Chap 3: Finite difference methods for PDEs

3.1 Basic notions of discretization

While we are interested in the continuum limit of PDEs, discretization of some kind is required for most numerical approaches

$$
\begin{array}{lccc}
 & X_0 & & Y_0 \\
 & \cup & & \cup \\
L: & X & & Y \\
 & (u & \longmapsto & Lu = S) \\
 & \downarrow D_h^X & & \downarrow D_h^Y \\
L_h: & X_h & \longrightarrow & Y_h
\end{array}
$$

continuum problem: $\boxed{Lu = S}$ continuum PDE

typically: $X_0 = C^0(\Omega)$, $\Omega \subseteq \mathbb{R}^n$

$$X = \left\{ u \in C^2(\Omega) \,\middle|\, \sup_{\Omega} |Lu| < \infty \right\} \subset X_0$$

regularity depends on differential operator L and the type of solution to be found

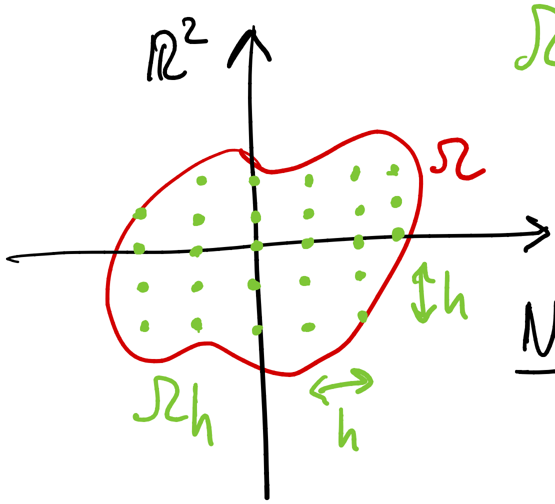
$$Y \subseteq Y_0 = C^0(\Omega)$$

Examples: $L = \Delta, \Box, \partial_t - \partial_x f(\cdot), \ldots$

discretized problem: $\boxed{L_h u_h = S_h}$ discretized PDE

typically: $X_h =$ "$\{$grid functions$\}$"

$$= \{ v : \Omega_h \to \mathbb{R}^n \}$$



$\Omega_h =$ "grid", e.g.: $\Omega \cap h\mathbb{Z}^n$

$h =$ discretization parameter

Notation:

$x \to x_i \in \{x_1, \ldots, x_{N_x}\}$

$t \to t_n \in \{t_1, \ldots, t_{N_t}\}$

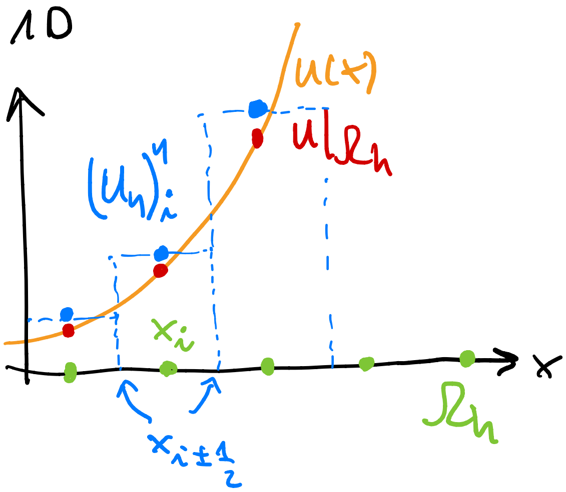
$v \in X_h : v_i^n \equiv v(x_i, t_n)$

D_h^X : discretization operator

$$D_h^X : X \to X_h$$

$$u \mapsto D_h^X u \equiv u_h$$

$$D_h^X u = u_h \quad \underline{\text{for example}} \quad \begin{cases} u|_{\Omega_h}, \text{ i.e. } (u_h)_i^n = u(x_i, t_n) & \text{finite difference methods} \\ (u_h)_i^n = \underbrace{\frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u(x, t_n)\,dx}_{\text{spatial average across } \Delta x} & \text{finite volume methods} \end{cases}$$



Notation: $x_{i\pm\frac{1}{2}} = x_i \pm \frac{\Delta x}{2}$

Example: $\Omega = \mathbb{R} \times (0, \infty)$, $h = \Delta x = \Delta t$

$$\Omega_h = \{(hi, hn) \mid i \in \mathbb{Z}, n \in \mathbb{N}\}$$

Notation: $(x_i, t_n) \equiv (hi, hn)$

$$v_i^n \equiv v(x_i, t_n), \quad v \in X_h$$

3.2 Finite difference approximations

Discretize continuum PDE by discrete grid:

- finite differences: point values at grid points (or staggered)
- finite volume: discrete values $\hat{=}$ averages over finite volumes

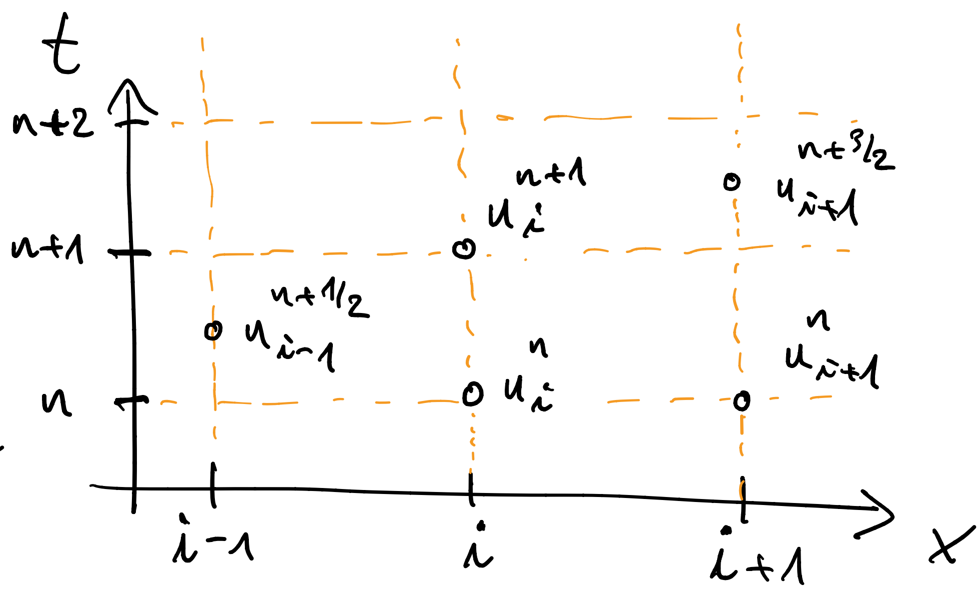
(→ later)

Consider $u = u(x,t)$ Grid $\Omega_h = \{(i\Delta x, n\Delta t) \mid i,n \in \mathbb{Z}\}$

<u>Notation:</u>

$u(x_i, t^n) = u_i^n$

$u\left(x_i + \frac{\Delta x}{2}, t^n\right) = u^n_{i+\frac{1}{2}}$



Note: in a finite difference approximation (FDA) of a PDE, different functions can be defined at different grid locations in which case the grid is called "staggered"

→ "staggered grid function"

e.g. u defined at $i + \frac{\Delta x}{2}$ → $u^n_{i+\frac{1}{2}}$

3.2.1 Partial derivatives / differential operators

Consider Taylor expansion:

$$u^n_{i+1} = u^n_i + \left.\frac{\partial u}{\partial x}\right|_i \Delta x + \frac{1}{2}\left.\frac{\partial^2 u}{\partial x^2}\right|_i (\Delta x)^2 + O((\Delta x)^3)$$

$$\left[u(x_i + \Delta x) = u(x_i) + \sum_{k=1}^{n} \frac{(\Delta x)^k}{k!} u^{(k)}(x_i) + O((\Delta x)^{n+1}) \right]$$

solve for $\frac{\partial u}{\partial x}$:

$$\left.\frac{\partial u}{\partial x}\right|_j = \frac{u^n_{i+1} - u_i}{\Delta x} - \frac{1}{2}\left.\frac{\partial u^2}{\partial x^2}\right|_i \Delta x + O((\Delta x)^2)$$

(*)

$$\rightsquigarrow \quad \boxed{D_1^+ u = \frac{u_{i+1}^n - u_i^n}{\Delta x}}$$

"forward-difference approximation"

"stencil size": max distance to neighboring grid points involved

accuracy / truncation error:

$$D_1^+ u - \frac{\partial u}{\partial x} \overset{(*)}{=} \frac{1}{2} \left.\frac{\partial u^2}{\partial x^2}\right|_i \Delta x + \mathcal{O}((\Delta x)^2) = \mathcal{O}(\Delta x)$$

$\rightsquigarrow$ first-order accurate

Similarly:

$$\boxed{D_1^- u = \frac{u_i^n - u_{i-1}^n}{\Delta x}}$$

"backward-difference approximation"

Improve accuracy:

$$\Delta x_i^+ \equiv x_{i+1} - x_i, \qquad \Delta x_i^- \equiv x_i - x_{i-1}$$

(can differ in general)

$$\rightsquigarrow \quad \Delta x_i \equiv \frac{1}{2}(\Delta x_i^+ + \Delta x_i^-)$$

forward: $u_{i+1}^n = u_i^n + \left.\frac{\partial u}{\partial x}\right|_i \Delta x_i^+ + \frac{1}{2}\left.\frac{\partial^2 u}{\partial x^2}\right|_i (\Delta x_i^+)^2 + \ldots$

backward: $u_{i-1}^n = u_i^n + \left.\frac{\partial u}{\partial x}\right|_i (-\Delta x_i^-) + \frac{1}{2}\left.\frac{\partial^2 u}{\partial x^2}\right|_i (\Delta x_i^-)^2 + \ldots$

$\downarrow$ substract

$$u_{i+1}^n - u_{i-1}^n = \left.\frac{\partial u}{\partial x}\right|_i (\Delta x_i^+ + \Delta x_i^-) + \frac{1}{2}\left.\frac{\partial^2 u}{\partial x^2}\right|_i \left[(\Delta x_i^+)^2 - (\Delta x_i^-)^2\right] + \ldots$$

$$\rightsquigarrow \left.\frac{\partial u}{\partial x}\right|_i = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x_i} - \frac{1}{2}\left.\frac{\partial^2 u}{\partial x^2}\right|_i \frac{(\Delta x_i^+)^2 - (\Delta x_i^-)^2}{2\Delta x_i} + \mathcal{O}\left((\Delta x_i)^2\right)$$

On a <u>uniform grid</u>:

$$\boxed{D_1^0 u = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} + \mathcal{O}((\Delta x)^2)}$$

"<u>centered difference approximation</u>"

<u>Remark</u>: By virtue of Taylor expansion we assume that the function u is continuously differentiable. Note that the limit

$$\lim_{\Delta x \to 0} \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}$$

can exist even if u is not differentiable
(e.g. $u(x,t) = |x| \leadsto \lim_{\Delta x \to 0} D_1^0 u = 0$)
We shall explore discontinuous solutions later.

Higher-order derivatives:

intuitively: repeatedly apply 1st order derivatives

Examples:

- $$\frac{\partial^2 u}{\partial x^2} \approx D_1^+ D_1^- u = \frac{1}{\Delta x}\left(D_1^+ u_i^n - D_1^+ u_{i-1}^n\right)$$
$$= \frac{1}{\Delta x}\left(\frac{u_{i+1}^n - u_i^n}{\Delta x} - \frac{u_i^n - u_{i-1}^n}{\Delta x}\right)$$
$$= \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2}$$
$$\equiv D^2 u$$

- $$\frac{\partial^2 u}{\partial x^2} \approx D_1^- D_1^+ u = \dots = D^2 u$$

- $$\frac{\partial^2 u}{\partial x^2} \approx D_{1/2}^0 D_{1/2}^0 u = \frac{1}{\Delta x} D_0^{1/2}\left(u_{i+1/2}^n - u_{i-1/2}^n\right)$$

$$= \frac{1}{\Delta x}\left(\frac{u_{i+1}^n - u_i^n}{\Delta x} - \frac{u_i^n - u_{i-1}^n}{\Delta x} \right)$$

$$= D^2 u$$

<u>General approach:</u>

<u>goal:</u> compute approximation to $u^{(k)} = \frac{\partial^k u}{\partial x^k}$ based on a stencil of $n \geq k+1$ points $x_1, \dots, x_n$

<u>Example:</u> want one-sided approximation to $\left.\frac{\partial u}{\partial x}\right|_i$ based on $u_i^n, u_{i-1}^n, u_{i-2}^n$

$\rightsquigarrow$ Taylor expansion

$$\to \quad u_{i-1} = u_i - u^{(1)}\big|_i \Delta x + \frac{1}{2} u^{(2)}\big|_i (\Delta x)^2 - \frac{1}{6} u^{(3)}\big|_i (\Delta x)^3 + \mathcal{O}(\Delta x^4)$$

$$\to \quad u_{i-2} = u_i - u^{(1)}\big|_i (2\Delta x) + \frac{1}{2} u^{(2)}\big|_i (2\Delta x)^2 - \frac{1}{6} u^{(3)}\big|_i (2\Delta x)^3 + \mathcal{O}(\Delta x^4)$$

Ansatz:

$$D_2 u\big|_i = c_1 u_i + \underline{c_2 u_{i-1}} + \underline{c_3 u_{i-2}}$$

$$= c_1 u_i + c_2 u_i - c_2 u^{(1)} \Delta x + c_2 \frac{1}{2} u^{(2)} (\Delta x)^2$$

$$- c_2 \frac{1}{6} u^{(3)} (\Delta x)^3$$

$$+ c_3 u_i - c_3 u^{(1)} (2\Delta x) + c_3 \frac{1}{2} u^{(2)} (2\Delta x)^2$$

$$- c_3 \frac{1}{6} u^{(3)} (\Delta x)^3 \quad + \ldots$$

$$= (c_1 + c_2 + c_3)\, u_i - (c_2 + 2c_3)\, \Delta x\, u^{(1)}$$

$$+ \frac{1}{2}(c_2 + 4c_3)(\Delta x)^2 u^{(2)} - \frac{1}{6}(c_2 + 8c_3)(\Delta x)^3 u^{(3)}$$

$$+ \ldots$$

note: coefficients are of the form

$$\frac{1}{(\ell-1)!} \sum_{m=1}^{n} c_m (x_m - x_i)^{\ell-1}$$

Need to agree with $u^{(1)}\big|_i$ to highest order possible:

$$\rightsquigarrow \quad \left.\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ -(c_2 + c_3)\Delta x &= 1 \\ \frac{1}{2}(c_2 + 4c_3)(\Delta x)^2 &= 0 \end{aligned}\right\} \qquad \begin{aligned} c_1 &= \frac{3}{2\Delta x} \\ c_2 &= -\frac{2}{\Delta x} \\ c_3 &= \frac{1}{2\Delta x} \end{aligned}$$

(Note: higher coefficients 0 would lead to over-determined system)

$$\rightsquigarrow \quad \boxed{D_2 u\big|_i = \frac{1}{2\Delta x}\left(3u_i - 4u_{i-1} + u_{i-2}\right)}$$

Truncation error:

$$D_2 u|_i - u^{(1)}|_i = -\frac{1}{6}(c_2 + 8c_3)(\Delta x)^3 u^{(3)}|_i + \dots$$

$$= -\frac{1}{6}\left(-\frac{2}{\Delta x} + \frac{4}{\Delta x}\right)(\Delta x)^3 u^{(3)}|_i + \dots$$

$$= -\frac{1}{3}(\Delta x)^2 u^{(3)}|_i + \dots = \mathcal{O}(\Delta x^2)$$

$\rightsquigarrow$ 2nd order accurate

General procedure: Consider stencil $\{x_\ell\}_{\ell=1,\dots,n}$ around point $\bar{x} = x_i$ (may or may not be part of stencil), $\boxed{n \geq k+1}$. Consider n Taylor series $\ell = 1,\dots,n$:

$$(*) \quad \underline{u(x_\ell)} - \underbrace{\underline{u(x_i)}}_{u_i^n} = u^{(1)}|_i \underline{(x_\ell - x_i)} + \dots + \frac{1}{k!}\underline{(x_\ell - x_i)^k} u^{(k)}|_i + \mathcal{O}(h^p)$$

known

$\rightsquigarrow$ linear system of n equations in k unknowns

Ansatz: $u^{(k)}|_i = c_1 u(x_1) + \dots + c_n u(x_n) + \mathcal{O}(\Delta x^p)$

know that we can choose (see above)

$$(**) \quad \frac{1}{(\ell-1)!} \sum_{m=1}^{n} c_m (x_m - x_i)^{\ell-1} = \begin{cases} 1, & \ell - 1 = k \\ 0 & \text{otherwise} \end{cases}$$

$\ell = 1,\dots,n$

If points $\{x_\ell\}_{\ell=1,\dots,n}$ are distinct, $(**)$ is an $n\times n$ non-singular system and thus has a unique solution. Can write:

$$\underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ & (a_{\ell m})_{\substack{\ell=2,\dots,n\\ m=1,\dots,n}} = \frac{1}{(\ell-1)!}(x_m-x_i)^{\ell-1} & \end{pmatrix}}_{\equiv A\,\in\,\mathrm{Mat}(n\times n)} \underbrace{\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}}_{c} = \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{b} \begin{array}{l} \left.\begin{array}{l} \\ \\ \end{array}\right\} \text{cancellation of lower order terms} \\ \leftarrow k+1 \\ \left.\begin{array}{l} \\ \\ \end{array}\right\} \text{cancellation of higher order terms} \end{array}$$

<u>Remarks:</u> 1) if $n \leq k$ there are too few points in the stencil, then RHS b and solution c both zero.

2) Accuracy: RHS $\rightsquigarrow$ in Taylor expansions $(*)$:

$$\left(\sum_{m=1}^{n} c_m (x_m-x_i)^{\ell-1}\right) u^{(\ell-1)}\Big|_i \overset{\ell-1\neq k}{=} 0$$

$\nearrow$ $\underline{\ell-1<k}$: necessary to cancel lower order terms and get $O(\Delta x)$ accuracy

$\searrow$ $\underline{\ell-1>k}$:

cancel higher
order terms and
obtain higher than
1st order accuracy

⇝ procedure is at least

$\mathcal{O}((\Delta x)^p)$, $p = n-k$ accurate

$p > n-k$ achievable if additional higher order terms cancel (e.g. centered differences)

$\boxed{\text{General: increasing stencil size increases accuracy}}$

3.2.2 Sample Discretizations

1) 1D wave equation, standard $O(\Delta x^2)$

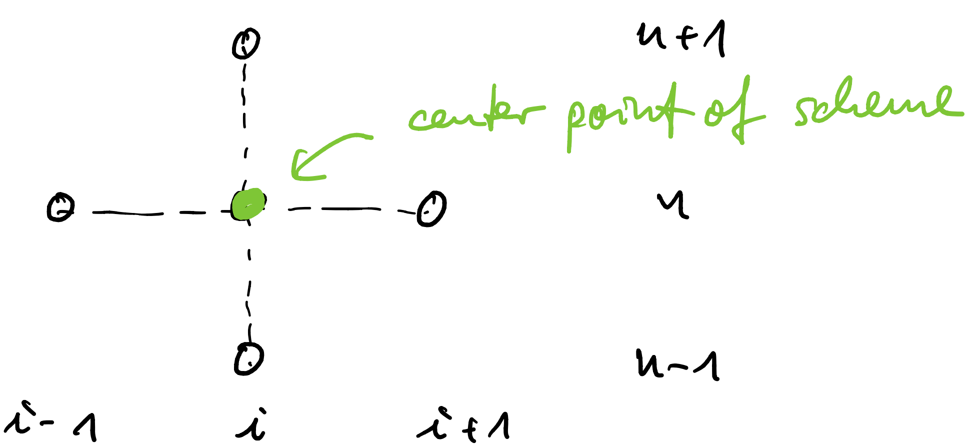
$$\text{IBVP}\begin{cases} u_{tt} - c^2 u_{xx} = 0, \quad 0 \le x \le 1, \ t \ge 0 \\ \left.\begin{aligned} u(x,0) &= u_0(x) \\ u_t(x,0) &= v_0(x) \end{aligned}\right\} \text{ initial data} \\ u(0,t) = u(1,t) = 0 \} \text{ fixed (Dirichlet) boundary conditions} \end{cases}$$

Consider standard centered $O(h^2)$ discretization in space & time

$$\left.\frac{\partial^2 u}{\partial t^2}\right|_i^n = \frac{u_i^{n+1} - 2u_i^n + u_i^{n-1}}{(\Delta t)^2} + O(\Delta t^2)$$

$$\left.\frac{\partial^2 u}{\partial x^2}\right|_i^n = \frac{u_{i+1}^n - 2u_i + u_{i-1}^n}{(\Delta x)^2} + O(\Delta x^2)$$

stencil:



$\leadsto$ FDA approximation to $\mathcal{O}(\Delta x^2, \Delta t^2)$:

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

$$\left.\begin{aligned} u_i^0 &= u(x_i, 0) \\ u_i^1 &= v(x_i, 0)\Delta t \end{aligned}\right\} \text{discretized initial data}$$

$$\underbrace{u_1^n}_{x=0} = \underbrace{u_I^n}_{x=1} = 0 \quad \Big\} \text{ discretized boundary condition}$$

explicitly solve for u^{n+1}:

$$u_i^{n+1} = 2u_i^n - u_i^{n-1} + \bigg(\underbrace{\frac{c\Delta t}{\Delta x}}_{\equiv\lambda}\bigg)^2 \left(u_{i+1}^n - 2u_i^n + u_{i-1}^n\right)$$

$\leadsto$ linear system for unknowns $\{u_i^{n+1}\}_{i=1,\dots,I}$

write: $A\, u^{n+1} = b$

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} u_1^{n+1} \\ \vdots \\ u_I^{n+1} \end{pmatrix} = \begin{pmatrix} 2u_1^n - u_1^{n+1} + \lambda^2(\;) \\ \vdots \\ 2u_I^n - u_I^{n+1} + \lambda^2(\;) \end{pmatrix}$$

accuracy: $\mathcal{O}(\Delta x^2, \Delta t^2)$

A: diagonal $\leadsto$ explicit scheme
(will see: stable if $\lambda = \frac{c\Delta t}{\Delta x} \leq 1$ CFL condition)

2) 1D diffusion equation, Crank–Nicholson

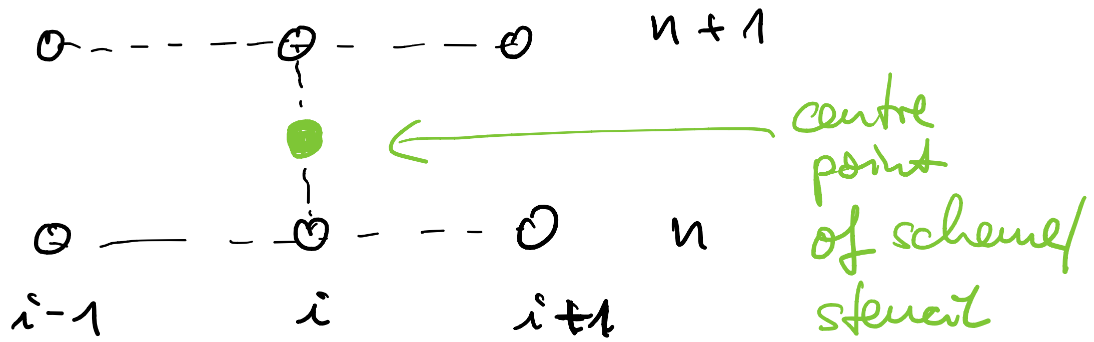
$$\text{IBVP} \quad \begin{cases} u_t - \sigma u_{xx} = 0, & 0 \leq x \leq 1, \ t \geq 0 \\ u(x,0) = u_0(x) \\ u(0,t) = u(1,t) = 0 \end{cases}$$

Consider: centered in time & space
(minimizes truncation error, minimizes instabilities)

$$u_t \Big|_i^{n+\frac{1}{2}} = \frac{u_i^{n+1} - u_i^n}{\Delta t} + \mathcal{O}(\Delta t^2)$$

$$u_{xx} \Big|_i^n = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2} + \mathcal{O}(\Delta x^2)$$

stencil



need to apply time averaging on RHS:

$$T^{n+\frac{1}{2}}(v) = \frac{1}{2}\left(v^{n+1} + v^{n}\right) = v^{n+\frac{1}{2}} + O(\Delta t^2)$$

$\rightsquigarrow$ FDA approximation to $O((\Delta x)^2, (\Delta t)^2)$:

at $t: n+\frac{1}{2}$

$$\frac{u_i^{n+1} - u_i^{n}}{\Delta t} = \sigma T^{n+\frac{1}{2}}\left(u_{xx}\big|_i^n\right) \qquad i = 2, \dots, I$$

$$= \sigma \frac{1}{2}\left[\frac{u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}}{(\Delta x)^2} + \frac{u_{i+1}^{n} - 2u_i^{n} + u_{i-1}^{n}}{(\Delta x)^2}\right]$$

$\rightsquigarrow$ linear system for $\{u_i^{n+1}\}_{i=2,\dots,I-1}$

write as

$$A u^{n+1} \equiv a_+ u_{i+1}^{n+1} + a_0 u_i^{n+1} + a_- u_{i-1}^{n+1} = b, \quad i = 2, \dots, I-1$$

$$\begin{pmatrix} \frac{1}{\Delta t^2} + \frac{1}{\Delta x^2} & -\frac{1}{\Delta x^2} & & 0 \\ -\frac{1}{\Delta x^2} & \ddots & \ddots & \\ & \ddots & \ddots & -\frac{1}{\Delta x^2} \\ 0 & & -\frac{1}{\Delta x^2} & \frac{1}{\Delta t^2} + \frac{1}{\Delta x^2} \end{pmatrix} \begin{pmatrix} u_2^{n+1} \\ \vdots \\ \vdots \\ u_{I-1}^{n+1} \end{pmatrix} = \begin{pmatrix} \left(\frac{1}{\Delta t^2} + \frac{1}{\Delta x^2}\right) u_i^n + \\ \frac{1}{2(\Delta x)^2}\left(u_{i+1}^n + u_{i-1}^n\right) \end{pmatrix}$$

$\leadsto$ tridiagonal matrix $\leadsto$ implicit scheme, couples unknowns u_{i+1}, u_i, u_{i-1} at advanced time level $n+1$!

accuracy: $\mathcal{O}(\Delta x^2, \Delta t^2)$

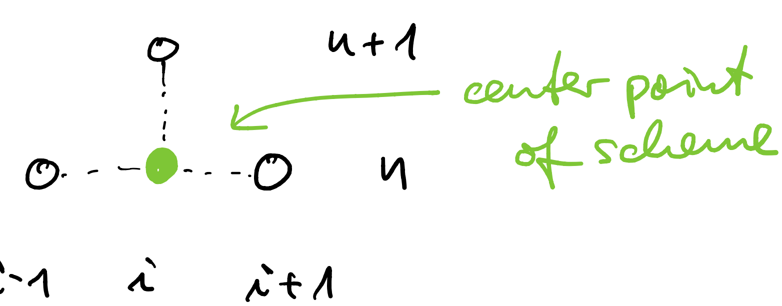
3) 1D advection equation

$$\text{IBVP} \quad \begin{cases} u_t + a u_x = 0 & 0 \le x \le 1, \ t \ge 0 \\ u(x,0) = u_0(x) \\ u(0,t) = u(1,t) = 0 \end{cases}$$

(i) Consider simple forward in time, centered in space FDA:

$$u_x\Big|_i^n \approx \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}, \quad u_t\Big|_i^n = \frac{u_i^{n+1} - u_i^n}{\Delta t}$$

stencil



$$\leadsto \frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{a}{2\Delta x}\left(u_{i+1}^n - u_{i-1}^n\right) = 0$$

$$\Leftrightarrow \quad u_i^{n+1} = u_i^n - \underbrace{\frac{a\Delta t}{2\Delta x}}_{\equiv \lambda}\left(u_{i+1}^n - u_{i-1}^n\right)$$

linear system for u^{n+1}: $\quad A u^{n+1} = b$

$$\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \begin{pmatrix} u_1^{n+1} \\ \vdots \\ u_I^{n+1} \end{pmatrix} = \begin{pmatrix} u_i^n - \lambda\left(u_{i+1}^n - u_{i-1}^n\right) \end{pmatrix}$$

$\rightsquigarrow$ explicit scheme $\qquad \mathcal{O}(\Delta t, \Delta x^2)$

(will see: unstable for any fixed $\frac{\Delta t}{\Delta x}$)

(ii) Lax – Friedrichs method

replace $u_i^n \to \frac{1}{2}\left(u_{i-1}^n + u_{i+1}^n\right)$

$$\rightsquigarrow \quad u_i^{n+1} = \frac{1}{2}\left(u_{i+1}^n + u_{i-1}^n\right) - \frac{a\Delta t}{2\Delta x}\left(u_{i+1}^n - u_{i-1}^n\right)$$

accuracy: $\mathcal{O}(\Delta t, \Delta x^2)$

(will see: stable if $\left|\frac{a\Delta t}{\Delta x}\right| \leq 1$)

3.3 Consistency, stability, convergence

General problem: Does the solution to the discrete problem provide an approximation to the continuous problem?

$\|u-u_h\| \xrightarrow{h\to 0} 0$ for some appropriate norm $\|\cdot\|$?

⇝ need to understand how u_h varies with h

⇝ need basic concepts of numerical analysis

Problem setting: Consider numerical scheme of the form

$$\boxed{u^{n+1} = G(S_+, S_-)\, u^n}$$

column vector containing sufficient unknowns to write problem in 1st order in time form

update operator, polynomial in S_+, S_-, not necessarily linear

$S_+ u_i = u_{i+1}$

$S_- u_i = u_{i-1}$

Note: a large class of problems can be written in this form

Example 1: $u_t + a u_x = 0$

with $u_i^{n+1} = u_i^n - a\frac{\Delta t}{\Delta x}(u_i^n - u_{i-1}^n)$

$\leadsto G(S_+, S_-) = \left(1 - a\frac{\Delta t}{\Delta x}\right)\mathbb{I} + \frac{a\Delta t}{\Delta x} S_-$

Example 2: rewriting <u>multi-level schemes</u>

$$u_{tt} - c^2 u_{xx} = 0$$

with: $u_i^{n+1} = 2u_i^n - u_i^{n-1} + \lambda^2(u_{i+1}^n - 2u_i^n + u_{i-1}^n)$

(3-level scheme) $\quad \lambda = \frac{c\Delta t}{\Delta x}$

$\leadsto$ define auxiliary variables

$$u_i^{n+1} = 2u_i^n - v_i^n + \lambda^2(\dots)$$

$$v_i^{n+1} = u_i^n$$

$\leadsto u^n \equiv (u_1^n, v_1^n, \dots, u_I^n, v_I^n)$

Example 3: <u>Implicit schemes</u> are of the form $L(S_+, S_-)u^{n+1} = Q u^n$

$\leadsto$ consider: $u^{n+1} = L^{-1} Q u^n \equiv G u^n$

Definition (Stability): Consider above setting, let $T > 0$ be fixed. The numerical scheme is called <u>stable</u> wrt. the norm $\|\cdot\|$ if there exists constants $c(T), \beta, \tau$ such that

$$\|u^n\| \leq c(T)\, e^{\beta \frac{n\Delta t}{T}} \|u^0\| \qquad \forall\ 0 \leq \Delta t < \tau,\ n \leq \frac{T}{\Delta t}$$

Definition (truncation error, consistency):

Consider above setting. The numerical scheme is said to be <u>consistent</u> of order (q,p) wrt. the norm $\|\cdot\|$ if for any exact solution v of the PDE

$$\|\tau^n\| \equiv \frac{1}{\Delta t}\left\|v(\cdot, t^{n+1}) - G\, v(\cdot, t^n)\right\| = \left[\mathcal{O}(\Delta t^q) + \mathcal{O}(\Delta x^p)\right]$$

τ^n: "<u>local truncation error</u>"

Definition (convergence): Consider above setting. The numerical scheme is <u>convergent</u> of order (q,p) wrt. to the norm $\|\cdot\|$ if for any exact solution v of the PDE

$$\|E^n\| \equiv \|v(\cdot,t^n)-u^n\| = \mathcal{O}(\Delta t^q) + \mathcal{O}(\Delta x^p)$$

uniformly for all $n \in \mathbb{N}$.

$E^n = v^n - u^n$: "global error"

Theorem (Lax theorem):

(Lax & Richtmyer 1956, Comm. Pure Appl. Math. 9, 267)

Consider above setting and assume that G is linear in u^n. If the numerical scheme is stable and consistent of order (q,p) and the initial data are approximated consistently, i.e. $\|u^0 - v(\cdot,0)\| = \mathcal{O}(\Delta x^p)$, then the scheme is convergent of order (q,p), i.e.

$$\|u^n - v(\cdot,t^n)\| = \mathcal{O}(\Delta t^q) + \mathcal{O}(\Delta x^p)$$

uniformly $\forall\ n \leq \frac{T}{\Delta t}$ and v exact solution to the PDE.

Remark: the above definition & the Lax theorem require the IVP to be well-posed, i.e. that a solution v exists and that it is unique. More on this later.

Proof: Let $t^n = n\Delta t \le T$.

stability: $\|u^n\| = \|G^n u^0\| \le c(T) e^{\beta \frac{n\Delta t}{T}} \|u^0\|$

$$\to \|G^n\| \le c(T) e^{\beta \frac{n\Delta t}{T}}$$

consistency: $v^n \equiv v(\cdot, t^n) = G v^{n-1} + \Delta t \tau^n$

with $\|\tau^n\| = \mathcal{O}(\Delta x^p) + \mathcal{O}(\Delta t^q)$

$\rightsquigarrow$ global error:

$$E^n = u^n - v^n = G u^{n-1} - G v^{n-1} - \Delta t \tau^n$$

$$\overset{\text{linearity}}{=} G E^{n-1} - \Delta t \tau^n$$

$$= G^2 E^{n-2} - \Delta t G \tau^{n-1} - \Delta t \tau^n$$

$$= G^n E^0 - \Delta t \sum_{j=0}^{n-1} (G)^j \tau^j$$

$\downarrow$ consistency of initial data approx. $\|E^0\| = \mathcal{O}(\Delta x^p)$

$$\rightsquigarrow \|E^n\| \leq \Delta t \sum_{j=0}^{n-1} \|(G)^j \tau^j\| + \mathcal{O}(\Delta x^p)$$

$\|\tau\| \equiv \max_{j \leq n-1} \|\tau^j\|$

$$\leq \Delta t \|\tau\| \sum_{j=0}^{n-1} c(T)\, e^{\beta \frac{j \Delta t}{T}} + \mathcal{O}(\Delta x^p)$$

$$\leq \|\tau\| \underbrace{\Delta t\, n}_{=T} c(T)\, e^{\beta \frac{n \Delta t}{T}} + \mathcal{O}(\Delta x^p)$$

$\hookrightarrow 1 + \left(\beta \frac{n \Delta t}{T}\right) + \frac{1}{2}(\ )^2 \ldots$

$$= \mathcal{O}(\Delta t^q) + \mathcal{O}(\Delta x^p)$$

$\square$

Remarks:

1) The converse is also true ($\rightsquigarrow$ "equivalence theorem"), but requires more work.
For practical purposes we will not need the converse statement.

2) For non-linear operators G the statement is false, in general.

3.4 Stability analysis & CFL cond.

Problem: how to tell a scheme converges?

most often don't know exact solution which is why a numerical approach is chosen in the first place!

$\rightsquigarrow$ use Lax theorem, consistency & stability easier to show, but need sufficient criterion for stability

Theorem (von Neumann stability condition):

A finite difference scheme

vector of unknowns $\leftarrow$ $$u^{n+1} = G(S_+, S_-)\, u^n$$

linear in u is stable w.r.t. to the L^2-norm if and only if there exist constants α, τ such that

$$|g(\xi)| \leq 1 + \alpha \Delta t$$

for all $\xi \in \mathbb{R}$ and $0 \leq \Delta t \leq \tau$. Here, $g(\xi)$ is the <u>amplification factor</u> ("symbol of G")

$\sqrt{-1}$

$$g(\xi) \equiv G(e^{-i\xi}, e^{i\xi}).$$

<u>Proof:</u> Consider discrete Fourier transform of $u \equiv (u_j)_{j \in \mathbb{N}}$:

$$\hat{u}(\xi) = \sum_j u_j e^{ij\xi} \qquad 0 \leq \xi \leq 2\pi$$

$i \equiv \sqrt{-1}$

Observe:

$$\widehat{S_+ u}(\xi) = \sum_j S_+ u_j \, e^{ij\xi} = \sum_j u_{j+1} e^{ij\xi}$$

$$\overset{\ell = j+1}{=} \sum_\ell u_\ell e^{i(\ell-1)\xi} = e^{-i\xi} \sum_j u_j e^{ij\xi}$$

$$= e^{-i\xi} \hat{u}(\xi)$$

$$\widehat{S_- u}(\xi) = e^{i\xi} \hat{u}(\xi)$$

$$\Rightarrow \hat{u}^{n+1} = G(e^{-i\xi}, e^{i\xi}) \hat{u}^n$$

$$= g(\xi) \hat{u}^n \qquad (*)$$

Discrete L^2 norms:

$$\|u^n\|_2^2 = \Delta x \sum_j (u_j^n)^2$$

periodicity

$$\|\hat{u}^n\|_2^2 = \int_0^{2\pi} |\hat{u}^n(\xi)|^2 \, d\xi$$

Parseval property:

$$\|\hat{u}\|_2^2 = \int_0^{2\pi} |\hat{u}(\xi)|^2 d\xi = \int_0^{2\pi} \Big|\sum_j u_j e^{ij\xi}\Big|^2 d\xi$$

$(e^{ij\xi})_{j\in\mathbb{N}}$ orthogonal

$$= \int_0^{2\pi} \sum_j |u_j e^{ij\xi}|^2 d\xi \overset{|e^{ij\xi}|=1}{=} \int_0^{2\pi} \sum_j u_j^2 \, d\xi$$

$$= 2\pi \sum_j u_j^2 = \frac{2\pi}{\Delta x} \|u\|_2^2$$

① assume $g(\xi) \leq 1 + \alpha \Delta t$:

$$\leadsto \|u^{n+1}\|_2^2 = \frac{\Delta x}{2\pi} \underbrace{\int_0^{2\pi} |\hat{u}^{n+1}(\xi)|^2 d\xi}_{\|\hat{u}^{n+1}\|_2^2}$$

$$\overset{(*)}{=} \frac{\Delta x}{2\pi} \int_0^{2\pi} |g(\xi)|^2 \, |\hat{u}^n(\xi)|^2 \, d\xi$$

$$\leq \frac{\Delta x}{2\pi} (1 + \alpha \Delta t)^2 \underbrace{\int_0^{2\pi} |\hat{u}^n(\xi)|^2 d\xi}_{\|\hat{u}^n\|_2^2}$$

$$= (1+\alpha\Delta t)^2 \, \|u^n\|_2^2$$

$$\leq (e^{\alpha\Delta t})^2 \, \|u^n\|_2^2$$

$$\Rightarrow \quad \|u^n\|_2^2 \leq (e^{\alpha\Delta t})^{2n} \, \|u^0\|_2^2$$

scheme is stable!

② the reverse: assume $|g(\xi)| \leq 1+\alpha\Delta t$ violated, show scheme cannot be stable

$\rightsquigarrow$ for any $\alpha > 0$, $\tau > 0$ $\exists \xi$ and $0 \leq \Delta t < \tau$ such that $|g(\xi)| > 1+\alpha\Delta t$ for a finite region $\xi \in I_\alpha = [\xi_{min}, \xi_{max}]$

Consider $u^n = G u^{n-1}$ for $1 \leq n \leq \frac{T}{\Delta t}$ and without loss of generality consider initial conditions that vanish outside I_α i.e. $\hat{u}^0 = 0$ on $\mathbb{R} \setminus I_\alpha$.

$$\rightsquigarrow |\hat{u}^n(\xi)| = |g^n(\xi)\,\hat{u}^0(\xi)| > (1+\alpha\Delta t)^n \, |\hat{u}^0(\xi)|$$

on I_α

Then:

$$\|G^n\| \, \|u^0\|_2 \geq \|G^n u^0\|_2 = \|u^n\|_2 \underset{\uparrow}{>} (1+\alpha\Delta t)^n \|u^0\|_2$$

Parseval property

$$\Rightarrow \|G^n\| \geq (1+\alpha\Delta t)^n$$

Now assume scheme is stable

$$\Rightarrow \|G^n\| \leq C_0 \quad \forall\ 1 \leq n \leq \frac{T}{\Delta t}$$

Thus for all $\alpha, \tau > 0$ $\exists\ \Delta t \in (0,\tau)$ such that

$$(1+\alpha\Delta t)^n \leq \|G^n\| \leq C_0 \quad \forall\ 0 \leq n \leq \frac{T}{\Delta t}$$

Now choose $n = \frac{T}{\Delta t}$

$$\leadsto \left(1 + \alpha \frac{T}{n}\right)^n \leq C_0$$

choose α sufficiently large such that $\frac{e^{\alpha T}}{2} > C_0$

choose τ sufficiently small such that

$$\left(1 + \frac{\alpha T}{n}\right)^n > \frac{2}{3} e^{\alpha T} \quad \text{for } n = \frac{T}{\Delta t},\ \Delta t \in (0,\tau)$$

(n sufficiently large)

$\left(\text{remember: } \lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n = e^x\right)$

Then:

$$\frac{1}{2}e^{\alpha T} > \left(1+\frac{\alpha T}{n}\right)^n \geq \frac{2}{3}e^{\alpha T}$$

contradiction.

$\square$

Remarks:

1) Note the theorem makes use of the fact that boundedness of the L^2-norm of u^n follows from boundedness of L^2-norm of $\hat{u}^n$ via Parseval's identity. The latter is much easier to show since while all $\{u^n_j\}_{j\in\mathbb{N}}$ are coupled via the difference equations, each $\hat{u}(\xi)$ satisfies (*) and is thus decoupled from all other wavenumbers ξ !

2) One can make the replacement

$$|g(\xi)| \leq e^{\alpha \Delta t}$$

and the theorem still holds.

Note: $|g(\xi)| \leq (1+\alpha\Delta t)^n \overset{n=\frac{T}{\Delta t}}{=} \left(1+\frac{\alpha T}{n}\right)^n$

$$\xrightarrow[]{\substack{n\to\infty \\ (\Delta t\to\infty)}} \quad e^{\alpha T} \qquad \text{(for } T>0 \text{ fixed)}$$

3) The stability bound allows for exponential growth of modes as, for example, required for

$$u_t + a u_x = u$$

but sometimes it is not appropriate, e.g., for

$$u_t + a u_x = \varepsilon u_{xx} \quad , \quad \varepsilon > 0$$

(all modes are damped)

Central in space, forward in time discretization leads to: $\quad |g| \le 1 + \frac{1}{2}\frac{a^2}{\varepsilon}\Delta t$

$\leadsto$ concept of "<u>strict / strong stability</u>":
if $|\hat{u}(\xi, t+\Delta t)| \le e^{\kappa \Delta t} |\hat{u}(\xi, t)| \quad \forall \xi$ (cont. problem),
require $g(\xi) \le e^{\kappa \Delta t} \quad \forall \xi$ (discrete problem)

4) If one wants $\|u^n\| \le c \|u^0\|$ (cf. 3))
for stability (cf. Def. in Sec. 3.3),
the requirement on $g(\xi)$ in the theorem sharpens to

$$|g(\xi)| \le 1 \quad \forall \xi$$

5) alternative estimation of amplification factor

1) $\Rightarrow$ it suffices to consider a arbitrary single wave number ξ

$$u_j^n = e^{i\xi x_j}$$

to evaluate amplification factor.

6) 3+-level schemes: either

(i) use method in 5) $\to$ e.g. leapfrog scheme

(ii) amplification matrix analysis:

convert multi-level scheme to effective 2-level scheme as discussed before

$$u^{n+1} = G\, u^n$$

all unknowns $\to$
$$\begin{pmatrix} u^{n+1} \\ v^{n+1} \end{pmatrix} = G \begin{pmatrix} u^n \\ v^n \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} \hat{u}^{n+1}(\xi) \\ \hat{v}^{n+1}(\xi) \end{pmatrix} = G(e^{-i\xi}, e^{i\xi}) \begin{pmatrix} \hat{u}^n(\xi) \\ \hat{v}^n(\xi) \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} g_{11}(\xi) & g_{12}(\xi) \\ g_{21}(\xi) & g_{22}(\xi) \end{pmatrix}}_{g(\xi) = \text{amplification matrix}} \begin{pmatrix} \hat{u}^n \\ \hat{v}^n \end{pmatrix}$$

Assume that $g(\xi)$ has a complete set of (orthonormal) <u>eigenvectors</u> $g_\ell(\xi)$ and corresponding <u>eigenvalues</u> $\lambda_\ell(\xi)$, and suppose there exist constants $\alpha, \tau > 0$ such that

$$\boxed{|\lambda_\ell(\xi)| \leq (1 + \alpha \Delta t), \quad \ell = 1, 2}$$

for all $\xi \in \mathbb{R}$ and $0 \leq \Delta t \leq \tau$.

Again bound $\|u^{n+1}\|_2$ by $\|\hat{u}^{n+1}\|_2$ using Parseval's property:

$$\leadsto \quad \hat{u}^n = g^n(\xi)\, \hat{u}^0$$

$$= g^n(\xi) \sum_{\ell=1}^{2} c_\ell^0\, g_\ell(\xi) \quad \text{here}$$

$$= \sum_\ell c_\ell^0\, \lambda_\ell^n\, g_\ell(\xi)$$

$$\rightsquigarrow \|u^n\|_2^2 = \frac{\Delta x}{2\pi} \underbrace{\int_0^{2\pi} |\hat{u}^n(\xi)|^2 \, d\xi}_{\|\hat{u}^n\|_2^2}$$

$$\overset{\text{see above}}{=} \frac{\Delta x}{2\pi} \int_0^{2\pi} \Big| \sum_{i=1}^{2} \lambda_i^n c_i^0 g_i(\xi) \Big|^2 d\xi$$

$$\leq \frac{\Delta x}{2\pi} (1+\alpha \Delta t)^{2n} \overbrace{\int_0^{2\pi} |\hat{u}^0(\xi)|^2 d\xi}^{\|\hat{u}^0(\xi)\|_2^2}$$

$$= (1+\alpha\Delta t)^{2n} \|u^0\|_2^2$$

$$\leq (e^{\alpha \Delta t})^{2n} \|u^0\|_2^2 \quad \text{stable !}$$

Example: stability of advection equation

with scheme:

$$u_j^{n+1} = u_j^n - a\lambda\left[\theta u_{j+1}^n + (1-2\theta)u_j^n - (1-\theta)u_{j-1}^n\right]$$

$$= \left[(1-a\lambda(1-2\theta))I - a\lambda\theta S_+ + a\lambda(1-\theta)S_-\right] u_j^n$$

$$\equiv G(S_+, S_-)\, u_j^n, \qquad \lambda = \frac{\Delta t}{\Delta x}$$

$\theta \in [0,1]$, where

$$\left.\begin{array}{ll} \theta = 0: & \underline{\text{backward}} \\ \theta = \frac{1}{2}: & \underline{\text{centered}} \\ \theta = 1: & \underline{\text{forward}} \end{array}\right\} \text{difference in space}$$

$$\leadsto g(\xi) = Q(e^{-i\xi}, e^{i\xi})$$

$$= 1 - a\lambda(1-2\theta) - a\lambda\theta e^{-i\xi} + a\lambda(1-\theta)e^{i\xi}$$

$$= \left[1 - a\lambda(1-2\theta) + a\lambda(1-2\theta)\cos\xi\right]$$

$$+ i(a\lambda\sin\xi)$$

$$\Rightarrow |g(\xi)|^2 = 1 - 2a\lambda(1-\cos\xi)\left\{1 - 2\theta - \lambda a\left[1 + 2\theta(\theta-1) \times (1-\cos\xi)\right]\right\}$$

Assume $a > 0$.

$\rightsquigarrow$ $\underline{\theta = 1:}$

$$|g(\xi)|^2 = 1 - 2a\lambda(1-\cos\xi)\{-1-\lambda a\}$$

$$= 1 + 2a\lambda\underbrace{(1-\cos\xi)}_{\in[0,2]}(1+\lambda a)$$

$$\leq 1 + 4a\lambda(1+\lambda a) \rightarrow \underline{\text{unstable}}$$

$\underline{\theta = \frac{1}{2}:}$ $$|g(\xi)|^2 = 1 + 2\lambda^2 a^2 (1-\cos\xi)\underbrace{\left\{1 - \tfrac{1}{2}(1-\cos\xi)\right\}}_{\in[0,1]}$$

$$\leq 1 + \lambda^2 a^2 \rightarrow \underline{\text{unstable}}$$

$\underline{\theta = 0:}$ $$|g(\xi)|^2 = 1 - a\lambda\underbrace{(1-\cos\xi)}_{\in[0,2]}(1-\lambda a)$$

$$\leq 1 - 2a\lambda(1-\lambda a)$$

$\rightarrow$ $\underline{\text{stable if } 1 - \lambda a > 0}$

$$\boxed{C = a\lambda = \frac{a\Delta t}{\Delta x} < 1}$$

$\underline{\text{CFL}}$

$\underline{\text{condition}}$

(Courant, Friedrichs, Lewy)

$\underline{\text{Courant}}$

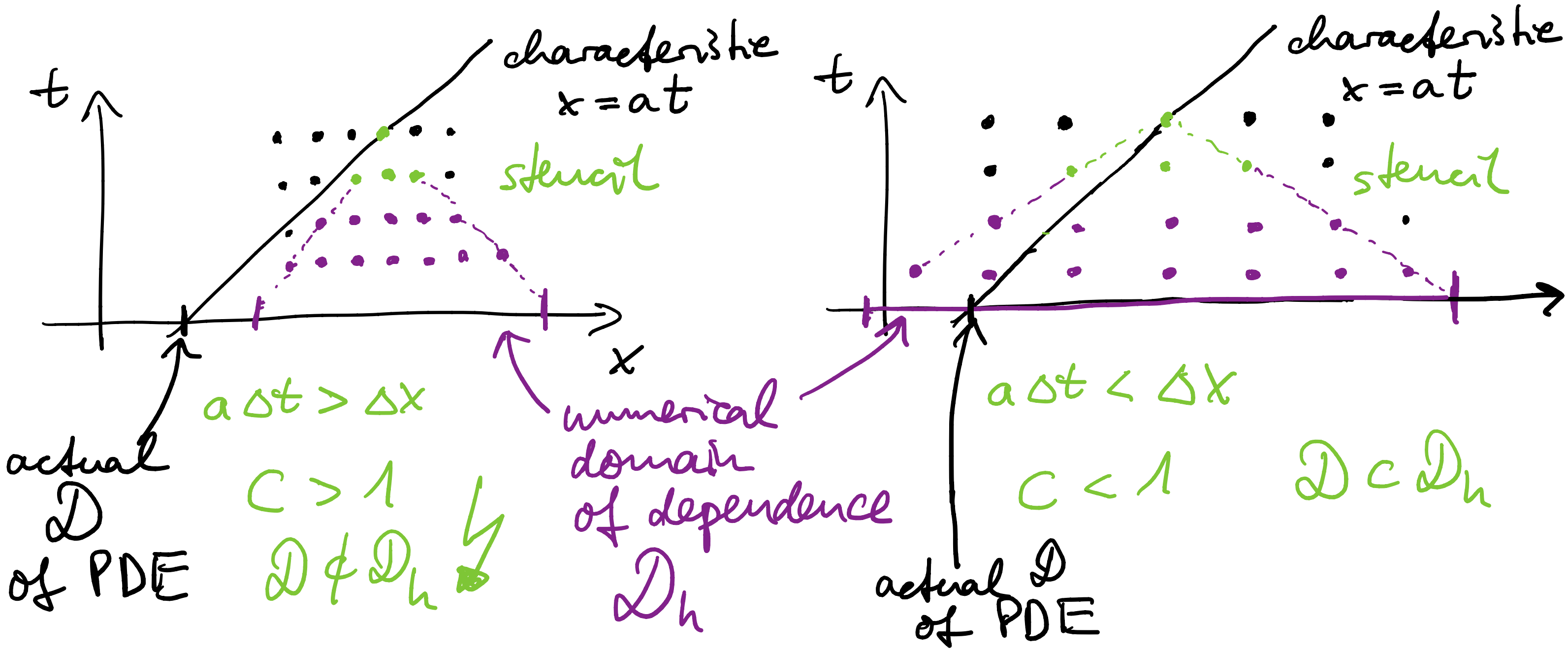
$\underline{\text{number}}$

Remark: The previous result $\frac{a\Delta t}{\Delta x} < 1$ is one special case of a more general principle, a necessary but not sufficient condition for stability & convergence:

The CFL condition: A numerical method can be convergent only if its numerical domain of dependence contains the true domain of dependence of the PDE, at least in the limit as Δt and Δx go to zero.

(Courant, Friedrichs & Lewy, Math. Ann. 100, 32 (1928))

Advection equation:



Remarks: 1) For linear hyperbolic systems with m eigenvalues/wavespeeds $\lambda^1, \dots, \lambda^m$ (cf. discussion on domain of dependence):

$$C = \frac{\Delta t}{\Delta x} \max_p |\lambda^p| \leq 1$$

Typically use explicit methods with C fixed.

For non-linear conservation equation

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

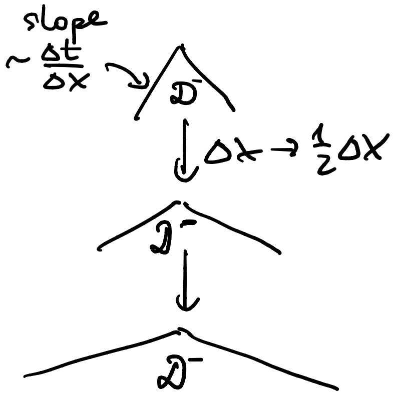
generalize to:

$$C = \sup \frac{\Delta t |f'(u)|}{\Delta x} < 1$$

2) A wider stencil $\Rightarrow$ less restrictive CFL condition on Δt

3) Parabolic equations: diffusion equation severe constraints on explicit method from CFL, since $\mathcal{D} = (-\infty, \infty)$.
(infinite propagation speed)

$\leadsto$ take $\Delta t = O(\Delta x^2)$ then



explicit method captures $\mathcal{D}$ in limit $\Delta t, \Delta x \to 0$

(as we refine $\Delta x \to \frac{1}{2}\Delta x \Rightarrow \Delta t \to \frac{1}{4}\Delta t$
$\Rightarrow$ "slope of $\mathcal{D}^-$" widens: $\frac{\Delta t}{\Delta x} \to \frac{1}{2}\frac{\Delta t}{\Delta x} \xrightarrow{\Delta x \to 0} 0 \Rightarrow \mathcal{D}_h \xrightarrow{\Delta x \to 0} (-\infty, \infty)$)

better: <u>use implicit method</u>

$\to \mathcal{D}_h = \mathcal{D} = (-\infty, \infty)$. The Crank-Nicholson method, for example, satisfies CFL condition for any Δt (the tridiagonal matrix couples all points such that solution at a given point depends on all other points)

3.5 Diffusion & dispersion

Stability & accuracy are not the only aspects to be considered when evaluating a numerical scheme!

→ Discretization introduces additional terms not present in continuum PDE

Consider advection equation and let $v(x,t)$ denote an exact solution of the discretized equ

Example 1: Upwind method

$$u_i^{n+1} = u_i^n - \frac{a\Delta t}{\Delta x}\left(u_i^n - u_{i-1}^n\right), \quad a>0.$$

$$\Big\downarrow \text{ v satisfies exactly}$$

$$v(x,t+\Delta t) = v(x,t) - \frac{a\Delta t}{\Delta x}\left[v(x,t) - v(x-\Delta x,t)\right]$$

$$\Big\downarrow \text{ Taylor series}$$

$$\left(v_t + \frac{1}{2}\Delta t\, v_{tt} + \frac{1}{6}(\Delta t)^2 v_{ttt} + \dots\right) + a\left(v_x - \frac{1}{2}\Delta x\, v_{xx} + \frac{1}{6}(\Delta x)^2 v_{xxx} + \dots\right) = 0$$

$$\Leftrightarrow \boxed{v_t + a v_x = \frac{1}{2}\left(a\Delta x\, v_{xx} - \Delta t\, v_{tt}\right) - \frac{1}{6}\left[a(\Delta x)^2 v_{xxx} + (\Delta t)^2 v_{ttt}\right] + \dots}$$

different from original continuum PDE!

PDE for $v(x,t)$

Take $\frac{\Delta t}{\Delta x}$ = fixed.

$\rightsquigarrow$ 1st order: recover original advection equ

2nd order:

$$v_t + a v_x = \frac{1}{2}\left(a\Delta x\, v_{xx} - \Delta t\, v_{tt}\right) \qquad (*)$$

$$\downarrow \partial_t$$

$$v_{tt} = -a v_{xt} + \frac{1}{2}\left(a\Delta x\, v_{xxt} - \Delta t\, v_{ttt}\right)$$

$$\partial_x$$

$$v_{xt} = -a v_{xx} + \frac{1}{2}\left(a\Delta x\, v_{xxx} - \Delta t\, v_{ttx}\right)$$

$$v_{tt} = a^2 v_{xx} + O(\Delta t)$$

$$\downarrow \text{in } (*)$$

advection-diffusion equation

$$\boxed{v_t + a v_x = \frac{1}{2} a \Delta x \Big(1 - \underbrace{\frac{a \Delta t}{\Delta x}}_{\lambda\text{: Courant number}}\Big) v_{xx}}$$

Note: 1) u_i^n provide <u>2nd order accurate</u> approximation to the true solution of <u>this equation</u>, but only <u>1st order</u> accurate solution to <u>advection equ</u>

2) the diffusion coefficient

$$v_d = \frac{1}{2} a \Delta x (1 - \lambda)$$

is called <u>numerical diffusion</u>. Observe:

$v_d = 0 \iff \lambda = 1$ (marginally stable)

$v_d > 0 \iff \lambda < 1$ (CFL condition)

Upwind method is <u>dissipative & stable</u>!

$v_d < 0 \iff \lambda > 1$ ill-posed backward heat equ with exponentially growing solutions

Upwind method unstable!

Note the relation between properties of modified equ (PDE for v) and stability of continuum PDE

Example 2: Lax-Wendroff method

2nd order accurate 2-level 3-point method

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

$\downarrow$ similar analysis in terms of exact solution $v(x,t)$ to discretized problem

$$\boxed{v_t + a v_x = -\frac{1}{6} a (\Delta x)^2 \left[1 - \left(\frac{a\Delta t}{\Delta x}\right)^2\right] v_{xxx}}$$

PDE for $v(x,t)$

Note: 1) u_n^i $\begin{cases} \text{3rd-order accurate sol. to this equ.} \\ \text{2nd-order " " " advection eq.} \end{cases}$

2) v_{xxx} term leads to dispersive behavior (see below)

General analysis: PDE dispersion relation

previous examples suggest PDE for exact solution of discretized problem of the general form

$$v_t + a_1 v_x + a_2 v_{xx} + a_3 v_{xxx} + a_4 v_{xxxx} + \dots = 0$$

$\Big\downarrow$ Fourier transformation $\quad \hat{v}(\xi,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} v(x,t) e^{-i\xi x} dx$

$$\hat{v}_t(\xi,t) = -a_1 i\xi \hat{v}(\xi,t) + a_2 \xi^2 \hat{v}(\xi,t) + a_3 i\xi^3 \hat{v}(\xi,t)$$
$$- a_4 \xi^4 \hat{v}(\xi,t) + \dots$$

$$\Rightarrow \quad \hat{v}(\xi,t) = e^{-i\omega t}\, \hat{\eta}(\xi) \quad \left(\hat{\eta}(\xi) = \hat{v}(\xi,0)\right)$$

with $\boxed{\omega(\xi) = a_1\xi + ia_2\xi^2 - a_3\xi^3 - ia_4\xi^4 - \dots}$

dispersion relation of the PDE

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{\eta}(\xi)\, e^{i(\xi x - \omega(\xi)t)}\, d\xi$$

Note: 1) dispersion relation can be found via ansatz $\quad u(x,t) = e^{-i\omega t} e^{i\xi x}$

(different wave numbers ξ decouple)

2) $$e^{-i\omega t} = \underbrace{e^{(a_2\xi^2 - a_4\xi^4 + \dots)t}}_{\text{diffusive term}} \times \underbrace{e^{i(a_1\xi - a_3\xi^3 + \dots)t}}_{\text{dispersive term}}$$

(exponential growth or decay) (waves travel at different speed)

Upwind: diffusive

Lax-Wendroff: (purely) dispersive

in general: may be combination of both

3) purely dispersive $\iff$ $\omega(\xi)$ real $\forall \xi \in \mathbb{R}$

$$\rightsquigarrow u(x,t) \overset{1)}{=} e^{i\xi\left(x - \underbrace{\frac{\omega}{\xi}}_{\text{phase velocity}} t\right)}$$

phase velocity (propagation speed of this mode)

$$\boxed{c_p = \frac{\omega(\xi)}{\xi}}$$

$$\boxed{c_g = \frac{d\omega(\xi)}{d\xi}}$$ group velocity

In general: $a \neq c_g \neq c_p$

Conclusion:

- schemes in which leading error term is composed of 2nd order spatial derivatives show numerical diffusion / numerical viscosity
- schemes in which leading error term is composed of 3rd order spatial derivatives introduce numerical dispersion