Goal

Part 1: Explicat Runge Kutta methods ode's (1/3 course)

Part 2: Num. approx (2/3 — pde's

Course: lessons (pictures boards)

exam: 80% final exam

20% graded exercises

3 groups, 1 graded ex for one group

4 graded ex for each student

3 validated

2023: 50 students 4 failed

24 40 1

Why pde's
IPCC Reports 1992 →

Part 1 : explicit RK ode's

Find $\vec{y}(t) \in \mathbb{R}^d$ such that

$$(1)\begin{cases} \vec{y}'(t) = \vec{f}(t, \vec{y}(t)) & \leftrightarrow t_0 \\ \vec{y}(t_0) = \vec{y}_0 \end{cases}$$

where $\vec{y}_0 \in \mathbb{R}^d$ and $\vec{f} : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$, $(t, \vec{z}) \to \vec{f}(t, \vec{z})$

are given

$$\begin{cases} y_1'(t) = f_1(t, y_1(t), \dots, y_d(t)) \\ \vdots \\ y_d'(t) = f_d(\quad\quad\quad\quad\quad) \end{cases}$$

Ex : Newton's eq : $\begin{cases} m\,\vec{X}''(t) = \vec{F}(t, \vec{X}(t), \vec{X}'(t)) \\ \vec{X}(0) = \vec{X}_0 \quad \vec{X}'(0) = \vec{V}_0 \end{cases}$

$$\vec{X}'(t) = \vec{V}(t) \qquad \vec{X}(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \quad \vec{V}(t) = \begin{pmatrix} V_1(t) \\ V_2(t) \\ V_3(t) \end{pmatrix}$$

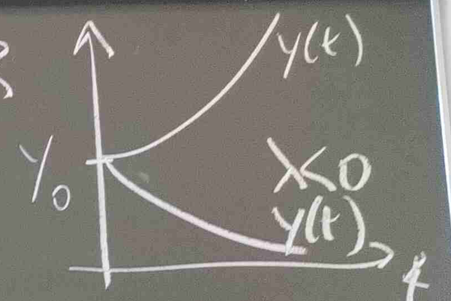
$$\begin{cases} \vec{V}'(t) = \frac{1}{m}\vec{F}(t, \vec{X}(t), \vec{V}(t)) \\ \vec{X}'(t) = \vec{V}(t) \end{cases}$$

$$\vec{y}(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \\ V_1(t) \\ V_2(t) \\ V_3(t) \end{pmatrix}$$

Examples $d = 1$

$$\begin{cases} y'(t) = \lambda y(t) & \lambda \in \mathbb{R} \\ y(0) = y_0 \end{cases}$$

$$y(t) = y_0 e^{\lambda t}$$

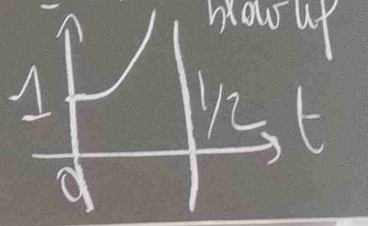


$$\begin{cases} y'(t) = (y(t))^3 \\ y(0) = 1 \end{cases}$$

$$y(t) = \frac{1}{\sqrt{1-2t}}$$

$y' y^n = 1 \quad n = -3$

$\frac{1}{n+1} y^{n+1} = t + c$

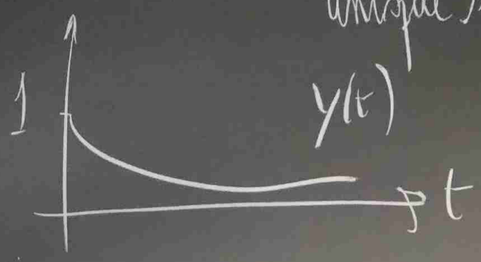


$y'(t) = -(y(t))^3$

$y(0) = 1$

$y' y^n = -1 \quad n = -3$

$y(t) = \frac{1}{\sqrt{1+2t}}$



unique sol.

$y' y = -y^4$

$\frac{1}{2}\frac{d}{dt}(y(t))^2 \leq 0$

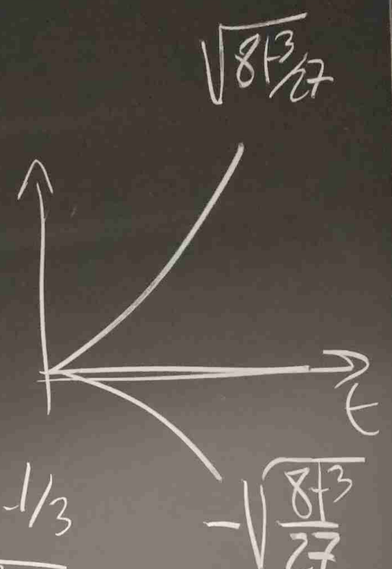
$(y(t))^2 \leq (y(0))^2$

$y'(t) = (y(t))^{1/3}$

$y(0) = 0$

3 solutions: $y' y^n = 1 \quad n = -1/3$

$y(t) = 0$ and $\pm\sqrt{\frac{8t^3}{27}}$



$y'(t) = 3y(t) - 3t$

$y(0) = \alpha$

numerically ill-posed

$y(t) = (\alpha - 1/3)e^{3t} + t + 1/3$

$\alpha = 1/3 \quad y(10) = 10 + 1/3$

$\alpha = 1/3 + 10^{-6} \quad y(10) = 10^{-6} e^{30} + 10 + 1/3$

$\simeq 10^7 !$

Thm:

- **local existence**, if $f \in C^1$ in the neighb of t_0, then there exists a unique sol of (1) in the neighb of t_0.
- **global existence**, if f is lipschitz cont. with respect to the second variable.

$\exists L>0 \ \forall t, \bar{z}_1, \bar{z}_2 \quad \|f(t,z_1) - f(t,z_2)\| \leq L\|z_1 - z_2\|$

then there exists a unique sol of (1) $\forall t$

Euler scheme ($d=1$, $t_0=0$, final time T)

h time step $h = T/N$

$t_n = nh \quad n = 0, 1, \ldots, N$

let y_n be an approx. of $y(t_n)$

Computed using Euler explicit scheme

$$\frac{y_{n+1} - y_n}{h} = f(t_n, y_n)$$

$y_{n+1} = y_n + h f(t_n, y_n)$ explicit formula

Compute the error at final time $E_N = y(t_N) - y_N$

$y(t_n)$ satisfies $y'(t_n) = f(t_n, y(t_n))$

if $y \in C^2[0,T]$ then

$\exists\, t_n < s_n < t_{n+1}$; $y(t_{n+1}) = y(t_n) + h y'(t_n) + \frac{h^2}{2} y''(s_n)$.

$$y'(t_n) = \frac{y(t_{n+1}) - y(t_n)}{h} - \frac{h}{2} y''(s_n)$$

$$\frac{y(t_{n+1}) - y(t_n)}{h} = f(t_n, y(t_n)) + \frac{h}{2} y''(s_n)$$

$$\frac{E_{n+1} - E_n}{h} = f(t_n, y(t_n)) - f(t_n, y_n) + \frac{h}{2} y''(s_n)$$

$$E_{n+1} = E_n + h\left(f(t_n, y(t_n)) - f(t_n, y_n)\right) + \frac{h^2}{2} y''(s_n)$$

$$|E_{n+1}| \le |E_n| + h\left|f(t_n, y(t_n)) - f(t_n, y_n)\right| + Ch^2$$

where $C = \frac{1}{2} \max_{0 \le s \le T} |y''(s)|$

If f is lipschitz cont w.r second variable

then

$$|E_{n+1}| \le |E_n|(1 + hL) + Ch^2$$

$$\le |E_{n-1}|(1+hL)^2 + Ch^2\left(1 + (1+hL)\right)$$

$$\le \vdots$$

$$\le |E_0|(1+hL)^{n+1} + Ch^2\left(1 + (1+hL) + \cdots + (1+hL)^n\right)$$

$$|E_N| \le \overset{0}{Ch^2}\left(1 + (1+hL) + \cdots + (1+hL)^N\right)$$

$a, b \in \mathbb{R}$, $a^N - b^N = (a-b)(a^{N-1} + a^{N-2}b + \cdots + ab^{N-2} + b^{N-1})$

$a = 1 \quad b = 1 + hL$

$$|E_N| \le Ch^2 \frac{(1+hL)^N - 1}{1 + hL - 1}$$

$$\le \frac{Ch}{L}\left((1+hL)^N - 1\right)$$

$x > 0 \quad e^x = 1 + x + \frac{x^2}{2} e^{\xi} \quad 0 < \xi < x$

thus $e^x \geq 1+x$ and

$$|\underset{N}{E}| \leq \frac{Ch}{L}\left(e^{hLN} - 1\right)$$

$$= \frac{Ch}{L}\left(e^{LT} - 1\right)$$

Lemma: If $y \in C^2[0,T]$, iff f is Lipschitz cont w.r to the second var.

then $\exists C > 0$ $\forall 0 < h < T$

$$|y(t_N) - y_N| \leq Ch$$

Proof: $C = \frac{1}{2}\max_{0 \leq s \leq T} |y''(s)| \times \frac{1}{L}\left(e^{LT} - 1\right)$

Numexperiment; choose f, y_0 therefore $y \in C^2[0,T]$, final error is divided by 2 if h is.

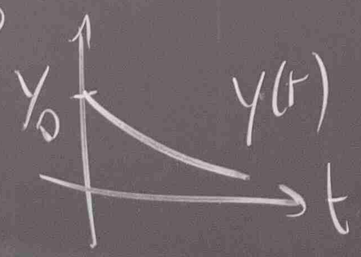
Issues: • if T, L are large then e^{LT} is large

• implicit scheme? stability?

• This estimate is not optimal for the linear ode:

$$\begin{cases} y'(t) = \lambda y(t) & \lambda < 0 \\ y(0) = y_0 \end{cases}$$

$y(t) = y_0 e^{\lambda t}$



lemma:

- Euler implicit: $\frac{y_{n+1}-y_n}{h} = \lambda y_{n+1}$ then

$$\exists C>0 \; \forall y_0 \in \mathbb{R} \; \forall T>0 \; \forall h>0 \; \forall \lambda<0 \quad |y(t_N)-y_N| \leq CT\lambda^2 h |y_0|$$

(no restriction)

- Euler explicit $\frac{y_{n+1}-y_n}{h} = \lambda y_n$ then

$$\exists C>0 \; \forall y_0 \in \mathbb{R} \; \forall T>0 \; \forall \lambda<0 \; \forall 0<h<\frac{-2}{\lambda}$$

$$|y(t_N)-y_N| \leq CT\lambda^2 |y_0| h$$

Stability Condition

Eulero esplicito:

$$y(t_N) = y_0 e^{\lambda N h}$$

$$y_{n+1} = y_n (1 + h\lambda)$$

$$y_N = y_0 (1 + h\lambda)^N$$

$$E_N = y_0 \left(e^{\lambda N h} - (1 + h\lambda)^N \right)$$

$a^N - b^N = (a-b)(a^{N-1} + \dots + b^{N-1})$

$$E_N = y_0 \left(e^{\lambda h} - (1 + h\lambda) \right) \left(e^{\lambda h (N-1)} + \dots + (1 + h\lambda)^{N-1} \right)$$

$x > 0 \quad e^{-x} = 1 - x + \frac{x^2}{2} e^{-s} \quad 0 < s < x$

$$\left| e^{-x} - (1 - x) \right| = \left| \frac{x^2}{2} e^{-s} \right| \leq \frac{x^2}{2}$$

$$|E_N| \leq |y_0| \frac{h^2 \lambda^2}{2} \left(\left| e^{\lambda h (N-1)} \right| + \left| e^{\lambda h (N-1)} (1 + h\lambda) \right| + \dots + \left| (1 + h\lambda)^{N-1} \right| \right)$$

If $|1 + \lambda h| \leq 1$ then $|E_N| \leq |y_0| \frac{h^2 \lambda^2}{2} N = |y_0| \frac{T \lambda^2}{2} h$ we proved the lemma with $C = \frac{1}{2}$

$-1 \leq 1 + \lambda h \leq 1$ ← true

$\lambda < 0$

true if $h \leq -\frac{2}{\lambda}$

Exercise 2: $\vec{y}'(t) + A\vec{y}(t) = 0 \quad A \in \mathbb{R}^{N \times N}$ sym. pos. def.

$$\vec{y}'(t) = \vec{f}(t, \vec{y}(t)) \qquad t > t_0$$
$$\vec{y}(t_0) = \vec{y}_0 \qquad \vec{y}(t) \in \mathbb{R}^d$$

Euler scheme, order 1

$t_1 = t_0 + h$ first time step

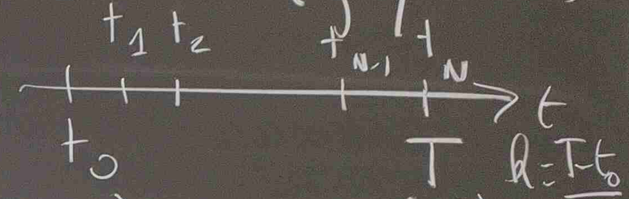
y_1 approx $y(t_1)$

$$y_1 = y_0 + h f(t_0, y_0)$$

$$y(t_0 + h) = y(t_0) + h \underbrace{f(t_0, y(t_0))}_{y'(t_0)} + O(h^2)$$

$$y(t_1) - y_1 = O(h^2)$$
local error
$(f \in \mathcal{C}^2)$

$h \leftarrow$



$h = \frac{T - t_0}{N}$

$$y(t_N) - y_N = O(h)$$
final error

higher order : Runge Kutta order 4

Runge, order 2

$$y'(t_0 + h/2) = f\left(t_0 + h/2, y(t_0 + h/2)\right)$$

$$\frac{y(t_1) - y(t_0)}{h} = f\left(t_0 + h/2, y(t_0 + h/2)\right) + O(h^2)$$
$y \in \mathcal{C}^3$

$$y(t_0 + h) = y(t_0 + h/2) + \frac{h}{2} y'(t_0 + h/2) + \frac{h^2}{4.2} y''(t_0 + h/2) + \frac{h^3}{8.6} y'''(\alpha)$$

$$y(t_0) = y(t_0 + h/2) - \frac{h}{2} y'(t_0 + h/2) + \text{———} - \frac{h^3}{8.6} y'''(\beta)$$

$t_0 + h/2 < \alpha < t_0 + h$

$t_0 < \beta < t_0 + h/2$

$$y(t_0+h) - y(t_0) = h\, y'(t_0 + h/2) + O(h^3)$$

$$\frac{y(t_1)-y(t_0)}{h} = f\left(t_0 + h/2,\ y(t_0) + \frac{h}{2} f(t_0, y_0)\right) + O(h^2)$$

$$y(t_1) = y(t_0) + h\left[f\left(t_0 + h/2,\ y(t_0) + \frac{h}{2} f(t_0, y_0)\right)\right] + O(h^3)$$

$$y_1 = y_0 + h\, f\left(t_0 + h/2,\ y_0 + \frac{h}{2} f(t_0, y_0)\right)$$

$$y(t_1) - y_1 = O(h^3)$$

$$y(t_N) - y_N = O(h^2)$$

This can be written in the Runge Kutta (R
ous follows:

$$\vec{k}_1 = \vec{f}(t_0, \vec{y}_0)$$ (evaluating $f(t_0, y_0)$ is costly, $d = 10^6$)

$$\vec{k}_2 = \vec{f}\left(t_0 + h/2,\ \vec{y}_0 + \frac{h}{2}\vec{k}_1\right)$$

$$\vec{y}_1 = \vec{y}_0 + h\,\vec{k}_2$$

General RK scheme, s stages, explicit:

$$k_1 = f(t_0, y_0)$$

$$k_2 = f(t_0 + c_2 h,\ y_0 + h a_{21} k_1)$$

$$\vdots$$

$$k_s = f(t_0 + c_s h,\ y_0 + h(a_{s1} k_1 + \cdots + a_{s,s-1} k_{s-1}))$$

$$y_1 = y_0 + h(b_1 k_1 + \cdots + b_s k_s)$$

table

$c_1 = 0$	0	(0)
c_2	a_{21}	
$\vdots$		0
c_s	a_{s1} $\cdots\cdots$	$a_{s,s-1}$ 0
	b_1	b_{s-1} b_s

the a_{ij} matrix is lowertriangular
0 on the diagonal, explicit scheme

Ex: Euler

0	
	1

Runge

1/2	1/2	
	0	1

RK4 (order 4) explicit

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
	1/6	4/6	4/6	1/6

Remarks:

- there is no RK method of order 5 with 5 stages!
- the table

c_1	a_{11}	⋯	a_{1s}
⋮	⋮		⋮
c_s	a_{s1}	⋯	a_{ss}
	b_1	⋯	b_s

defines an implicit RK method with s stages

Ex: Crank Nicolson (order 2)

$$y_1 = y_0 + \frac{h}{2}\left(f(t_0, y_0) + f(t_1, y_1)\right)$$

$$\left(y(t_1) = y(t_0) + \int_{t_0}^{t_1} f(t, y(t))\,dt\right)$$

$k_1 = f(t_0, y_0)$

~~$k_2 = f(t_0 + h, y_1)$~~ $k_2 = f(t_0 + h, y_0 + h(\frac{1}{2}k_1 + \frac{1}{2}k_2))$

$y_1 = y_0 + h(\frac{1}{2}k_1 + \frac{1}{2}k_2)$

nonlinear eq. to obtain k_2

0	0	0	
1	1/2	1/2	
	1/2	1/2	

Rem: • there are not Runge Kutta methods of order 5 with 5 stages!

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ c_2 & & & \\ \vdots & & & \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array}$$

defines an implicit RK method

Ex: Crank-Nicolson

$$Y_1 = Y_0 + \frac{h}{2}\left(f(t_0, Y_0) + f(t_1, Y_1)\right)$$

$k_1 = f(t_0, Y_0)$

~~$k_2 = f(t_0 + h, Y_1)$~~

$Y_1 = Y_0 + h\left(\frac{1}{2}k_1 + \frac{1}{2}k_2\right)$

$k_2 = f\left(t_0 + h, Y_0 + h\left(\frac{1}{2}k_1 + \frac{1}{2}k_2\right)\right)$ $k_2 \rightarrow$ solve a nonlin eq.

Def: the s stages RK method (impl. or expl.) is of order p if the error after on step is

$$\underbrace{Y(t_0 + h) - Y_1}_{\text{local error}} = O(h^{p+1})$$

Question: • how to choose the c_i, a_{ij}, b_i so that the error is of order $p = 4$ (Thm 1)

• if the local error (error after on step) is $O(h^{p+1})$ is the final error $O(h^p)$? (Thm 2)

Thm 1: $y' = f(y)$ $(y'(t) = f(y(t))$ autonomous case
reason why ex)

$1 \le p \le 4$ $\quad y(t_0 + h) - y_1 = O(h^{p+1})$ if and only if

$\sum_{i=1}^{s} b_i = 1$ $\quad p=1$

$\sum_{i,j=1}^{s} b_i a_{ij} = 1/2$ $\quad p=2$

$\sum_{i,j,k=1}^{s} b_i a_{ij} a_{ik} = 1/3$ $\quad \sum_{i,j,k=1}^{s} b_i a_{ij} a_{jk} = 1/6$ $\quad p=3$

$\sum_{i,j,k,l=1}^{s} b_i a_{ij} a_{ik} a_{il} = 1/4$

$\sum_{i,j,k,l=1}^{s} b_i a_{ij} a_{jk} a_{il} = 1/8$

$\sum_{i,j,k,l} b_i a_{ij} a_{jk} a_{jl} = 1/12$

$\sum_{i,j,k,l} b_i a_{ij} a_{jk} a_{kl} = 1/24$

$p=4$

These conditions are valid for impl. and expl RK methods

When $c_1 = 0$ and $c_i = \sum_{j=1}^{s} a_{ij}$ $i = 1, \dots, s$ (the $y' = f(y)$ assumption can be justified)

these conditions simplify to

$\sum_i b_i = 1$	$p=1$
$\sum_i b_i c_i = 1/2$	$p=2$
$\sum_i b_i c_i^2 = 1/3$	$\sum_{ij} b_i a_{ij} c_j = 1/6$ $p=3$
$\sum_i b_i (c_i)^3 = 1/4$	
$\sum_{ij} b_i a_{ij} c_j c_i = 1/8$	
$\sum_{ij} b_i a_{ij} c_j^2 = 1/12$	
$\sum_{ijk} b_i a_{ij} a_{jk} c_k = 1/24$	

It can be checked that for $s = 4$

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
	1/6	2/6	2/6	1/6

satisfies these conditions

no scheme order 5 with 5 stages

Proof thm 1:

order 1: $y(t_0+h) = y(t_0) + h y'(t_0) + O(h^2)$ (1)

$= y_0 + h f(y_0) + O(h^2)$

$y_1(h) = y_0 + h \underbrace{\sum_{i=1}^{s} b_i k_i(h)}_{\phi(h)}$

$k_i(h) = f\left(y_0 + \underbrace{h \sum_{j=1}^{s} a_{ij} k_j(h)}_{\psi_i(h)}\right)$ $i = 1, \dots, s$

$$y_1(h) = y_0 + \phi(0) + h\phi'(0) + O(h^2) \quad (2)$$

$$\phi'(h) = \sum_{i=1}^{s} b_i k_i(h) + h \sum_{i=1}^{s} b_i k_i'(h)$$

$$\phi'(0) = \sum_{i=1}^{s} b_i f(y_0)$$

$$(1)-(2) \cdot \vec{y}(t_0+h) - \vec{y}_1(h) = h\underbrace{\left(1 - \sum_{i=1}^{s} b_i\right)}_{=0} \vec{f}(y_0) + O(h^2)$$

order 2

$$y(t_0+h) = y(t_0) + hy'(t_0) + \frac{h^2}{2} y''(t_0) + O(h^3)$$

$$y_i'(t) = f(y_1(t), \dots, y_d(t)) \quad i = 1, \dots, d$$

$$y_i''(t) = \sum_{j=1}^{d} \frac{\partial f}{\partial y_j}(\vec{y}(t))\, y_j'(t)$$

$$\vec{y}'' = f'(y)\vec{y}' \qquad f'(y): \mathbb{R}^d \to \mathbb{R}^d, \quad \vec{u} \to f'(y)\vec{u}$$

$$\left(f'(y)\vec{u}\right)_i = \sum_{j=1}^{d} \frac{\partial f}{\partial y_j}(y) u_j$$

$$y''(t_0) = f'(y_0) f(y_0)$$

$\in \mathbb{R}^d$ — matrix — $\in \mathbb{R}^d$

order 2 cond $p=2$:

$$\vec{y}_1(t_1) = \vec{y}(t_0) + h\vec{y}'(t_0) + \frac{h^2}{2}\vec{y}''(t_0) + O(h^3)$$

$$y_i'(t) = f_i(y_1(t), \dots, y_d(t)) \quad i = 1, \dots, d$$

$$y_i''(t) = \sum_{j=1}^{d} \frac{\partial f_i}{\partial y_j}(y(t))\, y_j'(t)$$

$$y'' = f'(y)\, y' \qquad f'(y) : \mathbb{R}^d \to \mathbb{R}^d \text{ linear}$$

$$\vec{u} \to f'(y)\vec{u} \text{ form}$$

$y'' \in \mathbb{R}^d$, $f'(y) \in \mathbb{R}^{d\times d}$, $y' \in \mathbb{R}^d$

$$(f'(y)u)_i = \sum_{j=1}^{d} \frac{\partial f_i}{\partial y_j} u_j$$

$$\underbrace{y''(t_0)}_{\text{vector}} = \underbrace{f'(y_0)}_{\text{matrix}}\underbrace{f(y_0)}_{\text{vector}}$$

$$y(t_1) = y_0 + h f(y_0) + \frac{h^2}{2} f'(y_0) f(y_0)$$

$$y_1(h) = y_0 + \phi(0) + h\phi'(0) + \frac{h^2}{2}\phi''(0) + O(h^3)$$

$$\phi''(h) = 2\sum_{i=1}^{s} b_i k_i'(h) + h\sum_{i=1}^{s} b_i k_i''(h)$$

$$k_i''(h) = f'(y_0 + \psi_i(h))\,\psi_i''(h)$$

$$\psi_i'(h) = \sum_{j=1}^{s} a_{ij}\, k_j(h) + h \sum_{j=1}^{s} a_{ij}\, k_j'(h)$$

$$\psi_i'(0) = \sum_{j=1}^{s} a_{ij}\, k_j(0) = \sum_{j=1}^{s} a_{ij}\, f(y_0)$$

$$\phi''(0) = 2 \sum_{i=1}^{s} b_i\, k_i'(0) = 2 \sum_{i=1}^{s} b_i\, f'(y_0) \sum_{j=1}^{s} a_{ij}\, f(y_0)$$

$$= 2 \sum_{i,j=1}^{s} b_i a_{ij}\, f'(y_0) f(y_0)$$

order 2 condition ($p = 2$):

$$1 - 2 \sum_{i,j=1}^{s} b_i a_{ij} = 0$$

$p = 3$:

$$y''' = f''(y)(y', y') + f'(y)\, y''$$

bilinear form — matrix vectors

$$f''(y) : \mathbb{R}^d \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$$

$$(u, v) \longrightarrow f''(y)(u, v)$$

$$1 \le i \le d \qquad \left(f''(y)(u,v)\right)_i = \sum_{j,k=1}^{d} \frac{\partial^2 f_i}{\partial y_j \partial y_k}(y(t))\, u_j v_k$$

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \qquad y(t) \in \mathbb{R}^d$$

RK

y_1 approx of $y(t_0+h)$

$$k_i = f\left(t_0 + c_i h,\ y_0 + h \sum_{j=1}^{s} a_{ij} k_j\right) \qquad i = 1, \dots, s$$

if $s = i-1$ explicit RK

$$y_1 = y_0 + h \sum_{i=1} b_i k_i$$

- order conditions c_i, a_{ij}, b_i so that $\underbrace{y(t_1) - y_1}_{\text{local error}} = \underbrace{O(h^{p+1})}_{\text{order } p}$

thm 1

- adaptive time step h ?
- error at final time $y(t_n) - y_n = O(h^p)$

thm 2

Adaptive time step (heuristic)

y_1 approx of $y(t_0+h)$ with an order p method : $y(t_1)-y_1 = O(h^{p+1})$

$\hat{y}_1$ — $p-1$ — : $y(t_1)-\hat{y}_1 = O(h^p)$

$$\underbrace{\|y_1-\hat{y}_1\|}_{\text{computable}} - \underbrace{\|y(t_0+h)-y_1\|}_{O(h^{p+1})} \leq \underbrace{\|y(t_0+h)-\hat{y}_1\|}_{O(h^p)} \leq \underbrace{\|y(t_0+h)-y_1\|}_{O(h^{p+1})} + \underbrace{\|y_1-\hat{y}_1\|}_{\text{computable}}$$

therefore, $\|y_1-\hat{y}_1\|$ is a computable approx of $\|y(t_0+h)-\hat{y}_1\|$ (provided h is small enough)

Given h, y_0, compute y_1 and $\hat{y}_1$

$\|(y_1-\hat{y}_1)(h)\| = Ch^p$

We can estimate $C = \dfrac{\|(y_1-\hat{y}_1)(h)\|}{h^p}$

We are looking for h_{opt} such that

$(y_1-\hat{y}_1)(h_{opt}) \simeq$ TOL given tolerance (TOL = 0,001)

We can then compute h_{opt}

$\|(y_1-\hat{y}_1)(h_{opt})\| = Ch_{opt}^p = \dfrac{\|(y_1-\hat{y}_1)(h)\|}{h^p} h_{opt}^p$

$= TOL$

$$h_{opt} = h\left(\frac{TOL}{\|(y_1-\hat{y}_1)(h)\|}\right)^{1/p}$$

Algorithm:

For each time, given h and y_0 (should be $y_n \approx y(t_n)$)

- compute $y_1, \hat{y}_1, \|y_1 - \hat{y}_1\|$, $h_{opt} = h\left(\frac{TOL}{\|(y_1 - \hat{y}_1)(\eta)\|}\right)^{1/p}$
- set $h = h_{opt}$
- If $\|y_1 - \hat{y}_1\| \leq TOL$ go to next timestep
- Else, redo time with new h

Ex: y_1: Runge scheme, Explicit RK order 2

$$k_1 = f(t_0, y_0)$$
$$k_2 = f(t_0 + h/2, y_0 + \tfrac{h}{2} k_1)$$
$$y_1 = y_0 + h k_2$$

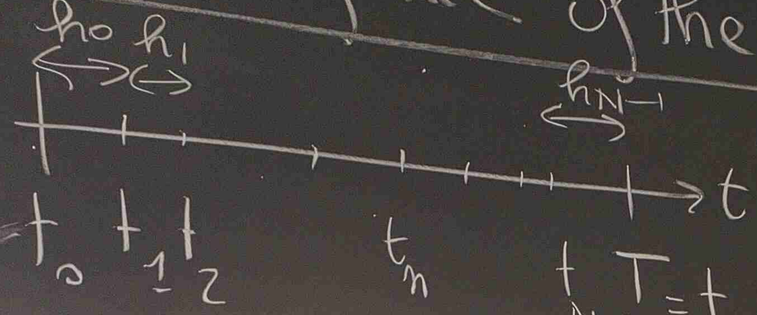
0	
1/2	1/2
	0 1

$\hat{y}_1$: Euler explicit $\quad k_1 = f(t_0, y_0)$, $\hat{y}_1 = y_0 + h k_1$

$$y_1 - \hat{y}_1 = h(k_2 - k_1)$$

matlab: ode45, $p = 4$

Convergence of the ERK



$t_{n+1} = t_n + h_n \quad n = 0, 1, \dots, N-1$

$h_{max} = \max(h_0, \dots, h_{N-1})$

The ERK can be written as:

$$Y_{n+1} = Y_n + h_n \phi(t_n, Y_n, h_n)$$

ϕ will be defined later

then, for an

if h_{max} is small enough

$\| y(t_n$

- Prove thm 2
- Explicit ϕ,

Proof:

Thm 2: (convergence)

- Assume that the local error is of order p:

$\exists C > 0 \quad \forall h \le h_{max} \quad \forall t_0 \le t \le T$

locally $\| y(t+h) - (y(t) + h\,\phi(t, y(t), h)) \| \le C h^{p+1}$

- Assume that ϕ is Lipschitz cont with respect to the second variable: $\exists \Lambda > 0 \quad \forall h \le h_{max} \quad \forall (t, z_1)$ and (t, z_2) in the neighborhood of $(t, y(t)) \quad \forall t_0 \le t \le T$

$\|\phi(t,z_1,h)-\phi(t,z_2,h)\| \le \Lambda \|z_1-z_2\|$ (*)

then, for any $n=1,\dots,N$:

if h_{max} is small enough $\|y(t_n)-y_n\| \le \frac{C}{\Lambda}\left(e^{\Lambda(t_n-t_0)}-1\right)h_{max}^p$

- Prove thm 2
- Explicit ϕ, prove (*)

Proof: $y(t_{n+1})-y_{n+1} = y(t_n+h_n)-\left(y_n+h_n\phi(t_n,y_n,h_n)\right)$

$= y(t_n+h_n)-\left(y(t_n)+h_n\phi(t_n,y(t_n),h_n)\right)$

$+\left(y(t_n)+h_n\phi(t_n,y(t_n),h_n)\right)$

$-\left(y_n+h_n\phi(t_n,y_n,h_n)\right)$

Therefore $\|y(t_{n+1})-y_{n+1}\| \le C h_n^{p+1} + \|y(t_n)-y_n\|\left(1+h_n\Lambda\right)$

$x>0 \quad 0<\theta<x$

$e^x = 1+x+\frac{x^2}{2}e^\theta$

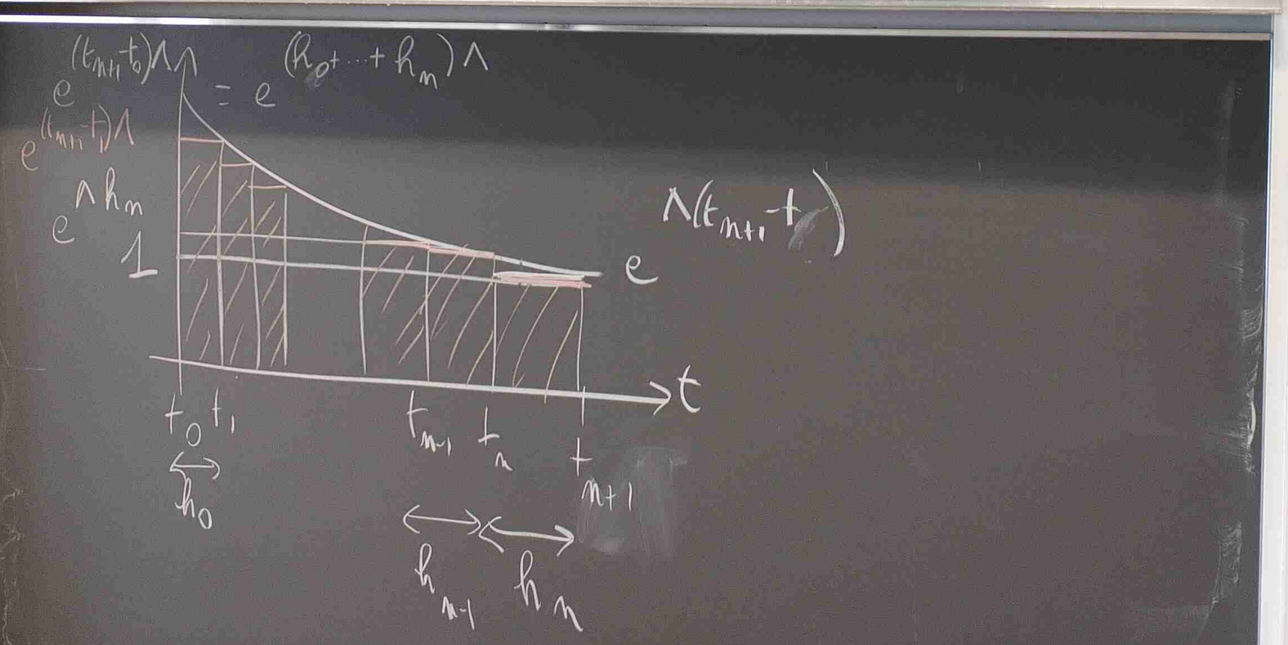
$e^x \ge 1+x$

$$\|y(t_{n+1}) - y_{n+1}\| \leq C h_n^{p+1} + C h_{n-1}^{p+1} e^{h_n \wedge} + \|y(t_{n-1}) - y_{n-1}\| e^{(h_n + h_{n-1})\wedge}$$

$$\leq C h_n^{p+1} + C h_{n-1}^{p+1} e^{h_n \wedge} + \dots + C h_0^{p+1} e^{(h_n + \dots + h_1)\wedge} + \underbrace{\|y(t_0) - y_0\|}_{h_n \wedge 0} e^{(h_n + \dots + h_0)\wedge}$$

$$\leq C h_{max}^p \left(h_n + h_{n-1} e^{\wedge} + \dots + h_0 e^{(h_n + \dots + h_1)\wedge} \right)$$

$$\leq C h_{max}^p \int_{t_0}^{t_{n+1}} e^{\wedge (t_{n+1} - t)} \, dt \; \frac{1}{\wedge} \left(e^{\wedge (t_{n+1} - t_0)} - 1 \right)$$



$$\|\phi(t,z_1,h)-\phi(t,z_2,h)\| \le \Lambda \|z_1-z_2\| \quad (*)$$

Def of ϕ? Does it satisfy the lipschitz cond?

$$k_i = f\left(t_0+c_i h,\ y_0 + h\sum_{j=1}^{i-1} a_{ij} k_j\right)$$

$$y_1 = y_0 + h\sum_{i=1}^{s} b_i k_i$$

$$\phi(t,z,h) = \sum_{i=1}^{s} b_i k_i(t,z,h)$$

$$k_i(t,z,h) = f\left(t+c_i h,\ z + h\sum_{j=1}^{i-1} a_{ij} k_j(t,z,h)\right)$$

$$k_1(t,z,h) = f(t,z)$$

$i = 2,\dots,s$

Lemma: if f is locally lipschitz cont. with respect to the second var.

$\exists L>0$ $\forall (t,z_1)$ and (t,z_2) in the neighb of $(t,y(t))$:

$$\|f(t,z_1)-f(t,z_2)\| \le L\|z_1-z_2\|$$

then there exists Λ (depending on $L, c_i, a_{ij}, b_i, h_{max}$) such that

$\forall t_0 \le t \le T$ $\forall h \le h_{max}$ $\forall (t,z_1)(t,z_2)$ in the neighb. $(t,y(t))$ we have

$$\|\phi(t,z_1,h)-\phi(t,z_2,h)\| \le \Lambda\|z_1-z_2\|$$

Proof: $\|\phi(t,z_1,h) - \phi(t,z_2,h)\| \leq \sum_{i=1}^{s} |b_i| \|k_i(t,z_1,h) - k_i(t,z_2,h)\|$

$k_i(t,z,h)$ locally lipschitz w.r. to z?

$\|k_1(t,z_1,h) - k_1(t,z_2,h)\| = \|f(t,z_1) - f(t,z_2)\|$

$\leq L\|z_1 - z_2\|$

Proceed with $k_2, k_3, \dots, k_s$.

Part 2: Finite differences for pde's

1. Introduction, motivation
2. Elliptic problems → laplace eq
3. Parabolic → heat
4. Hyperbolic → transport wave

Part 2: Finite Differences pde's

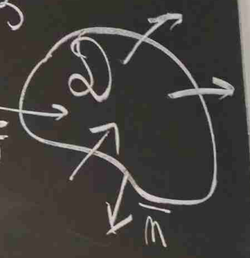
1 - Motivation
2 - Elliptic
3 - Parabolic
4 - Hyperbolic

1 - 1992 first IPCC report (1972) global warming

$\vec{v} : \mathbb{R}^3 \to \mathbb{R}^3$ $\mathcal{C}^1$

Any domain $\mathcal{D} \subset \mathbb{R}^3$

$\int_{\partial\mathcal{D}} \vec{v} \cdot \vec{n}\, ds = 0$



divergence thm $\int_{\mathcal{D}} \operatorname{div} \vec{v}\, dx = 0 \quad \forall \mathcal{D}$

thus $\operatorname{div} \vec{v} = 0$ in $\mathbb{R}^3$

Electrostatics $\vec{E}$ electrical field
ρ density of charges

$\forall \mathcal{D}$ domain of $\mathbb{R}^3$: $\int_{\partial\mathcal{D}} \vec{E} \cdot \vec{n}\, ds = \int_{\mathcal{D}} \rho\, dx$

$$\int_{\mathcal{D}} \operatorname{div} \vec{E}\, dx = \int_{\mathcal{D}} \rho\, dx \quad \forall \mathcal{D}$$

$$\operatorname{div} \vec{E} = \rho$$

$$\operatorname{curl} \vec{E} = \vec{0}$$

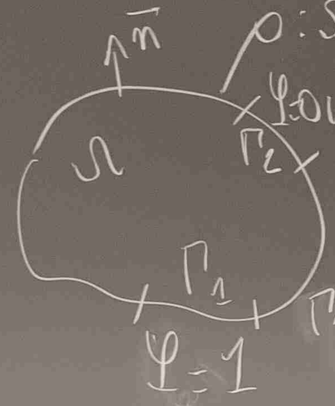
$$\vec{\nabla} \times \vec{E} = \vec{0} \rightarrow \vec{E} = -\vec{\nabla}\varphi$$

$$\vec{\nabla} = \begin{pmatrix} \partial/\partial x_1 \\ \partial/\partial x_2 \\ \partial/\partial x_3 \end{pmatrix}$$

$$-\operatorname{div} \vec{\nabla}\varphi = \rho$$

Pbm to be solved

data : Ω comput. domain $\subset \mathbb{R}^3$



φ unknown, $\varphi : \Omega \to \mathbb{R}$ such that

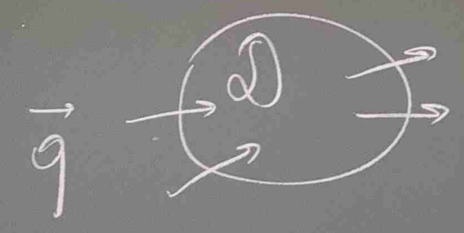
$$\begin{cases} -\Delta\varphi = \rho & \text{in } \Omega \\ \varphi = 1 & \text{on } \Gamma_1 \\ \varphi = 0 & \text{on } \Gamma_2 \\ \nabla\varphi \cdot n = 0 & \text{on } \Gamma_3 = \partial\Omega \setminus (\Gamma_1 \cup \Gamma_2) \end{cases}$$

model problem : $d = 1, 2, 3$

$$\begin{cases} -\Delta u = f & \Omega =]0,1[^d \\ u = 0 & \partial\Omega \end{cases}$$

section 2

heat transfer :



$\vec{q}$: heat flux

f : heat source

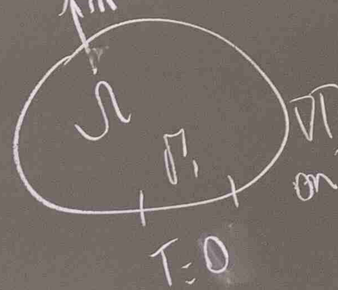
$$\int_{\partial\mathcal{D}} \vec{q} \cdot \vec{n}\, ds = \int_{\mathcal{D}} f\, dx \quad \forall \mathcal{D}$$

$$\operatorname{div} \vec{q} = f$$

but $\vec{q} = -k\vec{\nabla}T$ T : temperature, k : diffusion coeff

$$-\operatorname{div}(k\vec{\nabla}T) = f$$

Pbm to be solved



$\nabla T \cdot n = 0$ on $\partial\Omega \setminus \Gamma_1$

$$\begin{cases} -\operatorname{div}(k\vec{\nabla}T) = f & \text{in } \Omega \\ T = 0 & \text{on } \Gamma_1 \\ k\vec{\nabla}T \cdot \vec{n} = 0 & \text{on } \partial\Omega \setminus \Gamma_1 \end{cases}$$

heat equation (heat transfer with time)

$$\frac{d}{dt}\int_\Omega \underbrace{\rho c_p T}_{\text{enthalpy}}\,dx + \int_{\partial\Omega} \underbrace{-k\nabla T\cdot n}_{\text{heat flux}}\,ds = \int_\Omega f\,dx \qquad \forall \Omega$$

rate of change energy

$$\frac{\partial}{\partial t}(\rho c_p T) - \operatorname{div}(k\nabla T) = f \quad \text{in } \mathbb{R}^3$$

$\rho c_p = 1$ $k = 1$ $T = u(x,t)$ $\frac{\partial u}{\partial t} - \Delta u = f$ heat eq. (parabolic)

hyperbolic: wave eq.

Electromagnetics in vaccum

$$\vec{\nabla}\times\vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

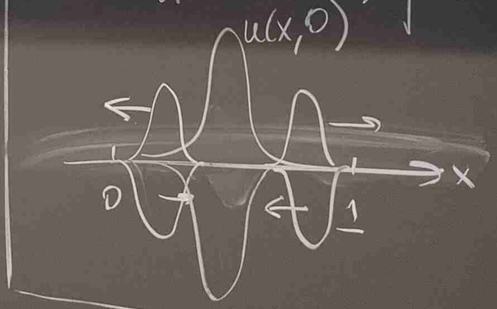
$$\vec{\nabla}\times\vec{B} = \mu_0\varepsilon_0\frac{\partial \vec{E}}{\partial t}$$

$$\vec{\nabla}\times\vec{\nabla}\times\vec{E} = -\frac{\partial}{\partial t}\vec{\nabla}\times\vec{B} = -\mu_0\varepsilon_0\frac{\partial^2\vec{E}}{\partial t^2}$$

since $\operatorname{div}\vec{E} = 0$

$$\mu_0\varepsilon_0\frac{\partial^2\vec{E}}{\partial t^2} - \Delta\vec{E} = \vec{0}$$

$d = 1$ $\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$ elastic rope

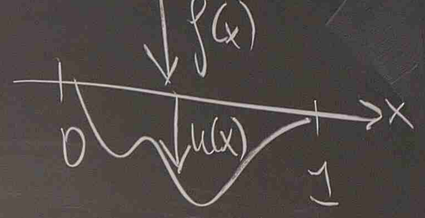


2. Elliptic problems

d = 1 model problem

given $f:[0,1]\to\mathbb{R}$ find $u:[0,1]\to\mathbb{R}$ such that

$$\begin{cases} -u''(x) = f(x) & 0<x<1 \\ u(0)=0 \quad u(1)=0 \end{cases}$$



Remarks:

* exact solution

$$u(x) = x\int_0^1 dt\int_0^t f(s)\,ds - \int_0^x dt\int_0^t f(s)\,ds$$

if $f\in C^0[0,1]$ then $u\in C^2[0,1]$

$$u(x) = (1-x)\int_0^x t f(t)\,dt + x\int_x^1 (1-t)f(t)\,dt$$

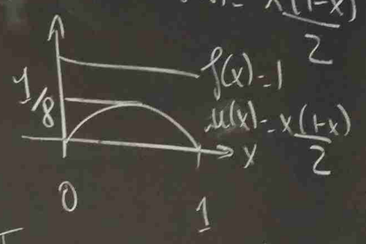
thus $u(x) = \int_0^1 G(x,t) f(t)\,dt$

$G(x,t) = (1-x)t \quad 0 \le t \le x$

$= x(1-t) \quad x \le t \le 1$

if $f \ge 0$ then $u \ge 0$ and $\|u\|_{L^\infty(0,1)} \le \frac{1}{8}\|f\|_{L^\infty(0,1)}$

* $f(x) = 1$ $u(x) = \frac{x(1-x)}{2}$



* Fourier series: if $f(x) = \sum_{k=1}^{N} a_k \sin(k\pi x)$

then $u(x) = \sum_{k=1}^{N} \frac{a_k}{k^2\pi^2} \sin(k\pi x)$

* eigenvalues $u(x) \not\equiv 0$

$\begin{cases} -u''(x) = \lambda u(x) \\ u(0) = 0 \quad u(1) = 0 \end{cases}$

$u(x) = \sin(k\pi x) \quad k = 1, 2, \dots$

$\lambda = k^2\pi^2$

* if we multiply $-u'' = f$ by u, $\int_0^1 dx$, int. by parts, use $u(0) = u(1) = 0$

$\int_0^1 (u'(x))^2\,dx = \int_0^1 f(x)u(x)\,dx$

deformation work forces

uniqueness: assume that u_1 and u_2 are solutions

$-(u_1 - u_2)''(x) = 0 \quad 0 < x < 1$

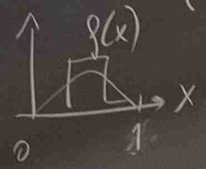
$(u_1 - u_2)(0) = 0 \quad (u_1 - u_2)(1) = 0$ then

$\int_0^1 \left((u_1 - u_2)'(x)\right)^2 dx = 0 \quad (u_1 - u_2)(x) = cst$

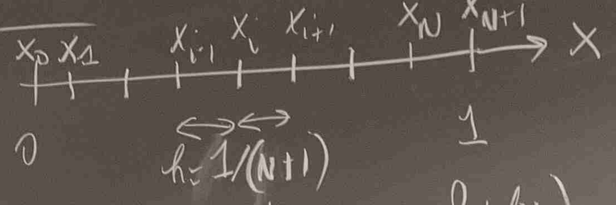
but since $(u_1 - u_2)(0) = 0$

$cst = 0$

existance in $\{v \in L^2(0,1),\ v' \in L^2(0,1),\ v(0)=0=v(1)\}$



Finite differences



$x_i = ih \quad i = 0, \dots, N+1$

Goal: to compute approximations u_i of $u(x_i)$ $i = 1, \dots, N$

using Taylor expansions

$x_{i+1} = x_i + h \qquad x_{i-1} = x_i - h \qquad i = 1, \dots, N$

assume that $u \in \mathcal{C}^4[0,1]$, $\forall i = 1, \dots, N$, $\exists\ x_i < \alpha_i < x_{i+1}$ such that $x_{i-1} < \beta_i < x_i$

$$u(x_{i+1}) = u(x_i) + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u^{(4)}(\alpha_i)$$

$$u(x_{i-1}) = u(x_i) - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u^{(4)}(\beta_i)$$

$$u''(x_i) = \frac{u(x_{i-1}) - 2u(x_i) + u(x_{i+1})}{h^2} + r_i \quad \text{with } |r_i| \leq \frac{1}{12}\max_{0 \leq x \leq 1}|u^{(4)}(x)|\,h^2$$

Since $-u''(x_i) = f(x_i) \qquad i = 1, \dots, N$

Then $\frac{-u(x_{i-1}) + 2u(x_i) - u(x_{i+1})}{h^2} = f(x_i) + r_i$

The scheme is obtained replacing $u(x_i)$ by u_i and removing r_i:

$$\begin{cases} -\frac{u_{i-1} + 2u_i - u_{i+1}}{h^2} = f(x_i) & i = 1, \dots, N \\ u_0 = 0 \quad u_{N+1} = 0 \end{cases}$$

This corresponds to solving the linear system

$$A\vec{u} = \vec{f}$$

$$\vec{f} = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{pmatrix} \quad \vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} \quad A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & (0) \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2 \end{pmatrix}$$

tridiagonal matrix

known — unknown

Lemma: A is symmetric positive definite, therefore

$$\forall \vec{f} \in \mathbb{R}^N \ \exists!\, \vec{u} \in \mathbb{R}^N \text{ s.t. } A\vec{u} = \vec{f}$$

Proof: $\forall \vec{v} \in \mathbb{R}^N$, $\vec{v}^T A \vec{v} = (\vec{v}, A\vec{v})$

$$(v_1 \dots v_N) \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}$$

$$= \frac{1}{h^2} \left(2v_1^2 - 2v_1 v_2 + 2v_2^2 + \dots - 2v_N v_{N-1} + 2v_N^2 \right)$$

$$= \frac{1}{h^2} \left(v_1^2 + (v_1 - v_2)^2 + \dots + (v_N - v_{N-1})^2 + v_N^2 \right) \geq 0$$

$$= 0 \iff (v_1 = v_2 = \dots v_N = 0)$$

invertible $\operatorname{Ker} A = 0$

$A\vec{v} \Rightarrow \vec{v}^T A \vec{v} = 0 \Leftrightarrow \vec{v} = \vec{0}$

Convergence? $u_i \xrightarrow[h \to 0]{N \to \infty} u(x_i)$?

$\vec{U} = \begin{pmatrix} u(x_1) \\ \vdots \\ u(x_N) \end{pmatrix}$ $A\vec{U} = \vec{f} + \vec{r}$ $\vec{r} = \begin{pmatrix} r_1 \\ \vdots \\ r_N \end{pmatrix}$

error $\vec{U} - \vec{u}$ satisfies $A(\vec{U} - \vec{u}) = \vec{r}$ (*)

Consistency

Lemma (stability): $A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & (0) \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ (0) & & -1 & 2 \end{pmatrix}$

$\forall \vec{g} \in \mathbb{R}^N$ let $\vec{v} \in \mathbb{R}^N$ s.t. $A\vec{v} = \vec{g}$

then $\|\vec{v}\|_\infty \leq \frac{1}{8} \|\vec{g}\|_\infty$

$\left(\|\vec{v}\|_\infty = \max_{1 \leq i \leq N} |v_i| \right)$

use this lemma in (*):

$$\max_{1 \leq i \leq N} |u(x_i) - u_i| = \|\vec{U} - \vec{u}\|_\infty \leq \frac{1}{8} \|\vec{r}\|_\infty \leq \boxed{\frac{1}{8} \frac{1}{12} \max_{0 \leq x \leq 1} |u^{(4)}(x)|} h^2$$

C

Thm (convergence in $\|\cdot\|_\infty$):

if $u \in \mathcal{C}^4[0,1]$ (true if $f \in \mathcal{C}^2[0,1]$) then $\exists C > 0 \; \forall 0 < h \leq \frac{1}{N+1}$

$$\max_{0 \leq x \leq 1} |u(x_i) - u_i| \leq C h^2$$

Rem: • C dep on u, not on h

• set u, compute f, then the error

$$\max_{1 \le i \le N} |u(x_i) - u_i|$$

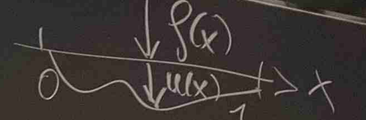
is (at least) divided by $2^2 = 4$ when h is divided by 2

$10^2 = 100$

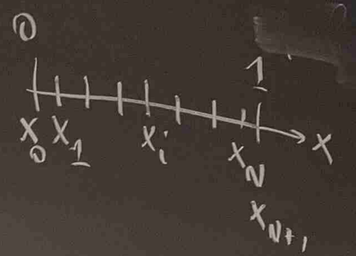
$N+1$ mult

2 - Elliptic pbms $d=1$

$$\begin{cases} -u''(x) = f(x) & 0<x<1 \\ u(0)=0 \quad u(1)=0 \end{cases}$$



u_i approx $u(x_i)$ $i=1,\dots,N$



FD scheme

$$A\vec{u} = \vec{f} \qquad \vec{f} = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{pmatrix} \qquad \vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix}$$

$x_i = ih$ $i = 0, 1, \dots, N, N+1$

$h = \frac{1}{N+1}$

$$A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} \qquad \vec{U} = \begin{pmatrix} u(x_1) \\ \vdots \\ u(x_N) \end{pmatrix}$$

$$A\vec{U} = \vec{f} + \vec{r}$$

$$\|\vec{r}\|_\infty \leq \frac{1}{12}\max_{0\leq x\leq 1}|u^{(4)}(x)|\,h^2$$

$$A(\vec{U}-\vec{u}) = \vec{r}$$

Lemma: $\forall \vec{g} \in \mathbb{R}^N$ let $\vec{v}$ s.t. $A\vec{v} = \vec{g}$ $\quad \|\vec{v}\|_\infty \leq \frac{1}{8}\|\vec{g}\|_\infty$

$$\|\vec{U}-\vec{u}\|_\infty \leq \frac{1}{8}\|\vec{r}\|_\infty \leq \frac{1}{8\cdot 12}\max_{0\leq x\leq 1}|u^{(4)}(x)|\,h^2 = \max_{1\leq i\leq N}|v_i|$$

Thm: If $u \in \mathcal{C}^4([0,1])$ $\exists C>0$ $\forall 0<h\leq 1$ $\max_{1\leq i\leq N}|u(x_i)-u_i| \leq Ch^2$

Proof lemma: uses 2 properties

- $\exists \bar{w}$ s.t $A\bar{w} = \bar{1}$ $\quad w_i = \frac{x_i(1-x_i)}{2}$

(intuition: $\begin{cases} -w''(x) = 1 & 0<x<1 \\ w(0) = 0 = w(1) \end{cases}$ $\quad w(x) = \frac{x(1-x)}{2}$)

$A\bar{w} = \bar{0}$ (remainder is 0)

$\bar{w} = \begin{pmatrix} w(x_1) \\ \vdots \\ w(x_N) \end{pmatrix}$

Can check it

- Positivity $(A\vec{z} \geqslant \vec{0}) \Rightarrow (\vec{z} \geqslant \vec{0})$

(at the cont. level: if $f \geqslant 0$ then $u \geqslant 0$)

Let us uses these 2 properties by def of $\|\cdot\|_\infty$

$-\|\vec{q}\|_\infty \leqslant q_i \leqslant \|\vec{q}\|_\infty \quad i = 1, \dots, N$

$-\vec{1}\,\|\vec{q}\|_\infty \leqslant \vec{q} \leqslant \vec{1}\,\|\vec{q}\|_\infty$

$(A\bar{w} = \vec{1})\|\vec{q}\|_\infty \quad A(\bar{w}\|\vec{q}\|_\infty) = \vec{1}\|\vec{q}\|_\infty$

$A(\bar{w}\|\vec{q}\|_\infty + \vec{v}) = \vec{1}\|\vec{q}\| + \vec{q} \geqslant \vec{0}$

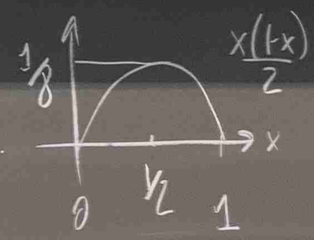
$A(\bar{w}\|\vec{q}\|_\infty - \vec{v}) = \vec{1}\|\vec{q}\|_\infty - \vec{q} \geqslant \vec{0}$

Using positivity:

$-\bar{w}\|\vec{q}\|_\infty \leqslant \vec{v} \leqslant \bar{w}\|\vec{q}\|_\infty$

$\|\vec{v}\|_\infty \leqslant \|\bar{w}\|\,\|\vec{q}\|_\infty$

$\leqslant \frac{1}{8}\|\vec{q}\|_\infty$



Proof of positivity: let $\vec{z} \in \mathbb{R}^N$ s.t. $A\vec{z} \geqslant \vec{0}$ we want to prove that $\vec{z} \geqslant \vec{0}$

let z_k s.t $z_k \leqslant z_j \quad j = 1, \dots, N \quad \left(z_k = \min_{1 \leqslant j \leqslant N} z_j\right)$

It suffices to prove $z_k \geqslant 0$

If $k=1$ $\quad 2z_1 - z_2 \geq 0$

$2z_1 \geqslant z_2 \geqslant z_1$

thus $z_1 \geqslant 0$

If $k=N$ $\quad 2z_N - z_{N-1} \geqslant 0$

$2z_N \geqslant z_{N-1} \geqslant z_N$

thus $z_N \geqslant 0$

If $2 \leq k \leq N-1$

$2z_k - z_{k-1} - z_{k+1} \geqslant 0$

$2z_k \geqslant z_{k-1} + z_{k+1} \geqslant z_{k-1} + z_k$

thus $z_k \geqslant z_{k-1}$ but z_k is the min!

therefore $z_k = z_{k-1}$, continue until reach index 1.

Remark:

- $\bar{A}\bar{w} \geqslant \bar{1}$ is sufficient in the proof
- Positivity if true in the following case:

  $a_{ii} > 0 \quad a_{ij} \leqslant 0 \quad j \neq i \quad a_{ii} \geqslant \sum_{j \neq i} |a_{ij}|$, strictly on at least one line

  A is irreducible (the graph of A is strongly connected)

Ex:

$$\begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 3 & -2 & & & \\ -1 & & -2 & & \\ & & & & \\ & & & & -2 \\ & & & -1 & 3 \end{pmatrix}$$

corresponds to $-u''(x) = f(x)$
$u(0) = 0$
$u(1) = 0$

$-u''(x) + u'(x) = f(x)$

not true

$$\begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & & & & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}$$

$\overline{1}$ is the kernel
corresponds to $-u''(x) = f(x)$
$u'(0) = 0 \quad u'(1) = 0$
if $u(x)$ is solution $u(x)+C$ is also

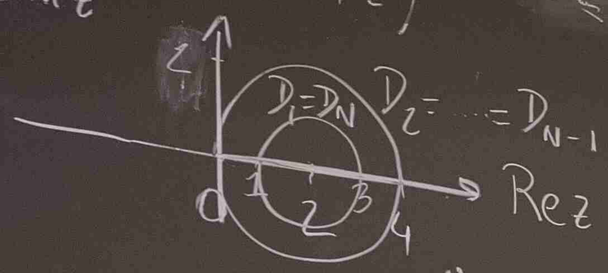
Another interesting result: Gerschgorin thm

$A \in \mathbb{R}^{N \times N} \quad A\overline{\Psi} = \lambda\overline{\Psi} \quad \overline{\Psi} \neq \overline{0}$

$\lambda \in \bigcup_{i=1}^{N} D_i \quad D_i$ Gerschgorin disk $D_i = \{z \in \mathbb{C};\ |a_{ii} - z| \leq \sum_{\substack{j=1 \\ j \neq i}}^{N} |a_{ij}|\}$

$$A = \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$$

$0 \leq \lambda \leq 4$



λ eigenvalue $\Rightarrow \lambda \in \bigcup_{i=1}^{N} D_i$

but A sym $\Rightarrow \lambda$ is real thus $0 < \lambda \leq 4$

Proof: $A\underline{\varphi} = \lambda \underline{\varphi}$

line i: $\left|(a_{ii} - \lambda)\varphi_i\right| = \left|\sum_{\substack{j=1 \\ j \neq i}}^{N} a_{ij}\varphi_j\right| \leq \sum_{j \neq i} |a_{ij}|\,|\varphi_j|$

let i be s.t. $|\varphi_i| \geqslant |\varphi_j| \quad j = 1, \dots, N$

$$|a_{ii} - \lambda| \leqslant \sum_{j \neq i} |a_{ij}| \underbrace{\frac{|\varphi_j|}{|\varphi_i|}}_{\leq 1} \qquad \lambda \in D_i \text{ for some } i$$

Thm (convergence in the $\|\cdot\|_2$ norm)

If $u \in \mathcal{C}^4[0,1]$ then $\exists C > 0$ $\forall\, 0 < h \leq 1$

$$\underbrace{\left(\sum_{i=1}^{N} (u(x_i) - u_i)^2\right)^{1/2}}_{\|\vec{v} - \vec{u}\|_2} \leqslant C h^{3/2} \qquad h = \frac{1}{N+1}$$

Proof:
- easy: we convergence $\|\cdot\|_\infty$ norm

$$\left(\sum_{i=1}^{N} (u(x_i) - u_i)^2\right)^{1/2} \leqslant \left(N \max_{1 \leq i \leq N} |u(x_i) - u_i|^2\right)^{1/2} \leqslant \sqrt{N}\, C h^2 \leqslant C h^{3/2}$$

$$C = \frac{1}{8} \frac{1}{12} \max_{0 \leq x \leq 1} |u^{(4)}(x)|$$

- using the eigenvalues of A

Since A is spd there exists an orthonormal basis of eigenvectors $\underline{\varphi}_i$ ($\underline{\varphi}_i^T \underline{\varphi}_j = \delta_{ij}$ $i,j = 1, \dots, N$) and strictly positive eigenvalues

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \dots \leqslant \lambda_N$$

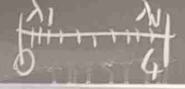
Moreover $\forall \vec{v} \in \mathbb{R}^N$

$$\lambda_1 \|\vec{v}\|_2^2 \leq \vec{v}^T A \vec{v} \leq \lambda_N \|\vec{v}\|_2^2$$

(expand $\vec{v}$ in the basis of $\vec{\varphi}_i$, use $A\vec{\varphi}_i = \lambda_i \vec{\varphi}_i$ $i = 1, \dots, N$)

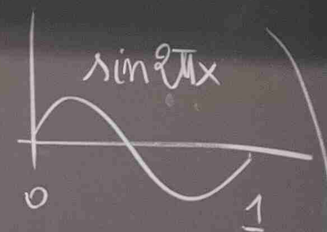
For the $\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$ matrix we have an explicit formula for the λ_i

$$\lambda_i = 2\left(1 - \cos\frac{\pi i}{N+1}\right) \quad i = 1, \dots, N$$



$$\frac{1}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} \quad \text{eigenvalues } \lambda_i = 2\left(1 - \cos\frac{\pi i}{N+1}\right)(N+1)^2$$

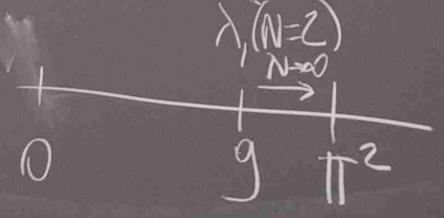
(intuition: $-\varphi''(x) = \lambda\varphi(x)$ $0 < x < 1$ $\varphi(0) = 0$ $\varphi(1) = 0$ $\quad \varphi_k(x) = \sin(k\pi x)$ $\quad \lambda_k = k^2\pi^2$ $\quad \lambda_1 = \pi^2$ $\quad k \to \infty$ $\lambda_k = +\infty$



back to

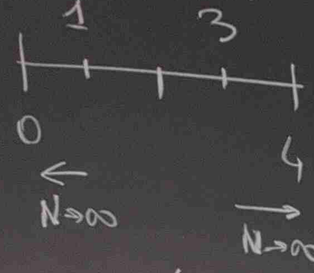
$$\lambda_1 = 2\left(1 - \cos\frac{\pi}{N+1}\right)(N+1)^2 \simeq \frac{\pi^2}{2(N+1)^2}(N+1)^2 \simeq \pi^2$$

in fact



$N=2 \quad A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$

$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ -1 \end{pmatrix}$



$A = 9 \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad \lambda_1 = 9 \qquad \lambda_2 = 27$

back to the proof of the theorem

$A(\vec{U}-\vec{u}) = \vec{r} \qquad \|\vec{r}\|_\infty \leq Ch^2$

scalar prod with $\vec{U}-\vec{u}$

$$\lambda_1 \|\vec{U}-\vec{u}\|_2^2 \leq (\vec{U}-\vec{u})^T A(\vec{U}-\vec{u}) = \vec{r}^T(\vec{U}-\vec{u}) \leq \|\vec{r}\|_2 \, \|\vec{U}-\vec{u}\|$$

$$\text{therefor } \|\vec{U}-\vec{u}\|_2 \leq \frac{\|\vec{r}\|_2}{\lambda_1} \leq \frac{\frac{1}{12}\max_{0\leq x\leq 1}|u^{(4)}(x)|\sqrt{N}\,h^2}{\lambda_1}$$

but $\lambda_1 \geqslant 9$ for $N \geqslant 2$ thus $\leq \frac{1}{12\cdot 9}\max_{0\leq x\leq 1}|u^{(4)}(x)|\,h^{3/2} \leq C h^{3/2}$

Thm: if $u \in \mathcal{C}^4[0,1]$ $\exists C>0$ $\forall 0<h\leq 1/3$

$$\left((\bar{U}-\bar{u})^T A(\bar{U}-\bar{u})\right)^{1/2} \leq C h^{3/2}$$

Proof:

$$A(\bar{U}-\bar{u}) = \bar{r}$$

$$(\bar{U}-\bar{u})^T A(\bar{U}-\bar{u}) = \bar{r}^T(\bar{U}-\bar{u}) \leq \|\bar{r}\|_2 \, \|\bar{U}-\bar{u}\|_2$$

$$\leq \frac{1}{12} \max_{0<x<1} |u^{(4)}(x)| \sqrt{N} \, \frac{1}{\sqrt{\lambda_1}} \left((\bar{U}-\bar{u})^T A(\bar{U}-\bar{u})\right)^{1/2}$$

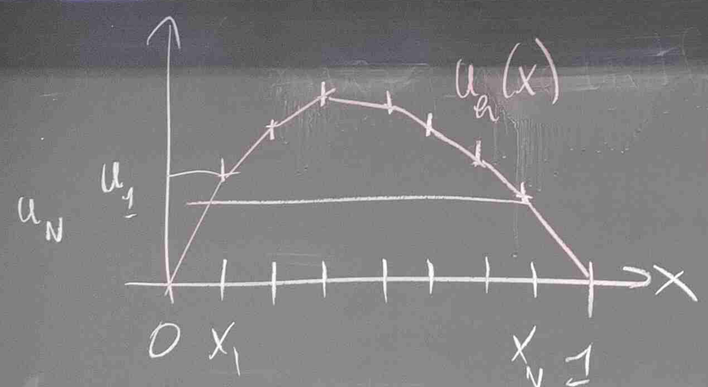
Link with finite elements

in the num approx of pde course:

$$\left(\int_0^1 (u(x)-u_h(x))^2 dx\right)^{1/2} \leq C h^2$$

and $\left(\int_0^1 (u'(x)-u_h'(x))^2 dx\right)^{1/2} \leq Ch$

What is u_h?



For the $-u''(x) = f(x)$ problem Finite Elements and Finite Differences lead to the same linear system $A\bar{u} = \vec{f}$

The two thm $\|\cdot\|_2$ correspond to these estimates

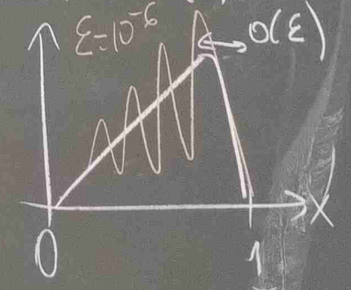
Extensions:

$$\begin{cases} -u''(x) + c(x)u(x) = f(x) & 0<x<1 \\ u(0) = 0 = u(1) \end{cases}$$

$c(x) \geq 0$ $(A\bar{w} \geq \bar{1})$

$$\begin{cases} -u''(x) = f(x) \\ u(0) = 0 \quad u'(1) = 0 \end{cases}$$

$$\begin{cases} \varepsilon u''(x) + u'(x) = 1 \\ u(0) = 0 \quad u(1) = 0 \end{cases} \qquad \varepsilon << 1$$



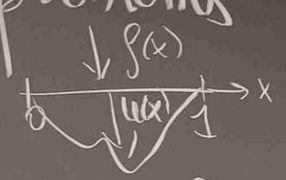
16/53 answers

1 3 8 4

Part 2 FD pde's

Elliptic problems

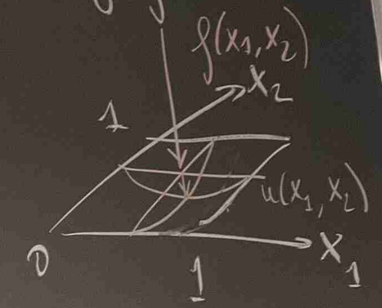
$d=1$



$-u''(x) = f(x)$

today $d=2,3$

$d=2$ def of a membrane



Given $f: [0,1]^2 \to \mathbb{R}$ find $u: [0,1]^2 \to \mathbb{R}$ s.t.

$$\begin{cases} -\Delta u(x_1,x_2) = f(x_1,x_2) & \forall (x_1,x_2) \in]0,1[^2 \\ u(x_1,x_2) = 0 & \forall (x_1,x_2) \in \partial(]0,1[^2) \end{cases}$$

$\Delta u = \operatorname{div} \operatorname{grad} u \qquad \vec{\nabla} u = \begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \frac{\partial u}{\partial x_2} \end{pmatrix}$

$= \vec{\nabla} \circ \vec{\nabla} u$

$= \frac{\partial^2 u}{\partial x_1} + \frac{\partial^2 u}{\partial x_2}$

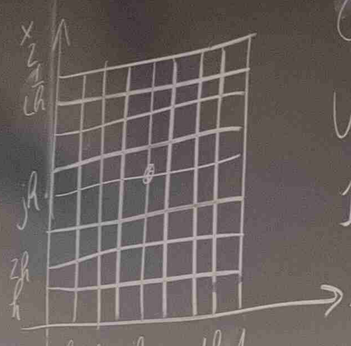
Finite Diff. grid

grid size $h = \frac{1}{L+1}$

Goal: to compute

U_{ij} approx of $u(ih, jh)$

$1 \le i, j \le L$



$$-\Delta u(ih, jh) = f(ih, jh) \qquad 1 \le i, j \le L$$

FD approx of $\frac{\partial^2 u}{\partial x_1}(ih, jh)$ $\frac{\partial^2 u}{\partial x_2}(ih, jh)$

$$\frac{\partial^2 u}{\partial x_1}(ih, jh) = \qquad (u \in \mathcal{C}^4([0,1]^2))$$

$$\frac{u((i-1)h, jh) - 2u(ih, jh) + u((i+1)h, jh)}{h^2} + O(h^2)$$

$$\frac{\partial^2 u}{\partial x_2}(ih, jh) = \frac{u(ih, (j-1)h) - 2u(ih, jh) + u(ih, (j+1)h)}{h^2} + O(h^2)$$

$$\frac{4u(ih, jh) - u((i-1)h, jh) - u((i+1)h, jh) - u(ih, (j-1)h) - u(ih, (j+1)h)}{h^2} = f(ih, jh) + O(h^2)$$

The scheme is obtained by replacing $u(ih, jh)$ by U_{ij} and forgetting the $O(h^2)$ term

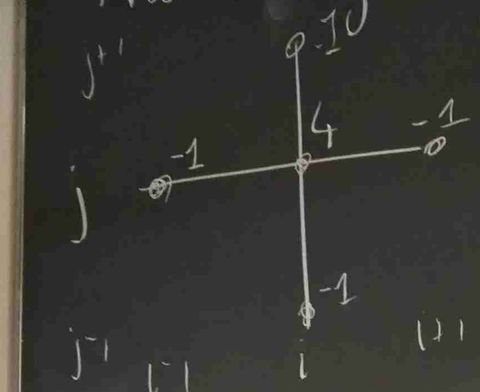
$$\frac{4U_{ij} - U_{i-1,j} - U_{i+1,j} - U_{i,j-1} - U_{i,j+1}}{h^2} = f(ih, jh) \qquad (1 \le i, j \le L)$$

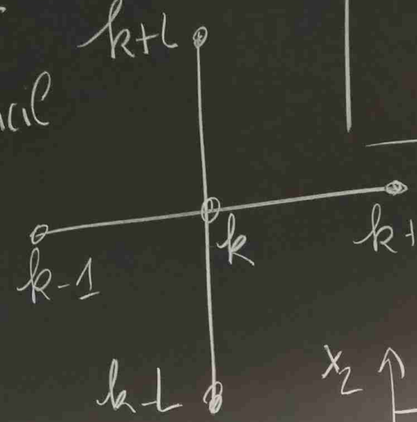
For all $i, j = 1, \dots, L$

Boundary conditions:

$U_{ij} = 0$ if $i = 0$ or $i = L+1$ or $j = 0$ or $j = L+1$

Numebering row by row
5 points stencil



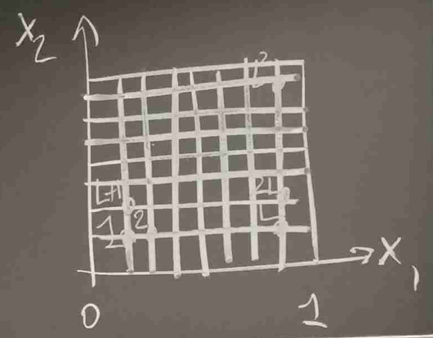


k=0
i=1,L
j=1,L
k=k+1

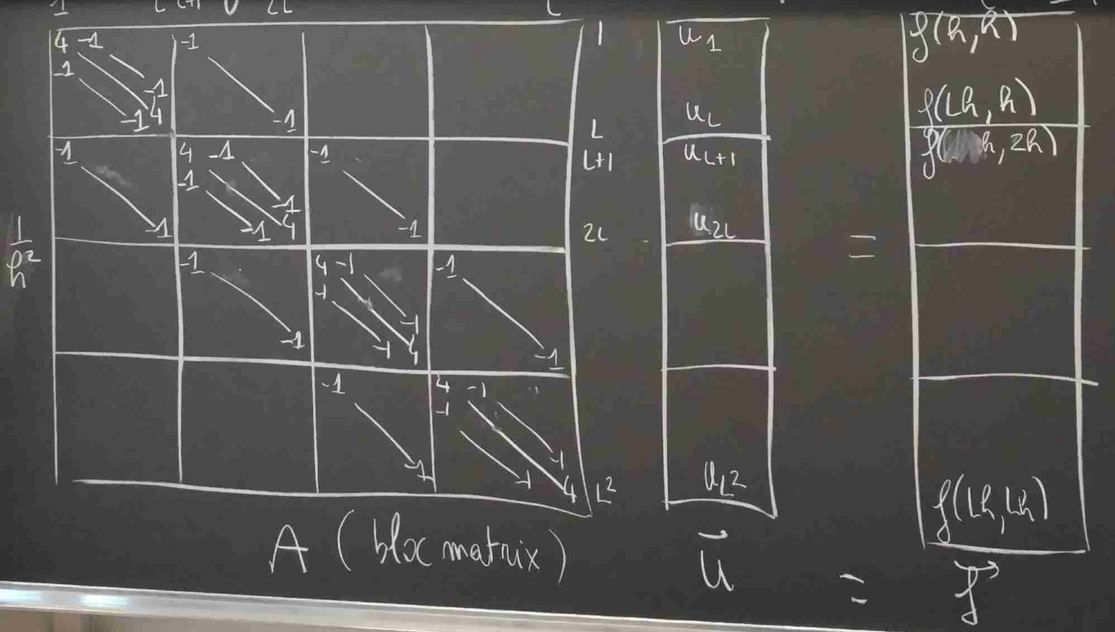
(i,j) numbering

$U_{ij} \rightarrow u_k$

k numbering
$k = 1, \dots, L^2$



Linear system to be solved (corresponding to the scheme)
(L = 4)



A (blx matrix)

$\vec{u}$ =

$\vec{f}$

If $u \in \mathcal{C}^4([0,1]^2)$ it can be proven that

$$\max_{1 \le i,j \le L} |u(ih, jh) - U_{ij}| \le C h^2$$ (C depends only on u)

(A is monotone, $A\vec{z} \ge \vec{0} \Rightarrow \vec{z} \ge \vec{0}$)

2 questions:

- implementation of $A\vec{u} = \vec{f}$ with matlab / octave / python (numpy)
- how to solve efficiently $A\vec{u} = \vec{f}$?

Implementation: using the kronecker product (kron in matlab)

$$A = \frac{1}{h^2}\left(D_L \otimes I_L + I_L \otimes D_L \right)$$

$$D_L = \begin{pmatrix} 2 & -1 & & (0) \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2 \end{pmatrix}_{L \times L} \qquad I_L = \begin{pmatrix} 1 & & (0) \\ & \ddots & \\ (0) & & 1 \end{pmatrix}_{L \times L}$$

$B, C \in \mathbb{R}^{L\times L}$ $B \otimes C \in \mathbb{R}^{L^2 \times L^2}$

$$B \otimes C = \begin{pmatrix} b_{11}C & b_{12}C & \cdots & b_{1L}C \\ \vdots & \ddots & & \vdots \\ b_{L1}C & \cdots & & b_{LL}C \end{pmatrix}$$

$$A = \frac{1}{h^2}\left(\begin{pmatrix} 2I_L & -I_L & & \\ -I_L & 2I_L & -I_L & \\ & -I_L & 2I_L & -I_L \\ & & -I_L & 2I_L \end{pmatrix} + \begin{pmatrix} D_L & & & \\ & D_L & & \\ & & D_L & \\ & & & D_L \end{pmatrix} \right)$$

$d=3$: $-\Delta u(x_1,x_2,x_3) = f(x_1,x_2,x_3) \quad \forall (x_1,x_2,x_3) \in \Omega = (0,1)^3$

$u(x_1,x_2,x_3) = 0 \quad \in \partial(0,1)^3$

$$\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_1^2}$$

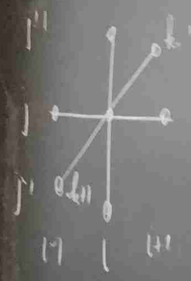
FD grid:

$U_{i,j,k}$ approx $u(ih, jh, kh)$;

$i,j,k = 1,\dots,L$

Scheme: $\dfrac{6U_{i,j,k} - U_{i-1,j,k} - U_{i+1,j,k} - U_{i,j-1,k} - U_{i,j+1,k} - U_{i,j,k-1} - U_{i,j,k+1}}{h^2} = f(ih, jh, kh)$

7 points stencil



$i,j,k = 1,\dots,L$

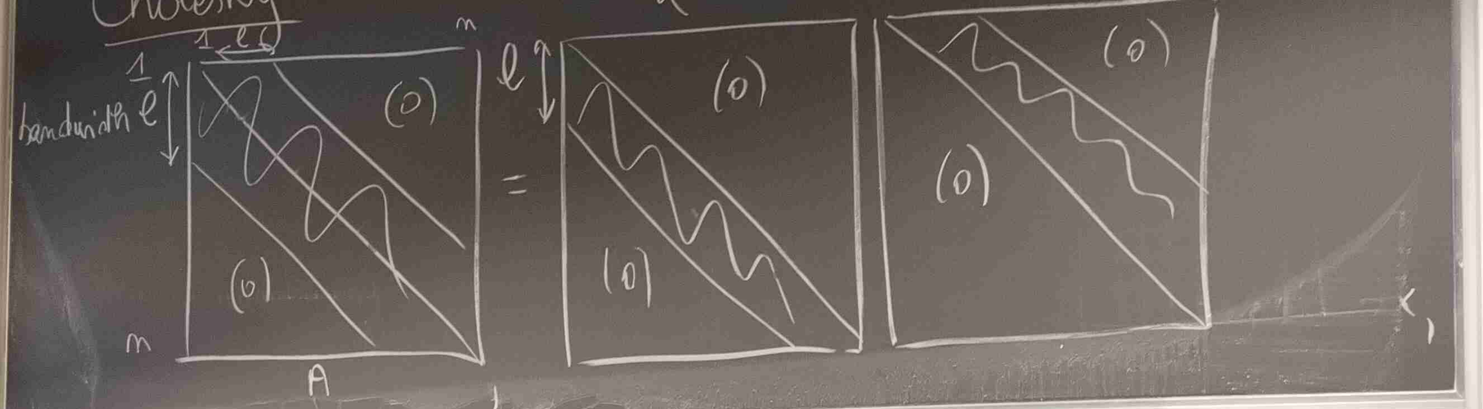
$U_{i,j,k} = 0$ if $i=0$ or $i=L+1$ or $j=0$ or $j=L+1$ or $k=0$ or $k=L+1$

numbering row by row, slice by slice

$A \in \mathbb{R}^{L^3}$ $A = \frac{1}{h^2}\left(D_L \otimes I_L \otimes I_L + I_L \otimes D_L \otimes I_L + I_L \otimes I_L \otimes D_L \right)$

Solve the linear system: direct methods (Cholesky)
or iterative methods (gradient, conjugate gradient)

Cholesky



unknown : n

memory : $n\ell$ (Cholesky dec. fills the band)

CPU(complexity) : $n\ell^2$ ($i=1,n$ $j=1,\ell$ $k=1,\ell$)

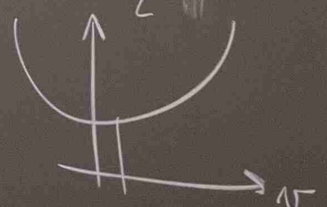
Gradient

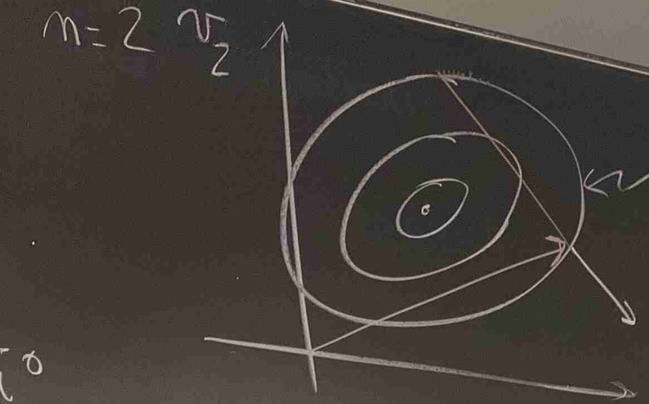
Aspd

$(A\bar{u}=\bar{f}) \Leftrightarrow (J(\bar{u}) \leq J(\bar{v})\ \forall \bar{v} \in \mathbb{R}^n)$

$J(\bar{v}) = \frac{1}{2}\bar{v}^T A \bar{v} - \bar{f}^T \bar{v}$

$n=1$ $\min \frac{1}{2} a v^2 - bv$





$n = 2$

$\vec{u}^k$ approx of $\vec{u}$, $A\vec{u} = \vec{f}$

$J'(\vec{v}) = A\vec{v} - \vec{f}$

$\vec{u}^0$

$k = 0, 1, 2, \dots$

$\vec{r}^k = \vec{f} - A\vec{u}^k$ (residual, stop if $\|\vec{r}^k\| < \epsilon$)

$\vec{u}^{k+1} = \vec{u}^k + \alpha^k \vec{r}^k$

α^k is such that $J(\vec{u}^k + \alpha^k \vec{r}^k) \leq J(\underbrace{\vec{u}^k + \alpha \vec{r}^k}_{\ell(\alpha)}) \quad \forall \alpha \in \mathbb{R}$

α^k is s.t. $\ell'(\alpha^k) = 0$

$\forall \alpha \in \mathbb{R} \quad \ell'(\alpha) = J'(\vec{u}^k + \alpha \vec{r}^k)^T \vec{r}^k$

$= \left(A(\vec{u}^k + \alpha \vec{r}^k) - \vec{f}\right)^T \vec{r}^k$

$$\alpha_k = \frac{\vec{r}_k^T \vec{r}_k}{\vec{r}_k^T A \vec{r}_k}$$

Can be shown $\|\vec{u}^{k+1} - \vec{u}\|_A \leq \|\vec{u}^k - \vec{u}\|_A \frac{K(A)-1}{K(A)+1}$

$\|\vec{v}\|_A^2 = \vec{v}^T A \vec{v}$

back to ou
it can be
(

$\varepsilon \leqslant (1$

$\frac{k}{12}$

$k =$

next time

where $K(A) = \frac{\lambda_n}{\lambda_1}$

$A\varphi_i = \lambda_i \varphi_i \quad i = 1, \dots, n$

$\varphi_i^T \varphi_j = \delta_{ij} \quad i,j = 1, \dots, n$

$0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$

back to our linear system $A\bar{u} = \bar{f} \quad A \in \mathbb{R}^{L^2 \times L^2}$

it can be shown that $K(A) = O(L^2)$

$\begin{cases} -\varphi''(x) = \lambda \varphi(x) \\ \varphi(0) = \varphi(1) = 0 \end{cases} \quad \varphi(x) = \sin(k\pi x) \quad k = 1, 2, \dots$

$\lambda_k = k^2 \pi^2$

then $\frac{K(A)-1}{K(A)+1} = \frac{1 - \frac{1}{K(A)}}{1 + \frac{1}{K(A)}} = 1 - O\left(\frac{1}{L^2}\right)$

number of iterations?

k s.t. $\|\bar{u}^k - \bar{u}\|_A = \varepsilon \|\bar{u}^0 - \bar{u}\|_A \quad \left(\varepsilon \ll \frac{1}{10^{-6}}\right)$

$\varepsilon \|\bar{u}^0 - \bar{u}\|_A \leq \|\bar{u}^0 - \bar{u}\|_A \cdot \underbrace{\left(\frac{K(A)-1}{K(A)+1}\right)^k}_{1 - O(L^2)}$

$\uparrow$

$= \|\bar{u}^k - \bar{u}\|_A$

$\varepsilon \leq \left(1 - O\left(\frac{1}{L^2}\right)\right)^k = 1 - O\left(\frac{k}{L^2}\right)$

$\frac{k}{L^2} \quad O(1-\varepsilon)$

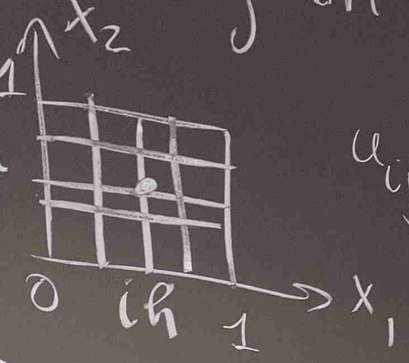
$k = O(L^2)$

next time: recap

Recap:

$-\Delta u = f$ in $]0,1[^d$ $\quad d=1,2,3$

$d=2$ $\quad h=\frac{1}{L+1}$



u_{ij} approx of $u(ih, jh)$

$1 \le i,j \le L$

$A\vec{u} = \vec{f}$ $\quad d=1 : A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}_L = T_L$

$\lambda_k = 2 - 2\cos\frac{k\pi}{L+1}$ $\quad k = 1, \dots, L$

$d=2$ $\quad A = \frac{1}{h^2}\left(T_L \otimes I_L + I_L \otimes T_L\right)$

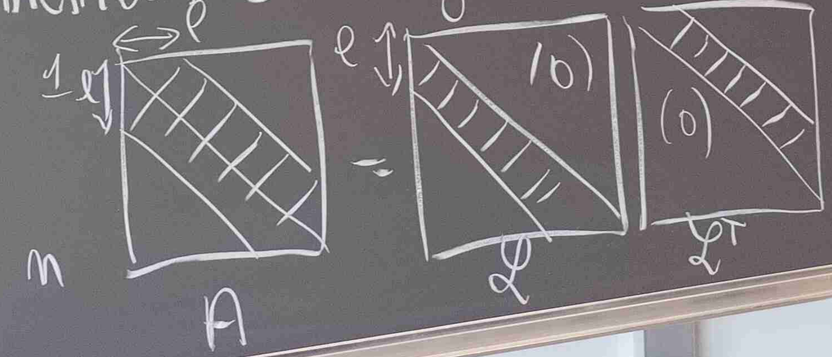
$d=3$ $\quad A = \frac{1}{h^2}\left(T_L \otimes I_L \otimes I_L + I_L \otimes T_L \otimes I_L + I_L \otimes I_L \otimes T_L\right)$

Direct or iterative methods to solve $A\vec{u} = \vec{f}$

Direct method: Cholesky dec

mem: $O(nl)$

CPU: $O(nl^2)$



$A \in \mathbb{R}^{m \times m}$ s.p.d

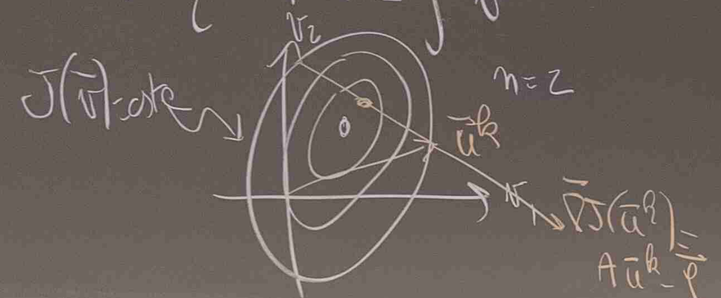
(When A is co
it can be sho
k s.t. $\|\vec{u}^k$
$k=0($
$k=0$
better than

Iterative methods : gradient descent

$\left(\vec{u} \text{ s.t. } A\vec{u} = \vec{f}\right) \Leftrightarrow \left(\vec{u} \text{ s.t } J(\vec{u}) \leq J(\vec{v}) \ \forall \vec{v} \in \mathbb{R}^n\right)$

$A \in \mathbb{R}^{n \times n}$ s.p.d

$$J(\vec{v}) = \frac{1}{2}\vec{v}^T A \vec{v} - \vec{f}^T \vec{v}$$



$\vec{u}^0 \in \mathbb{R}^n$, $\vec{r}^0 = \vec{f} - A\vec{u}^0$

$k = 0, 1, 2, \ldots$

$\vec{u}^{k+1} = \vec{u}^k + \alpha^k \vec{r}^k \quad \left(\alpha^k \text{ s.t } J(\vec{u}^k + \alpha^k \vec{r}^k) \leq J(\vec{u}^k + \alpha \vec{r}^k) \ \forall \alpha \in \mathbb{R}\right)$

$\vec{r}^{k+1} = \vec{f} - A\vec{u}^{k+1}$

if $\|\vec{r}^{k+1}\| \leq \varepsilon$ exit

$\|\vec{u}^{k+1} - \vec{u}\|_A \leq \|\vec{u}^k - \vec{u}\|_A \frac{K(A)-1}{K(A)+1}$ $\quad K(A) = \frac{\lambda_n}{\lambda_1}$

$A\vec{\varphi}_i = \lambda_i \vec{\varphi}_i \quad 0 < \lambda_1 \leq \ldots \leq \lambda_n$

(When A is coming from FD for $-\Delta u = f$ it can be shown that $K(A) = O(L^2)$)

k s.t. $\|\vec{u}^k - \vec{u}\|_A \leq \varepsilon \|\vec{u}^0 - \vec{u}\|$ ← doesn't depend on $\vec{u}^0$

$k = O(K(A))$

$k = O(L^2)$

better than gradient descent

Conjugate gradient

the direction of descent is not $\vec{r}^k$

($n = 2$, only 2 iterations

n ——— n ———)

$\vec{u}^0$, $\vec{r}^0 = \vec{f} - A\vec{u}^0 = \vec{d}^0$

$k = 0, 1, 2, \ldots$

$\vec{u}^{k+1} = \vec{u}^k + \alpha^k \vec{d}^k \quad \left(\alpha^k \text{ s.t } J(\vec{u}^k + \alpha^k \vec{d}^k) \leq J(\vec{u}^k + \alpha \vec{d}^k) \ \forall \alpha \in \mathbb{R}\right)$

$\vec{r}^{k+1} = \vec{f} - A\vec{u}^{k+1}$ $\quad$ if $\|\vec{r}^{k+1}\| \leq \varepsilon$ exit

$\vec{d}^{k+1} = \vec{r}^{k+1} + \beta^k \vec{d}^k \quad \left(\beta^k \text{ s.t. } \vec{d}^{k+1\,T} A \vec{d}^k = 0\right)$

it can be proven that

$$\|\vec{u}^{k+1} - \vec{u}\|_A \leq \|\vec{u}^k - \vec{u}\|_A \frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1}$$

therefore $k = O(L)$

Summarize these findings in a table

RAM: 3 Gb : $10^9 + 10^9 = 10^9$ 1 D.

GHz : 3 GHz

	Direct Cholesky	Iterat Conj. Grad
unknowns	$O(L^d)$	$O(L^d)$
Memory	$O(L^d L^{d-1})$	$O(L^d)$
CPU (Complexity)	$O(L^d L^{2(d-1)})$	$O(L^d L)$

d = 1	Cholesky
	$O(L^1)$
	$O(L^1)$
	$O(L^1)$

↑ optimal

d = 3	Chol	CG
	$O(L^3)$	$O(L^3)$
	$O(L^5)$	$O(L^3)$
	$O(L^7)$	$O(L^4)$

$O(L^3)$ optimal

multigrid achieves

d = 3 iterative method this should be used

Link between FD and Finite elements

- Variational (weak) formulation

$d=1$ $\int_0^1 \left(-u''(x) = f(x)\right) v(x)\,dx$ $0<x<1$

$v:[0,1]\to\mathbb{R}$

int. by parts $\int_0^1 u'(x)v'(x)\,dx - \left[u'(x)v(x)\right]_{x=0}^{x=1} = \int_0^1 f(x)v(x)\,dx$

Choose v s.t. $v(0)=0=v(1)$

Find $u\in V$ s.t $\int_0^1 u'(x)v'(x)\,dx = \int_0^1 f(x)v(x)\,dx$ $\forall v\in V$

$V=\{v\in L^2(0,1);\ v'\in L^2(0,1);\ v(0)=v(1)=0\}$ 1)

$\dim V=+\infty$

- $\varphi_1,\varphi_2,\dots,\varphi_N \in V$ linearly indep, 2)

$V_N = \operatorname{span}(\varphi_1,\dots,\varphi_N)$ $\dim V_N = N$

Find

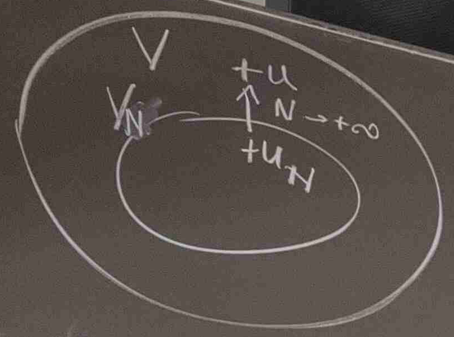
$u_N(x)$ =

$d=1$

$h=\frac{1}{N+1}$

F

t



Find $u_N \in V_N$ s.t. $\int_0^1 u_N'(x) v_N'(x)\,dx = \int_0^1 f(x) v_N(x)\,dx \quad \forall v_N \in V_N$

$u_N(x) = \sum_{j=1}^{N} \tilde{u}_j \varphi_j(x)$

$\in \mathbb{R}$, N unknowns

$v_N(x) = \varphi_i(x) \quad i = 1, \dots, N$

linear system

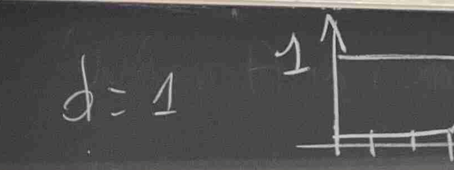
$\sum_{j=1}^{N} \tilde{u}_j \underbrace{\int_0^1 \varphi_i'(x)\varphi_j'(x)\,dx}_{A_{ij}} = \underbrace{\int_0^1 f(x)\varphi_i(x)\,dx}_{f_i} \quad i = 1, \dots, N$

$A\vec{\tilde{u}} = \vec{f}$

- Finite elements: φ_i?

Advantage

3D: triangle, tet



$d = 1$

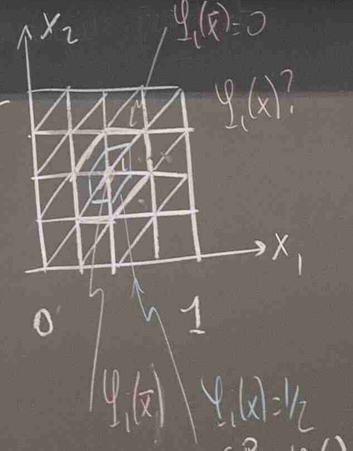
$h = \frac{1}{N+1}$

FD: $A\vec{u} = \vec{f}$ FE: $A\vec{\tilde{u}} = \vec{f}$

$f_i = f(x_i)$

trapeze formula to eval $\int_0^1 f(x)\varphi_i(x)\,dx$

2 equivalent linear systems $u_i = \tilde{u}_i$



$d = 2$

$\int_0^1 \int_0^1 f(x_1, x_2)\varphi_i(x_1, x_2)\,dx_1 dx_2$

eval. trapeze form.

FD $\Leftrightarrow$ FE

$\varphi_i \in C^0[0,1]^2$ $\varphi_i|_{\text{triangle}} \in \mathbb{P}_1$ $\varphi_i(\vec{x}) = 1$ $\varphi_i(\vec{x}) = 0$

satisfies:

$\frac{\partial u}{\partial t}(x,t) - \frac{\partial^2 u}{\partial x^2}(x,t)$

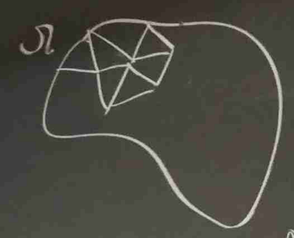
$u(x,0) = u_0(x)$

$u(0,t) = 0$ $u(1,t)$

boundary cond.

initial cond.

Advantage of FE



Part 2. FD pde's

1. Elliptic
2. Parabolic : the heat equation

Data : initial temp $u_0 : [0,1] \to \mathbb{R}$

heat source : $f : [0,1] \times [0,T] \to \mathbb{R}$

Unknown : $u \quad [0,1] \times [0,T] \to \mathbb{R}$

temperature $x, t \to u(x,t)$

u satisfies :

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) - \dfrac{\partial^2 u}{\partial x^2}(x,t) = f(x,t) & 0 < x \leq 1 \quad 0 \leq t \leq T \\ u(x,0) = u_0(x) & 0 < x \leq 1 \\ u(0,t) = 0 \quad u(1,t) = 0 & 0 < t \leq T \end{cases}$$

↕ boundary cond.

initial cond

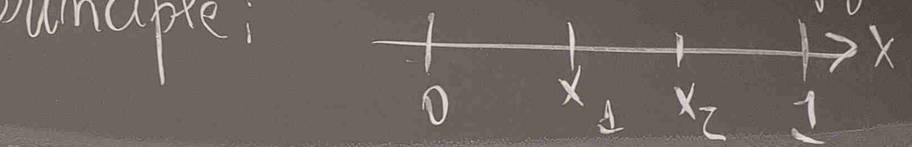
Remarks:

- Fourier series; if $u_0(x) = \sin(\pi x)$ $f(x,t) = \sin(\pi x)$

$$u(x,t) = \sin(\pi x)\, e^{-\pi^2 t} + \frac{1}{\pi^2}\left(1 - e^{-\pi^2 t}\right)\sin(\pi x)$$

First term in the Fourier expansion

- the heat corresponds to an energy conservation principle;



integr between $x = x_1$ and $x = x_2$

$$\frac{d}{dt}\underbrace{\int_{x_1}^{x_2} u(x,t)\,dx}_{\text{thermal energy}} - \underbrace{\left[\frac{\partial u}{\partial x}(x,t)\right]_{x=x_1}^{x=x_2}}_{\text{heat flux entering or leaving } (x_1,x_2)} = \underbrace{\int_{x_1}^{x_2} f(x,t)\,dx}_{\text{power of the heat source}}$$

- Positivity and max. principle

  If $\left.\begin{matrix} u_0(x) \geq 0 \\ f(x,t) \geq 0 \end{matrix}\right\}$ then $u(x,t) \geq 0$

  If $f(x,t) = 0$ then $\max_{0 \leq x \leq 1} |u(x,t)| \leq \max_{0 \leq x \leq 1} |u_0(x)|$

- A priori estimate: multiply the heat eq by $u(x,t)$

$$\frac{1}{2}\frac{d}{dt}\int_0^1 (u(x,t))^2 dx + \int_0^1 \left(\frac{\partial u}{\partial x}(x,t)\right)^2 dx = \int_0^1 f(x,t)u(x,t)dx \quad (*)$$

Cauchy-Schwarz ineq

$$\int_0^1 f(x,t)u(x,t)dx \leq \left(\int_0^1 (f(x,t))^2 dx\right)^{1/2} \left(\int_0^1 (u(x,t))^2 dx\right)^{1/2} \leq$$

Young's ineq $\quad ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$ $\quad (a+b)^2 \geq 0$

$$\int_0^1 f(x,t)u(x,t)dx \leq \frac{1}{2}\int_0^1 (f(x,t))^2 dx + \frac{1}{2}\int_0^1 (u(x,t))^2 dx$$

Poincaré ineq: $v \in C^1[0,1]$ $v(0) = v(1) = 0$ then

$$\int_0^1 (v(x))^2 dx \leq \int_0^1 (v'(x))^2 dx$$

$$\leq \frac{1}{2}\int_0^1 (f(x,t))^2 dx + \frac{1}{2}\int_0^1 \left(\frac{\partial u}{\partial x}(x,t)\right)^2 dx$$

$$\frac{1}{2}\frac{d}{dt}\int_0^1 (u(x,t))^2 dx + \frac{1}{2}\int_0^1 \left(\frac{\partial u}{\partial x}(x,t)\right)^2 dx \leq \frac{1}{2}\int_0^1 (f(x,t))^2 dx$$

int $t=0$ and $t=T$:

$$\int_0^1 (u(x,T))^2 dx + \int_0^T\int_0^1 \left(\frac{\partial u}{\partial x}(x,t)\right)^2 dx\,dt \leq \int_0^1 (u_0(x))^2 dx + \int_0^T\int_0^1 (f(x,t))^2 dx\,dt \quad (**)$$

Conseq: unicité

assume that there are

2 sol u_1 and u_2:

$$\frac{\partial}{\partial t}(u_1-u_2) - \frac{\partial^2 u}{\partial x^2}(u_1-u_2)=0$$

$$(u_1-u_2)(x,0) = 0$$

$$(u_1-u_2)(0,t) = 0 = (u_1-u_2)(1,t)$$

$$(\star\star): \int_0^1 \left((u_1-u_2)(x,T)\right)^2 dx \leq 0$$

Parabolic: the heat eq
hyperbolic: the wave eq, the transport eq.

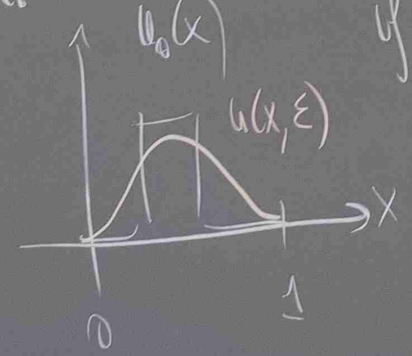
given $u_0 : [0,1] \to \mathbb{R}$ $f : [0,1] \times [0,T] \to \mathbb{R}$ find $u(x,t)$

s.t.

$$\begin{cases} \frac{\partial u}{\partial t}(x,t) - \frac{\partial^2 u}{\partial x^2}(x,t) = f(x,t) & 0<x<1,\ 0<t\leq T \\ u(0,t) = 0 = u(1,t) \\ u(x,0) = u_0(x) \end{cases}$$

Remarks:

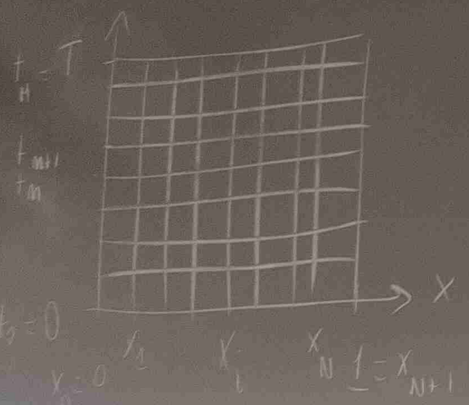
- explicit sol Fourier series
- Positivity, Max. principle
- A priori estimate
- Smoothness



if $u_0 \in L^2(0,1)$ then
$u \in C^\infty([0,T] \times [\varepsilon, +\infty])$ $\forall \varepsilon > 0$

Finite Differences

Implicit Euler scheme



M number of time steps

$\tau = \frac{T}{M}$ time step

$t_n = n\tau \quad n = 0, 1, \dots, M$

N number of internal points

$h = \frac{1}{N+1}$ space step

$x_i = ih \quad i = 0, \dots, N+1$

Goal: to compute u_i^n approx of $u(x_i, t_n)$

Start with $u_i^0 = u_0(x_i) \quad i = 1, \dots, N$

For a given n, how to compute u_i^{n+1} $i = 1, \dots, N$ from u_i^n $i = 1, \dots, N$?

Write the eq heat at x_i, t_{n+1}

$$\frac{\partial u}{\partial t}(x_i, t_{n+1}) - \frac{\partial^2 u}{\partial x^2}(x_i, t_{n+1}) = f(x_i, t_{n+1})$$

- Use a centered finite diff. to approx $\frac{\partial^2 u}{\partial x^2}(x_i, t_{n+1})$
- Use a backward —— f —— $\frac{\partial u}{\partial t}(x_i, t_{n+1})$

$$\frac{u(x_i, t_{n+1}) - u(x_i, t_n)}{\tau} - \frac{u(x_{i-1}, t_{n+1}) - 2u(x_i, t_{n+1}) + u(x_{i+1}, t_{n+1})}{h^2} = f(x_i, t_{n+1}) + r_i^n \quad \text{(remainder)}$$

$$|r_i^n| \leq \frac{1}{12} \max_{\substack{0 \leq x \leq 1 \\ 0 \leq t \leq T}} \left| \frac{\partial^6 u}{\partial x^6}(x,t) \right| h^2 + \frac{1}{2} \max_{\substack{0 \leq x \leq 1 \\ 0 \leq t \leq T}} \left| \frac{\partial^2 u}{\partial t^2}(x,t) \right| \tau, \text{ if } u \in \mathcal{C}^4 \text{ space and } \mathcal{C}^2 \text{ time}$$

The scheme is obtained replacing $u(x_i, t_n)$ by u_i^n and forgetting r_i^n

Scheme: $n = 0, 1, \ldots, M-1$

$$\frac{u_i^{n+1} - u_i^n}{\tau} - \frac{u_{i-1}^{n+1} - 2u_i^{n+1} + u_{i+1}^{n+1}}{h^2} = f(x_i, t_{n+1}) \quad i = 1, \ldots, N$$

$$u_0^{n+1} = 0 = u_{N+1}^{n+1}$$

u_i^n, $i = 1, \ldots, N$ are known

This corresponds to solving a linear system

$$\vec{u}^n = \begin{pmatrix} u_1^n \\ \vdots \\ u_N^n \end{pmatrix} \quad \vec{f}^{n+1} = \begin{pmatrix} f(x_1, t_{n+1}) \\ \vdots \\ f(x_N, t_{n+1}) \end{pmatrix} \quad A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & (0) \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ (0) & & -1 & 2 \end{pmatrix}$$

$$(I + \tau A)\vec{u}^{n+1} = \vec{u}^n + \tau \vec{f}^{n+1}$$

unique sol $(I + \tau A)$ is sym pos. def.

note that the exact solution sol

$$\vec{U}^{n+1} = \begin{pmatrix} u(x_1, t_{n+1}) \\ \vdots \\ u(x_N, t_{n+1}) \end{pmatrix} \text{ satisfies } (I + \tau A)\vec{U}^{n+1} = \vec{U}^n + \tau \vec{r}^n$$

$$\vec{r}^n = \begin{pmatrix} r_1^n \\ \vdots \\ r_N^n \end{pmatrix}$$

Stability and convergence in the $\|\cdot\|_\infty$ norm

$$\|\vec{v}\|_\infty = \max_{1\le i\le N} |v_i|$$

Lemma: stability (analogous of positivity and max principle)

$\forall h, \tau > 0$

- $\|\vec{u}^{n+1}\|_\infty \le \|\vec{u}^{n}\|_\infty + \tau \|\vec{g}^{n+1}\|_\infty$
- If $\vec{f}^{n+1}$ and $\vec{u}^{n} \ge \vec{0}$ then $\vec{u}^{n+1} \ge \vec{0}$

Proof:

- Consider index k such that $\underbrace{|u_k^{n+1}|}_{\|\vec{u}^{n+1}\|_\infty} \ge |u_i^{n+1}| \quad i = 1, \dots, N$

$$\left|\left(1+\frac{2\tau}{h^2}\right) u_k^{n+1}\right| = \left|\frac{\tau}{h^2}\left(u_{k-1}^{n+1} + u_{k+1}^{n+1}\right) + u_k^n + \tau f(x_k, t_{n+1})\right|$$

$$\left(1+\frac{2\tau}{h^2}\right) |u_k^{n+1}| \le \frac{\tau}{h^2}\left(|u_{k-1}^{n+1}| + |u_{k+1}^{n+1}|\right) + |u_k^n| + \tau |f(x_k, t_{n+1})|$$

$$\left(1+\frac{2\tau}{h^2}\right)|u_k^{n+1}| \leq \frac{\tau}{h^2}\left(|u_k^{n+1}| + |u_k^{n+1}|\right) + |u_k^n| + \tau|f(x_k, t_{n+1})| \qquad k=1,\dots,N$$

$$\|\vec{u}^{n+1}\|_\infty = |u_k^{n+1}| \leq |u_k^n| + \tau|f(x_k, t_{n+1})| \leq \|\vec{u}^n\|_\infty + \tau\|f^{n+1}\|_\infty$$

if $k=1$ $|u_{k-1}^{n+1}| = 0 \leq |u_k^{n+1}|$

if $k=N$ $|u_{k+1}^{n+1}| = 0 \leq |u_k^{n+1}|$

- Consider index k such that $u_k^{n+1} \leqslant u_i^{n+1}$ $i=1,\dots,N$

$$\left(1+\frac{2\tau}{h^2}\right)u_k^{n+1} = \frac{\tau}{h^2}\left(u_{k-1}^{n+1} + u_{k+1}^{n+1}\right) + u_k^n + \tau f(x_k, t_{n+1})$$

$\geqslant 0$ $\geqslant 0$

if $k=2,\dots,N-1$

$$\geqslant \frac{\tau}{h^2}\left(u_k^{n+1} + u_k^{n+1}\right)$$

$u_k^{n+1} \geqslant 0$

if $k=1$: $(u_0^{n+1} = 0)$

$$\left(1+\frac{2\tau}{h^2}\right)u_k^{n+1} = \frac{\tau}{h^2}u_{k+1}^{n+1} + u_k^n + \tau f(x_k, t_{n+1})$$

$$\left(1+\frac{\tau}{h^2}\right)u_k^{n+1} \geqslant u_k^n + \tau f(x_k, t_{n+1})$$

$\geqslant 0$ $\geqslant 0$

therefore $u_k^{n+1} \geqslant 0$

same if $k=N$

Thm. Convergence in $\|\cdot\|_\infty$ norm

If $u \in \mathcal{C}^4$ in space and $\mathcal{C}^2$ in time, there exist $C>0$ s.t. $\forall h>0 \forall \tau>0$

$$(T=M) \qquad \underbrace{\|\bar{U}^M - \bar{u}^M\|_\infty}_{\max_{i=1,\dots,N} |u(x_i,t^M)-u_i^M|} \leq C(h^2+\tau)$$

interpretation:

Choose u, select M and N

the error $\max_{i=1,\dots,N} |u(x_i,t^M)-u_i^M|$ is divided by 4 if N is multiplied by 2 and M is multiplied by 4

Proof We have

$$(I+\tau A)(\bar{U}^{n+1} - \bar{u}^{n+1}) = \bar{U}^n - \bar{u}^n + \tau \vec{r}^n \text{ with } \|\vec{r}^n\|_\infty \leq D(h^2+\tau)$$

$$\max\left(\underbrace{\frac{1}{12}\max_{x,t}\left|\frac{\partial^4 u}{\partial x^4}(x,t)\right|}_{\text{indep of } h \text{ and } \tau}, \frac{1}{2}\max_{x,t}\left(\frac{\partial^2 u}{\partial t^2}(x,t)\right| \right)$$

proceeding as in the lemma

$$\|\bar{U}^{n+1} - \bar{u}^{n+1}\|_\infty \leq \|\bar{U}^n - \bar{u}^n\|_\infty + \tau \|\vec{r}^n\|_\infty$$

$$\|\bar{U}^M - \bar{u}^M\|_\infty \leq \underbrace{\|\bar{U}^0 - \bar{u}^0\|_\infty}_{0} + \tau \sum_{n=0}^{M-1} \|\vec{r}^n\| \leq \underbrace{C}_{T} D(h^2+\tau) = TD(h^2+\tau)$$

Explicit scheme

$$\frac{\partial u}{\partial t}(x_i,t_n) - \frac{\partial^2 u}{\partial x^2}(x_i,t_n) = f(x_i,t_n)$$

- forward finite diff formula for $\frac{\partial u}{\partial t}(x_i,t_n)$
- centered for $\frac{\partial^2 u}{\partial x^2}(x_i,t_n)$

We obtain the following scheme

$$\begin{cases} \dfrac{u_i^{n+1}-u_i^n}{\tau} - \dfrac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{h^2} = f(x_i,t_n) \quad i=1,\dots,N \\ u_0^n = 0 = u_{N+1}^n \end{cases}$$

explicit scheme

$$\begin{cases} u_i^{n+1} = \left(1-\dfrac{2\tau}{h^2}\right)u_i^n + \dfrac{\tau}{h^2}\left(u_{i-1}^n + u_{i+1}^n\right) + \tau f(x_i,t_n) \quad i=1,\dots,N \\ u_0^n = 0 = u_{N+1}^n \end{cases}$$

$$\vec{u}^{n+1} = \left(I-\tau A\right)\vec{u}^n + \tau \vec{f}^n$$

$$\vec{u}^{n+1} = \left(I+\tau A\right)^{-1}\left(\vec{u}^n + \tau \vec{f}^{n+1}\right)$$

$\tau \ll 1$ $\quad (I+\tau A)^{-1} \sim I - \tau A$

$x \ll 1$ $\quad \frac{1}{1+x} \sim 1-x$

Stabilit
Lemma:

Proof $i=1,$

u_i^{n+1}

u_i^{n+1}

The exact solution satisfies

$$\vec{U}^{n+1} = (I - \tau A)\vec{U}^n + \tau \vec{f}^n + \tau \vec{R}^n$$

with $\|\vec{R}^n\|_\infty \leq D(h^2 + \tau)$ (D depends only on u)

D = same D as in the convergence proof

Stability of the explicit scheme $\|\cdot\|_\infty$ norm

Lemma:

If $\tau \leq \underbrace{\frac{h^2}{2}}_{\text{stability cond}}$ then we have

- $\|\vec{u}^{n+1}\|_\infty \leq \|\vec{u}^n\|_\infty + \tau \|\vec{f}^n\|_\infty$
- If $\vec{f}^n \geqslant \vec{0}$ and $\vec{u}^n \geqslant \vec{0}$ then $\vec{u}^{n+1} \geqslant 0$

Proof $i = 1, \dots, N$

$$u_i^{n+1} = \left(1 - \frac{2\tau}{h^2}\right) u_i^n + \frac{\tau}{h^2}\left(u_{i-1}^n + u_{i+1}^n\right) + \tau f(x_i, t_n)$$

$$|u_i^{n+1}| \leqslant \underbrace{\left|1 - \frac{2\tau}{h^2}\right|}_{= 1 - \frac{2\tau}{h^2} \text{ if } \left(1 - \frac{2\tau}{h} \geqslant 0\right) \Leftrightarrow \left(\tau \leq \frac{h^2}{2}\right)} |u_i^n| + \frac{\tau}{h^2}\left(|u_{i-1}^n| + |u_{i+1}^n|\right) + \tau |f(x_i, t_n)|$$

for $i = 2, \dots, N-1$

$$|u_i^{n+1}| \leqslant \left(1 - \frac{2\tau}{h^2} + \frac{2\tau}{h^2}\right) \underbrace{\max_{1 \leq j \leq N} |u_j^n|}_{= \|\vec{u}^n\|_\infty} + \tau \underbrace{\max_{1 \leq j \leq N} |f(x_j, t_n)|}_{\|\vec{f}^n\|_\infty}$$

Thm. Convergence in $\|\cdot\|_\infty$ norm

If $u \in \mathcal{C}^4$ in space and $\mathcal{C}^2$ in time, there exist $C>0$ s.t. $\forall h>0 \forall \tau>0$

$$\|\bar{U}^M - \bar{u}^M\|_\infty \leq C(h^2+\tau) \quad (\tau = \tfrac{T}{M})$$

$$\|\bar{U}^M - \bar{u}^M\|_\infty = \max_{i=1,\dots,N} |u(x_i, t^M) - u_i^M|$$

interpretation:

Choose u, select M and N

the error $\max_{i=1,\dots,N} |u(x_i, t^M) - u_i^M|$ is divided by 4 if M is multiplied by 4

N ——— 2

---

if $\lambda = 1 \quad \left(\tau \leq \frac{h^2}{2}\right)$

$$|u_i^{n+1}| \leq \left(1 - \frac{2\tau}{h^2}\right)|u_i^n| + \frac{\tau}{h^2}\left(0 + |u_{i+1}^n|\right) + \tau|f(x_i, t_n)|$$

$$\leq \left(1 - \frac{2\tau}{h^2} + \frac{\tau}{h^2}\right)\|\bar{u}^n\|_\infty + \tau\|f^n\|_\infty$$

$\frac{2\tau}{h^2} + \frac{\tau}{h^2} \leq 1$

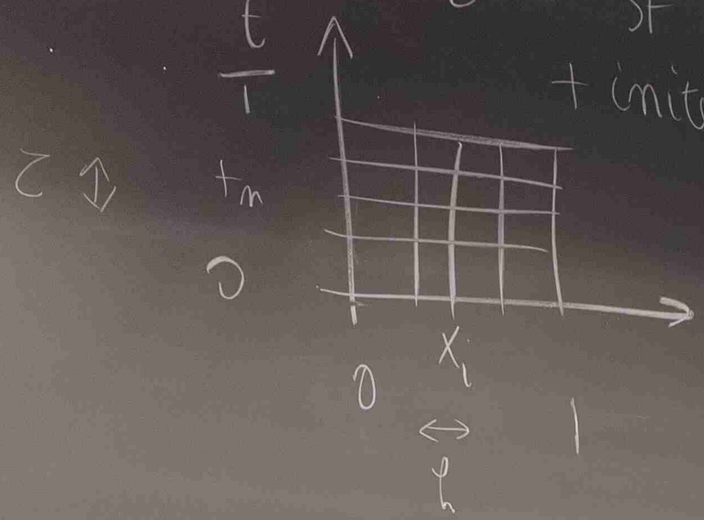
for all $i = 1, \dots, N$ we have $|u_i^{n+1}| \leq \|\bar{u}^n\|_\infty + \tau\|\bar{f}^m\|_\infty$

therefore $\|\bar{u}^{n+1}\|_\infty$ ———

Summary

$$\frac{\partial u}{\partial t}(x,t) - \frac{\partial^2 u}{\partial x^2}(x,t) = f(x,t) \qquad 0<x<1 \quad 0<t\leq T$$

+ initial and boundary cond



$x_i = i h = \frac{i}{N+1} \quad i = 1, \dots, N$

$t_n = n\tau \quad n = 0, 1, \dots, M$

$= n\frac{T}{M}$

Implicit scheme $(I + \tau A)\vec{u}^{n+1} = \vec{u}^n + \tau \vec{f}^{n+1}$

Explicit scheme $\vec{u}^{n+1} = (I - \tau A)\vec{u}^n + \tau \vec{f}^n$

here $\vec{u}^n = \begin{pmatrix} u_1^n \\ \vdots \\ u_N^n \end{pmatrix}$ approx of $\vec{U}^n = \begin{pmatrix} u(x_1, t_n) \\ \vdots \\ u(x_N, t_n) \end{pmatrix}$ $\vec{f}^n = \begin{pmatrix} f(x_1, t_n) \\ \vdots \\ f(x_N, t_n) \end{pmatrix}$

$$A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$$

IS: $(I+\tau A)\bar U^{n+1} = \bar U^n + \tau(\bar f^{n+1} + \bar r^{n+1})$

$\|\bar r^{n+1}\|_\infty \le C(h^2+\tau)$

C indep of h and τ (dep on $\frac{\partial^2 u}{\partial x^2}, \frac{\partial^4 u}{\partial x^4}$)

Ex: $\bar U^{n+1} = (I-\tau A)\bar U^n + \tau(\bar f^n + \bar r^n)$

$\|\bar r^n\| \le C(h^2+\tau)$

Stability: IS: $\forall h,\tau>0 \quad \|\bar u^{n+1}\|_\infty \le \|\bar u^n\|_\infty + \tau\|\bar f^{n+1}\|_\infty$

ES: if $\tau \le h^2/2$ ——— $\|\bar f^n\|_\infty$

Convergence:

IS: $(I+\tau A)(\bar U^{n+1} - \bar u^{n+1}) = \bar U^n - \bar u^n + \tau \bar r^{n+1}$

$\|\bar U^{n+1} - \bar u^{n+1}\|_\infty \le \|\bar U^n - \bar u^n\|_\infty + \tau\|\bar r^{n+1}\|_\infty$

$\|\bar U^M - \bar u^M\|_\infty \le C\underbrace{TM}_{T}(h^2+\tau)$

ES: if $\tau \le h^2/2$ $\quad \bar U^{n+1} - \bar u^{n+1} = (I-\tau A)(\bar U^n - \bar u^n) + \tau \bar r^n$

$\|\bar U^{n+1} - \bar u^{n+1}\|_\infty \le \|\bar U^n - \bar u^n\|_\infty + \tau\|\bar r^n\|_\infty$

$\|\bar U^M - \bar u^M\|_\infty \le CT(h^2+\tau)$

Stability and convergence in the $\|\cdot\|_2$ norm

$A\underline{\varphi}_i = \lambda_i \underline{\varphi}_i \quad i=1,\dots,N$

$\underline{\varphi}_i^T \underline{\varphi}_j = \delta_{ij} \quad i,j=1,\dots,N$

$\lambda_i = 2\left(1-\cos\frac{i\pi}{N+1}\right)(N+1)^2 \quad i=1,\dots,N$

$\left(-\varphi''(x) = \lambda\varphi(x),\ \varphi(0)=\varphi(1)=0\right.$

$\left.\varphi(x) = \sin(k\pi x) \quad \lambda = k^2\pi^2\right)$

$\lambda_1 = 2\left(1-\cos\frac{\pi}{N+1}\right)(N+1)^2$

$\sim 2\,\frac{\pi^2}{2(N+1)^2}(N+1)^2 = \pi^2$

$\lambda_N = 2\left(1-\cos\frac{N\pi}{N+1}\right)(N+1)^2$

$\sim 2\left(1+1-\frac{\pi^2}{2(N+1)^2}\right)(N+1)^2$

$\sim \frac{4}{h^2}$

Stability $\|\vec{v}\|_2^2 = \sum_{i=1}^{N} v_i^2$

Lemma:

a. $\forall B \in \mathbb{R}^{N\times N}$ symmetric, $B\vec{\Psi}_i = \mu_i \vec{\Psi}_i$ $i = 1, \dots, N$

$\forall \vec{v} \in \mathbb{R}^N \quad \|B\vec{v}\|_2 \le \max_{1\le i\le N} |\mu_i| \, \|\vec{v}\|_2$

b. IS: $\forall h, \tau > 0 \quad \|\vec{u}^{n+1}\|_2 \le \|\vec{u}^n\|_2 + \tau \|\vec{f}^{n+1}\|_2$

c. ES: if $\tau \le \frac{2}{\lambda_N}$ ——— $\|\vec{f}^m\|_2$

Remi $\frac{2}{\lambda_N} \sim \frac{2}{4/h^2} = \frac{h^2}{2}$

a. $\forall \vec{v} \in \mathbb{R}^N \quad \vec{v} = \sum_{i=1}^{N} v_i \vec{\Psi}_i$

$\|B\vec{v}\|_2^2 = \left\| \sum_{i=1}^{N} v_i \mu_i \vec{\Psi}_i \right\|_2^2$

$= \sum_{i=1}^{N} v_i^2 \mu_i^2$

$\le \max_{1\le i\le N} \mu_i^2 \underbrace{\sum_{i=1}^{N} v_i^2}_{\|\vec{v}\|_2^2}$

b. $\vec{u}^{n+1} = (\underline{I} + \tau A)^{-1} (\vec{u}^n + \tau \vec{f}^{n+1})$

$\|\vec{u}^{n+1}\|_2 \le \max_i \underbrace{\frac{1}{1+\tau\lambda_i}}_{\le 1 \text{ since } \lambda_i > 0} \left(\|\vec{u}^n\|_2 + \tau \|\vec{f}^{n+1}\|_2 \right)$

true for any $h, \tau > 0$

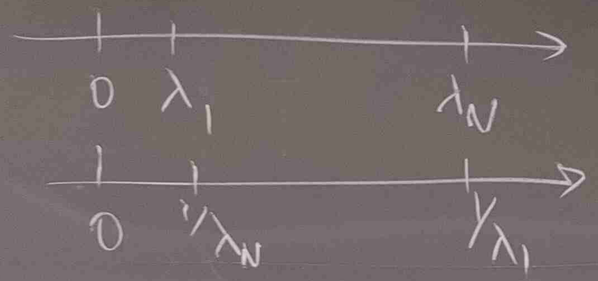
c. $\|\vec{u}^{n+1}\|_2 \le \max_{1\le i\le N} |1 - \tau\lambda_i| \, \|\vec{u}^n\|_2 + \tau \|\vec{f}^n\|_2$

$|1-\tau\lambda_i| \leq 1 \quad i = 1, \dots, N$?

$-1 \leq 1-\tau\lambda_i \leq 1$

$\tau\lambda_i \leq 2 \quad i = 1, \dots, N$

always true ($\forall h, \tau > 0$)



true if $\tau \leq \frac{2}{\lambda_N}$

If $\tau \leq \frac{2}{\lambda_N}$

$\|\bar{u}^{n+1}\|_2 \leq \|\bar{u}^n\|_2 + \tau\|\bar{f}^n\|_2$

Thm (Convergence in the $\|\cdot\|_2$ norm)

If $u \in \mathcal{C}^2$ in time and $\mathcal{C}^4$ in space

Then $\exists D > 0$ indep of h and τ s.t.

$$\|\bar{U}^M - \bar{u}^M\|_2 \leq D\,\frac{h^2+\tau}{h^{1/2}}$$

true $\forall h, \tau > 0$ for the Implicit scheme

true if $\tau \leq \frac{2}{\lambda_N}$ — Explicit —

Proof: IS; without using $\|\cdot\|_\infty$

$$\vec{U}^{n+1} - \vec{u}^{n+1} = (I+\tau A)^{-1}\left(\vec{U}^n - \vec{u}^n + \tau \vec{r}^{n+1}\right)$$

$$\|U^{n+1} - \vec{u}^{n+1}\|_2 \leq \|\vec{U}^n - \vec{u}^n\|_2 + \tau \|\vec{r}^{n+1}\|_2$$

$$|r_i^{n+1}| \leq C(h^2+\tau)$$

$$\sum_{i=1}^{N} |r_i^{n+1}|^2 \leq C^2 N (h^2+\tau)^2 \leq C^2 (N+1)(h^2+\tau)^2 = \frac{C^2(h^2+\tau)^2}{h}$$

$$\|\vec{r}^{n+1}\|_2 \leq C\frac{(h^2+\tau)}{h^{1/2}}$$

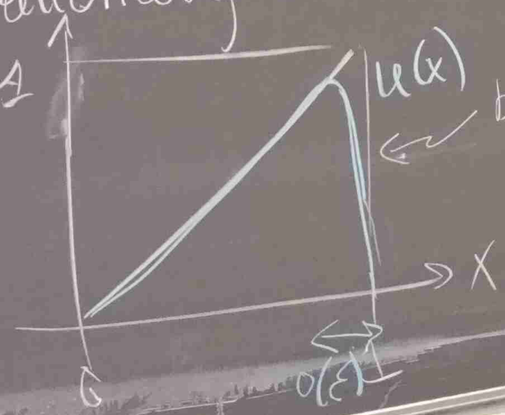
$$\|\vec{U}^M - \vec{u}^M\|_2 \leq + \tau \sum_{n=0}^{M-1} \|\vec{r}^{n+1}\|_2 \leq C\frac{(h^2+\tau)}{h^{1/2}} \underbrace{\tau M}_{T}$$

Extensions:

- advection diffusion: $\frac{\partial u}{\partial t} - \varepsilon \frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x} = f$

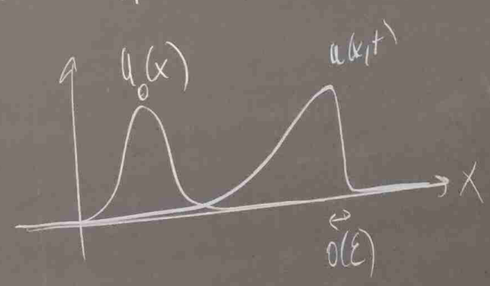
  ε small

  $f = 1$ stationary solution $-\varepsilon u''(x) + u'(x) = 1$, $u(0) = u(1) = 0$



- nonlinear advection diffusion

$$\frac{\partial u}{\partial t} - \varepsilon \frac{\partial^2 u}{\partial x^2} + u\frac{\partial u}{\partial x} = 0$$



wing

Navier-Stokes

$$\rho\left(\frac{\partial \vec{u}}{\partial t} + (\vec{u}\cdot\vec{\nabla})\vec{u}\right) - \mu\Delta\vec{u} + \vec{\nabla}p = \vec{f}$$

$$\operatorname{div}\vec{u} = 0$$

Supersonic flow

Mach 2

2000 km/h

shock wave

Mach 0,85

1000 km/h

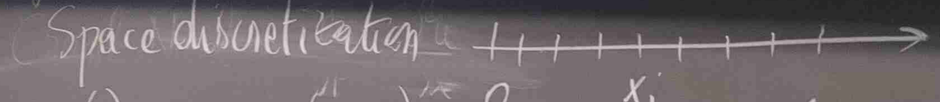
above Mach 1

below mach one

transonic flow

- high order schemes in time (RK4)

Space discretization



$x_i = ih \quad i = 0, \dots, N+1$

$u_i(t)$ approx $u(x_i, t)$

$$\frac{\partial}{\partial t}u(x_i,t) - \frac{\partial^2}{\partial x^2}u(x_i,t) = f(x_i,t) \quad i = 1,\dots,N$$

$$u_i'(t) - \frac{u_{i-1}(t) - 2u_i(t) + u_{i+1}(t)}{h^2} = f(x_i,t) \quad i = 1,\dots,N$$

$$\vec{u}'(t) + A\vec{u}(t) = \vec{f}(t) \qquad \vec{u}(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{pmatrix}$$

Order 2 schemes

- Implicit: Crank-Nicolson (order better Euler implicit and explicit)

$$\frac{\vec{u}^{n+1}-\vec{u}^{n}}{\tau} + A\,\frac{\vec{u}^{n}+\vec{u}^{n+1}}{2} = \frac{\vec{f}(t_n)+\vec{f}(t_{n+1})}{2}$$ order 2 in time

For instance if $\vec{f}(x,t)=0$ take the scalar product with $\vec{u}^{n+1}+\vec{u}^{n}$

$$\frac{1}{\tau}\left(\|\vec{u}^{n+1}\|_2^2 - \|\vec{u}^{n}\|_2^2\right) + \underbrace{\left(\vec{u}^{n+1}+\vec{u}^{n}\right)^t A\,\frac{\vec{u}^{n}+\vec{u}^{n+1}}{2}}_{\geq 0 \text{ A s.p.d.}} = 0$$

therefore $\|\vec{u}^{n+1}\|_2 \leq \|\vec{u}^{n}\|_2$

moreover the final error is $O\left(\frac{h^2+\tau^2}{h^{3/2}}\right)$ $\left(\text{if } \frac{\partial^3 u}{\partial t^3} \text{ is } \neq 0\right)$

- Explicit RK2

$$\vec{y}'(t) = \vec{f}(t,\vec{y}(t))$$
$$\vec{k}_1 = \vec{f}(t_0,\vec{y}_0)$$
$$\vec{k}_2 = \vec{f}(t_0+\tau/2,\ \vec{y}_0+\tau/2\,\vec{k}_1)$$
$$\vec{y}_1 = \vec{y}_0 + \tau\vec{k}_2$$

Here $(\vec{u}'(t) = \vec{f}(t) - A\vec{u}(t))$

Given $\vec{u}^n$:

- compute $\vec{k}_1 = \vec{f}(t_n) - A\vec{u}^n$

$$\vec{k}_2 = \vec{f}(t_n + \tau/2) - A\left(\vec{u}^n + \frac{\tau}{2}\vec{k}_1\right)$$

$$\vec{u}^{n+1} = \vec{u}^n + \tau\vec{k}_2$$

Stable under condition, final error is $O\left(\frac{h^2+\tau^2}{h^{1/2}}\right)$

- Can be extended to dimension 2 and 3 in space

$$\frac{\partial^2 u}{\partial x^2} \longrightarrow \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \text{ in 3d}$$

2d: $A = \frac{1}{h^2}\left(I_L \otimes D_L + D_L \otimes I_L\right)$ L = number of points in each direction

3d: $\left(I_L \otimes I_L \otimes D_L + \dots\right.$

$$\left(\begin{array}{ll} -\Delta\varphi = \lambda\varphi & \text{in } [0,1]^2 \\ \varphi = 0 & \text{on } \partial([0,1]^2) \end{array}\right.$$

$$\varphi(x,y) = \sin(k\pi x)\sin(l\pi y)$$

$$\lambda = (k^2+l^2)\pi^2$$

Next lesson

- Elliptic
- Parabolic
- Hyperbolic
  - transport
  - wave

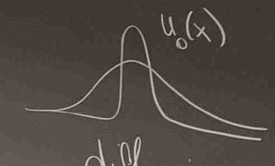
elliptic $\begin{cases} -u''(x) = f(x) & 0<x<1 \\ u(0)=0 \quad u(1)=0 \end{cases}$

$-\Delta u = f$ $(0,1)^d$
$u=0$ $\partial(0,1)^d$ $d=2,3$

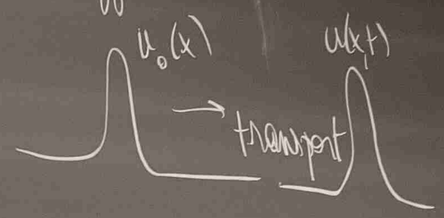
parabolic $\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = f$

$\frac{\partial u}{\partial t} - \Delta u = f$ heat diffusion

hyperbolic
- transport
- wave



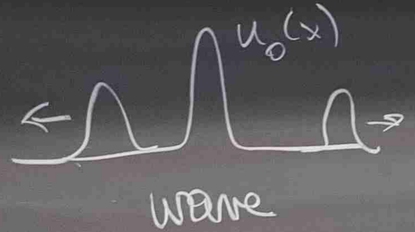
diffusion



$u(x,0) = u_0(x)$

Ex: u: conc of pol
c: transport
f: source of

Solution when



wave

---

transport

data: $u_0(x)$ initial cond $x \in \mathbb{R}$ $(0<x<1$ later$)$
$c(x,t)$ velocity of transport $t>0$
$f(x,t)$ source term

unknown: $u(x,t)$ $x \in \mathbb{R}$ $t>0$

Solution when

$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = f$

let $X(t)$ s.t. X'

then $\frac{d}{dt} u(x$

if u is sol of the

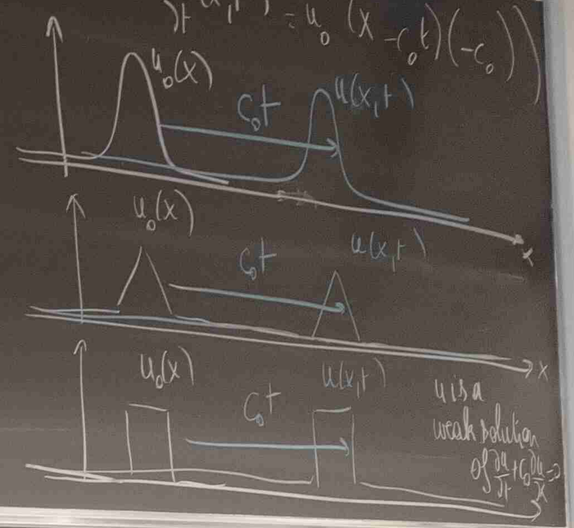
$$\begin{cases} \frac{\partial u}{\partial t}(x,t) + c(x,t)\frac{\partial u}{\partial x}(x,t) = f(x,t) & x \in \mathbb{R}\ t>0 \\ u(x,0) = u_0(x) & x \in \mathbb{R} \end{cases}$$

Ex: u: conc of pollutant in a river
c: transport velocity
f: source or sink

- Solution when $f(x,t)=0$ $c(x,t)=c_0=cst$

$$u(x,t) = u_0(x-c_0t)$$

(check: $\frac{\partial u}{\partial t}(x,t) = u_0'(x-c_0t)(-c_0)$)



Solution when c is $\mathcal{C}^1$

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = f$$

let $X(t)$ s.t $X'(t) = c(X(t),t)$ (trajectories)

then $\frac{d}{dt}u(X(t),t) = \frac{\partial u}{\partial t}(X(t),t) + \frac{\partial u}{\partial x}(X(t),t)\underbrace{X'(t)}_{c(X(t),t)}$

if u is sol of the transp. eq

$$\frac{d}{dt}u(X(t),t) = f(X(t),t)$$

$$u(X(t),t) = \underbrace{u(X(0),0)}_{f(x,t)=0} + \int_0^t f(X(s),s)\,ds$$

if $c(x,t) = x$ $X'(t) = X(t)$

$X(t) = X(0)e^t$ $t>0$

$u(\underbrace{X(0)e^t}_{x},t) = u_0(X(0))$

$u(x,t) = u_0(xe^{-t})$

- Conservative transport eq

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(cu) = f$$

$$\int_{x_1}^{x_2} \left(\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(cu)\right) dx = \int_{x_1}^{x_2} f\,dx$$

$$\frac{d}{dt}\int_{x_1}^{x_2} u(x,t)\,dx + \Big[c(x,t)u(x,t)\Big]_{x=x_1}^{x=x_2} = \int_{x_1}^{x_2} f(x,t)\,dx$$

amount of pollutant in (x_1, x_2) — flux of pollutant entering or leaving — source or sink of pollutant

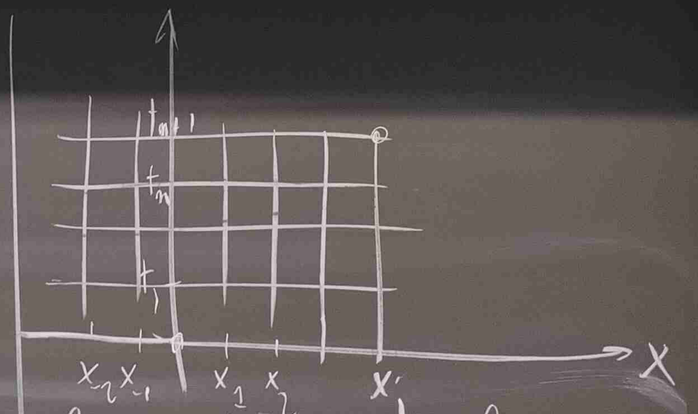
compressible flows

$$\rho\left(\frac{\partial \vec{u}}{\partial t} + (\vec{u}\cdot\vec{\nabla})\vec{u}\right) - \mu\Delta\vec{u} + \vec{\nabla}\vec{p} = \vec{f}$$

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho\vec{u}) = 0$$

$p(\rho)$

Numerical schemes



h space step $\quad x_j = jh \quad j \in \mathbb{Z}$

τ time step $\quad t_n = n\tau \quad n = 0, 1, 2$

goal: to compute u_j^n approx $u(x_j, t_n)$

time marching

$u_j^0 = u_0(x_j) \quad j \in \mathbb{Z}$

given $u_j^n \quad j \in \mathbb{Z}$, how to compute u_j^{n+1} ?

Centered scheme: (always instable !)

$$\underbrace{\frac{\partial u}{\partial t}(x_j, t_n)}_{\text{forward FDF}} + c(x_j, t_n) \underbrace{\frac{\partial u}{\partial x}(x_j, t_n)}_{\text{centered FDF}} = f(x_j, t_n) \qquad j \in \mathbb{Z}$$

explicit

scheme: $\dfrac{u_j^{n+1} - u_j^n}{\tau} + c(x_j, t_n) \dfrac{u_{j+1}^n - u_{j-1}^n}{2h} = f(x_j, t_n) \quad j \in \mathbb{Z}$

this scheme is unstable $\forall h, \tau > 0$

Case $f = 0 \quad c = c_0 = cst \quad u_0(x) = e^{imx}$ (Fourier series $\sum_m c_m e^{imx}$)

$$u_j^1 = u_j^0 - \frac{c_0 \tau}{2h}\left(u_{j+1}^0 - u_{j-1}^0\right) \quad j \in \mathbb{Z}$$

$$u_j^0 = e^{imjh}$$

$$u_j^1 = e^{imjh} - \frac{c_0 \tau}{2h}\left(e^{im(j+1)h} - e^{im(j-1)h}\right)$$

$$u_j^1 = e^{imjh}\left(1 - \frac{c_0\tau}{2h}\left(e^{imh} - e^{-imh}\right)\right)$$

$$= \frac{}{\left(1 - ic_0\frac{\tau}{h}\sin(mh)\right)}$$

induction

$$u_j^n = e^{imjh}\left(\frac{}{}\right)^n$$

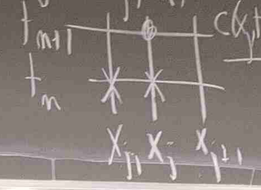
$$|u(x,t)| = |u_0(x - c_0 t)| = |e^{im(x-c_0t)}| = 1$$

$$|u_j^n| = |e^{imjh}|\left(1 + \left(\frac{c_0\tau}{h}\right)^2 \sin^2(mh)\right)^{\frac{n}{2}}$$

$n \to \infty \longrightarrow +\infty$ whenever $mh \neq k\pi$, $k \in \mathbb{Z}$

upwind scheme: $\left(\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = f\right)$

if $c(x_j, t_n) > 0$

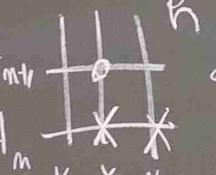


o : to be computed

x : used to approx $\frac{\partial u}{\partial x}(x_j, t_n)$

$$\frac{u_j^{n+1} - u_j^n}{\tau} + c(x_j,t_n)\frac{u_j^n - u_{j-1}^n}{h} = f(x_j,t_n) \quad j \in \mathbb{Z}$$

if $c(x_j,t_n) < 0$



$$\frac{u_j^{n+1} - u_j^n}{\tau} + c(x_j,t_n)\frac{u_{j+1}^n - u_j^n}{h} = f(x_j,t_n) \quad j \in \mathbb{Z}$$

if $f = 0$ $u_0(x) = e^{imx}$ $c(x,t) = c_0 > 0$

$$u_j^1 = u_j^0 - \frac{c_0\tau}{h}\left(u_j^0 - u_{j-1}^0\right) \quad j \in \mathbb{Z}$$

$$= e^{imjh} - \frac{c_0\tau}{h}\left(e^{imjh} - e^{im(j-1)h}\right)$$

$$= e^{imjh}\left(1 - \frac{c_0\tau}{h}\left(1 - e^{-imh}\right)\right)$$

$$u_j^n = \frac{}{}$$

$$|u_j^n| = |e^{imjh}|\underbrace{\left|1 - \frac{c_0\tau}{h}\left(1 - e^{-imh}\right)\right|^n}_{\leq 1 \text{ if } c_0\tau/h \leq 1}$$

This was stability in the ℓ^2 norm (L^2 norm)

Stability in the ∞ norm

Lemma: $f=0$ $c(x,t)=c_0>0$ $\dfrac{u_j^{n+1}-u_j^n}{\tau}+c_0\dfrac{u_j^n-u_{j-1}^n}{h}=0$ $j\in\mathbb{Z}$

If $\tau\le h/c_0$ then:

- If $u_j^n\ge 0\ \forall j\in\mathbb{Z}$ then $u_j^{n+1}\ge 0\ \forall j\in\mathbb{Z}$
- $\sup_{j\in\mathbb{Z}}|u_j^{n+1}|\le\sup_{j\in\mathbb{Z}}|u_j^n|$

Remark: if $c_0<0$ $\dfrac{u_j^{n+1}-u_j^n}{\tau}+c_0\dfrac{u_{j+1}^n-u_j^n}{h}=0$

L^∞ stable (satisfies the 2 above prop if $\tau\le\dfrac{h}{|c_0|}$)

Proof: $u_j^{n+1}=u_j^n\underbrace{\left(1-\frac{\tau c_0}{h}\right)}_{\text{if }\tau\le h/c_0\text{ this is}\ge 0}+\frac{\tau c_0}{h}u_{j-1}^n$ $j\in\mathbb{Z}$

- If $u_j^n\ge 0$ $j\in\mathbb{Z}$ then $u_j^{n+1}\ge 0\ \forall j\in\mathbb{Z}$
- $|u_j^{n+1}|\le\left|1-\frac{\tau c_0}{h}\right||u_j^n|+\left|\frac{\tau c_0}{h}\right||u_{j-1}^n|$

$$u_j^1 = e^{imjh}\left(1 - \frac{c_0\tau}{2h}\left(e^{imh} - e^{-imh}\right)\right)$$

$$= \left(1 - i\frac{c_0\tau}{h}\sin(mh)\right)e^{imjh}$$

induction

$$u_j^n = e^{imjh}\left(\quad\right)^n$$

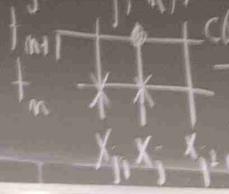
$$|u(x,t)| = |u_0(x - c_0 t)| = |e^{im(x-c_0t)}| = 1$$

$$|u_j^n| = |e^{imjh}|\left(1 + \left(\frac{c_0\tau}{h}\right)^2 \sin^2(mh)\right)^{\frac{n}{2}}$$

$n \to \infty \longrightarrow +\infty$ whenever $mh \neq k\pi$, $k \in \mathbb{Z}$

upwind scheme. $\left(\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = f\right)$

if $c(x_j, t_n) > 0$

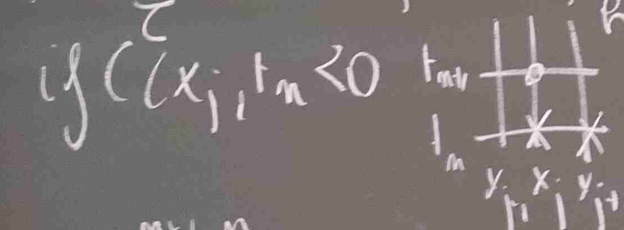


o : to be computed

x : used to approx $\frac{\partial u}{\partial x}(x_j, t_n)$

$$\frac{u_j^{n+1} - u_j^n}{\tau} + c(x_j, t_n)\frac{u_j^n - u_{j-1}^n}{h} = f(x_j, t_n) \quad j \in \mathbb{Z}$$

if $c(x_j, t_n) > 0$



$$\frac{u_j^{n+1} - u_j^n}{\tau} + c(x_j, t_n)\frac{u_{j+1}^n - u_j^n}{h} = f(x_j, t_n) \quad j \in \mathbb{Z}$$

if $f = 0$ $u_0(x) = e^{imx}$ $c(x,t) = c_0 > 0$

$$u_j^1 = u_j^0 - \frac{c_0\tau}{h}\left(u_j^0 - u_{j-1}^0\right) \quad j \in \mathbb{Z}$$

$$= e^{imjh} - \frac{c_0\tau}{h}\left(e^{imjh} - e^{im(j-1)h}\right)$$

$$= e^{imjh}\left(1 - \frac{c_0\tau}{h}\left(1 - e^{-imh}\right)\right)$$

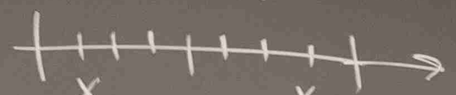
$$u_j^n = \left(\quad\right)^n$$

$$|u_j^n| = |e^{imjh}|\underbrace{\left|1 - \frac{c_0\tau}{h}\left(1 - e^{-imh}\right)\right|^n}_{\leq 1 \text{ if } c_0\tau/h \leq 1}$$

$\forall j \in \mathbb{Z} \quad |u_j^{n+1}| \leq \underbrace{\left(1 - \frac{c_0\tau}{h} + \frac{c_0\tau}{h}\right)}_{=1} \sup_{k\in\mathbb{Z}} |u_k^n|$

The upwind is stable, converges $O(h+\tau)$

What happens if $0 \leq x \leq 1$?



$x_j = jh \quad j = 0, 1, \dots, N+1 \quad h = \frac{1}{N+1}$

$c(0,t)$ $\rightleftharpoons$ $\leftrightharpoons$ $c(1,t)$

$\frac{\partial u}{\partial t}(x,t) + c(x,t)\frac{\partial u}{\partial x}(x,t) = f(x,t) \quad 0 < x < 1, \ t > 0$

$u(x,0) = u_0(x)$

if $c(0,t) > 0$ then $u(0,t) = 0$ (or something else)

if $c(1,t) < 0$ then $u(1,t) = 0$

the upwind scheme is the same but

if $c(x_1, t_n) \geq 0 \quad u_0^n = 0$

if $c(x_N, t_n) < 0 \quad u_{N+1}^n = 0$

scheme

Lemma: if

Extensions

$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}$

if $c(x_j, t_n)$

$u_j^{n+1} - u$

Stability (∞ norm) when $f(x,t)=0$ $c(x,t)=c_0>0$

upw. scheme: $\begin{cases} \frac{u_j^{n+1}-u_j^n}{\tau} + c_0 \frac{u_j^n - u_{j-1}^n}{h} = 0 & j=1,\dots,N+1 \\ \text{with } u_0^n = 0 \end{cases}$

Lemma: if $c_0\tau/h \le 1$ then

- if $u^n \ge 0$ $j=1,\dots,N+1$ then $u_j^{n+1} \ge 0$ $j=1,\dots,N+1$
- $\max_{1\le j\le N+1} |u_j^{n+1}| \le \max_{1\le j\le N+1} |u_j^n|$

Proof: $u_j^{n+1} = u_j^n \left(1 - c_0\frac{\tau}{h}\right) + c_0\frac{\tau}{h} u_{j-1}^n$

$j=1,\dots,N+1$

Remark: the scheme converges $O(h+\tau)$ if $\tau \le h/|c_0|$

Extensions (1d)

- $\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(cu) = f$

if $c(x_j,t_n) > 0$

$$\frac{u_j^{n+1}-u_j^n}{\tau} + \frac{c(x_j,t_n)u_j^n - c(x_{j-1},t_n)u_{j-1}^n}{h} = f(x_j,t_n)$$

- Higher order schemes

Lax-Wendroff $f=0$ $c(x,t)=c_0$

$$u(x_j,t_{n+1}) = u(x_j,t_n) + \tau\frac{\partial u}{\partial t}(x_j,t_n) + \tau^2/2\,\frac{\partial^2 u}{\partial t^2}(x_j,t_n) + O(\tau^3)$$

$$\frac{\partial u}{\partial t}(x, t_n) + c_0 \frac{\partial u}{\partial x}(x, t_n) = 0$$

$$\frac{u(x, t_{n+1}) - u(x, t_n)}{\tau} - \frac{\tau}{2} \underbrace{\frac{\partial^2 u}{\partial t^2}(x, t_n)}_{= c_0^2 \frac{\partial^2 u}{\partial x^2}(x, t_n)} + c_0 \frac{\partial u}{\partial x}(x, t_n) = 0 + O(\tau^2)$$

use centered finite difference formula for $\frac{\partial^2 u}{\partial x^2}$ and $\frac{\partial u}{\partial x}$ $\quad O(h^2 + \tau^2)$

stable if $\tau \leq \frac{h}{|c_0|}$

Hyperbolic

transport

wave

1d $x \in \mathbb{R}$ or $(0,1)$

2d

1d $x, y \in (0,1)^2$

transport 1d $\quad \frac{\partial u}{\partial t}(x,t) + c(x,t)\frac{\partial u}{\partial x}(x,t)$

if $c(ih, t_n) \geq 0$ $\quad \frac{u_i^{n+1} - u_i^n}{\tau} + c(jh, t_n)\frac{u_j^n - u_{j-1}^n}{h} = 0$

< 0 $\quad \frac{u_{i+1}^n - u_i^n}{h}$

upwind scheme : centered scheme + diffusion

$c(x,t) = c_0 > 0$

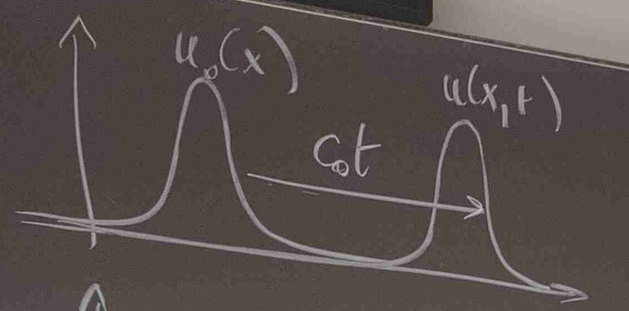
$$\frac{u_j^{n+1} - u_j^n}{\tau} + c_0 \frac{u_j^n - u_{j-1}^n}{h} = 0$$

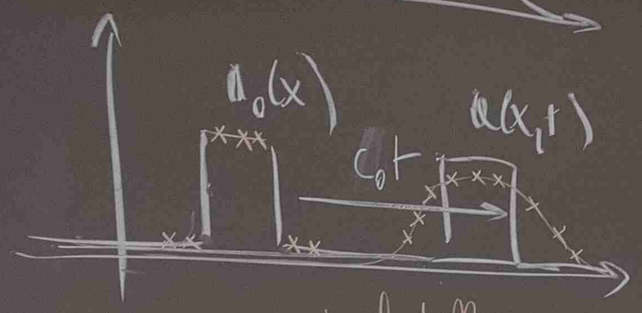
$$= c_0 \frac{u_{j+1}^n - u_{j-1}^n}{2h} - \frac{c_0 h}{2} \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{h^2}$$

centered for $c_0 \frac{\partial u}{\partial x}$

$\frac{\partial^2 u}{\partial x^2}$

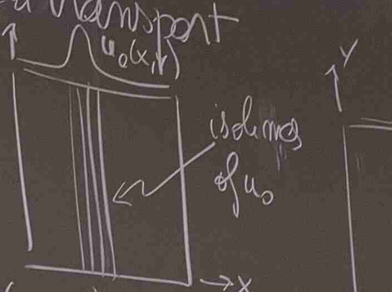
Why

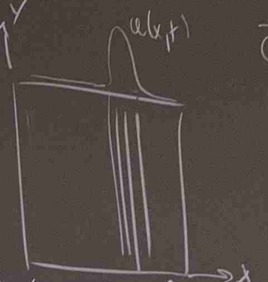




numerical diffusion

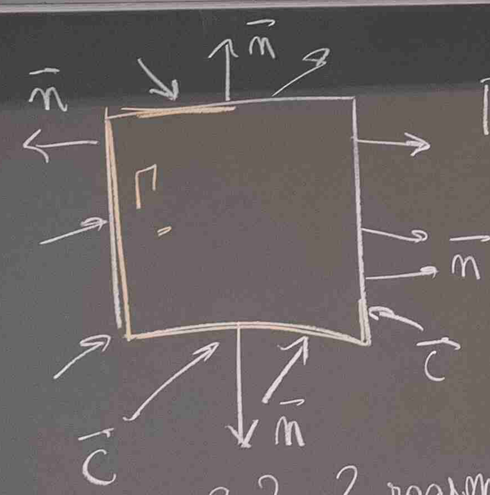
2d transport





$$\vec{c}(x,y,t) = \begin{pmatrix} c_1(x,y,t) \\ c_2(x,y,t) \end{pmatrix} \qquad \vec{\nabla} u = \begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \frac{\partial u}{\partial x_2} \end{pmatrix}$$

$$\begin{cases} \frac{\partial u}{\partial t}(x,y,t) + \vec{c}(x,y,t) \cdot \vec{\nabla} u(x,y,t) = f(x,y,t) & (x,y) \in (0,1)^2 \quad t > 0 \quad (1) \\ u(x,y,0) = u_0(x,y) \\ u(x,y,t) = 0 \quad (x,y) \in \Gamma_- \quad t > 0 \quad \Gamma_- = \{ \end{cases}$$



Γ_- = inflow

$= \{(x,y) \in \partial(0,1)^2 \,;\, \vec{c} \cdot \vec{n} < 0\}$

Why Γ_- ? 2 reasons

- Let $\vec{X}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ the trajectories of the particles

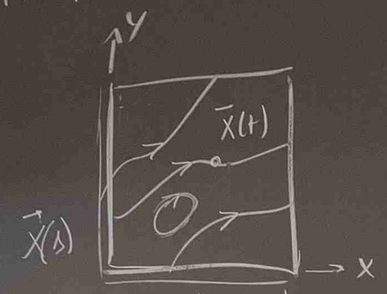
$$\vec{X}'(t) = \vec{c}(\vec{X}(t), t) \qquad (\vec{c} \in C^1)$$

then $\frac{d}{dt} u(\vec{X}(t), t)$ where $u(x,y,t)$ is sol of (1) satisfies

$$\frac{d}{dt} u(\vec{X}(t), t) = \frac{\partial u}{\partial t}(\vec{X}(t), t) + \underbrace{\vec{X}'(t)}_{= \vec{c}(\vec{X}(t), t)} \cdot \vec{\nabla} u(\vec{X}(t), t)$$

$$= f(\vec{X}(t), t)$$

so that $u(\bar{X}(t),t) = u(\bar{X}(s),s) + \int_s^t f(X(\mu),\mu)\,d\mu$



- A priori estimate: multiply (1) by u and $\int_0^1\int_0^1 dx\,dy$ $\left(\Omega=(0,1)^2,\ f=0,\ \operatorname{div}\bar{c}=0\right)$

$$\int_\Omega \frac{\partial u}{\partial t}u\,dx\,dy + \int_\Omega (\bar{c}\cdot\bar{\nabla}u)u\,dx\,dy = 0$$

$$\frac{1}{2}\frac{d}{dt}\int_\Omega u^2\,dx\,dy + ?$$

$$\operatorname{div}\left(\bar{c}\frac{u^2}{2}\right) = \frac{u^2}{2}\underbrace{\operatorname{div}\bar{c}}_{=0} + \bar{c}\cdot\bar{\nabla}\frac{u^2}{2} = (\bar{c}\cdot\bar{\nabla}u)u$$

$$\int_\Omega \operatorname{div}\left(\bar{c}\frac{u^2}{2}\right)dx = \int_\Omega (\bar{c}\cdot\bar{\nabla}u)u\,dx\,dy$$

Thus $\int_\Omega (\bar{c}\cdot\bar{\nabla}u)u\,dx\,dy$

$$= -\int_{\partial\Omega} \bar{c}\cdot\vec{n}\,\frac{u^2}{2}\,ds$$

$$= \int_{\Gamma_-} \underbrace{\bar{c}\cdot\vec{n}}_{\geq 0}\,\underbrace{\frac{u^2}{2}}_{=0}\,ds + \int_{\partial\Omega\setminus\Gamma_-} \underbrace{\bar{c}\cdot\vec{n}}_{\geq 0}\,\frac{u^2}{2}\,ds$$

≥ 0

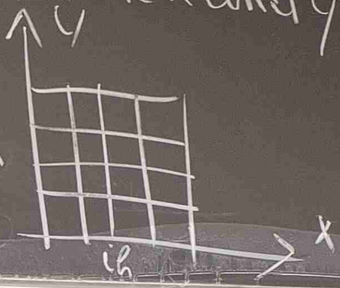
Therefore $\frac{d}{dt}\int_\Omega u^2\,dx\,dy \le 0$

that is $\int_\Omega u^2(x,y,t)\,dx\,dy \le \int_\Omega u_0^2(x,y)\,dx\,dy$

Scheme: upwind in the x and y directions

$u_{ij}^n \simeq u(ih, jh, t_n)$ $i,j = 1,\dots,N$

$h = \frac{1}{(N+1)}$ N points in x,y directions



if $c_1(ih, jh, t_n) > 0$ if $c_2(ih, jh, t_n) \gtrless 0$

$$\frac{u_{ij}^{n+1} - u_{ij}^n}{\tau} + c_1(ih, jh, t_n)\,\frac{u_{ij}^n - u_{i-1,j}^n}{h} + c_2(ih, jh, t_n)\,\frac{u_{ij}^n - u_{i,j-1}^n}{h} = f(ih, jh, t_n)$$

+ 3 other cases $\qquad \frac{u_{i,j+1}^n - u_{ij}^n}{h}$

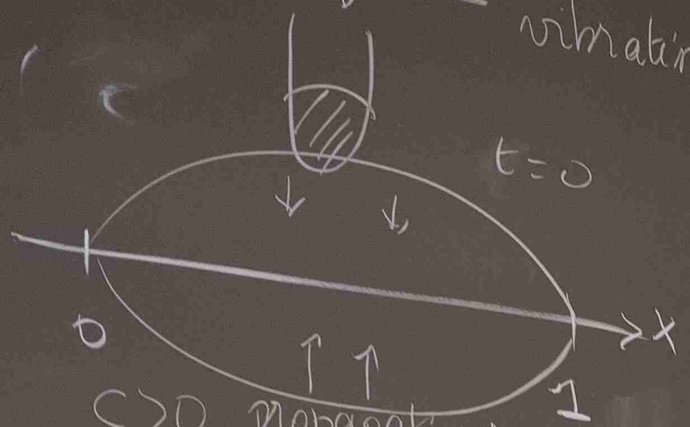
The 2d upwind scheme is stable under condition $O(h+\tau)$ $(\vec{c}(x,y,t) = \vec{c}_0 = \text{cst})$

Remarks:

- if $f = 0$ where u

wave propagation

vibrating string



$u_0(x)$ def. at $t=0$

$v_0(x)$ def. speed

$c>0$ propagation speed

$f(x,t)$ force on the string

$u(x,t)$ deformation

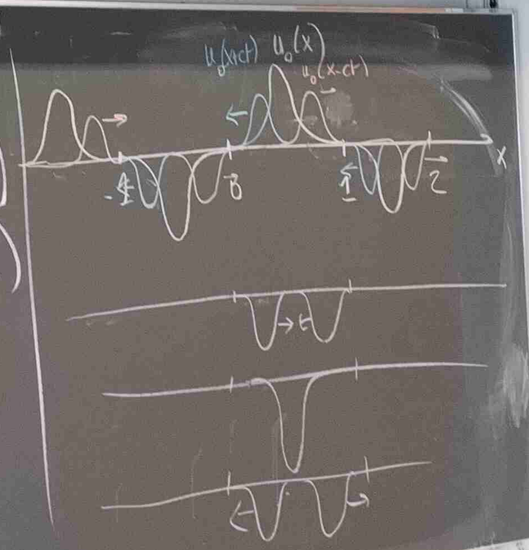
u such that

$$\begin{cases} \frac{\partial^2 u(x,t)}{\partial t^2} - c^2 \frac{\partial^2 u(x,t)}{\partial x^2} = f(x,t) & 0<x<1 \quad t>0 \\ u(0,t) = 0 = u(1,t) & t>0 \\ u(x,0) = u_0(x) & 0<x<1 \\ \frac{\partial u}{\partial t}(x,0) = v_0(x) & 0<x<1 \end{cases}$$

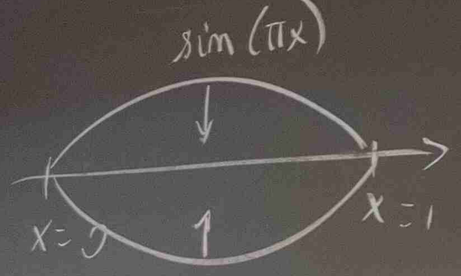
Remarks:

- if $f=0$ $v_0=0$ then $u(x,t) = \frac{1}{2}\left(u_0(x-ct)+u_0(x+ct)\right)$

  where u_0 is extended on the real line ($u_0: \mathbb{R}\to\mathbb{R}$)
  - 2-periodic
  - odd



- if $f=0$ $v_0=0$ / $u_0(x)=\sin(\pi x)$
then $u(x,t)=\sin(\pi x)\cos(\pi c t)$



first term in the Fourier series

- A priori error estimate (energy conservation)

Multiply the wave eq by $\frac{\partial u}{\partial t}(x,t)$ (velocity) and $\int_0^1 dx$:

$$\underbrace{\int_0^1 \frac{\partial^2 u}{\partial t^2}\frac{\partial u}{\partial t}\,dx}_{\frac{1}{2}\frac{d}{dt}\int_0^1\left(\frac{\partial u}{\partial t}\right)^2 dx} - c^2\underbrace{\int_0^1 \frac{\partial^2 u}{\partial x^2}\frac{\partial u}{\partial t}\,dx}_{\text{int. by parts w.r.t. } x} = \underbrace{\int_0^1 f\frac{\partial u}{\partial t}\,dx}_{0}$$

rate of change of kinetic energy

power forces

$(uv)' = u'v + uv'$

$u \to u'$

$(u'v)' = u''v + u'v'$

$$\frac{1}{2}\frac{d}{dt}\int_0^1 \left(\frac{\partial u}{\partial t}\right)^2 dx + c^2 \int_0^1 \frac{\partial u}{\partial x}\frac{\partial}{\partial x}\frac{\partial u}{\partial t}\,dx + \left[\frac{\partial u}{\partial x}(x,t)\,\frac{\partial u}{\partial t}(x,t)\right]_{x=0}^{x=1}$$

$$\underbrace{\frac{1}{2}\frac{d}{dt}\int_0^1 \left(\frac{\partial u}{\partial x}\right)^2 dx} = \int_0^1 f\frac{\partial u}{\partial t}\,dx$$

$$\frac{\partial u}{\partial t}(0,t) = \lim_{\varepsilon\to 0} \frac{u(0,t+\varepsilon)-u(0,t)}{\varepsilon} = 0$$

Finally $\frac{1}{2}\frac{d}{dt}\int_0^1 \left(\left(\frac{\partial u}{\partial t}\right)^2 + c^2\left(\frac{\partial u}{\partial x}\right)^2\right)dx = \int_0^1 f\frac{\partial u}{\partial t}\,dx$

def energy

if $f = 0$ then $\int_0^1 \left(\left(\frac{\partial u}{\partial t}(x,t)\right)^2 + c^2\left(\frac{\partial u}{\partial x}(x,t)\right)^2\right)dx$

$= \int_0^1 \left(\left(v_0(x)\right)^2 + c^2\left(u_0'(x)\right)^2\right)dx$

if $f \neq 0$ $\frac{1}{2}\frac{d}{dt}\underbrace{\int_0^1 \left(\left(\frac{\partial u}{\partial t}\right)^2 + c^2\left(\frac{\partial u}{\partial x}\right)^2\right)dx}_{y(t)} \overset{\text{CS.}}{\leq} \underbrace{\left(\int_0^1 f(x,t)^2\,dx\right)^{1/2}}_{c(t)} \left(\int_0^1 \left(\frac{\partial u}{\partial t}(x,t)\right)^2 dx\right)^{1/2}$

Space discret

$i = 1, \dots, N$

$h = \frac{1}{N+1}$

Differential

$\vec{u}(t) = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix}$

(2) $\begin{cases} \vec{u}''(t) + \\ \vec{u}(0) = \vec{u}_0 \\ \vec{u}'(0) = \vec{v}_0 \end{cases}$

$$\frac{1}{2}\frac{d}{dt}y(t) \le C(t)\,(y(t))^{1/2}$$

$$\frac{d}{dt}(y(t))^{1/2} \le C(t)$$

$$(y(t))^{1/2} \le (y(0))^{1/2} + \int_0^t C(s)\,ds$$

Space discretization

$i = 1, \dots, N$

$h = \frac{1}{N+1}$

wave eq at x_i $i = 1, \dots, N$

$$\frac{\partial^2 u}{\partial t^2}(x_i,t) - c^2\frac{\partial^2 u}{\partial x^2}(x_i,t) = f(x_i,t)$$

$$\frac{\partial^2 u}{\partial t^2}(x_i,t) - c^2\frac{u(x_{i-1},t) - 2u(x_i,t) + u(x_{i+1},t)}{h^2} = f(x_i,t) + O(h^2)$$

$u_i(t)$ approx of $u(x_i,t)$: scheme

$$\begin{cases} u_i''(t) - c^2\dfrac{u_{i-1}(t) - 2u_i(t) + u_{i+1}(t)}{h^2} = f(x_i,t) & i = 1, \dots, N \\ u_0(t) = 0 = u_{N+1}(t) \end{cases}$$

Differential system

$$\vec{u}(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{pmatrix} \qquad \vec{f}(t) = \begin{pmatrix} f(x_1,t) \\ \vdots \\ f(x_N,t) \end{pmatrix} \qquad A = \frac{c^2}{h^2}\begin{pmatrix} 2 & -1 & 0 \\ -1 & \ddots & \ddots \\ 0 & \ddots & -1 \\ & & -1 \quad 2 \end{pmatrix} \qquad \vec{u}_0 = \begin{pmatrix} u_0(x_1) \\ \vdots \\ u_0(x_N) \end{pmatrix} \qquad \vec{v}_0 = \begin{pmatrix} v_0(x_1) \\ \vdots \\ v_0(x_N) \end{pmatrix}$$

$$(2)\quad \begin{cases} \vec{u}''(t) + A\,\vec{u}(t) = \vec{f}(t) & t > 0 \\ \vec{u}(0) = \vec{u}_0 \\ \vec{u}'(0) = \vec{v}_0 \end{cases}$$

Time discretization: $t_n = n\tau$ $n = 0, 1, \dots$ $\vec{u}^n$ approx of $\vec{u}(t_n)$ $O(\tau^2)$

Explicit scheme: $\dfrac{\vec{u}^{n-1} - 2\vec{u}^n + \vec{u}^{n+1}}{\tau^2} + A\vec{u}^n = \vec{f}(t_n)$ stable if $\tau < \frac{2}{c}$

Implicit scheme: $\dfrac{\vec{u}^{n-1} - 2\vec{u}^n + \vec{u}^{n+1}}{\tau^2} + A\,\dfrac{\vec{u}^{n-1} + 2\vec{u}^n + \vec{u}^{n+1}}{4} = \vec{f}(t_n)$ stable $\forall\, h, \tau > 0$

origin of the implicit scheme

Write (2) as a first order system

$$\begin{cases} \vec{u}' = \vec{v} \\ \vec{v}' + A\vec{u} = \vec{f} \\ \vec{u}(0) = \vec{u}_0 \\ \vec{v}(0) = \vec{v}_0 \end{cases}$$

Crank-Nicolson ($O(\tau^2)$)

$$\frac{\vec{u}^{n+1} - \vec{u}^{n}}{\tau} = \frac{\vec{v}^{n} + \vec{v}^{n+1}}{2}$$

diff between 2 consec. time steps

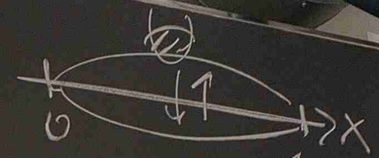
$$\frac{\vec{v}^{n+1} - \vec{v}^{n}}{\tau} + A\frac{\vec{u}^{n} + \vec{u}^{n+1}}{2} = \frac{\vec{f}^{n} + \vec{f}^{n+1}}{2}$$

$$\frac{\vec{u}^{n+1} - 2\vec{u}^{n} + \vec{u}^{n-1}}{\tau} = \frac{\vec{v}^{n+1} - \vec{v}^{n} + \vec{v}^{n} - \vec{v}^{n-1}}{2}$$

use to eliminate v^{n-1}, v^{n}, v^{n+1}

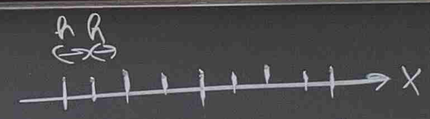
Wave eq

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2}(x,t) - c^2 \dfrac{\partial^2 u}{\partial x^2}(x,t) = f(x,t) & 0<x<1 \quad t>0 \\ u(0,t) = 0 \quad u(1,t) = 0 & t>0 \\ u(x,0) = u_0(x) \\ \dfrac{\partial u}{\partial t}(x,0) = v_0 \end{cases}$$



$f = 0 \quad \frac{1}{2}\frac{d}{dt} \int_0^1 \left(\left(\frac{\partial u}{\partial t}(x,t)\right)^2 + c^2 \left(\frac{\partial u}{\partial x}(x,t)\right)^2 \right) dx = 0$ (1)

Space discretization



$h = \frac{1}{N+1}$

$u(x_i,t)$ approx by $y_i(t)$

$x_i = ih \quad x_N = 1$

$\vec{u}''(t) + A\vec{u}(t) = \vec{f}(t)$

$$\vec{u}(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_N(t) \end{pmatrix} \qquad A = \frac{c^2}{h^2}\begin{pmatrix} 2 & -1 & (0) \\ -1 & \ddots & -1 \\ (0) & -1 & 2 \end{pmatrix} \qquad \vec{f} = \begin{pmatrix} f(x_1,t) \\ \vdots \\ f(x_N,t) \end{pmatrix}$$

$$\vec{u}_0 = \begin{pmatrix} u_0(x_1) \\ \vdots \\ u_0(x_N) \end{pmatrix} \qquad \vec{v}_0 = \begin{pmatrix} v_0(x_1) \\ \vdots \\ v_0(x_N) \end{pmatrix}$$

$\vec{u}(0) = \vec{u_0}$

$\vec{u}'(0) = \vec{v_0}$

Stability of

Take the s

$\vec{u}^{n+1} - \vec{u}^{n-1}$

$\vec{u}^{n+1} - \vec{u}^{n}$

$\vec{u}$

$\tau = \frac{1}{M}$

$\vec{u}^n$ approx

(therefore u_i^n

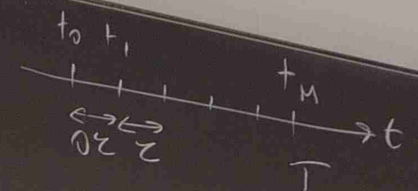
Implicit scheme

Explicat

Time discretization

$t_n = n\tau \quad n = 0, \dots, M$

$\tau = \frac{T}{M}$



$\vec{u}^n$ approx of $\vec{u}(t_n)$

(therefore u_i^n approx of $u(x_i, t_n)$)

Implicit scheme (Newmark)

$$\frac{\vec{u}^{n+1} - 2\vec{u}^n + \vec{u}^{n-1}}{\tau^2} + A\,\frac{\vec{u}^{n+1} + 2\vec{u}^n + \vec{u}^{n-1}}{4} = \vec{f}(t_n) \quad (2)$$

Explicit

same

$$+ A\vec{u}^n = \vec{f}(t_n) \quad (3)$$

Both scheme allows to compute $\vec{u}^2, \dots, \vec{u}^M$

(after, how to compute $\vec{u}^0, \vec{u}^1$?)

are consistent $O(h^2 + \tau^2)$

Question: are they stable?

We wish to reproduce (1) at the discrete level

L^2 Stability of the implicit scheme ($f = 0$)

Take the scalar prod of (2) with

$\vec{u}^{n+1} - \vec{u}^{n-1}$ (discrete counterpart of $\frac{d}{dt}$)

$$\left(\frac{\vec{u}^{n+1} - \vec{u}^n}{\tau^2} + \left(\frac{\vec{u}^n - \vec{u}^{n-1}}{\tau^2}\right)\right)^T \left(\vec{u}^{n+1} - \vec{u}^n + (\vec{u}^n - \vec{u}^{n-1})\right) + \left(A\,\frac{\vec{u}^{n+1} + \vec{u}^n + \vec{u}^n + \vec{u}^{n-1}}{4}\right)^T \left(\vec{u}^{n+1} + \vec{u}^n - (\vec{u}^n + \vec{u}^{n-1})\right) = 0$$

$\langle \vec{a}, \vec{b} \rangle_A = \vec{a}^T A \vec{b}$

$$\left(\vec{a} + \vec{b}\right)^T \left(\vec{a} - \vec{b}\right)$$

$\langle \vec{a}+\vec{b}, \vec{a}-\vec{b} \rangle_A$

$\|\vec{a}\|^2 - \|\vec{b}\|^2$

$\langle \vec{a}, \vec{a} \rangle_A - \langle \vec{b}, \vec{b} \rangle_A$

$$\left\| \frac{\vec{u}^{n+1} - \vec{u}^{n}}{\tau^2} \right\|^2 - \left\| \frac{\vec{u}^{n} - \vec{u}^{n-1}}{\tau^2} \right\|^2 + \frac{1}{4}\left(\vec{u}^{n} + \vec{u}^{n+1}\right)^T A\left(\vec{u}^{n} + \vec{u}^{n+1}\right) - \frac{1}{4}\left(\vec{u}^{n-1} + \vec{u}^{n}\right)^T A\left(\vec{u}^{n-1} + \vec{u}^{n}\right) = 0 \qquad (4)$$

Sum $n = 1, \dots, M-1$:

$$\left\| \frac{\vec{u}^{M} - \vec{u}^{M-1}}{\tau} \right\|^2 + \frac{1}{4}\left(\vec{u}^{M} + \vec{u}^{M-1}\right)^T A\left(\vec{u}^{M} + \vec{u}^{M-1}\right) = \left\| \frac{\vec{u}^{1} - \vec{u}^{0}}{\tau} \right\|^2 + \frac{1}{4}\left(\vec{u}^{1} + \vec{u}^{0}\right)^T A\left(\vec{u}^{1} + \vec{u}^{0}\right)$$

Discret counterpart of (1), no restriction on $h, \tau > 0$

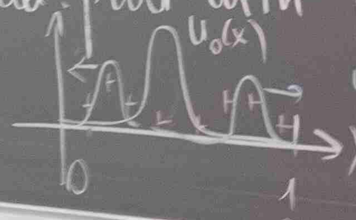
The scheme is stable $\forall\, h, \tau > 0$

Exercice: convergence proof

$\vec{U}(t_n) = \begin{pmatrix} \vdots \\ u(x_N, t_n) \end{pmatrix}$ exact sol. error $\vec{U}(t_n) - \vec{u}^n = \vec{E}^n$

$$\frac{\vec{E}^{n-1} - 2\vec{E}^{n} + \vec{E}^{n+1}}{\tau^2} + A\,\frac{\vec{E}^{n-1} + 2\vec{E}^{n} + \vec{E}^{n-1}}{4} = \vec{r}^{\,n} \qquad |r_i^n| \le C\left(h^2 + \tau^2\right)$$

Take the scal. prod with $\vec{E}^{n+1} - \vec{E}^{n-1}$



$$u(x,t) = \frac{1}{2}\left(u_0(x-ct) + u_0(x+ct)\right)$$

Stability of the explicit scheme

(3) is written as (2) + rhs

$$\frac{\vec{u}^{n+1} - 2\vec{u}^{n} + \vec{u}^{n-1}}{\tau^2} + A\,\frac{\vec{u}^{n-1} + 2\vec{u}^{n} + \vec{u}^{n+1}}{4} = A\,\frac{\vec{u}^{n-1} - 2\vec{u}^{n} - \vec{u}^{n+1}}{4}$$

Take the scalar prod with $\vec{u}^{n+1} - \vec{u}^{n-1}$

We obtain the lhs of (4) and the rhs is

$$\frac{1}{4}\left(\vec{u}^{n+1} - \vec{u}^{n} + (\vec{u}^{n} - \vec{u}^{n-1})\right)^T A\,\frac{\vec{u}^{n+1} - \vec{u}^{n} - (\vec{u}^{n} - \vec{u}^{n-1})}{4}$$

$$\langle \vec{a} + \vec{b}\,,\ \vec{a} - \vec{b} \rangle_A$$

$$= \frac{1}{4}(\vec{u}^{n+1} - \vec{u}^{n})^T A(\vec{u}^{n+1} - \vec{u}^{n}) - \frac{1}{4}(\vec{u}^{n} - \vec{u}^{n-1})^T A(\vec{u}^{n} - \vec{u}^{n-1})$$

Sum $n = 1, \dots, M-1$; we obtain

$$\left\| \frac{\vec{u}^M - \vec{u}^{M-1}}{\tau} \right\|^2 + \frac{1}{4}\left(\vec{u}^M + \vec{u}^{M-1}\right)^T A\left(\vec{u}^M + \vec{u}^{M-1}\right) + \frac{1}{4}\left(\vec{u}^1 - \vec{u}^0\right)^T A\left(\vec{u}^1 - \vec{u}^0\right)$$

$$= \left\| \frac{\vec{u}^1 - \vec{u}^0}{\tau} \right\|^2 + \frac{1}{4}\left(\vec{u}^1 + \vec{u}^0\right)^T A\left(\vec{u}^1 + \vec{u}^0\right) + \underbrace{\frac{1}{4}\left(\vec{u}^M - \vec{u}^{M-1}\right)^T A\left(\vec{u}^M - \vec{u}^{M-1}\right)}_{\text{bad term}}$$

$$\leq \frac{1}{4}\lambda_N \left\|\vec{u}^M - \vec{u}^{M-1}\right\|^2$$

A spd $A\vec{\varphi}_i = \lambda_i \vec{\varphi}_i$, $\vec{\varphi}_i^T \vec{\varphi}_j = \delta_{ij}$

$0 < \lambda_1 \leq \dots \leq \lambda_N$

$\forall \vec{v} \in \mathbb{R}^N$ $\lambda_1 \|\vec{v}\|^2 \leq \vec{v}^T A \vec{v} \leq \lambda_N \|\vec{v}\|^2$ $\left(\vec{v} = \sum_j \alpha_j \vec{\varphi}_j\right)$

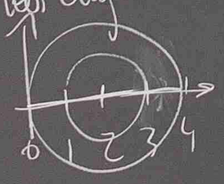
Therefore if $\frac{1}{\tau^2} - \frac{1}{4}\lambda_N \geq 0$ then

$$\left(\vec{u}^M + \vec{u}^{M-1}\right)^T A\left(\vec{u}^M + \vec{u}^{M-1}\right) \leq \frac{\|\vec{u}^1 - \vec{u}^0\|^2}{\tau^2} + \frac{1}{4}\left(\vec{u}^1 + \vec{u}^0\right)^T A\left(\vec{u}^1 + \vec{u}^0\right)$$ and the scheme is stable

Stability cond: $\tau^2 \leq \frac{4}{\lambda_N}$

Recall that

$\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$ has eigenvalues $0 < \lambda_i < 4$



$\lambda_i = 2 - 2\cos\frac{i\pi}{N+1}$ $\quad N \to \infty$

$\lambda_1 \to 0$ $\quad \lambda_N \to 4$

$$A = \frac{c^2}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}$$

$$\lambda_N \underset{N\to\infty}{\sim} \frac{4c^2}{h^2}$$

$\tau^2 \leq \frac{4}{4c^2}h^2$ N is large

$\tau^2 \leq \frac{h^2}{c^2}$ $\left(\tau \leq \frac{h}{c}\right)$ $\quad u(x,t) = \frac{1}{2}\left(u_0(x-ct) + u_0(x+ct)\right)$

(recall that $\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0$, $\tau \leq \frac{h}{c}$ upwind) $\quad u(x,t) = u_0(x-ct)$

$\left(\frac{\partial u}{\partial t} + c_0\frac{\partial u}{\partial x} = 0 \;,\; \frac{\partial^2 u}{\partial t^2} + c_0\frac{\partial^2 u}{\partial t \partial x} = 0,\; \frac{\partial^2 u}{\partial t \partial x} + c_0\frac{\partial^2 u}{\partial x^2} = 0,\; \frac{\partial^2 u}{\partial t^2} - c_0^2\frac{\partial^2 u}{\partial x^2} = 0,\; \frac{\partial u}{\partial t}(0) = -c_0\frac{\partial u}{\partial x}(0)\right)$

Exp: Set N, $h = \frac{1}{N+1}$

Vary τ $n = 1, \dots, M = 1000$

Observation: if $\tau \le \frac{h}{c}$ get some results (no explosion)

if $\tau > \frac{h}{c}$ may have explosion (NaN)

$|u_i^n| > e^{317}$

Extension:

- damping $\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} + \alpha \frac{\partial u}{\partial t} = 0 \quad \alpha > 0$ mult by $\frac{\partial u}{\partial t}$

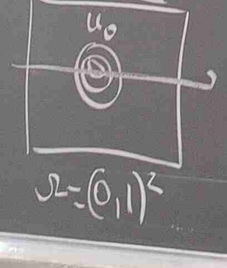
$$\frac{1}{2}\frac{d}{dt}\int_0^1 \left(\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial u}{\partial x}\right)^2\right) dx + \alpha \int_0^1 \left(\frac{\partial u}{\partial t}\right)^2 = 0$$

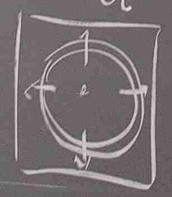
Therefore $\frac{d}{dt}\int_0^1 \left(\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial u}{\partial x}\right)^2\right) dx \le 0$ the energy is decreasing

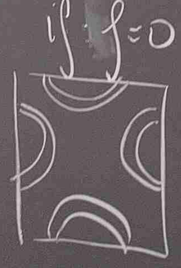
- 2, 3d: $\frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = f \qquad \Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$ 2d

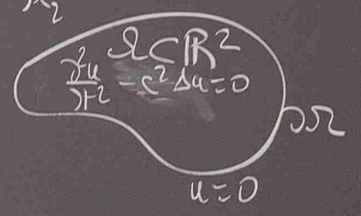
mult by $\frac{\partial u}{\partial t}$:

$$\frac{1}{2}\frac{d}{dt}\int_\Omega \left(\left(\frac{\partial u}{\partial t}\right)^2 + |\nabla u|^2\right) dx = 0$$









How to compute $\vec{u}^0$ and $\vec{u}^1$?

$$\vec{u}^0 = \begin{pmatrix} \vdots \\ u^0 \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ u_0(x_i) \\ \vdots \end{pmatrix}$$

$\vec{u}^1$? (use $\frac{\partial u}{\partial t}(x,0) = v_0(x)$)

$$u(x,\tau) = u(x,0) + \tau \frac{\partial u}{\partial t}(x,0) + \frac{\tau^2}{2}\frac{\partial^2 u}{\partial t^2}(x,0) + O(\tau^3) \quad \text{(Taylor)}$$

$$\tau v_0(x) + \frac{\tau^2}{2} c^2 \frac{\partial^2 u}{\partial x^2}(x,0) + \dots$$

Scheme: $u_i^1 = u_i^0 + \tau v_0(x_i) + \frac{\tau^2 c^2}{2} \frac{u_{i-1}^0 - 2u_i^0 + u_{i+1}^0}{h^2}$

This yields a final error $O(h^2 + \tau^2)$

$= \left\| \frac{\vec{u}^1 - \vec{u}^0}{\tau} \right\|$

Therefore

Stability

Recall

$A =$

ODE

$$\bar{y}'(t) = \bar{f}(t, \bar{y}(t))$$
$$\bar{y}(t_0) = \bar{y}_0 \in \mathbb{R}^d$$

Explicit Runge Kutta s stages

first time

$$y_1 \text{ approx } y(t_1) = y(t_0 + h)$$

$$k_1 = f(t_0, y_0)$$
$$k_2 = f(t_0 + c_2 h, y_0 + h a_{21} k_1)$$
$$\vdots$$
$$k_s = f(t_0 + c_s h, y_0 + h \sum_{i=1}^{s-1} a_{si} k_i)$$

$$y_1 = y_0 + h \sum_{i=1}^{s} b_i k_i$$

c_2	a_{21}			
c_3	a_{31}	$\cdots$	$a_{s,s-1}$	
	b_1	$\cdots$	b_{s-1}	b_s

c_i, b_i, a_{ij} ? $\quad |y(t_0+h) - y_1| \leq C h^{p+1}$

$p = 1, 2, 3, 4$

$\bar{y}'(t) = f(\bar{y}(t))$ implicit or explicit

$p = 1$: $y(t_0 + h) = y(t_0) + h y'(t_0) + O(h^2)$

$= y_0 + h f(y_0) + O(h^2)$

$$y_1(h) = y_0 + h \underbrace{\sum_{i=1}^{s} b_i k_i(h)}_{\Phi(h)}$$

$$k_i(h) = f\Big(y_0 + \underbrace{h \sum_{j=1}^{s} a_{ij} k_j(h)}_{\Psi_i(h)}\Big) \qquad i = 1, \dots, s$$

$\Phi(h)$

$\Phi'(0)$

$y(t_0 + h)$

t_1

if $|\Phi(t, z$

then $|y($

$y(t_{n+1}) - y_$

$$y_1(h) = y_0 + \underbrace{\phi(0)}_{0} + h\phi'(0) + O(h^2)$$

$$\phi'(h) = \sum_{i=1}^{s} b_i k_i(h) + h \sum_{i=1}^{s} b_i k_i'(h)$$

$$\phi'(0) = \sum_{i=1}^{s} b_i \, f(y_0)$$

$$y(\underbrace{t_0 + h}_{t_1}) - y_1(h) = \left(1 - \sum_{i=1}^{s} b_i\right) f(y_0) + O(h^2)$$

order 1 if $\sum b_i = 1$

$y' = f(y)$

$y'' = f'(y)y'$

$y''(t_0) = f'(y_0)f(y_0)$

Convergence ($h_n = h$ constant, $t_n = nh$)

$y_{n+1} = y_n + h\phi_f(t_n, y_n, h_n)$ Expl RK

$\left\| y(t+h) - \left(y(t) + h\phi(t, y(t), h)\right) \right\| \le Ch^{p+1}$

if $\left| \phi(t, z_1, h) - \phi(t, z_2, h) \right| \le \Lambda |z_1 - z_2|$

then $|y(t_n) - y_n| \le \dfrac{Ch^p}{\Lambda}\left(e^{\Lambda t_n} - 1\right)$

$$\left| y(t_{n+1}) - y_{n+1} \right| = \left| y(t_{n+1}) - \left(y_n + h\phi(t_n, y_n, h_n)\right) + \left(y(t_n) + h\phi(t_n, y(t_n), h)\right) - \left(y(t_n) + h\phi(t_n, y(t_n), h)\right) \right|$$

$|y(t_{n+1}) - y_{n+1}| \le Ch^{p+1} +$

$|y(t_n) - y_n|(1 + h\Lambda)$

$\le Ch^{p+1}\left(1 + (1+h\Lambda)\right) + \left(y(t_{n-1}) - y_{n-1}\right)(1+h\Lambda)^2$

$\le Ch^{p+1}\left(1 + (1+h\Lambda) + \dots + (1+h\Lambda)^n\right)$

$1+x+\dots+x^n = \frac{x^{n+1}-1}{x-1}$ | $x>0$ $1+x \le e^x$

$|y(t_{n+1}) - y_{n+1}| \le Ch^{p+1} \frac{(1+h\lambda)^{n+1}-1}{1+h\lambda-1}$

$\le \frac{Ch^p}{\lambda}\left(e^{\lambda(n+1)h} - 1\right)$

y_1 s.t. $|y(t_0+h) - y_1| = O(h^{p+1})$

$\hat{y}_1$ — $|y(t_0+h) - \hat{y}_1| = O(h^p)$

$|y_1 - \hat{y}_1| = |y(t_0+h) - y_1|$

$\le \underbrace{O(h^p)}$ $|y(t_0+h) - \hat{y}_1|$

$\le \underbrace{|y(t_0+h) - y_1|}_{O(h^{p+1})} + \underbrace{|y_1 - \hat{y}_1|}_{\text{Computable}}$

$|y_1 - \hat{y}_1|$ is a computable approx of $|y(t_0+h) - \hat{y}_1|$

(provided h is small enough)

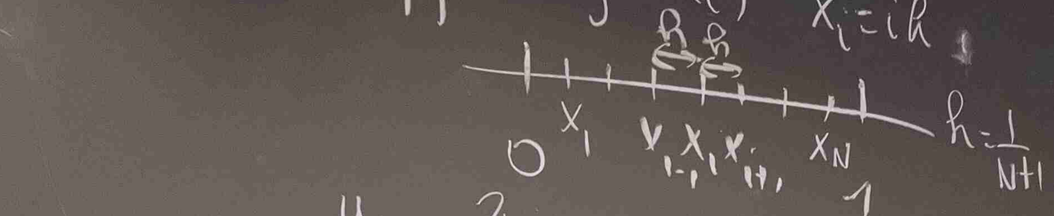
h given compute $|y_1(h) - \hat{y}_1(h)| = Ch^p$ → approx of C

h_{opt} such that $|y_1(h_{opt}) - \hat{y}_1(h_{opt})| = Tol = Ch_{opt}^p \to h_{opt}$

Elliptic

$$\begin{cases} -u''(x) = f(x) & 0<x<1 \\ u(0)=0 \\ u(1)=0 \end{cases}$$

u_i approx of $u(x_i)$ $\quad x_i = ih$



$$-\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} = f(x_i)$$

$A\vec{u} = \vec{f}$

$\vec{U} = \begin{pmatrix} u(x_1) \\ \vdots \\ u(x_N) \end{pmatrix}$ $\quad A\vec{U} = \vec{f} + \vec{r}$

$|r_i| \leq \frac{1}{12} \max_{0\leq x\leq 1} |u^{(4)}(x)| \, h^2$

$A(\vec{U} - \vec{u}) = \vec{r}$

$\|\vec{v}\|_\infty = \max_{1\leq i\leq N} |v_i|$

Lemma: $A\vec{v} = \vec{g} \quad \|\vec{v}\|_\infty \leq \frac{1}{8}\|\vec{g}\|_\infty$

$$A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{pmatrix}$$

- $A\vec{w} = \vec{1}$ $\quad (-w''(x) = 1)$ $\quad w(x) = \frac{x(1-x)}{2}$

  $w_i = w(x_i)$

- $(A\vec{z} \geq \vec{0}) \Rightarrow (\vec{z} \geq \vec{0})$

$\|\vec{U} - \vec{u}\|_\infty \leq \frac{1}{8}\|\vec{r}\|_\infty$

$\leq Ch^2$

$\|\vec{U} - \vec{u}\|_2^2 = \sum_{i=1}^{N} (u(x_i) - u_i)^2$

$\leq NCh^4$

$\leq Ch^3$

$h = \frac{1}{N+1}$

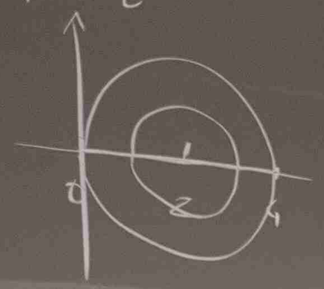
$N \leq N+1 = \frac{1}{h}$

$A\vec{\varphi}_i = \lambda_i \vec{\varphi}_i \qquad \vec{\varphi}_i^T \vec{\varphi}_j = \delta_{ij} \quad i,j = 1,\dots,N$

$0 < \lambda_1 \le \lambda_2 \le \dots \le \lambda_N$

$\forall \vec{v} \in \mathbb{R}^N \qquad \lambda_1 \|v\|_2^2 \le \vec{v}^T A \vec{v} \le \lambda_N \|v\|_2^2$

$A = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{pmatrix}$



$\vec{v}^T A \vec{v} = v_1^2 + (v_1 - v_2)^2 + \dots + (v_{N-1} - v_N)^2 + v_N^2$

$0 \le \lambda_i \le 4 \qquad 0 < \lambda_i \le 4$

$A = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{pmatrix}$

$\lambda_i = 2\left(1 - \cos\frac{i\pi}{N+1}\right)(N+1)^2$

$\lambda_1 = 2\left(1 - \cos\frac{\pi}{N+1}\right)(N+1)^2$

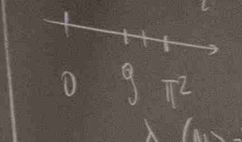
$= 2\left(\frac{\pi^2}{2(N+1)^2} + o\left(\frac{1}{N^2}\right)\right)(N+1)^2$

$\sim \pi^2$

$\varphi(x) = \sin(\pi x) \qquad -\varphi'' = \pi^2 \varphi$

$N = 2 \qquad A = 9\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$

$\lambda_1 = 9 \qquad \lambda_2 = 27$



$\lambda_1 \ (N \ge 2)$

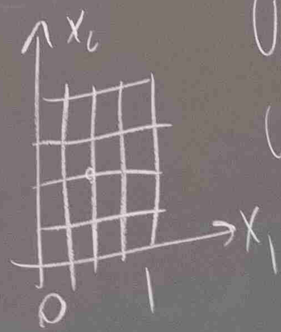
$\lambda_N \sim \frac{4}{h^2} = 4(N+1)^2$

$A(\vec{U} - \vec{u}) = \vec{r}$ recall prod $\vec{U} - \vec{u}$

$\lambda_1 \|\vec{U} - \vec{u}\|_2^2 \le (\vec{U} - \vec{u})^T A(\vec{U} - \vec{u}) = \vec{r}^T(\vec{U} - \vec{u}) \le \|\vec{r}\|_2 \|\vec{U} - \vec{u}\|_2$

$\|\vec{U} - \vec{u}\|_2 \le \frac{\|\vec{r}\|_2}{\lambda_1} \le 9\|\vec{r}\|_2 \le 9Ch^{3/2}$

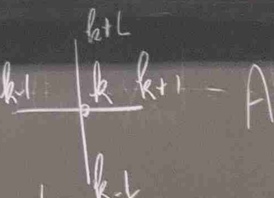
$\begin{cases} -\Delta u = f & \text{in } \Omega = (0,1)^2 \\ u = 0 & \text{on } \partial\Omega \end{cases}$



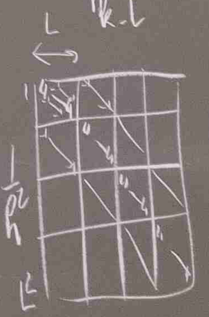
$U_{i,j}$ approx $u(ih, jh)$

$U_{i,j} \to u_k$

$i,j = 1,\dots,L$



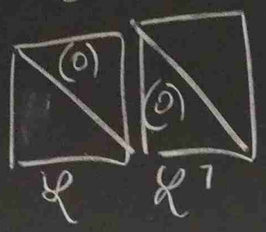
$A\vec{u} = f \quad \vec{u} \in \mathbb{R}^{L^2}$



$A = \frac{1}{h^2}\left(D_L \otimes I_L + I_L \otimes D_L\right)$

$D_L = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{pmatrix}$

How to solve $A\bar{u} = \bar{f}$?

- $A = \mathcal{L}\mathcal{L}^T$:  Cholesky

- gradient $\left(\min_{\bar{v}} \left(\frac{1}{2} \bar{v}^T A \bar{v} - \bar{f}^T \bar{v} \right) \right)$

  gradient $A\bar{v} - \bar{f}$

$$\bar{v}^{k+1} = \bar{v}^k + \alpha_k (\bar{f} - A\bar{v}^k)$$

- Conjugate gradient

gradient number of it

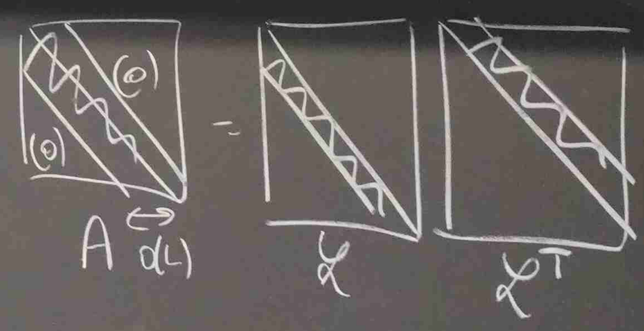
$$\left(\frac{\lambda_N}{\lambda_1} \right)^2$$

Conj. gradient

$$\frac{\lambda_N}{\lambda_1} \sim O(L)$$

$d = 2$

	unknowns	Mem	CPU
Cholesky	$O(L^2)$	$O(L^2 L)$	$O(L^2 L^2)$
Conj grad	$O(L^2)$	$O(L^2)$	$O(L^2 L)$

L^2 

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = f \qquad u$$

$$\frac{1}{2}\frac{d}{dt}\int_0^1 (u(x,t))^2 dx + \int_0^1 \left(\frac{\partial u}{\partial x}(x,t)\right)^2 dx = \int_0^1 f(x,t)u(x,t)\,dx$$

if $f(x,t) \geq 0$ and $u(x,0) \geq 0$ then $u(x,t) \geq 0$

Schemes : implicit $\dfrac{\bar{u}^{n+1} - \bar{u}^n}{\tau} + A\bar{u}^{n+1} = \bar{f}(t_{n+1})$

explicit $\dfrac{\bar{u}^{n+1} - \bar{u}^n}{\tau} + A\bar{u}^n = \bar{f}(t_n)$

$$\bar{u}^n = \begin{pmatrix} u_1^n \\ \vdots \\ u_N^n \end{pmatrix} \qquad u_i^n \simeq u(x_i, t_n) \qquad A = \frac{1}{h^2}\begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} \qquad \bar{f}(t_n) = \begin{pmatrix} f(x_1, t_n) \\ \vdots \\ f(x_N, t_n) \end{pmatrix}$$

Implicit : $\|\bar{u}^{n+1}\|_\infty \leq \|\bar{u}^n\|_\infty + \tau\|\vec{f}(t_{n+1})\|_\infty$

Explicit : $\|f(t_n)\|_\infty$ if $\tau \leq \frac{h^2}{2}$

Let k be s.t. $\|\vec{u}^{n+1}\|_\infty = |u_k^{n+1}|$

$$\left(1+\frac{2\tau}{h^2}\right)|u_k^{n+1}| = \frac{\tau}{h^2}\left(u_{k-1}^{n+1} + u_{k+1}^{n+1}\right) + |u_k^n| + \tau|f(x_k,t_{n+1})|$$

$$\left(1+\frac{2\tau}{h^2}\right)|u_k^{n+1}| \leq \frac{\tau}{h^2}\left(|u_{k-1}^{n+1}| + |u_{k+1}^{n+1}|\right) + |u_k^n| + \tau|f(x_k,t_{n+1})|$$

$$\left(1+\frac{2\tau}{h^2}\right)|u_k^{n+1}| \leq \frac{2\tau}{h^2}|u_k^{n+1}| + |u_k^n| + \tau|f(x_k,t_{n+1})|$$

$$|u_k^{n+1}| \leq |u_k^n| + \tau|f(x_k,t_{n+1})|$$

explicit $1 - \frac{2\tau}{h^2} \geq 0$

$\|\cdot\|_2$ norm

Implicit: $\|\vec{u}^{n+1}\|_2 \leq \|\vec{u}^n\|_2 + \tau\|f(t_{n+1})\|_2$

Explicit: $\|f(t_n)\|_2$ if $\tau \leq \frac{h^2}{2}$

$$\vec{u}^{n+1} = (I+\tau A)^{-1}\left(\vec{u}^n + \tau\vec{f}(t_{n+1})\right)$$

$$\|\vec{u}^{n+1}\|_2 = \|\quad\|_2$$

$(I+\tau A)^{-1}$ eigenvalues are $\frac{1}{1+\tau\lambda_i} \leq 1$

$$\|\vec{u}^{n+1}\|_2 \leq \|\vec{u}^n\|_2 + \tau\|\vec{f}(t_{n+1})\|_2$$

$\|B\vec{v}\|_2 \leq \max_{1\leq i\leq N}|\mu_i|\,\|\vec{v}\|_2 \quad \forall \vec{v}\in\mathbb{R}^N$

$\vec{E}^{n+1} = \vec{U}^{n+1} - \vec{u}^{n+1}$

$$\frac{\vec{E}^{n+1} - \vec{E}^n}{\tau} + A\vec{E}^{n+1} = \vec{r}^{n+1}$$

$$\vec{U}^{n+1} = \begin{pmatrix} u(x_1,t_{n+1}) \\ \vdots \\ u(x_N,t_{n+1}) \end{pmatrix}$$

$|r_i^{n+1}| \leq C(h^2+\tau)$

$\|\vec{E}^{n+1}\|_\infty \leq \|E^n\|_\infty + \tau\|\vec{r}^{n+1}\|_\infty \quad \forall n = 0,\dots,M-1$