/2}, x > x_{j+1/2}} \tilde{u}(x, t_n).$$

Note that in the case of a piecewise constant approximation these limits are simply $u_L = \bar{u}_j^n$ and $u_R = \bar{u}_j^{n+1}$, and the solution to the Riemann problem is constant on the time interval $[t_n, t_n + \Delta t]$, provided the CFL condition $\Delta t \max(|f'(u_R)|, |f'(u_L)|) \leq \Delta x$ is satisfied. Note this is not true for higher order reconstructions.

To obtain higher order accuracy in time, we often use a predictor-corrector scheme such as the mid-point Runge-Kutta method in place of the first-order Euler step (70):

$$\begin{aligned} \bar{u}_{j+1/2} &= \bar{u}_j^n - \frac{\Delta t}{2\Delta x}\left(\tilde{F}_{j+1/2}^n - \tilde{F}_{j-1/2}^n\right) \\ \bar{u}_{j+1} &= \bar{u}_j^n - \frac{\Delta t}{\Delta x}\left(\tilde{F}_{j+1/2}^{n+1/2} - \tilde{F}_{j-1/2}^{n+1/2}\right) \end{aligned} \tag{71}$$

where

$$\tilde{F}_{j+1/2}^{n+1/2} = f(u^*(x_{j+1/2}, t_{n+1/2}^+)).$$

Oscillations and TVD methods

Recall from Figure 11 that one major problem with second order methods is that they tend to generate unphysical oscillations near discontinuities. One way to control these oscillation is to use a hybrid method where the second order reconstruction is *limited* to regions where the solution, u, is smooth and use a first order method in regions where u presents discontinuities of some sort. To find an intelligent way to do so, we first introduce a mathematical tool to control those oscillations. This is achieved by the total variation, which is introduced next.

Definition 6 *The total variation (TV) of a real valued function f is given by*

$$TV(f) = \sup \sum_{j=-\infty}^{+\infty} |f(\xi_{j+1}) - f(\xi_j)|$$

where the supremum is taken on all subdivisions $\cdots < \xi_{-1} < \xi_0 < \xi_1 < \cdots$ of the real line.

To help develop some intuition, we list a few properties of TV.

- If $f(x)$ is monotonic (i.e, non-increasing or non-decreasing) on some interval $[a, b]$, then the total variation of f on $[a, b]$ is given by

$$\mathrm{TV}_{[a,b]}(f) = |f(b) - f(a)|$$

- If $f(x)$ is piecewise constant on the line segments $(x_{j-1/2}, x_{j+1/2})$, then

$$\mathrm{TV}(f) = \sum_{j=-\infty}^{+\infty} |f(x_{j+1}) - f(x_j)|,$$

  i.e, the sum of all the jumps of f.

- If $f(x)$ is piecewise linear on the line segments $(x_{j-1/2}, x_{j+1/2})$, then

$$\mathrm{TV}(f) = \sum_{j=-\infty}^{+\infty} |f(x^-_{j+1/2}) - f(x^+_{j-1/2})| + \sum_{j=-\infty}^{+\infty} |f(x^+_{j+1/2}) - f(x^-_{j+1/2})|,$$

  i.e, the sum of the variations within each one of the line segments plus all the jumps across the interfaces.

- If f is differentiable, then

$$\mathrm{TV}(f) = \int_{-\infty}^{+\infty} |f'(x)| dx.$$

Since the exact solution to the advection equation simply propagates at a constant speed, its TV doesn't change with time. Moreover, it can be shown that the total variation for an entropic solution for a conservation law, in general, doesn't increase with time. It can decrease after a shock but never increases. However, this is not always the case for the numerical solution. Clearly the oscillations generated by the second order schemes in Figure 11 do increase the TV of the numerical solution. A reasonable way to attempt to avoid those unphysical oscillations, is thus to require that the total variation of the numerical solution does not increase.

Definition 7 *The numerical scheme*

$$\bar{u}_{j+1} = \bar{u}^n_j - \frac{\Delta t}{\Delta x}\left(\tilde{F}^n_{j+1/2} - \tilde{F}^n_{j-1/2}\right)$$

is said to be **total variation diminishing** *(TVD) if*

$$TV(\bar{u}^{n+1}) \leq TV(\bar{u}^n). \tag{72}$$

For instance we can show that the average step in the Reconstruct-Solve-Average algorithm does actually diminish total variation. Therefore, when the slopes are set to zeros, $\sigma_j = 0$, the upwind scheme for the advection equation is TVD.

Definition 8 *A numerical scheme*

$$u_j^{n+1} = H(u_{j-k}^n, u_{j-k+1}^n, \cdots, u_{j+l}^n)$$

is said to be monotone if

$$\frac{\partial H}{\partial u_\eta} \geq 0, \ \eta = j-k, \cdots, j+l.$$

One important property of monotone schemes is that they do not create new local maxima or minima:

$$\max u_j^{n+1} \leq \max u_j^n, \text{ and } \min u_j^{n+1} \geq \min u_j^n$$

As such monotonic schemes are TVD. On the other hand TVD schemes are monotonicity preserving. A numerical scheme is said to be monotonicity preserving if monotonic data remains monotonic:

$$\cdots u_{j-1}^n \leq u_j^n \leq u_{j+1}^n \leq \cdots \implies \cdots u_{j-1}^{n+1} \leq u_j^{n+1} \leq u_{j+1}^{n+1} \leq \cdots$$

We have the following general statement

$$\text{monotonic schemes} \subset \text{TVD schemes} \subset \text{monotonicity preserving schemes.}$$

It is easy to show that, under the CFL condition, $a\Delta t \leq \Delta x$, Godunov's method for the advection equation is monotonic. In fact, for $a > 0$, we have

$$u_j^{n+1} = H(u_j^n, u_{j-1}^n) \equiv (1-\mu)u_j^n + \mu u_{j-1}^n.$$

Therefore it is TVD as noted above. Unfortunately, it is not possible to construct linear second order schemes which are monotonic, according to the following theorem due to Godunov. In agreement with the fact that second order schemes exhibit oscillations.

Theorem 6 (Godunov) *A linear monotonic scheme is at most first order accurate.*

In order to achieve high order schemes without unphysical oscillatory behaviour we need to make schemes which are intrinsically nonlinear. A family of such schemes are the hybrid schemes mentioned above, which are second order in smooth regions and only first order near shocks, to avoid oscillations.

To guarantee that our hybrid numerical scheme remains TVD and second order accurate in smooth regions, we must choose the slopes σ_j so that the TV of the reconstructed function is not larger than that of the discrete cell averages $\bar{u}_j^n$. This is achieved by defining the slope σ_j as a non-linear function of the data $\bar{u}_j^n$ in such a way to control the total variation. The associated non-linear function is called a limiter and methods based on this idea are known as slope-limiter methods, first introduced by van Leer when he derived his famous **MUSCL** schemes (Monotonic Upstream-Centered Schemes).

Perhaps the simplest choice of a limiter, which guarantees second order accuracy in regions where u is smooth and satisfies the TVD property at the same time, is the *minmod* limiter:

$$\sigma_j = \text{minmod}\left(\frac{\bar{u}_{j+1}^n - \bar{u}_j^n}{\Delta x}, \frac{\bar{u}_j^n - \bar{u}_{j-1}^n}{\Delta x}\right) \tag{73}$$

where

$$\text{minmod}(a,b) = \begin{cases} a & \text{if } |a| < |b| \text{ and } ab > 0 \\ b & \text{if } |b| < |a| \text{ and } ab > 0 \\ 0 & \text{if } ab \leq 0. \end{cases}$$

Note that when a, b have the same sign, the minmod function chooses the one which is smaller in magnitude and when $ab \leq 0$, it returns zero. Instead of using the downwind slope, yielding the Lax-Wendroff scheme, or the upwind slope, leading to the Beam-Warming scheme, the minmod limiter chooses the one which is smaller in magnitude. When the upwind and downwind slopes have opposite sign, the reconstructed cell value is kept constant. This latter must correspond to a local minimum or a local maximum, and the minmod limiter tends to preserve the local extrema, hence avoiding to create overshootings and undershootings near those extrema. Thus it does not increase TV and does not generate oscillations.

A more popular choice of limiter, due to van Leer, is the **MC-limiter** (monotonized-centered limiter):

$$\sigma_j = \text{minmod}\left(\frac{\bar{u}^n_{j+1} - \bar{u}^n_{j-1}}{2\Delta x}, 2\frac{\bar{u}^n_{j+1} - \bar{u}^n_j}{\Delta x}, 2\frac{\bar{u}^n_j - \bar{u}^n_{j-1}}{\Delta x}\right)$$

Note, the MC-limiter chooses the smallest in magnitude between the centred-differences, corresponding to Fromm's method, and twice the upwind or the downwind formulas, and returns zeros when there is a change is sign. This limiter yields highly accurate centered slopes in smooth regions and sharper resolution near discontinuities.

Flux-limiters

Using piecewise linear reconstruction for the advection equation yields (see (68) and (69))

$$\bar{u}^{n+1}_j = \bar{u}^n_j - \frac{\Delta t}{\Delta x}(F^n_{j+1/2} - F^n_{j-1/2})$$

where

$$F^n_{j+1/2} = \begin{cases} a\bar{u}^n_j + \dfrac{a}{2}(\Delta x - a\Delta t)\sigma_j & \text{if } a > 0 \\ \\ a\bar{u}^n_{j+1} - \dfrac{a}{2}(\Delta x + a\Delta t)\sigma_{j+1} & \text{if } a < 0. \end{cases}$$

Introducing the negative and positive parts of a

$$a^+ = \max(a, 0), \quad a^- = \min(a, 0)$$

the flux $F_{j+1/2}$ can be rewritten in a compact form as

$$F_{j+1/2} = a^-\bar{u}_{j+1} + a^+\bar{u}_j + \frac{1}{2}|a|\left(1 - \frac{\delta t|a|}{\Delta x}\right)\delta_j, \tag{74}$$

where δ_j is the jump in $\bar{u}$ across the upwind interface; $\delta_j = \bar{u}_j - \bar{u}_{j-1}$ if $a > 0$ and $\delta_j = \bar{u}_{j+1} - \bar{u}_j$ if $a < 0$.

We introduce the upwind side of j as

$$J = \begin{cases} j-1 & \text{if } a > 0 \\ j+1 & \text{if } a < 0 \end{cases}$$

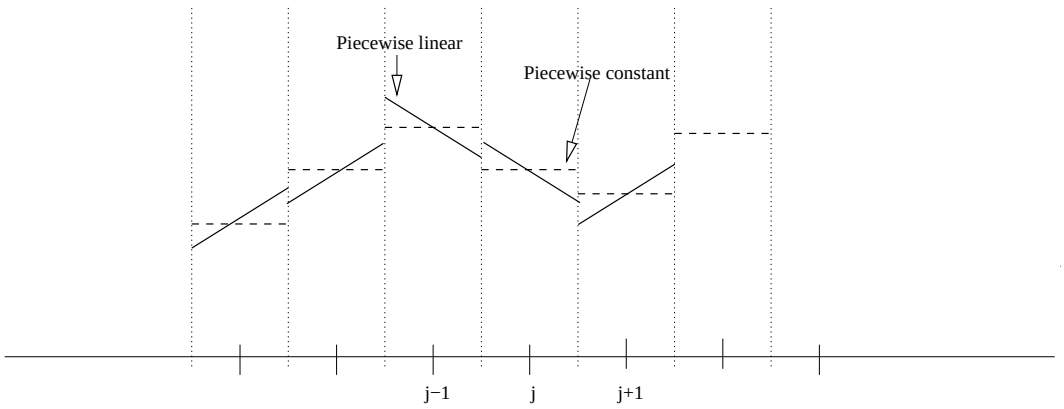

Figure 14: Piecewise constant (dashed) and piecewise linear reconstructions.

and the jump

$$\Delta\bar{u}_j = \bar{u}_{j+1} - \bar{u}_j.$$

We define

$$\theta = \frac{\Delta\bar{u}_J}{\Delta\bar{u}_j}.$$

and let

$$\delta_j = \phi(\theta)\Delta\bar{u}_j$$

in (74), where ϕ is a non-linear function of θ, called a flux-limiter, which is defined so that the resulting scheme is second order accurate in smooth regions and TVD at the same time.

Some linear choices of ϕ yield the popular second order (linear) schemes:

$$\begin{aligned}
\phi(\theta) &= 0 : && \text{upwind} \\
\phi(\theta) &= 1 : && \text{Lax-Wendroff} \\
\phi(\theta) &= \theta : && \text{Beam-Warming} \\
\phi(\theta) &= \frac{1}{2}(1+\theta) : && \text{Fromm,}
\end{aligned}$$

which are of course not TVD according to Godunov's theorem. While some clever non-linear choices yield **high resolution**–TVD methods. The following are the most popular ones:

$$\begin{aligned}
\phi(\theta) &= \max(0, \min(1, \theta)) : && \text{minmod} \\
\phi(\theta) &= \max(0, \min((1+\theta)/2, 2, 2\theta)) : && \text{MC-limiter} \\
\phi(\theta) &= \max(0, \min(1, 2\theta), \min(2, \theta)) : && \text{superbee} \\
\phi(\theta) &= \frac{1+\theta}{1+|\theta|} : && \text{van Leer.}
\end{aligned}$$

They all have their strengths and weaknesses, depending on the application. Those are well documented in the literature (see LeVeque's book: finite volume methods for hyperbolic problems, for examples). The performance of each of the choices of limiters listed above is shown on the different panels in Figure 15, for both the smooth and the non-smooth/square wave. We see that in general the TVD limiters capture very well both the smooth solution (with an accuracy comparable to the second order schemes) and the non-smooth square wave. The second order methods have an oscillatory behaviour near the discontinuity and the upwind method is too diffusive.

References

[1] Crispin W. Gardiner, 2004, *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences*, Springer.

[2] Randall J. LeVeque, 2002, Finite Volume Methods for Hyperbolic Problems (Cambridge Texts in Applied Mathematics),Cambridge University Press.

[3] Dale Durran, 1998, Numerical Methods for Wave Equations in Geophysical Fluid Dynamics, Springer.

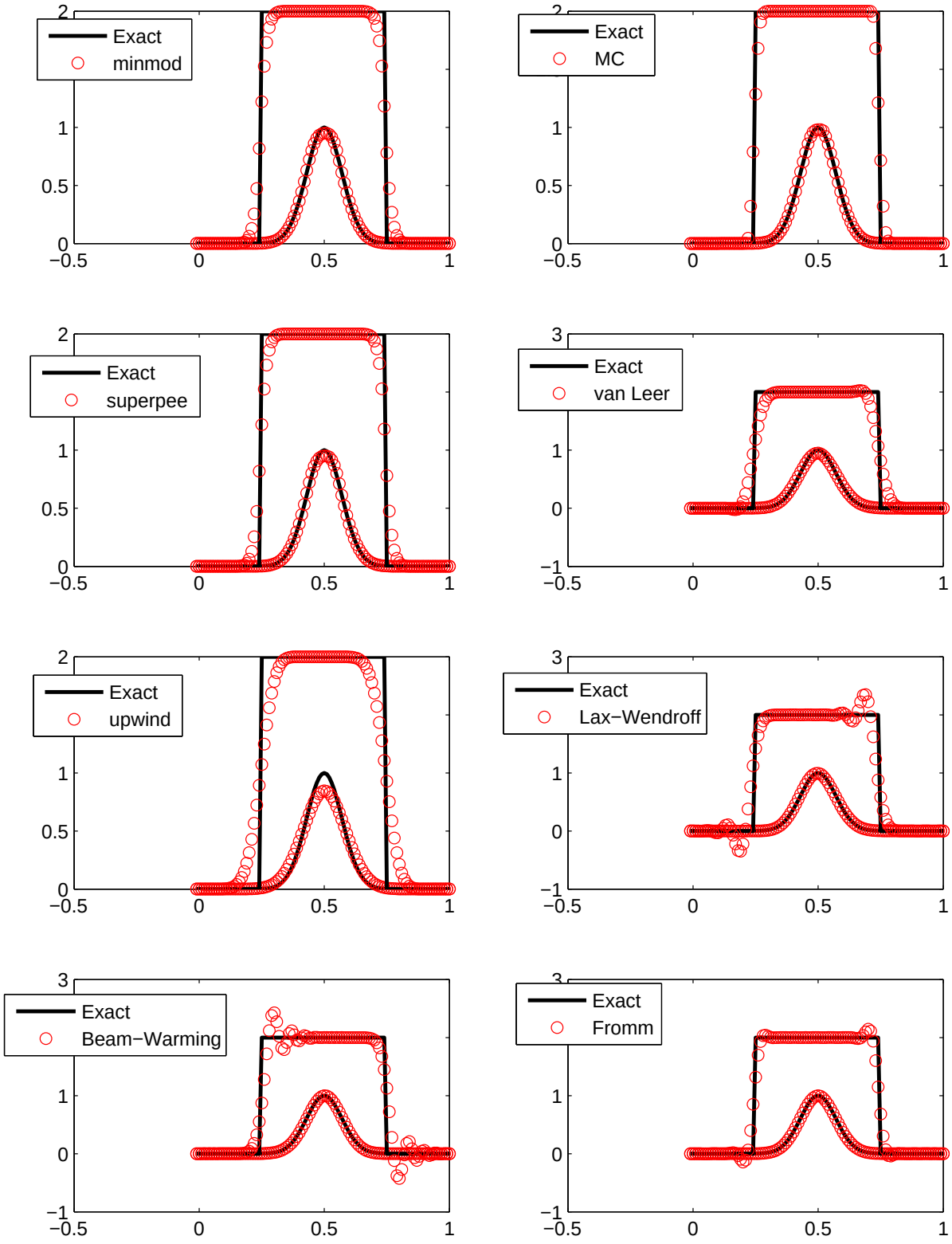

Figure 15: Performance of the different limiter functions listed in the text for both the smooth and non-smooth data.