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

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

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

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

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

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

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

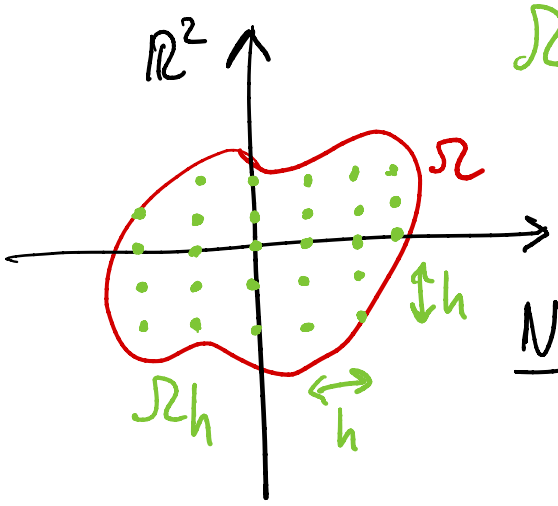
Examples: $L = \Delta,\ \Box,\ \partial_t - \partial_x f(\cdot),\ \dots$

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

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

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

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



h = discretization parameter

Notation: $x \to x_i \in \{x_1, \dots, x_{N_x}\}$

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

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

D_h^X : discretization operator

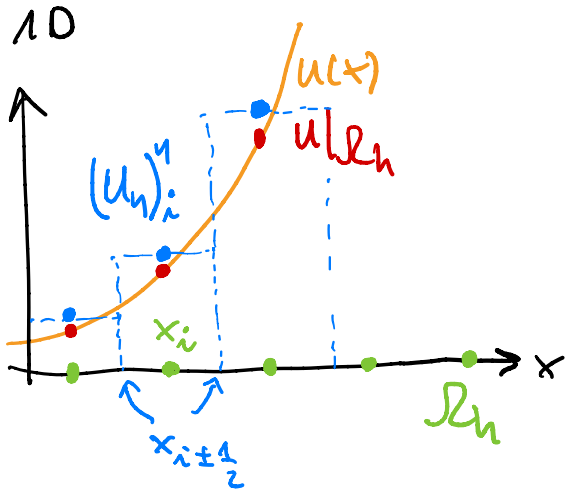
$D_h^X : X \to X_h$

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

$D_h^x u = u_h$ for example

$$\begin{cases} u|_{\Omega_h}, \text{ i.e. } (u_h)_i^n = u(x_i, t_n) & \text{finite difference methods} \\ (u_h)_i^n = \dfrac{1}{\Delta x}\displaystyle\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u(x, t_n)\,dx & \text{finite volume methods} \end{cases}$$

spatial average across Δx



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:

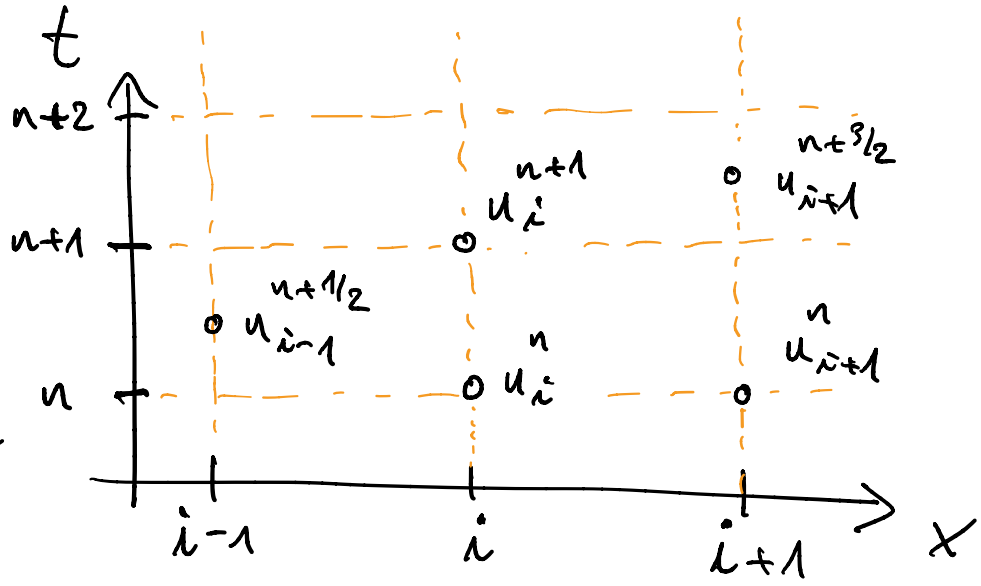
- finite differences: point values at grid points (or staggered)
- finite volume: discrete values $\hat{=}$ averages over finite volumes (→ later)

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

Notation:

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

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



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" and "staggered grid function"

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

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:

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

$\rightsquigarrow$ 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)

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

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

backward: $u^n_{i-1} = u^n_i + \left.\frac{\partial u}{\partial x}\right|_i (-\Delta x_i^-) + \frac{1}{2}\left.\frac{\partial^2 u}{\partial x^2}\right|_i (\Delta x_i^-)^2 + \dots$

↓ substract

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

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

On a uniform grid:

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

"centered difference approximation"

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

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

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

Higher-order derivatives:

intuitively: repeatedly apply 1st order derivatives

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

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

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

$$\equiv D^2 u$$

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

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

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

$$= D^2 u$$

General approach:

goal: compute approximation to $u^{(k)} = \frac{\partial^k u}{\partial x^k}$

based on a stencil of $n \geq k+1$ points

$x_1, \dots, x_n$

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

→ 1D Taylor expansion

→ $u_{i-1} = u_i - \left.u^{(1)}\right|_i \Delta x + \frac{1}{2}\left.u^{(2)}\right|_i (\Delta x)^2 - \frac{1}{6}\left.u^{(3)}\right|_i (\Delta x)^3 + O(\Delta x^4)$

→ $u_{i-2} = u_i - \left.u^{(1)}\right|_i (2\Delta x) + \frac{1}{2}\left.u^{(2)}\right|_i (2\Delta x)^2 - \frac{1}{6}\left.u^{(3)}\right|_i (2\Delta x)^3 + O(\Delta x^4)$

Ansatz:

$$\left.D_2 u\right|_i = c_1 u_i + c_2 u_{i-1} + c_3 u_{i-2}$$

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

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

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

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

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

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

$$+ \dots$$

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 $u^{(1)}|_i$ to highest order possible:

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

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

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

Truncation error:

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

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

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

$\rightsquigarrow$ 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), $n \geq k+1$. Consider n Taylor series $\ell = 1, \dots, n$:

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

known

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

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

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

$$\underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ & (a_{lm})_{\substack{l=2,\dots,n \\ m=1,\dots,n}} = \frac{1}{(l-1)!}(x_m - x_i)^{l-1} & \end{pmatrix}}_{\equiv A \in Mat(n\times n)} \underbrace{\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}}_{c} = \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{b}$$

cancellation of lower order terms

← k+1

cancellation of higher order terms

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

2) **Accuracy:** RHS $\neq 0$ in Taylor expansions (*):

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

$l-1 < k$: necessary to cancel lower order terms and get $O(\Delta x)$ accuracy

$l-1 > k$:

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

⇝ procedure is at least

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

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

General: increasing stencil size increases accuracy

3.2.2 Sample Discretizations

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

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

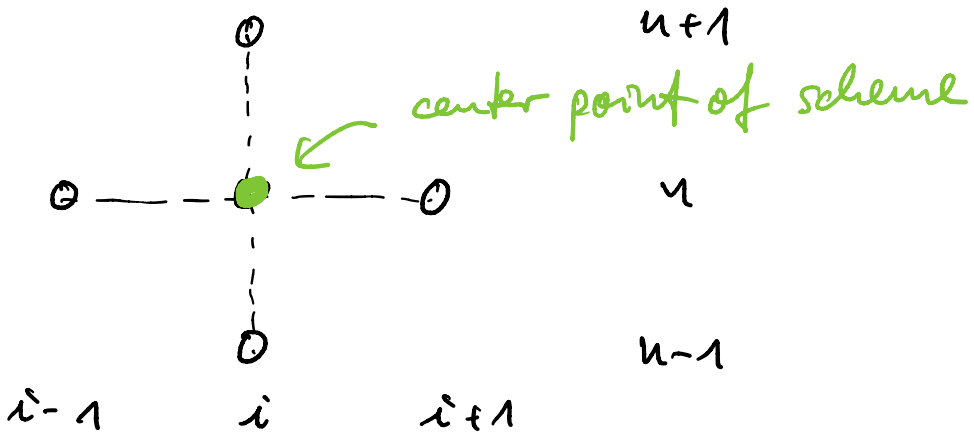
$$u(0,t) = u(1,t) = 0 \quad \text{fixed (Dirichlet) boundary conditions}$$

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

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

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

stencil:



1D FDA approximation to $O(\Delta x^2, \Delta t^2)$:

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

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

$$\left.\begin{aligned} u_1^n &= u_I^n = 0 \\ &\searrow x=0 \quad \searrow x=1 \end{aligned}\right\} \text{discretized boundary condition}$$

explicitly solve for u^{n+1}:

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

→ 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: $O(\Delta x^2, \Delta t^2)$

A: diagonal ⇒ explicit scheme

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

2) 1D diffusion equation, Crank-Nicholson

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

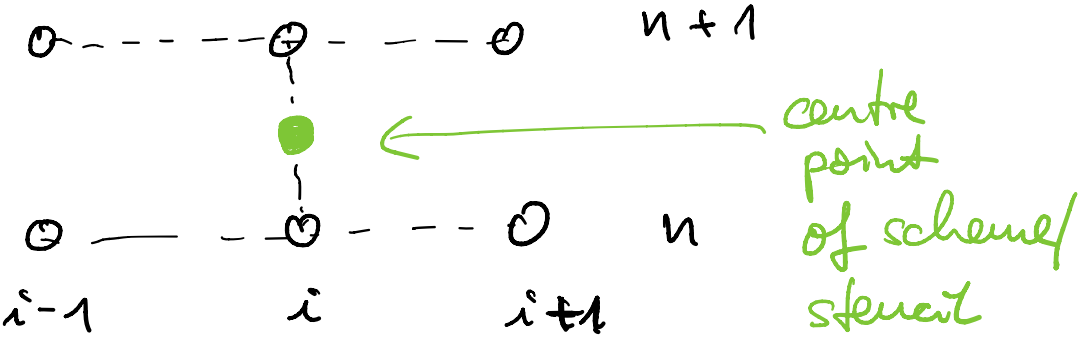
Consider: centered in time & space

(minimizes truncation error, minimizes instabilities)

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

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

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 $O((\Delta t)^2, (\Delta x)^2)$:

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

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

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

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

write as

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

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

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

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

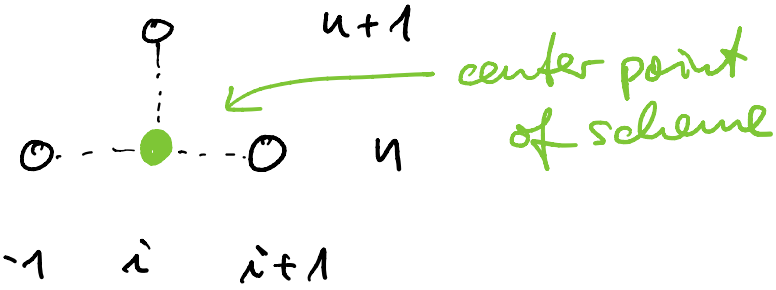
3) 1D advection equation

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

(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



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

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

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

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

⇒ explicit scheme $\quad O(\Delta t, \Delta x^2)$

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

(ii) Lax–Friedrichs method

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

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

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

(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

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

- u^{n+1}: column vector containing sufficient unknowns to write problem in 1st order in time form
- $G(S_+, S_-)$: 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 $u_i^{n+1} = u_i^n - a\frac{\Delta t}{\Delta x}(u_i^n - u_{i-1}^n)$

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

Example 2: rewriting multi-level schemes

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

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) $\quad \lambda = \frac{c\Delta t}{\Delta x}$

$\rightsquigarrow$ define auxiliary variables

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

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

$$\rightsquigarrow u^n \equiv (u_1^n, v_1^n, \cdots, u_I^n, v_I^n)$$

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

$\rightsquigarrow$ 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 stable wrt. the norm $\|\cdot\|$ if there exists constants $c(T), \beta, \tau$ such that

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

Definition (truncation error, consistency): Consider above setting. The numerical scheme is said to be consistent of order (q, p) wrt. the norm $\|\cdot\|$ if for any exact solution v of the PDE

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

τ: "local truncation error"

Definition (convergence): Consider above setting. The numerical scheme is convergent of order (q, p) wrt. to the norm $\|\cdot\|$ if

for any exact solution v of the PDE

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

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. $\|u^0 - v(\cdot, 0)\| = O(\Delta x^p)$, then the scheme is convergent of order (q, p), i.e.

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

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 $t^n = n\Delta t \leq T$.

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

$$\rightarrow \|G^n\| \leq c(T)\, e^{\beta \frac{n\Delta t}{T}}$$

consistency:

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

with $\|R\| = O(\Delta x^p) + O(\Delta t^q)$

no global error:

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

linearity $\rightarrow$

$$= G\, E^{n-1} - \Delta t\, R$$

$$= G^2 E^{n-2} - \Delta t\, G R - \Delta t\, R$$

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

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

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

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

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

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

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

□

Remarks: 1) The converse is also true (as "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!

⇒ use Lax theorem, consistency & stability easier to show, but need sufficient criterion for stability

Theorem (von Neumann stability condition):

A finite difference scheme

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

(u: vector of unknowns)

linear in u is stable wrt. 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 $0 \le \Delta t \le \tau$. Here, $g(\xi)$ is the <u>amplification factor</u> ("symbol of G")

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

($i \equiv \sqrt{-1}$)

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

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

Observe:

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

$\ell = j+1$

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

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

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

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

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

Discrete L^2 norms:

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

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

periodicity

Parseval property:

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

$$\overset{(e^{ij\xi})_{j\in\mathbb{N}} \text{ orthogonal}}{=} \int_0^{2\pi} \sum_j |u_j e^{ij\xi}|^2 \, d\xi \overset{|e^{ij\xi}|=1}{=} \int_0^{2\pi} \sum_j u_j^2 \, d\xi$$

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

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

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

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

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

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

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

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

scheme is stable!

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

$\Rightarrow$ 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 $\mathbb{R} \setminus I_\alpha$.

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

on I_α

Then:

$$\|G^n\| \|u^0\|_2 \geq \|G^n u^0\|_2 = \|u^n\|_2 \geq (1+\alpha\Delta t)^n \|u^0\|_2$$

(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 $\alpha, T > 0$ $\exists\ \Delta t \in (0, T)$ 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 $n = \frac{T}{\Delta t}$

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

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

choose T sufficiently small such that

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

(n sufficiently large)

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

Then:

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

contradiction.

□

Remarks: 1) If one wants $\|u^n\| \leq c(T) \|u^0\|$ for stability (cf. Def. in Sec. 3.3), the requirement on $g(\xi)$ in the theorem sharpens to

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

2) 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 $\{u_j^n\}_{j \in \mathbb{N}}$ are coupled via the difference equations, each $\hat{u}(\xi)$ satisfies (*) and is thus decoupled from all other wavenumbers ξ !