Numerical Analysis for Numerical Relativists

Matthew W. Choptuik
CIAR Cosmology and Gravity Program
Department of Physics and Astronomy
University of British Columbia
Vancouver BC

VII Mexican School on Gravitation and Mathematical Physics
Relativistic Astrophysics and Numerical Relativity
Playa del Carmen, Quintana Roo, Mexico

November 26 – December 2, 2006

http://laplace.physics.ubc.ca/∼matt/Teaching/06Mexico/mexico06.pdf

Numerical Analysis for Numerical Relativists: Summary

- Basic Finite Difference Techniques for Time Dependent PDEs
- Basic Finite Difference Techniques for Time Independent PDEs

Numerical Analysis for Numerical Relativists
(Some Of) What WON'T Be Covered

- Other discretrization techniques (spectral, finite-element)
- Multi-dimensional problems, except for 2D elliptic
- Important mathematical issues (well posedness, hyperbolicity, $\cdots$

Basic Finite Difference Techniques for Time Dependent PDEs: References

- Mitchell, A. R., and D. F. Griffiths, **The Finite Difference Method in Partial Differential Equations**, New York: Wiley (1980)

- Richtmeyer, R. D., and Morton, K. W., **Difference Methods for Initial-Value Problems**, New York: Interscience (1967)

- H.-O. Kreiss and J. Oliger, **Methods for the Approximate Solution of Time Dependent Problems**, GARP Publications Series No. 10, (1973)

- Gustatsson, B., H. Kreiss and J. Oliger, **Time-Dependent Problems and Difference Methods**, New York: Wiley (1995)

Basic Finite Difference Techniques for Time Dependent PDEs: Outline

- Preliminaries
- Types of IVPs (by example)
- Basic Concepts, Definitions & Techniques
- Sample Discretizations / FDAs
- The 1-D Wave Equation in More Detail
- Stability Analysis
- Dispersion and Disspation
- The Leap-Frog Scheme
- Error Analysis and Convergence Tests
- Dispersion and Dissipation in FDAs

Basic Finite Difference Techniques for Time Dependent PDEs: Preliminaries

- Can divide time-dependent PDEs into two broad classes:

  1. **Initial-value Problems (Cauchy Problems)**, spatial domain has no boundaries (either infinite or "closed"—e.g. "periodic boundary conditions"
  2. **Initial-Boundary-Value Problems**, spatial domain *finite*, need to specify boundary conditions

- **Note:** Even if *physical* problem is really of Type 1, finite computational resources $\longrightarrow$ finite spatial domain $\longrightarrow$ approximate as Type 2; will hereafter loosely refer to either type as an IVP.

- *Working Definition:* **Initial Value Problem**

  - State of physical system arbitrarily (usually) specified at some initial time $t = t_0$.
  - Solution exists for $t \geq t_0$; uniquely determined by equations of motion (EOM) and boundary conditions (BCs).

Issues in Finite Difference (FD) Approximation of IVPs

- Discretization (Derivation of FDA's)
- Solution of algebraic systems resulting from discretization
- Consistency
- Accuracy
- Stability
- Converegence
- Dispersion / Dissipation
- Treatment of Non-linearities
- Computational cost—expect $O(N)$ work ($N \equiv$ number of "grid points" (discrete events at which approximate solution is computed)

Types of IVPs (by example)

- In the following three examples, u is always a function of one space and one time variable, i.e. $u \equiv u(x,t)$.
- Such a problem is often referred to as "1-d" by numericists: time dimension is implicit
- Will also use the subscript notation for partial differentiation, e.g. $u_t \equiv \partial_t u$.

Types of IVPs (by example)

- Wave and "Wave-Like" ("Hyperbolic"): The 1-d Wave Equation

$$\begin{aligned} u_{tt} &= c^2 u_{xx} \qquad c \in \mathbf{R}, \\ u(x,0) &= u_0(x) \\ u_t(x,0) &= v_0(x) \end{aligned} \tag{1}$$

Types of IVPs (by example)

- Diffusion ("Parabolic"): The 1-d Diffusion Equation

$$\begin{aligned} u_t &= \sigma u_{xx} \qquad \sigma \in \mathbf{R}, \quad \sigma > 0. \\ u(x,0) &= u_0(x) \end{aligned} \tag{2}$$

Types of IVPs (by example)

- Schrödinger: The 1-d Schrödinger Equation

$$\begin{aligned} i\psi_t &= -\frac{\hbar}{2m}\psi_{xx} + V(x,t)\psi \qquad \psi \in \mathbf{C} \\ \psi(x,0) &= \psi_0(x) \end{aligned} \tag{3}$$

  - **Note:** Although $\psi(x,t)$ is *complex* in this case, can rewrite (3) as a *system* of 2 coupled scalar, real-valued equations.

Some Basic Concepts, Definitions and Techniques

- Will be considering the finite-difference approximation (FDA) of PDEs-0—will generally be interested in the continuum limit, where the *mesh spacing*, or *grid spacing*, usually denoted h, tends to 0.
- Because any specific calculation must necessarily be performed at some specific, *finite* value of h, we will also be (extremely!) interested in the way that our discrete solution varies as a function of h.
- Will *always* view h as the basic "control" parameter of a typical FDA.
- Fundamentally, for sensibly constructed FDAs, we expect the error in the approximation to go to 0, as h goes to 0.

Some Basic Concepts, Definitions and Techniques

- Let
$$Lu = f \tag{4}$$
denote a general *differential* system.

- For simplicity, concreteness, can think of $u = u(x,t)$ as a single function of one space variable and time,

- Discussion applies to cases in more independent variables ($u(x,y,t)$, $u(x,y,z,t)$ $\cdots$ etc.), as well as multiple *dependent* variables ($u = \mathbf{u} = [u_1, u_2, \cdots, u_n]$).

- In (4), L is some differential operator (such as $\partial_{tt} - \partial_{xx}$) in our wave equation example), u is the unknown, and f is some specified function (frequently called a *source* function) of the independent variables.

Some Basic Concepts, Definitions and Techniques

- Here and in the following, will *sometimes* be convenient use notation where a superscript h on a symbol indicates that it is discrete, or associated with the FDA, rather than the continuum.

- With this notation, we will generically denote an FDA of (4) by

$$L^h u^h = f^h \tag{5}$$

where u^h is the discrete solution, f^h is the specified function evaluated on the finite-difference mesh, and L^h is the finite-difference approximation of L.

Residual

- Note that another way of writing our FDA is

$$L^h u^h - f^h = 0 \tag{6}$$

- Often useful to view FDAs in this form for following reasons
  - Have a canonical view of what it means to solve the FDA—"drive the left-hand side to 0".
  - For iterative approaches to the solution of the FDA (which are common, since it may be too expensive to solve the algebraic equations directly), are naturally lead to the concept of a *residual*.
  - Residual is simply the level of "non-satisfaction" of our FDA (and, indeed, of any algebraic expression).
  - Specifically, if $\tilde{u}^h$ is some approximation to the true solution of the FDA, u^h, then the residual, r^h, associated with $\tilde{u}^h$ is just

$$r^h \equiv L^h \tilde{u}^h - f^h \tag{7}$$

- Leads to the view of a convergent, iterative process as being one which "drives the residual to 0".

Truncation Error

- *Truncation error*, τ^h, of an FDA is defined by

$$\tau^h \equiv L^h u - f^h \tag{8}$$

  where u satisfies the continuum PDE (4).

- Note that the *form* of the truncation error can always be computed (typically using Taylor series) from the finite difference approximation and the differential equations.

Convergence

- Assume FDA is characterized by a *single* discretization scale, h,
- we say that the approximation *converges* if and only if

$$u^h \to u \quad \text{as} \quad h \to 0. \tag{9}$$

- In practice, convergence is clearly our chief concern as numerical analysts, particularly if there is reason to suspect that the solutions of our PDEs are good models for real phenomena.
- Note that this is believed to be the case for many interesting problems in general relativistic astrophysics—the two black hole problem being an excellent example.

Consistency

- Assume FDA with truncation error τ^h is characterized by a single discretization scale, h,
- Say that the FDA is *consistent* if

$$\tau^h \to 0 \quad \text{as} \quad h \to 0. \tag{10}$$

- Consistency is obviously a necessary condition for convergence.

Order of an FDA

- Assume FDA is characterized by a single discretization scale, h
- Say that the FDA is p-*th order accurate* or simply p-*th order* if

$$\lim_{h \to 0} \tau^h = O(h^p) \qquad \text{for some integer } p \tag{11}$$

Solution Error

- Solution error, e^h, associated with an FDA is defined by

$$e^h \equiv u - u^h \tag{12}$$

Relation Between Truncation Error and Solution Error

- Common to tacitly assume that

$$\tau^h = O(h^p) \qquad \longrightarrow \qquad e^h = O(h^p)$$

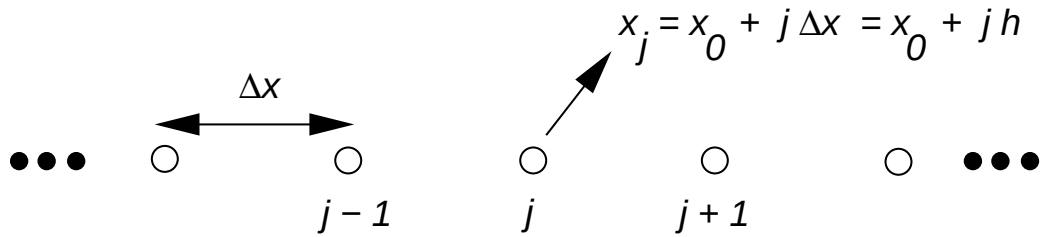
- Assumption is often warranted, but is extremely instructive to consider *why* it is warranted and to investigate (following Richardson 1910 (!)) in some detail the *nature* of the solution error.

- Will return to this issue in more detail later.

Deriving Finite Difference Formulae

- Essence of finite-difference approximation of a PDE:
  - Replacement of the continuum by a discrete lattice of grid points
  - Replacement of derivatives/differential operators by finite-difference expressions
- Finite-difference expressions (finite-difference quotients) approximate the derivatives of functions at grid points, using the grid values themselves. All operators and expressions needed here can easily be worked out using Taylor series techniques.
- Example: Consider task of approximating the first derivative $u_x(x)$ of a function $u(x)$, given a discrete set of values $u_j \equiv u(jh)$

Deriving Finite Difference Formulae



- One-dimensional, uniform finite difference mesh.
- Note that the spacing, $\Delta x = h$, between adjacent mesh points is *constant*.
- In notes, tacitly assume that the origin, x_0, of coordinate system is $x_0 = 0$.

Deriving Finite Difference Formulae

- Given the three values $u(x_j - h), u(x_j)$ and $u(x_j + h)$, denoted u_{j-1}, u_j, and u_{j+1} respectively, can compute an $O(h^2)$ approximation to $u_x(x_j) \equiv (u_x)_j$ as follows

- Taylor expanding, have

$$\begin{aligned} u_{j-1} &= u_j - h(u_x)_j + \frac{1}{2}h^2(u_{xx})_j - \frac{1}{6}h^3(u_{xxx})_j + \frac{1}{24}h^4(u_{xxxx})_j + O(h^5) \\ u_j &= u_j \\ u_{j+1} &= u_j + h(u_x)_j + \frac{1}{2}h^2(u_{xx})_j + \frac{1}{6}h^3(u_{xxx})_j + \frac{1}{24}h^4(u_{xxxx})_j + O(h^5) \end{aligned}$$

- Now seek a linear combination of u_{j-1}, u_j, and u_{j+1} which yields $(u_x)_j$ to $O(h^2)$ accuracy, i.e. we seek c_-, c_0 and c_+ such that

$$c_- \, u_{j-1} + c_0 \, u_j + c_+ \, u_{j+1} = (u_x)_j + O(h^2)$$

Deriving Finite Difference Formulae

- Results in a system of three linear equations for u_{j-1}, u_j, and u_{j+1}:

$$\begin{aligned} c_- + c_0 + c_+ &= 0 \\ -hc_- + hc_+ &= 1 \\ \frac{1}{2}h^2 c_- + \frac{1}{2}h^2 c_+ &= 0 \end{aligned}$$

  which has the solution

$$\begin{aligned} c_- &= -\frac{1}{2h} \\ c_0 &= 0 \\ c_+ &= +\frac{1}{2h} \end{aligned}$$

- Thus, $O(h^2)$ FDA for the first derivative is

$$\frac{u(x+h) - u(x-h)}{2h} = u_x(x) + O(h^2) \tag{13}$$

Deriving Finite Difference Formulae

- May not be obvious *a priori*, that the truncation error of approximation *is* $O(h^2)$
- Naive consideration of the number of terms in the Taylor series expansion which can be eliminated using 2 values (namely $u(x+h)$ and $u(x-h)$) suggests that the error might be $O(h)$.
- Fact that the $O(h)$ term "drops out" a consequence of the *symmetry*, or *centering* of the stencil: common theme in such FDA, called *centred* difference approximations
- Using same technique, can easily generate $O(h^2)$ expression for the *second* derivative, which uses the same difference stencil as the above approximation for the first derivative.

$$\frac{u(x+h) - 2u(x) + u(x-h)}{h^2} = u_{xx}(x) + O(h^2) \tag{14}$$

- *Exercise:* Compute the precise form of the $O(h^2)$ terms in expressions (13) and (14).

Sample Discretizations / FDAs

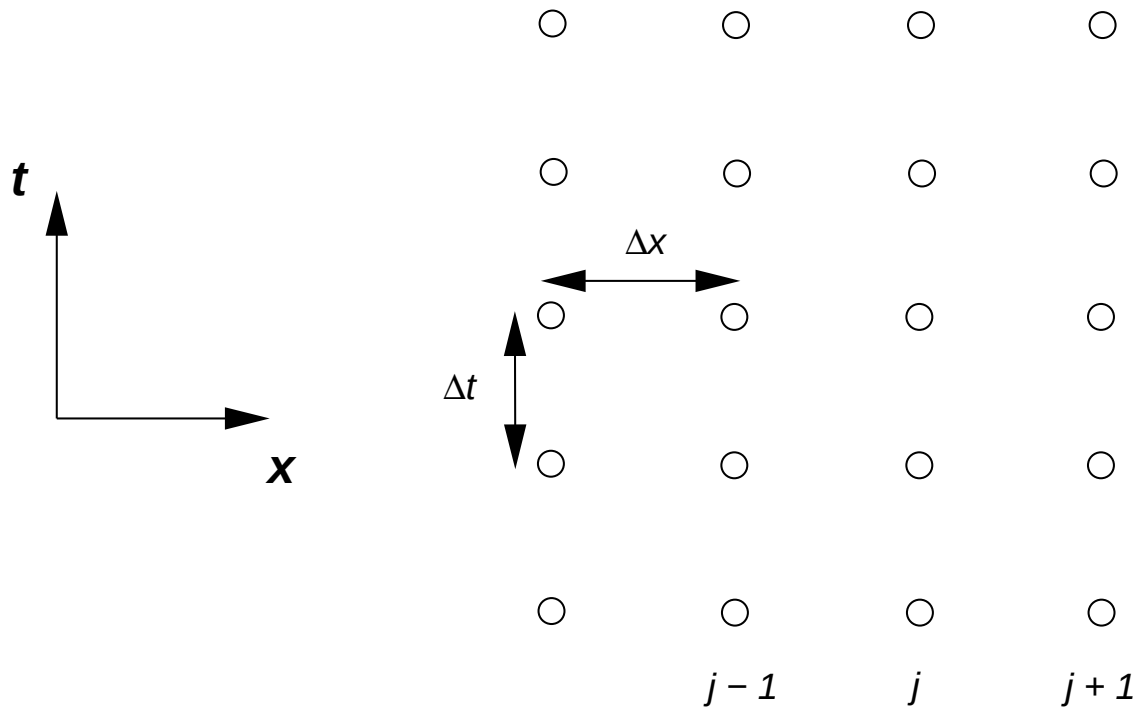
- 1-d Wave equation with fixed (Dirichlet) boundary conditions

$$\begin{aligned} u_{tt} &= u_{xx} \qquad (c=1) \qquad 0 \le x \le 1; \quad t \ge 0 \qquad (15) \\ u(x,0) &= u_0(x) \\ u_t(x,0) &= v_0(x) \\ u(0,t) &= u(1,t) = 0 \qquad (16) \end{aligned}$$

- Introduce discrete domain (uniform grid) (x_j, t^n)

$$\begin{aligned} t^n &\equiv n\,\triangle t\,, \qquad n = 0, 1, 2, \cdots \\ x_j &\equiv (j-1)\,\triangle x\,, \qquad j = 1, 2, \cdots J \\ u_j^n &\equiv u(n\,\triangle t\,, (j-1)\,\triangle x\,) \\ \triangle x &= (J-1)^{-1} \\ \triangle t &= \lambda\,\triangle x \qquad \lambda \equiv \text{“Courant number”} \end{aligned}$$

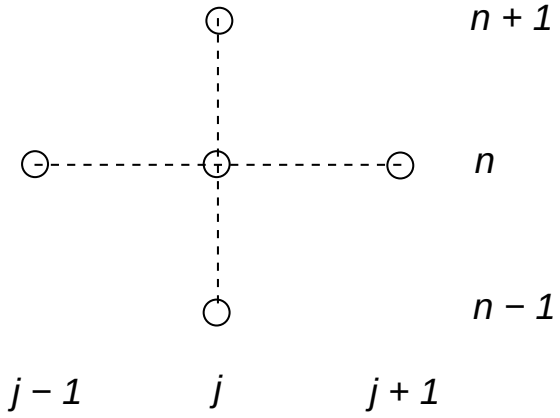
Uniform Grid for 1-D Wave Equation



- When solving wave equations using FDAs, typically keep λ constant when $\triangle x$ varied.
- FDA will always be characterized by a *single* discretization scale, h.

$$\begin{aligned} \triangle x &\equiv h \\ \triangle t &\equiv \lambda h \end{aligned}$$

Stencil for "Standard" $O(h^2)$ Approximation of 1-D Wave Equation



FDA for 1-D Wave Equation

- Discretized Interior equation

$$
\begin{aligned}
(\triangle t)^{-2}\left(u_j^{n+1}-2u_j^n+u_j^{n-1}\right) &= (u_{tt})_j^n+\frac{1}{12}\triangle t^2\,(u_{tttt})_j^n+O(\triangle t^4)\\
&= (u_{tt})_j^n+O(h^2)\\
(\triangle x)^{-2}\left(u_{j+1}^n-2u_j^n+u_{j-1}^n\right) &= (u_{xx})_j^n+\frac{1}{12}\triangle x^2\,(u_{xxxx})_j^n+O(\triangle x^4)\\
&= (u_{xx})_j^n+O(h^2)
\end{aligned}
$$

Putting these two together, get $O(h^2)$ approximation

$$
\frac{u_j^{n+1}-2u_j^n+u_j^{n-1}}{\triangle t^2}=\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{\triangle x^2}\qquad j=2,3,\cdots,J-1 \qquad (17)
$$

- Scheme such as (17) often called a *three level scheme* since couples *three "time levels"* of data (i.e. unknowns at three distinct, discrete times t^{n-1}, t^n, t^{n+1}.

FDA for 1-D Wave Equation

- Discretized Boundary conditions

$$u_1^{n+1} = u_J^{n+1} = 0$$

- Discretized Initial conditions
  - Need to specify *two* "time levels" of data (effectively $u(x,0)$ and $u_t(x,0)$), i.e. we must specify

$$u_j^0 \quad , \quad j = 1, 2, \cdots, J$$
$$u_j^1 \quad , \quad j = 1, 2, \cdots, J$$

    ensuring that the initial values are compatible with the boundary conditions.

- Can solve (17) *explicitly* for u_j^{n+1}:

$$u_j^{n+1} = 2u_j^n - u_j^{n-1} + \lambda^2 \left(u_{j+1}^n - 2u_j^n + u_j^{n-1} \right) \tag{18}$$

FDA for 1-D Wave Equation

- Also note that (18) is actually *linear system* for the unknowns $u_j^{n+1},\ j = 1, 2, \cdots, J$; in combination with the discrete boundary conditions can write

$$\mathbf{A}\ \mathbf{u}^{n+1} = \mathbf{b} \tag{19}$$

  where **A** is a *diagonal* $J \times J$ matrix and $\mathbf{u}^{n+1}$ and $\mathbf{b}$ are vectors of length J.

- Such a difference scheme for an IVP is called an *explicit* scheme.

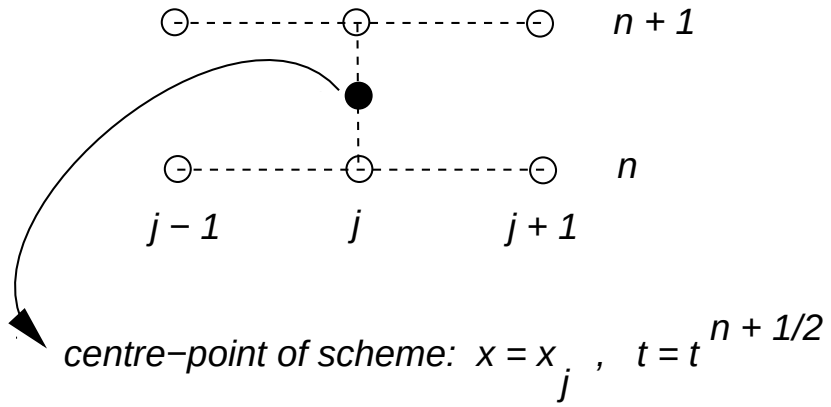
Sample Discretizations / FDAs

- 1-d Diffusion equation with Dirichlet boundary conditions

$$\begin{aligned} u_t &= u_{xx} \qquad (\sigma = 1) \qquad 0 \leq x \leq 1; \quad t \geq 0 \qquad (20) \\ u(x,0) &= u_0(x) \\ u(0,t) &= u_(1,t) = 0 \end{aligned}$$

- Use same discrete domain (grid) as for the 1-d wave equation.

Crank-Nicholson Stencil for $O(h^2)$ Approximation of 1-D Diffusion Equation



FDA for 1-D Diffusion Equation: Crank-Nicholson

- Scheme illustrates a useful "rule of thumb": *Keep the difference scheme "centred"*
  - centred in time, centred in space
  - minimizes truncation error for given h
  - tends to minimize instabilities

- Discretization of time derivative:

$$\begin{aligned} \triangle t^{-1}\left(u_j^{n+1} - u_j^n\right) &= (u_t)_j^{n+\frac{1}{2}} + \frac{1}{24}\triangle t^2 (u_{ttt})_j^{n+\frac{1}{2}} + O(\triangle t^4) \quad (21) \\ &= (u_t)_j^{n+\frac{1}{2}} + O(\triangle t^2) \end{aligned}$$

- $O(h^2)$ second-derivative operator:

$$D_{xx}\, u_j^n \equiv \triangle x^{-2}\left(u_{j+1}^n - 2u_j^n + u_{j+1}^n\right) \quad (22)$$

$$D_{xx} = \partial_{xx} + \frac{1}{12}\triangle x^2\, \partial_{xxxx} + O(\triangle x^4) \quad (23)$$

FDA for 1-D Diffusion Equation: Crank-Nicholson

- (Forward) Time-averaging operator, μ_t:

$$\mu_t \, u_j^n \equiv \frac{1}{2}\left(u_j^{n+1} + u_j^{n-1}\right) = u_j^{n+\frac{1}{2}} + \frac{1}{8}\triangle t^{\,2}\,(u_{tt})_j^{n+\frac{1}{2}} + O(\triangle t^{\,4}) \quad (24)$$

$$\mu_t = \left[I + \frac{1}{8}\triangle t^{\,2}\,\partial_{tt} + O(\triangle t^{\,4}) \right]_{t=t^{n+1/2}} \quad (25)$$

  where I is the identity operator.

- Assuming that $\triangle t = O(\triangle x) = O(h)$, is easy to show (*exercise*) that

$$\mu_t \left[D_{xx}\, u_j^n \right] = (u_{xx})_j^{n+\frac{1}{2}} + O(h^2)$$

- Putting above results together, get ($O(h^2)$) Crank-Nicholson approximation of (20):

$$\frac{u_j^{n+1} - u_j^n}{\triangle t} = \mu_t \left[D_{xx}\, u_j^n \right] \quad (26)$$

FDA for 1-D Diffusion Equation: Crank-Nicholson

- Written out in full, have

$$\frac{u_j^{n+1} - u_j^n}{\triangle t} = \frac{1}{2}\left[\frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{\triangle x^{\,2}} + \frac{u_{j+1}^{n} - 2u_j^{n} + u_{j-1}^{n}}{\triangle x^{\,2}}\right] \qquad j = 2, 3, \cdots, J \tag{27}$$

- Can rewrite (27) in the form

$$a_+\, u_{j+1}^{n+1} + a_0\, u_j^{n+1} + a_-\, u_{j-1}^{n+1} = b_j \qquad j = 2, 3, \cdots, J-1 \tag{28}$$

where

$$\begin{aligned} a_+ &\equiv -\frac{1}{2}\triangle x^{\,-2} \\ a_0 &\equiv \triangle t^{\,-1} + \triangle x^{\,-2} \\ a_- &\equiv -\frac{1}{2}\triangle x^{\,-2} \\ b_j &\equiv \left(\triangle t^{\,-1} - \triangle x^{\,-2}\right) u_j^n + \frac{1}{2}\triangle x^{\,-2}\left(u_{j+1}^n + u_{j-1}^n\right) \end{aligned}$$

FDA for 1-D Diffusion Equation: Crank-Nicholson

- Along with the BCs ($u_1^{n+1} = u_J^{n+1} = 0$), again have linear system of the form

$$\mathbf{A}\ \mathbf{u}^{n+1} = \mathbf{b}$$

  for the "unknown vector" $\mathbf{u}^{n+1}$.

- This time, matrix **A**, is *not* diagonal: scheme is called *implicit*—i.e. the scheme *couples* unknowns at the *advanced* time level, $t = t^{n+1}$.

- **A** is a *tridiagonal* matrix: all elements A_{ij} for which $j \neq i+1, i$ or $i-1$ vanish.

- Solution of tridiagonal systems can be performed very efficiently using special purpose routines (such as `DGTSV` in `LAPACK`)

- Specifically, the operation count for solution of (27) is $O(J)$.

Sample Discretizations / FDAs

- 1-d Schrödinger equation

- In analogy with diffusion equation, can immediately write down the Crank-Nicholson scheme for Schrödinger equation (3):

$$i\,\frac{\psi_j^{n+1}-\psi_j^n}{\triangle t} = -\frac{\hbar}{2m}\mu_t\left[D_{xx}\,\psi_j^n\right] + V(x_j)\,\mu_t\psi_j^n \tag{29}$$

- In this case get a *complex* tridiagonal system, which can also be solved in $O(J)$ time, using, for example, the `LAPACK` routine `ZGTSV`.

The 1-D Wave Equation in More Detail

- Recall "standard" $O(h^2)$ discretization:

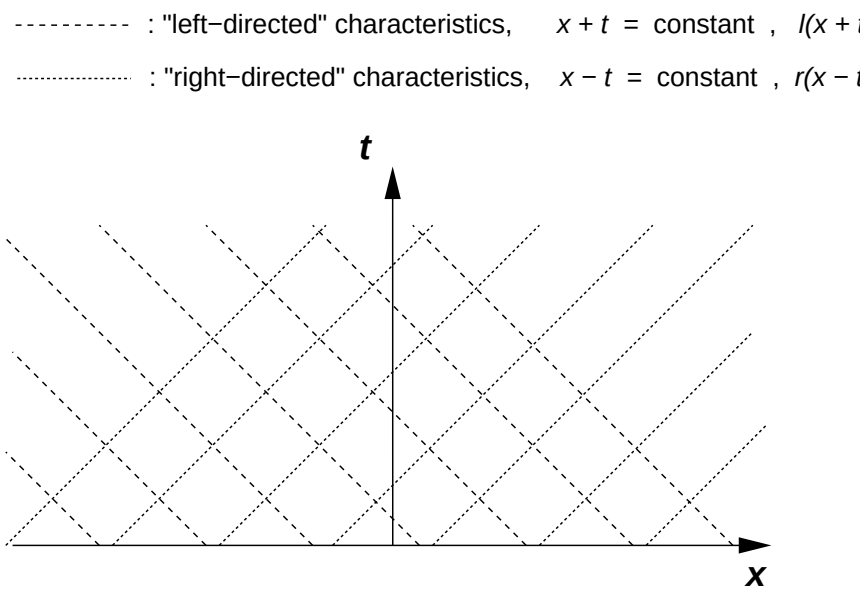
$$\begin{aligned} u_j^{n+1} &= 2u_j^n - u_j^{n-1} + \lambda^2 \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n \right), \qquad j = 2, 3, \cdots, J-1 \\ u_1^{n+1} &= u_J^{n+1} = 0 \end{aligned}$$

- To initialize the scheme, need to specify u_j^0 and u_j^1: equivalent (in the limit $h \to 0$) to specifying $u(x,0)$ and $u_t(x,0)$.

- First consider continuum case; for sake of presentation, assume solution of a true IVP on an unbounded domain; i.e. wish to solve

$$u_{tt} = u_{xx} \qquad -\infty < x < \infty \quad , \quad t \geq 0 \tag{30}$$

The 1-D Wave Equation in More Detail



- General solution of (30) is a superposition of an arbitrary *left-moving* profile ($v = -c = -1$), and an arbitrary *right-moving* profile ($v = +c = +1$); i.e.

$$u(x,t) = \ell(x+t) + r(x-t) \tag{31}$$

where

ℓ : constant along "left-directed" characteristics

r : constant along "right-directed" characteristics

The 1-D Wave Equation in More Detail

- Observation provides alternative way of specifying initial values—often convenient in practice.
- Rather than specifying $u(x,0)$ and $u_t(x,0)$ directly, specify *initial* left-moving and right-moving parts of the solution, $\ell(x)$ and $r(x)$.
- Specifically, set

$$u(x,0) = \ell(x) + r(x) \tag{32}$$

$$u_t(x,0) = \ell'(x) - r'(x) \equiv \frac{d\ell}{dx}(x) - \frac{dr}{dx}(x) \tag{33}$$

- Return now to the solution of finite-differenced version of the wave equation
- Clearly, given initial data (32–33), can trivially initialize u_j^0 with *exact* values, but can only approximately initialize u_j^1.
- Question: *How accurately must one initialize the advanced values to ensure second order ($O(h^2)$) accuracy of the difference scheme?*

The 1-D Wave Equation in More Detail

- Brief, heuristic answer to this question (can be more rigorously justified):
- Have $\triangle x = O(h)$, $\triangle t = O(h)$ and the FDA is $O(h^2)$. Since the scheme is $O(h^2)$, expect

$$u_{\text{exact}}(x,t) - u_{\text{FD}}(x,t) = O(h^2)$$

  for arbitrary, *fixed, FINITE* t.
- But number of time steps required to integrate to time t is $O(\triangle t^{-1}) = O(h^{-1})$.
- Thus, per-time-step error must be $O(h^2)/O(h^{-1}) = O(h^3)$, so require

$$(u_{\text{FD}})_j^1 = (u_{\text{exact}})_j^1 + O(h^3)$$

The 1-D Wave Equation in More Detail

- Can readily accomplish this using

  1. Taylor series
  2. Equation of motion to rewrite higher time derivatives in terms of spatial derivatives:

$$u_j^1 = u_j^0 + \triangle t\,(u_t)_j^0 + \frac{1}{2}\triangle t^2\,(u_{tt})_j^0 + O(\triangle t^3) \tag{34}$$

$$= u_j^0 + \triangle t\,(u_t) + \frac{1}{2}\triangle t^2\,(u_{xx})_j^0 + O(\triangle t^3) \tag{35}$$

which, using results from above, can be written as

$$u_j^1 = (\ell + r)_j + \triangle t\,(\ell' - r')_j + \frac{1}{2}\triangle t^2\,(\ell'' + r'')_j \tag{36}$$

Stability Analysis

- One of the most frustrating/fascinating features of FD solutions of time dependent problems: discrete solutions often "blow up"—e.g. floating-point overflows are generated at some point in the evolution
- 'Blow-ups" can sometimes be caused by legitimate (!) "bugs"—i.e. an incorrect implementation—at other times it is simply the *nature of the FD scheme* which causes problems.
- Are thus lead to consider the *stability* of solutions of difference equations
- Again consider the 1-d wave equation (15)
- Note that it is a *linear, non-dispersive* wave equation
- Thus the "size" of the solution does *not* change with time:

$$\|u(x,t)\| \sim \|u(x,0)\|\,, \tag{37}$$

where $\|\cdot\|$ is an suitable norm, such as the L_2 norm:

$$\|u(x,t)\| \equiv \left(\int_0^1 u(x,t)^2\,dx\right)^{1/2}\,. \tag{38}$$

Stability Analysis

- Will use the property captured by (37) as working definition of stability.
- In particular, if you believe (37) is true for the wave equation, then you believe the wave equation is stable.
- Fundamentally, if FDA approximation *converges*, then expect the same behaviour for the difference solution:

$$\|u_j^n\| \sim \|u_j^0\| \,. \tag{39}$$

- FD solution constructed by *iterating in time*, generating

$$u_j^0,\ u_j^1,\ u_j^2,\ u_j^3,\ u_j^4,\ \cdots$$

in succession, using the FD equation

$$u_j^{n+1} = 2u_j^n - u_j^{n-1} + \lambda^2 \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n \right) \,.$$

Stability Analysis

- *Not* guaranteed that (39) holds for all values of $\lambda \equiv \triangle t \,/\, \triangle x$.
- For certain λ, have
$$\|u_j^n\| \gg \|u_j^0\| \,,$$
and for those λ, $\|u^n\|$ *diverges* from u, even (especially!) as $h \to 0$—that is, the difference scheme is *unstable*.
- For many wave problems (including all linear problems), given that a FD scheme is *consistent* (i.e. so that $\hat{\tau} \to 0$ as $h \to 0$), *stability is the necessary and sufficient condition for convergence* (Lax's theorem).

Heuristic Stability Analysis

- Write general time-dependent FDA in the form

$$\mathbf{u}^{n+1} = \mathbf{G}[\mathbf{u}^n]\,, \tag{40}$$

- $\mathbf{G}$ is some *update operator* (linear in our example problem)
- $\mathbf{u}$ is a column vector containing sufficient unknowns to write the problem in first-order-in-time form.
- Example: introduce new, auxiliary set of unknowns, v_j^n, defined by

$$v_j^n = u_j^{n-1}\,,$$

then can rewrite differenced-wave-equation (17) as

$$u_j^{n+1} = 2u_j^n - v_j^n + \lambda^2\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right)\,, \tag{41}$$

$$v_j^{n+1} = u_j^n\,, \tag{42}$$

Heuristic Stability Analysis

- Thus with

$$\mathbf{u}^n = [u_1^n, v_1^n,\ u_2^n, v_2^n, \cdots u_J^n, v_J^n]\,,$$

(for example), (41-42) is of the form (40).

- Equation (40) provides compact way of describing the FDA solution.

- Given initial data, $\mathbf{u}^0$, solution after n time-steps is

$$\mathbf{u}^n = \mathbf{G}^n \mathbf{u}^0\,, \tag{43}$$

where $\mathbf{G}^n$ is the n-th power of the matrix $\mathbf{G}$.

- Assume that $\mathbf{G}$ has a complete set of orthonormal eigenvectors

$$\mathbf{e}_k, \quad k = 1, 2, \cdots J\,,$$

and corresponding eigenvalues

$$\mu_k, \quad k = 1, 2, \cdots J\,,$$

Heuristic Stability Analysis

- Thus have

$$\mathbf{G}\,\mathbf{e}_k = \mu_k\,\mathbf{e}_k, \quad k = 1, 2, \cdots J\,.$$

- Can then write initial data as (spectral decomposition):

$$\mathbf{u}^0 = \sum_{k=1}^{J} c_k^0\,\mathbf{e}_k\,,$$

  where the c_k^0 are coefficients.

- Using (43), solution at time-step n is

$$\mathbf{u}^n = \mathbf{G}^n \left(\sum_{k=1}^{J} c_k^0\,\mathbf{e}_k \right) \tag{44}$$

$$= \sum_{k=1}^{J} c_k^0\,(\mu_k)^n\,\mathbf{e}_k\,. \tag{45}$$

Heuristic Stability Analysis

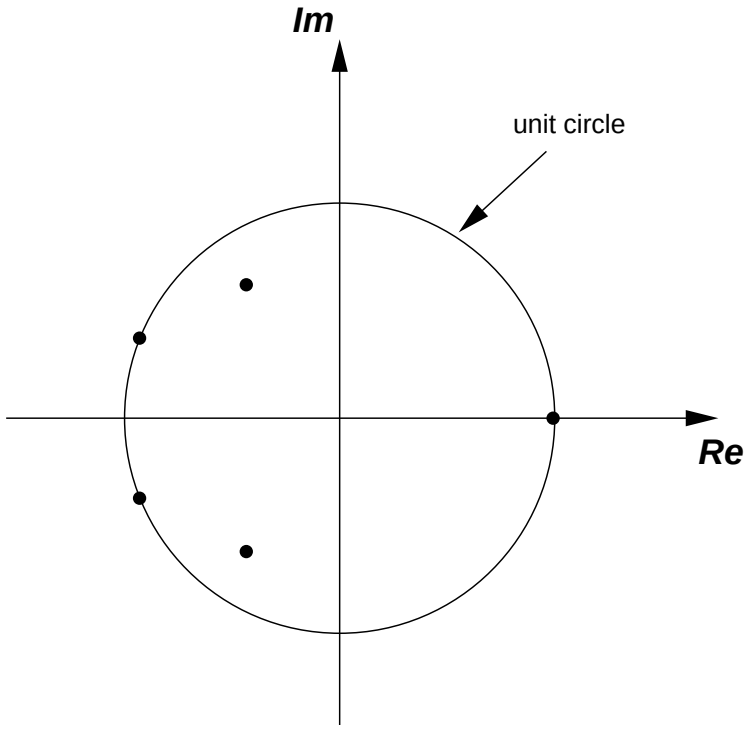
- If difference scheme is to be stable, must have

$$|\mu_k| \leq 1 \qquad k = 1, 2, \cdots J \tag{46}$$

(Note: μ_k will be complex in general, so $|\mu|$ denotes the complex modulus, $|\mu| \equiv \sqrt{\mu\mu^\star}$).

- Geometric interpretation: eigenvalues of the update matrix must lie on or within the unit circle

Heuristic Stability Analysis



- Schematic illustration of location in complex plane of eigenvalues of update matrix **G**.
- In this case, all eigenvalues (dots) lie on or within the unit circle, indicating that the corresponding finite difference scheme is stable.

Von-Neumann (Fourier) Stability Analysis

- Von-Neumann stability analysis based on the ideas sketched above
- Also assumes that the difference equation is linear with constant coefficients, and that the boundary conditions are periodic
- Can then use Fourier analysis: difference operators in real-space variable $x \longrightarrow$ algebraic operations in Fourier-space variable k
- Schematically, instead of writing

$$\mathbf{u}^{n+1}(x) = \mathbf{G}[\mathbf{u}^n(x)]\,,$$

consider the Fourier-domain equivalent:

$$\tilde{\mathbf{u}}^{n+1}(k) = \tilde{\mathbf{G}}[\tilde{\mathbf{u}}^n(k)]\,,$$

where k is the wave-number (Fourier-space variable) and $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{G}}$ are the Fourier-transforms of $\mathbf{u}$ and $\mathbf{G}$, respectively.

Von-Neumann (Fourier) Stability Analysis

- Specifically, define the Fourier-transformed grid function via

$$\tilde{\mathbf{u}}^n(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ikx}\, \mathbf{u}^n(x)\, dx\,. \tag{47}$$

- For a general difference scheme, will find that

$$\tilde{\mathbf{u}}^{n+1}(k) = \tilde{\mathbf{G}}(\xi)\, \tilde{\mathbf{u}}^n(k)\,,$$

  where $\xi \equiv kh$,

- Will have to show that $\tilde{\mathbf{G}}(\xi)$'s eigenvalues lie within or on the unit circle for all conceivable ξ.

- Appropriate range for ξ is

$$-\pi \le \xi \le \pi\,,$$

  since shortest wavelength representable on a uniform mesh with spacing h is $\lambda = 2h$ (Nyquist limit), corresponding to a maximum wave number $k = (2\pi)/\lambda = \pm\pi/h$.

Von-Neumann (Fourier) Stability Analysis

- Consider the application of the Von-Neumann stability analysis to our current model problem.

- First define (non-divided) difference operator D^2

$$D^2\, u(x) = u(x+h) - 2u(x) + u(x-h)\,.$$

- Suppress the spatial grid index and write the first-order form of the difference equation (41-42) as

$$\begin{aligned} u^{n+1} &= 2u^n - v^n + \lambda^2\, D^2\, u^n\,, \\ v^{n+1} &= u^n\,, \end{aligned}$$

or

$$\begin{bmatrix} u \\ v \end{bmatrix}^{n+1} = \begin{bmatrix} 2+\lambda^2\, D^2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}^{n}\,. \tag{48}$$

Von-Neumann (Fourier) Stability Analysis

- Need to know the action of D^2 in Fourier-space.
- Using inverse transform have

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx}\, \tilde{u}(k)\, dk\,,$$

so

$$\begin{aligned} D^2\, u(x) &= u(x+h) - 2u(x) + u(x-h) \\ &= \int_{-\infty}^{+\infty} \left(e^{ikh} - 2 + e^{-ikh}\right) e^{ikx}\, \tilde{u}(k)\, dk \\ &= \int_{-\infty}^{+\infty} \left(e^{i\xi} - 2 + e^{-i\xi}\right) e^{ikx}\, \tilde{u}(k)\, dk\,. \end{aligned}$$

Von-Neumann (Fourier) Stability Analysis

- Consider quantity $-4\sin^2(\xi/2)$:

$$\begin{aligned} -4\sin^2\frac{\xi}{2} &= -4\left(\frac{e^{i\xi/2}-e^{-i\xi/2}}{2i}\right)^2 \\ &= \left(e^{i\xi/2}-e^{-i\xi/2}\right)^2 = e^{i\xi}-2+e^{-i\xi}\,, \end{aligned}$$

so

$$D^2\,u(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(-4\sin^2\frac{\xi}{2}\right)e^{ikx}\,\tilde{u}(k)\,dk\,.$$

- In summary, under Fourier transformation, have

$$\begin{aligned} \mathbf{u}(x) &\longrightarrow \tilde{\mathbf{u}}(k)\,, \\ D^2\mathbf{u}(x) &\longrightarrow -4\sin^2\frac{\xi}{2}\tilde{\mathbf{u}}(k)\,. \end{aligned}$$

Von-Neumann (Fourier) Stability Analysis

- Use this result in the Fourier transform of (48): need to compute the eigenvalues of

$$\begin{bmatrix} 2 - 4\lambda^2 \sin^2(\xi/2) & -1 \\ 1 & 0 \end{bmatrix},$$

- Then must determine conditions so eigenvalues lie on or within the unit circle.

- Characteristic equation (roots are eigenvalues) is

$$\begin{vmatrix} 2 - 4\lambda^2 \sin^2(\xi/2) - \mu & -1 \\ 1 & -\mu\,, \end{vmatrix} = 0$$

or

$$\mu^2 + \left(4\lambda^2 \sin^2 \frac{\xi}{2} - 2\right)\mu + 1 = 0\,.$$

- Equation has roots

$$\mu(\xi) = \left(1 - 2\lambda^2 \sin^2 \frac{\xi}{2}\right) \pm \left(\left(1 - 2\lambda^2 \sin^2 \frac{\xi}{2}\right) - 1\right)^{1/2}.$$

Von-Neumann (Fourier) Stability Analysis

- Now need to find sufficient conditions for

$$|\mu(\xi)| \leq 1,$$

or equivalently

$$|\mu(\xi)|^2 \leq 1.$$

- Can write

$$\mu(\xi) = (1 - Q) \pm ((1 - Q)^2 - 1)^{1/2},$$

where the quantity, Q

$$Q \equiv 2\lambda \sin^2 \frac{\xi}{2},$$

is *real* and *non-negative* ($Q \geq 0$).

Von-Neumann (Fourier) Stability Analysis

- Now two cases to consider:

  1. $(1-Q)^2 - 1 \leq 0$,
  2. $(1-Q)^2 - 1 > 0$.

- First case: $((1-Q)^2-1)^{1/2}$ is purely imaginary, so have

$$|\mu(\xi)|^2 = (1-Q)^2 + (1-(1-Q)^2) = 1\,.$$

- Second case, $(1-Q)^2 - 1 > 0 \longrightarrow (1-Q)^2 > 1 \longrightarrow Q > 2$, so have

$$1 - Q - ((1-Q^2)-1)^{1/2} < -1\,,$$

- Thus in this case, stability criterion will *always* be violated.

Von-Neumann (Fourier) Stability Analysis

- Conclude that necessary condition for Von-Neumann stability is

$$(1-Q)^2 - 1 \leq 0 \longrightarrow (1-Q)^2 \leq 1 \longrightarrow Q \leq 2\,.$$

- Since $Q \equiv 2\lambda \sin^2(\xi/2)$ and $\sin^2(\xi/2) \leq 1$, must have

$$\lambda \equiv \frac{\triangle t}{\triangle x} \leq 1\,,$$

  for stability of scheme (17).

- Condition is often called the CFL condition—after Courant, Friedrichs and Lewy who derived it in 1928

- This type of instability has "physical" interpretation, often summarized by the statement *the numerical domain of dependence of an explicit difference scheme must contain the physical domain of dependence.*

Dispersion and Dissipation

- Consider an even simpler model "wave equation", so-called *advection*, or *color* equation:

$$\begin{aligned} u_t &= a\,u_x \quad (a>0) \qquad -\infty < x < \infty \quad , \quad t \geq 0 \\ u(x,0) &= u_0(x) \end{aligned} \tag{49}$$

  which has the exact solution

$$u(x,t) = u_0(x+at) \tag{50}$$

- Another example of a non-disspative, non-dispersive partial differential equation.

- Recall what "non-dispersive" means: note that (49) admits "normal mode" solutions:

$$u(x,t) \sim e^{ik(x+at)} \equiv e^{i(kx+\omega t)}$$

  where $\omega \equiv ka$ is the *dispersion relation*, and

$$\frac{d\omega}{dk} \equiv \text{ speed of propagation of mode with wave number } k$$

Dispersion and Dissipation

- In current case

$$\frac{d\omega}{dk} = a = \text{constant}$$

- means that all modes propagate at the same speed: precisely what is meant by "non-dispersive".

- Further, if general initial profile, $u_0(x)$, is resolved into "normal-mode" (Fourier) components, find that the magnitudes of the components are preserved in time, i.e. equation (49) is also *non-dissipative*.

- Ideally would like FD solutions to have the same properties—i.e. to be dissipationless and dispersionless,

- In general, will not be (completely) possible

- Will return to the issue of dissipation and dispersion in FDAs of wave problems later

The Leap-Frog Scheme

- First note that advection equation is a good prototype for the general hyperbolic *system*:

$$\mathbf{u}_t = \mathbf{A}\mathbf{u}_x \tag{51}$$

where **u**(x,t) is the n-component *solution vector*:

$$\mathbf{u}(x,t) = [u_1(x,t),\, u_2(x,t),\, \cdots u_n(x,t)] \tag{52}$$

and the $n \times n$ matrix $\mathbf{A}$ has distinct real eigenvalues

$$\lambda_1,\, \lambda_2,\, \cdots \lambda_n$$

so that, for example, there exists a similarity transformation $\mathbf{S}$ such that

$$\mathbf{SAS}^{-1} = \mathrm{diag}(\lambda_1,\, \lambda_2,\, \cdots \lambda_n)$$

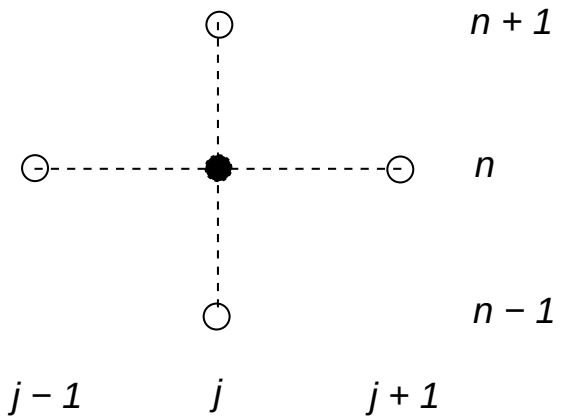
The Leap-Frog Scheme

- Leap-frog scheme is a commonly used finite-difference approximation for hyperbolic systems.
- For simple scalar ($n = 1$) advection problem (49):

$$u_t = a\, u_x$$

an appropriate stencil is as follows

The Leap-Frog Scheme



- Stencil (molecule/star) for leap-frog scheme as applied to scale advection equation
- Central grid point has been filled in this figure to emphasize that the corresponding unknown, u_j^n, does not appear in the local discrete equation at that grid point (hence the term "leap-frog")

The Leap-Frog Scheme

- Apply usual $O(h^2)$ approximations to ∂_x and ∂_t: leap-frog (LF) scheme is

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\triangle t} = a\,\frac{u_{j+1}^n - u_{j-1}^n}{2\triangle x} \tag{53}$$

  or explicitly

$$u_j^{n+1} = u_j^{n-1} + a\lambda\left(u_{j+1}^n - u_{j-1}^n\right) \tag{54}$$

  where

$$\lambda \equiv \frac{\triangle t}{\triangle x}$$

  is the *Courant number* as previously.

- *Exercise:* Perform a von Neumann stability analysis of (53) thus showing that $a\lambda \leq 1$ (i.e. the CFL condition) is necessary for stability.

The Leap-Frog Scheme

- LF scheme (53) is a *three-level* method.
- As in treatment of wave equation, $u_{tt} = u_{xx}$ using the "standard scheme", need to specify

$$u_j^0 \quad , \quad u_j^1 \qquad j = 1, 2, \cdots J$$

  to "get the scheme going"
- I.e. need to specify *two* numbers per spatial grid point.
- Contrast to continuum case where need to specify only *one* number per x_j, namely $u_0(x_j)$.
- Again, initialization of u_j^0 is trivial, given the (continuum) initial data $u_0(x)$,
- Again, need u_j^1 to $O(\triangle t^{\,3}) = O(h^3)$ accuracy for $O(h^2)$ global accuracy.
- Conside two possible approaches

The Leap-Frog Scheme

- *Approach 1: Taylor Series:* Developmentis parallel to that for the wave equation.

- Have

$$u_j^1 = u_j^0 + \triangle t \, (u_t)_j^0 + \frac{1}{2} \triangle t^2 \, (u_{tt})_j^0 + O(\triangle t^2)$$

- From equation of motion $u_t = a u_x$, get

$$u_{tt} = (u_t)_t = (a u_x)_t = a \, (u_t)_x = a^2 u_{xx}.$$

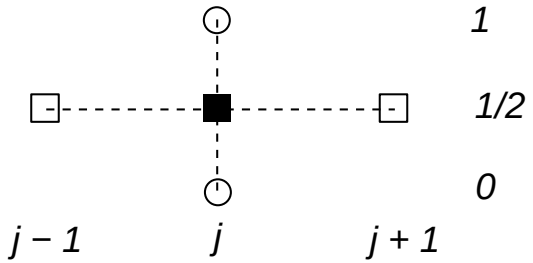
so initialization formula is

$$u_j^1 = u_j^0 + \triangle t \, (u_0')_j^0 + \frac{1}{2} \triangle t^2 \, (a^2 u_0'')_j^0 + O(\triangle t^3) \tag{55}$$

The Leap-Frog Scheme

- *Approach 2: Self-Consistent Iterative Approach:*
- Idea here is to initialize u_j^1 from u_j^0 and a version of the discrete equations of motion which introduces "ficticious" half-time-level

The Leap-Frog Scheme



- Stencil for initialization of leap-frog scheme for to (49).
- Note the introduction of the "fictitious" half-time level $t = t^{1/2}$ (squares).

The Leap-Frog Scheme

- Applying leap-frog scheme on the stencil in figure, have have

$$\frac{u_j^1 - u_j^0}{\triangle t} = a \frac{u_{j+1}^{\frac{1}{2}} - u_{j-1}^{\frac{1}{2}}}{2 \triangle x}$$

or, explicitly solving for u_j^1:

$$u_j^1 = u_j^0 + \frac{1}{2}\lambda \left(u_{j+1}^{\frac{1}{2}} - u_{j-1}^{\frac{1}{2}} \right)$$

- Straightforward to show that in order to retain $O(h^2)$ accuracy of the difference scheme, need "fictitious-time" values, $u_j^{1/2}$, accurate to $O(h^2)$ (i.e.can neglect terms which are of $O(h^2)$).

- In particular, *define* $u_j^{1/2}$, via

$$u_j^{\frac{1}{2}} = \frac{u_j^1 + u_j^0}{2}$$

The Leap-Frog Scheme

- Amounts to defining the half-time values via linear interpolation in the advanced and retarded unknowns will retain second-order accuracy.

- Are thus led to the following initialization algorithm expressed in pseudo-code (note, all loops over `j` are implicit:)

```
u[0,j] := u_0(x_j)
u[1,j] := u_0(x_j)
DO
   usave[j] := u[1,j]
   u[1/2,j] := (u[1,j] + u[0,j]) / 2

   u[1,j]   := u[0,j] + (lambda / 2) * (u[1/2,j+1] - u[1/2,j-1])

UNTIL  norm(usave[j] - u[1,j]) < epsilon
```

Error Analysis and Convergence Tests

- Discussion here applies to essentially *any* continuum problem which is solved using FDAs on a *uniform* mesh structure.
- In particular, applies to the treatment of ODEs and elliptic problems
- For such problems convergence is often easier to achieve due to fact that the FDAs are typically intrinsically stable
- Also note that departures from non-uniformity in the mesh do not, in general, complete destroy the picture: however, do tend to distort it in ways that are beyond the scope of these notes.
- Difficult to overstate importance of convergence studies

Sample Analysis: The Advection Equation

- Again consider the solution of advection equation, but this time impose periodic boundary conditions on our spatial domain

$$0 \leq x \leq 1$$

with $x = 0$ and $x = 1$ identified

$$\begin{aligned} u_t &= a\, u_x \quad (a > 0) \qquad 0 \leq x \leq 1, \quad t \geq 0 \\ u(x,0) &= u_0(x) \end{aligned} \tag{56}$$

- Note that initial conditions $u_0(x)$ must be compatible with periodicity, i.e must specify *periodic* initial data.

- Again, given initial data, $u_0(x)$, can immediately write down the full solution

$$u(x,t) = u_0(x + a\, t \bmod 1) \tag{57}$$

where $\bmod$ is the modulus function which "wraps" $x + a\ t$, $t > 0$ onto the unit interval.

Sample Analysis: The Advection Equation

- Due to the simplicity and solubility of this problem, will see that can perform a rather complete closed-form ("analytic") treatment of the convergence of simple FDAs of (56).

- Point of the exercise, however, is *not* to advocate parallel closed-form treatments for more complicated problems.

- Rather, key idea to be extracted that, in principle (always), and in practice (almost always, i.e. I've never seen a case where it *didn't* work, but then there's a lot of computations I haven't seen):

  The error, e^h, of an FDA is no less computable than the solution, u^h itself.

- Has widespread ramifications, one of which is that there is no excuse for publishing solutions of FDAs without error bars, or their equivalents!

Sample Analysis: The Advection Equation

- First introduce some difference operators for the usual $O(h^2)$ centred approximations of ∂_x and ∂_t:

$$D_x\, u_j^n \equiv \frac{u_{j+1}^n - u_{j-1}^n}{2\,\triangle x} \tag{58}$$

$$D_t\, u_j^n \equiv \frac{u_j^{n+1} - u_j^{n-1}}{2\,\triangle t} \tag{59}$$

- Again take

$$\triangle x \equiv h \qquad \triangle t \equiv \lambda\,\triangle x = \lambda\, h$$

  and hold λ fixed as h varies, so that, as usual, FDA is characterized by the single scale parameter, h.

- First key idea behind error analysis: want to view the solution of the FDA as a *continuum* problem,

- Hence express both the difference operators and the FDA solution as asymptotic series (in h) of differential operators, and continuum functions, respectively.

Sample Analysis: The Advection Equation

- Have the following expansions for D_x and D_t:

$$D_x = \partial_x + \frac{1}{6}h^2\,\partial_{xxx} + O(h^4) \tag{60}$$

$$D_t = \partial_t + \frac{1}{6}\lambda^2 h^2\,\partial_{ttt} + O(h^4) \tag{61}$$

- In terms of the general, abstract formulation discussed earlier, have

$$\begin{aligned}
L\,u - f = 0 \quad &\Longleftrightarrow \quad (\partial_t - a\,\partial_x)\,u = 0 \\
L^h u^h - f^h = 0 \quad &\Longleftrightarrow \quad (D_t - a\,D_x)\,u^h = 0 \\
L^h u - f^h \equiv \tau^h \quad &\Longleftrightarrow \quad (D_t - a\,D_x)\,u \equiv \tau^h = \frac{1}{6}h^2\left(\lambda^2\partial_{ttt} - a\,\partial_{xxx}\right)u + O(h^4)
\end{aligned}$$

Sample Analysis: The Advection Equation

- Second key idea behind error analysis: *The Richardson ansatz:* Appeal to L.F. Richardson's old observation (*ansatz*), that the solution, u^h, of *any* FDA which

  1. Uses a uniform mesh structure with scale parameter h,
  2. Is completely centred

  should have the following expansion in the limit $h \to 0$:

$$u^h(x,t) = u(x,t) + h^2 e_2(x,t) + h^4 e_4(x,t) + \cdots \tag{65}$$

- Here u is the continuum solution, while e_2, e_4, $\cdots$ are (continuum) *error functions* which *do not depend on* h.

- The Richardson expansion (65), is *the* key expression from which almost all error analysis of FDAs derives.

Sample Analysis: The Advection Equation

- In the case that the FDA is *not* completely centred, we will have to modify the *ansatz*.

- In particular, for first order schemes, will have

$$u^h(x,t) = u(x,t) + he_1(x,t) + h^2 e_x(x,t) + h^3 e_3(x,t) + \cdots \tag{66}$$

- Also Note that Richardson *ansatz* (65) is completely compatible with the assertion discussed in (), namely that

$$\tau^h = O(h^2) \quad \longrightarrow \quad e^h \equiv u - u^h = O(h^2) \tag{67}$$

- However, Richardson form (65) contains much more information than "second-order truncation error should imply second-order solution error"

- Richardson form dictates the precise form of the h dependence of u^h.

Sample Analysis: The Advection Equation

- Given the Richardson expansion, can proceed with error analysis.
- Start from the FDA, $L^h u^h - f^h = 0$, and replace both L^h and u^h with continuum expansions:

$$L^h u^h = 0 \quad \longrightarrow \quad (D_t - a\, D_x)\left(u + h^2 e_2 + \cdots\right) = 0$$

$$\longrightarrow \quad \left(\partial_t + \frac{1}{6}\lambda^2 h^2 \partial_{ttt} - a\,\partial_x - \frac{1}{6} a h^2\, \partial_{xxx} + \cdots\right)\left(u + h^2 e_2 + \cdots\right)$$

- Now demand that terms in (68) vanish order-by-order in h
- At $O(1)$ (zeroth-order), have

$$(\partial_t - a\,\partial_x)\, u = 0 \tag{69}$$

which is simply a statement of the *consistency* of the difference approximation.

Sample Analysis: The Advection Equation

- More interestingly, at $O(h^2)$ (second-order), find

$$(\partial_t - a\,\partial_x)\,e_2 = \frac{1}{6}\left(a\partial_{xxx} - \lambda^2\partial_{ttt}\right)u \tag{70}$$

- View u as a "known" function, then this is simply a PDE for the leading order error function, e_2.

- Moreover, the PDE governing e_2 is of *precisely* the same nature as the original PDE (49).

Sample Analysis: The Advection Equation

- In fact, can *solve* (70) for e_2.
- Given the "natural" initial conditions

$$e_2(x,0) = 0$$

(i.e. we initialize the FDA with the exact solution so that $u^h = u$ at $t = 0$), and defining $q(x+at)$:

$$q(x+at) \equiv \frac{1}{6} a \left(1 - \lambda^2 a^2\right) \partial_{xxx} u(x,t)$$

have

$$e_2(x,t) = t\, q(x + at \bmod 1) \tag{71}$$

- Note that, as is typical for leap-frog, we have *linear* growth of the finite difference error with time (to leading order in h).

Sample Analysis: The Advection Equation

- Also note that analysis can be extended to higher order in h—what results, then, is an entire *hierarchy* of differential equations for u and the error functions e_2, e_4, e_6, $\cdots$.

- Indeed, useful to keep following view in mind:

  When one solves an FDA of a PDE, one is *not* solving some system which is "simplified" relative to the PDE, rather, one is solving a much *richer* system consisting of an (infinite) hierarchy of PDEs, one for each function appearing in the Richardson expansion (65).

Convergence Tests

- In general case we will not be able to solve the PDE governing u, let alone that governing e_2—otherwise we wouldn't be considering the FDA in the first place!
- Is precisely in this instance where the true power of Richardson's observation is evident!
- The key observation is that starting from (65), and computing FD solutions using the same initial data, but with differing values of h, can learn a great deal about the error in FD approximations.
- The whole game of investigating the manner in which a particular FDA converges or doesn't (i.e. looking at what happens as one varies h) is known as *convergence testing*.
- Important to realize that there are no hard and fast rules for convergence testing; rather, one tends to tailor the tests to the specifics of the problem at hand, and, being largely an empirical approach, one gains experience and intuition as one works through more and more problems.
- However, the Richardson expansion, in some form or other, *always* underlies convergence analysis of FDAs.

Convergence Tests

- A simple example of a convergence test, and one commonly used in practice is as follows.
- Compute three distinct FD solutions u^h, u^{2h}, u^{4h} at resolutions h, $2h$ and $4h$ respectively, but using the same initial data (as naturally expressed on the 3 distinct FD meshes).
- Also assume that the finite difference meshes "line up", i.e. that the $4h$ grid points are a subset of the $2h$ points which are a subset of the h points
- Thus, the $4h$ points constitute a common set of events (x_j, t^n) at which specific grid function values can be directly (i.e. no interpolation required) and meaningfully compared to one another.

Convergence Tests

- From the Richardson *ansatz* (65), expect:

$$\begin{aligned} u^h &= u + h^2 e_2 + h^4 e_4 + \cdots \\ u^{2h} &= u + (2h)^2 e_2 + (2h)^4 e_4 + \cdots \\ u^{4h} &= u + (4h)^2 e_2 + (4h)^4 e_4 + \cdots \end{aligned}$$

- Then compute a quantity $Q(t)$, which will call a *convergence factor*, as follows:

$$Q(t) \equiv \frac{\|u^{4h} - u^{2h}\|_x}{\|u^{2h} - u^h\|_x} \tag{72}$$

where $\|\cdot\|_x$ is any suitable discrete spatial norm, such as the ℓ_2 norm, $\|\cdot\|_2$:

$$\|u^h\|_2 = \left(J^{-1} \sum_{j=1}^{J} \left(u_j^h\right)^2 \right)^{1/2} \tag{73}$$

- Subtractions in (72) can be taken to involve the sets of mesh points which are common between u^{4h} and u^{2h}, and between u^{2h} and u^h.

Convergence Tests

- Is simple to show that, if the FD scheme is converging, then should find:

$$\lim_{h \to 0} Q(t) = 4. \tag{74}$$

- In practice, can use additional levels of discretization, $8h$, $16h$, etc. to extend this test to look for "trends" in $Q(t)$ and, in short, to convince oneself (and, with luck, others), that the FDA really *is* converging.

- Additionally, once convergence of an FDA has been established, then point-wise subtraction of any two solutions computed at different resolutions, immediately provides an estimate of the level of error in both.

- For example, if one has u^h and u^{2h}, then, again by the Richardson *ansatz* have

$$\begin{aligned} u^{2h} - u^h &= \left(\left(u + (2h)^2 e_2 + \cdots \right) - \left(u + h^2 e_2 + \cdots \right) \right) && (75) \\ &= 3h^2 e_2 + O(h^4) \sim 3e^h \sim \frac{3}{4} e^{2h} && (76) \end{aligned}$$

Richardson Extrapolation

- *Richardson extrapolation:* Richardson's observation (65) also provides the basis for all the techniques of *Richardson extrapolation*

- Solutions computed at different resolutions are linearly combined so as to *eliminate* leading order error terms, providing more accurate solutions.

- As an example, given u^h and u^{2h} which satisfy (65), can take the linear combination

$$\bar{u}^h \equiv \frac{4u^h - u^{2h}}{3} \tag{77}$$

which, by (65), is easily seen to be $O(h^4)$, i.e. *fourth*-order accurate!

$$\begin{aligned} \bar{u}^h &\equiv \frac{4u^h - u^{2h}}{3} = \frac{4\left(u + h^2 e_2 + h^4 e_4 + \cdots\right) - \left(u + 4h^2 e_2 + 16h^4 e_4 + \cdots\right)}{3} \\ &= -4h^4 e_4 + O(h^6) = O(h^4) \end{aligned} \tag{78}$$

Richardson Extrapolation

- When it works, Richardson extrapolation has an almost magical quality about it
- However, generally have to start with fairly accurate (on the order of a few %) solutions in order to see the dramatic improvement in accuracy suggested by (78).
- Still a struggle to achieve that sort of accuracy (i.e. a few %) for *any* computation in many areas of numerical relativity/astrophysics, techniques based on Richardson extrapolation have not had a major impact in this context.

Independent Residual Evaluation

- Question that often arises in convergence testing: is the following:

  "OK, you've established that u^h is converging as $h \to 0$, but how do you know you're converging to u, the solution of the continuum problem?"

- Here, notion of an independent residual evaluation is very useful.

- Idea is as follows: have continuum PDE

$$Lu - f = 0 \tag{79}$$

  and FDA

$$L^h u^h - f^h = 0 \tag{80}$$

- Assume that u^h is apparently converging from, for example, computation of convergence factor (72) that looks like it tends to 4 as h tends to 0.

- However, do not know if we have derived and/or implemented our discrete operator L^h correctly.

Independent Residual Evaluation

- Note that implicit in the "implementation" is the fact that, particularly for multi-dimensional and/or implicit and/or multi-component FDAs, considerable "work" (i.e. analysis and coding) may be involved in setting up and solving the algebraic equations for u^h.

- As a check that solution *is* converging to u, consider a *distinct* (i.e. independent) discretization of the PDE:

$$\tilde{L}^h \tilde{u}^h - f^h = 0 \tag{81}$$

- Only thing needed from this FDA for the purposes of the independent residual test is the new FD operator $\tilde{L}^h$.

- As with L^h, can expand $\tilde{L}^h$ in powers of the mesh spacing:

$$\tilde{L}^h = L + h^2 E_2 + h^4 E_4 + \cdots \tag{82}$$

where E_2, E_4, $\cdots$ are higher order (involve higher order derivatives than L) differential operators.

Independent Residual Evaluation

- Now simply apply the new operator $\tilde{L}^h$ to our FDA u^h and investigate what happens as $h \to 0$.

- If u^h *is* converging to the continuum solution, u, will have

$$u^h = u + h^2 e_2 + O(h^4) \tag{83}$$

and will compute

$$\begin{aligned} \tilde{L}^h u^h &= \left(L + h^2 E_2 + O(h^4)\right)\left(u + h^2 e_2 + O(h^4)\right) && (84)\\ &= Lu + h^2(E_2\, u + L\, e_2) && (85)\\ &= O(h^2) && (86) \end{aligned}$$

- That is $\tilde{L}^h u^h$ will be a residual-like quantity that converges quadratically as $h \to 0$.

Independent Residual Evaluation

- Conversely, assume there is a problem in the derivation and/or implementation of $L^h u^h = f^h = 0$, but there is still convergence; i.e. for example,

$$u^{2h} - u^h \to 0 \quad \text{as} \quad h \to 0 \tag{87}$$

- Then must have something like

$$u^h = u + e_0 + h e_1 + h^2 e_2 + \cdots \tag{88}$$

  where crucial fact is that the error must have an $O(1)$ component, e_0.

- In this case, will compute

$$\begin{aligned} \tilde{L}^h u^h &= \left(L + h^2 E_2 + O(h^4)\right)\left(u + e_0 + h e_1 + h^2 e_2 + O(h^4)\right) \\ &= Lu + Le_0 + hLe_1 + O(h^2) \\ &= Le_0 + O(h) \end{aligned}$$

- Unless one is *extraordinarily* (un) lucky, and $L\,e_0$ vanishes, will *not* observe the expected convergence

Independent Residual Evaluation

- Instead, will see $\tilde{L}^h u^h - f^h$ tending to a *finite* ($O(1)$) value—a sure sign that something is wrong.
- Possible problem: might have slipped up in our implementation of the "independent residual evaluator", $\tilde{L}^h$
- In this case, results from test will be ambigous at best!
- However, a key point here is that because $\tilde{L}^h$ is only used *a posteriori* on a computed solution (never used to compute $\tilde{u}^h$, for example) it is a relatively easy matter to ensure that $\tilde{L}^h$ has been implemented in an error-free fashion (perhaps using symbolic manipulation facilities).
- Also, many of the restrictions commonly placed on the "real" discretization (such as stability and the ease of solution of the resulting algebraic equations) do not apply to $\tilde{L}^h$.
- Finally, note that although have assumed in the above that L, L^h and $\tilde{L}^h$ are *linear*, the technique of independent residual evaluation works equally well for non-linear problems.

Dispersion and Dissipation in FDAs

- Again consider the advection model problem, $u_t = a\,u_x$, but now discretize only in space (semi-discretization) using the usual $O(h^2)$ centred difference approximation:

$$u_t = a\,D_x\,u \equiv a\,\frac{u_{j+1} - u_{j-1}}{2\,\triangle x} \tag{89}$$

- Look for normal-mode solutions to (89) of the form

$$u = e^{ik\,(x + a'\,t)}$$

  where the "discrete phase speed", a', is to be determined.

- Substitution of this *ansatz* in (89) yields

$$ika'u = \frac{a\,(2i\sin(k\,\triangle x\,))}{2\,\triangle x}u$$

Dispersion and Dissipation in FDAs

- Solving for the discrete phase speed, a', find

$$a' = a\frac{\sin(k\,\triangle x\,)}{k\,\triangle x} = a\frac{\sin\xi}{\xi}$$

where have defined the dimensionless wave number, ξ:

$$\xi \equiv k\,\triangle x$$

- In *low frequency* limit, $\xi \to 0$, have expected result:

$$a' = a\frac{\sin\xi}{\xi} \to a$$

so that low frequency components propagate with the correct phase speed, a.

Dispersion and Dissipation in FDAs

- However, in *high frequency* limit, $\xi \to \pi$, have

$$a' = a\frac{\sin\xi}{\xi} \to 0 \quad !!$$

- Highest frequency components of the solution don't propagate at all!
- This is typical of FDAs of wave equations, particularly for relatively low-order schemes: propagation of high frequency components of the difference solution is essentially completely wrong.
- Arguably then, can be little harm in attenuating (dissipating) these components
- In fact, since high frequency components are potentially troublesome (particularly *vis a vis* non-linearities and the treatment of boundaries), is often advantageous to use a *dissipative* difference scheme.

Dispersion and Dissipation in FDAs

- Some FDAs are naturally dissipative (Lax-Wendroff scheme, for example), while others, such as leap-frog, are not.
- For leap-frog-based scheme, one idea is to add dissipative terms to the method, but in such a way as to retain $O(h^2)$ accuracy of the scheme.
- Consider leap-frog scheme as applied to the advection model problem:

$$u_j^{n+1} = u_j^{n-1} + a\lambda\left(u_{j+1}^n - u_{j-1}^n\right)$$

- Add dissipation to the scheme by modifying it as follows:

$$u_j^{n+1} = u_j^{n-1} + a\lambda\left(u_{j+1}^n - u_{j-1}^n\right) - \frac{\epsilon}{16}\left(u_{j+2}^{n-1} - 4u_{j+1}^{n-1} + 6u_j^{n-1} - 4u_{j-1}^{n-1} + u_{j-2}^{n-1}\right)$$

where ϵ is an adjustable, non-negative parameter.

Dispersion and Dissipation in FDAs

- Note that

$$\begin{aligned} u_{j+2}^{n-1} - 4u_{j+1}^{n-1} + 6u_j^{n-1} - 4u_{j-1}^{n-1} + u_{j-2}^{n-1} &= \triangle x^4 (u_{xxxx})_j^{n-1} + O(h^6) \\ &= \triangle x^4 (u_{xxxx})_j^{n} + O(h^5) = O(h^4) \end{aligned}$$

- Thus, added term does not change leading order truncation error, which is is $O(\triangle t^3) = O(h^3)$ per step

- Von Neumann analysis of modified scheme shows that, in addtion to the CFL condition $\lambda \leq 1$, must have $\epsilon < 1$ for stability, and, further, that the per-step amplification factor for a mode with wave number ξ is, to leading order

$$1 - \epsilon \sin^4 \frac{\xi}{2}$$

- Thus the addition of the dissipation term is analagous to the use of an explicit "high frequency filter" (low-pass filter), which has a fairly sharp rollover as $\xi \rightarrow \pi$.

Dispersion and Dissipation in FDAs

- Advantage to the use of explicit dissipation techniques (versus, for example, the use of an intrinsically dissipative scheme): amount of dissipation can be controlled by tuning the dissipation parameter.