

Hw2-A

Abhay Kumar CS/MTH-714 HW2

A.

(a) w_1, w_2, \dots, w_n are orthogonal

$$\langle w_i, w_j \rangle = 0 \quad \text{whenever } i \neq j$$

$$v \in \text{span}\{w_1, \dots, w_n\}$$

$$\Rightarrow v = c_1 w_1 + c_2 w_2 + \dots + c_n w_n \quad \text{where } c_1, \dots, c_n \text{ are constants.}$$

$$v^T w_1 = c_1 w_1^T w_1 = c_1 \|w_1\|^2$$

$$\Rightarrow c_1 = \frac{v^T w_1}{\|w_1\|^2}$$

$$\text{similarly, for general } j \quad v^T w_j = c_j \|w_j\|^2 \Rightarrow c_j = \frac{\langle v, w_j \rangle}{\|w_j\|^2}$$

$$v = \sum_{j=1}^n c_j w_j = \sum_{j=1}^n \frac{\langle v, w_j \rangle}{\|w_j\|^2} w_j$$

(b)

$$p_0 = r_0$$

$$p_n = r_n - \sum_{j=0}^{n-1} \frac{\langle r_n, p_j \rangle_A}{\|p_j\|_A^2} \quad \text{for } 1 \leq n \leq n^*-1$$

$n^* \leq N$: # of iterations to convergence

$$A \in \mathbb{R}^{N \times N}$$

(i)

If there is gradient along only n^* directions, & other dimensions don't have any function update

along these directions -

$$n^* < N \quad n^* = \# \text{ of iterations to converge}$$

for $n = n^* + 1$

$$P_n = r_n - \sum_{j=0}^{n-1} \frac{\langle r_n, p_j \rangle_A}{\|p_j\|_A^2} = 0 \quad \begin{array}{l} \text{(No update if} \\ r_n = 0 \\ \text{i.e. no residue,)} \end{array}$$

$$(ii) \quad \langle P_n, p_j \rangle_A = 0 \quad \text{for } 0 \leq j < n \leq n^* - 1$$

$$\begin{array}{l} n=L: \\ \Rightarrow 0 \leq j < L \end{array} \quad \langle p_L, p_j \rangle_A = \langle p_1, p_0 \rangle_A = 0$$

as p_1 & p_0 are A -conjugate.

Assume it true for $n = \tilde{n} < n \leq n^* - 1$

$$\forall 0 \leq j < \tilde{n}: \quad \langle p_{\tilde{n}}, p_j \rangle_A = 0$$

$$\text{i.e. } \langle p_{\tilde{n}}, p_0 \rangle_A = 0$$

$$\langle p_{\tilde{n}}, p_1 \rangle_A = 0$$

\vdots

$$\langle p_{\tilde{n}}, p_{\tilde{n}-1} \rangle_A = 0$$

$$\text{for } n = \tilde{n} + 1; \quad \boxed{0 \leq j < \tilde{n} + 1}$$

$$\langle p_n, p_j \rangle_A = \langle p_{\tilde{n}+1}, p_j \rangle_A$$

$\underbrace{\hspace{10em}}$

$$p_{\tilde{n}+1} = \left(r_{\tilde{n}+1} - \sum_{k=0}^{\tilde{n}} \frac{\langle r_{\tilde{n}+1}, p_k \rangle_A}{\|p_k\|_A^2} p_k \right)$$

$$\langle p_{\tilde{n}+1}, p_j \rangle_A = \langle \left(r_{\tilde{n}+1} - \sum_{k=0}^{\tilde{n}} \frac{\langle r_{\tilde{n}+1}, p_k \rangle_A}{\|p_k\|_A^2} p_k \right), p_j \rangle_A$$

$$= \langle r_{\tilde{n}+1}, p_j \rangle_A - \left\langle \sum_{k=0}^{\tilde{n}} \frac{\langle r_{\tilde{n}+1}, p_k \rangle_A}{\|p_k\|_A^2} p_k, p_j \right\rangle$$

$$= \langle r_{\tilde{n}+1}, p_j \rangle_A - 0$$

$$= 0 \quad (\text{as } r_{\tilde{n}+1} \text{ is orthogonal to previous update directions})$$

$$\forall 0 \leq j < \tilde{n}+1$$

proved

(C) $A \in \mathbb{R}^{N \times N}$ sym. p.d.

\mathbb{R}^N : orthonormal basis :- $\phi_1, \phi_2, \dots, \phi_N$
of eigenvectors.

$$A\phi_n = \lambda_n \phi_n \quad \& \quad \langle \phi_n, \phi_j \rangle = \delta_{nj}$$

$$\text{given: } \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$$

$$(i) \quad \langle Av, \omega \rangle \quad \text{we can write } v = \sum_{i=1}^N c_i \phi_i \quad \&$$

$$\omega = \sum_{i=1}^N d_i \phi_i$$

$$\text{we can get } c_i = \langle v, \phi_i \rangle, \quad d_i = \langle \omega, \phi_i \rangle$$

$$\begin{aligned}
\left\langle A \left(\sum_{i=1}^N c_i \phi_i \right), \sum_{j=1}^N d_j \phi_j \right\rangle &= \left\langle \sum_{i=1}^N \langle v, \phi_i \rangle \lambda_i \phi_i, \sum_{j=1}^N \langle \omega, \phi_j \rangle \phi_j \right\rangle \\
&= \sum_{i=1}^N \lambda_i \langle v, \phi_i \rangle \langle \omega, \phi_i \rangle \quad \left[\begin{array}{l} \text{using} \\ \langle \phi_i, \phi_j \rangle = \delta_{ij} \end{array} \right] \\
&= \sum_{n=1}^N \lambda_n \langle v, \phi_n \rangle \langle \phi_n, \omega \rangle
\end{aligned}$$

(ii) TP: $\lambda_n > 0 \quad \forall 1 \leq n \leq N$.

Since A is pos. defⁿ matrix

$$v^T A v > 0 \quad \forall v \in \mathbb{R}^n - \{ \vec{0} \}$$

using part (i),

$$v^T A v = \sum_{n=1}^N \lambda_n \underbrace{\langle v, \phi_n \rangle}_{=c_n} \underbrace{\langle \phi_n, v \rangle}_{=c_n}$$

$$= \sum_{n=1}^N \lambda_n c_n^2 > 0 \quad \forall v \in \mathbb{R}^n \text{ \& } v \neq \vec{0}$$

we can always select v s.t. $c_j \neq 0$ & $c_n = 0 \quad \forall n \neq j$

$$\Rightarrow \lambda_j > 0 \quad \forall j = [1, n]$$

$$\Rightarrow \boxed{\lambda_n > 0} \quad \text{for } 1 \leq n \leq N.$$

$$\begin{aligned}
\text{(iii)} \quad \langle A v, v \rangle &= \sum_{n=1}^N \lambda_n \langle v, \phi_n \rangle \langle \phi_n, v \rangle \\
&= \sum_{n=1}^N \lambda_n v^T \phi_n \phi_n^T v = \sum_{n=1}^N \lambda_n \|v\|^2
\end{aligned}$$

$$\lambda_1 \leq \lambda_n \leq \lambda_N \quad \forall 1 \leq n \leq N$$

$$\Rightarrow \lambda_1 \|v\|^2 \leq \lambda_n \|v\|^2 \leq \lambda_N \|v\|^2$$

$$\textcircled{8} \quad \lambda_1 \|v\|^2 \leq \langle Av, v \rangle \leq \lambda_N \|v\|^2$$

$$(iv) \text{ TP: } \|Av\| \leq \lambda_N \|v\|$$

$$v = \sum_{n=1}^N \langle v, \phi_n \rangle \phi_n$$

$$\begin{aligned} Av &= \sum_{n=1}^N \langle v, \phi_n \rangle A \phi_n \\ &= \sum_{n=1}^N \lambda_n \langle v, \phi_n \rangle \phi_n \end{aligned}$$

$$\begin{aligned} \|Av\|^2 &= (Av)^T (Av) = \left(\sum_{n=1}^N \lambda_n \langle v, \phi_n \rangle \phi_n \right)^T \left(\sum_{n=1}^N \lambda_n \langle v, \phi_n \rangle \phi_n \right) \\ &= \sum \lambda_n^2 (\langle v, \phi_n \rangle)^2 \\ &= \sum \lambda_n^2 (\underbrace{v^T \phi_n}_{})^2 \leq \lambda_N^2 \sum_{n=1}^N (\underbrace{v^T \phi_n}_{})^2 \\ &\leq \lambda_N^2 \|v\|^2 \quad \downarrow \text{ using } \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N \\ &\quad \underbrace{\sum_{n=1}^N (v^T \phi_n)^2}_{= \|v\|^2} \end{aligned}$$

$$\Rightarrow \|Av\| \leq \lambda_N \|v\|$$

$$(d) \quad p_{n+1} = (1 + \beta_n) p_n - \alpha_n A p_n - \beta_{n-1} p_{n-1} \quad 1 \leq n \leq n^* - 2$$

$$p_{n+1} = r_{n+1} - \sum_{j=0}^n \frac{\langle r_{n+1}, p_j \rangle_A}{\|p_j\|_A^2} p_j$$

$$w_n = A p_n$$

$$\alpha_n = \frac{r_n^T r_n}{p_n^T w_n} = \frac{r_n^T r_n}{p_n^T A p_n}$$

$$\beta = \frac{r_{n+1}^T r_{n+1}}{r_n^T r_n}$$

$$x_{n+1} = x_n + \alpha_n p_n$$

$$\begin{aligned} r_{n+1} &= r_n - \alpha_n w_n \\ &= r_n - \alpha_n A p_n \end{aligned}$$

$$r_n = \frac{r_{n+1} - r_{n-1}}{r_n + r_{n-1}}$$

$$p_n = r_n + \beta_{n-1} p_{n-1}$$

~~$$\begin{aligned}
 p_{n+1} &= r_n - \alpha_n A p_n - \sum_{j=0}^n \frac{\langle (r_n - \alpha_n A p_n), p_j \rangle_A}{\|p_j\|_A^2} p_j \\
 &= -\alpha_n A p_n + r_n - \underbrace{\sum_{j=0}^{n-1} \frac{\langle r_n, p_j \rangle_A}{\|p_j\|_A^2} p_j}_{= p_n} - \frac{\langle r_n, p_n \rangle_A}{\|p_n\|_A^2} p_n + \sum_{j=0}^n \frac{\langle \alpha_n A p_n, p_j \rangle_A}{\|p_j\|_A^2} p_j \\
 &= p_n - \alpha_n A p_n + \frac{\langle r_n, p_n \rangle_A}{\|p_n\|_A^2} p_n + \frac{\alpha_n \langle A p_n, p_n \rangle_A}{\|p_n\|_A^2} p_n
 \end{aligned}$$~~

True for $1 \leq n \leq n^*-2$; as the below relation depends on p_{n+1} .

$$p_{n+1} = r_{n+1} + \beta_n p_n$$

$$= \underbrace{r_n - \alpha_n A p_n}_{p_n} + \beta_n p_n$$

$$= p_n - \beta_{n-1} p_{n-1} - \alpha_n A p_n + \beta_n p_n$$

$$= (1 + \beta_n) p_n - \alpha_n A p_n - \beta_{n-1} p_{n-1}$$

$$\Rightarrow \boxed{p_{n+1} = (1 + \beta_n) p_n - \alpha_n A p_n - \beta_{n-1} p_{n-1}} ; \text{ for } 1 \leq n \leq n^*-2$$

(c) $A \in \mathbb{R}^{N \times N}$ non-singular \Rightarrow invertible

$$A^N = \text{lin. comb}^n \{I, A, A^2, \dots, A^{n-1}\}$$

Cayley-Hamilton Theorem :-

$$\begin{aligned}
 & A \in \mathbb{R}^{n \times n}, \quad I_n : \text{Identity matrix } n \times n \text{ size} \\
 & \text{Characteristic polynomial of } A \text{ is defined as} \\
 & \quad P(\lambda) = \det(\lambda I_n - A) \\
 & P(\lambda) = \lambda^n + C_{n-1} \lambda^{n-1} + \dots + C_1 \lambda + C_0 \\
 & \text{We can create analogous polynomial } P(A) \text{ in the matrix} \\
 & \quad A \text{ instead of the scalar variable } \lambda, \\
 & P(A) = A^n + C_{n-1} A^{n-1} + \dots + C_1 A + C_0 I_n.
 \end{aligned}$$

$$\begin{aligned}
 & p(t) = \det(A - tI) \\
 & \Rightarrow P(A) = 0 \\
 & P(t) = t^N + C_1 t^{N-1} + \dots + C_N
 \end{aligned}
 \quad
 \begin{bmatrix}
 a_{11}-t & a_{12} & \dots & a_{1n} \\
 a_{21} & a_{22}-t & & \vdots \\
 \vdots & & \ddots & \\
 a_{n1} & \dots & & a_{nn}-t
 \end{bmatrix}$$

$$\begin{aligned}
 P(A) &= A^N + C_1 A^{N-1} + \dots + C_N I = 0 \\
 \Rightarrow A^N &= -[C_N I + C_{N-1} A + \dots + C_1 A^{N-1}] \\
 & \text{(where } C_1, \dots, C_N \text{ are constant scalars.)}
 \end{aligned}$$

$$= \text{lin comb}^n \text{ of } \{I, A, A^2, A^3, \dots, A^{N-1}\}$$

$$\text{and } \lambda \neq 0 \quad \Rightarrow \quad \dots$$

(+)

$$Au = f$$

$$\Leftrightarrow u = u + \alpha(f - Au)$$

Richardson iteration:-

$$u_{n+1} = u_n + \alpha(f - Au_n)$$

(i) error, $e_n = u_n - u$

$$\begin{aligned} e_{n+1} &= u_{n+1} - u \\ &= u_n + \alpha(f - Au_n) - u \\ &= e_n + \alpha(f - Au_n) \\ &= e_n + \alpha(Au - Au_n) \\ &= e_n - \alpha A e_n \\ &= (I - \alpha A) e_n. \end{aligned}$$

(ii) $\|e_{n+1}\| \leq \rho \|e_n\|$

from (i) $e_{n+1} = (I - \alpha A) e_n$

$$\|e_{n+1}\| = \|(I - \alpha A) e_n\| \quad \text{--- } \textcircled{1}$$

using result in part (c) $\|Av\| \leq \lambda_{\max} \|v\|$

$\textcircled{1}$ becomes

$$\|e_{n+1}\| \leq |\lambda_{\max}(I - \alpha A)| \|e_n\|$$

$$\lambda_{\max}(I - \alpha A) = \left| \max_{1 \leq j \leq N} (1 - \alpha \lambda_j) \right| = \rho$$

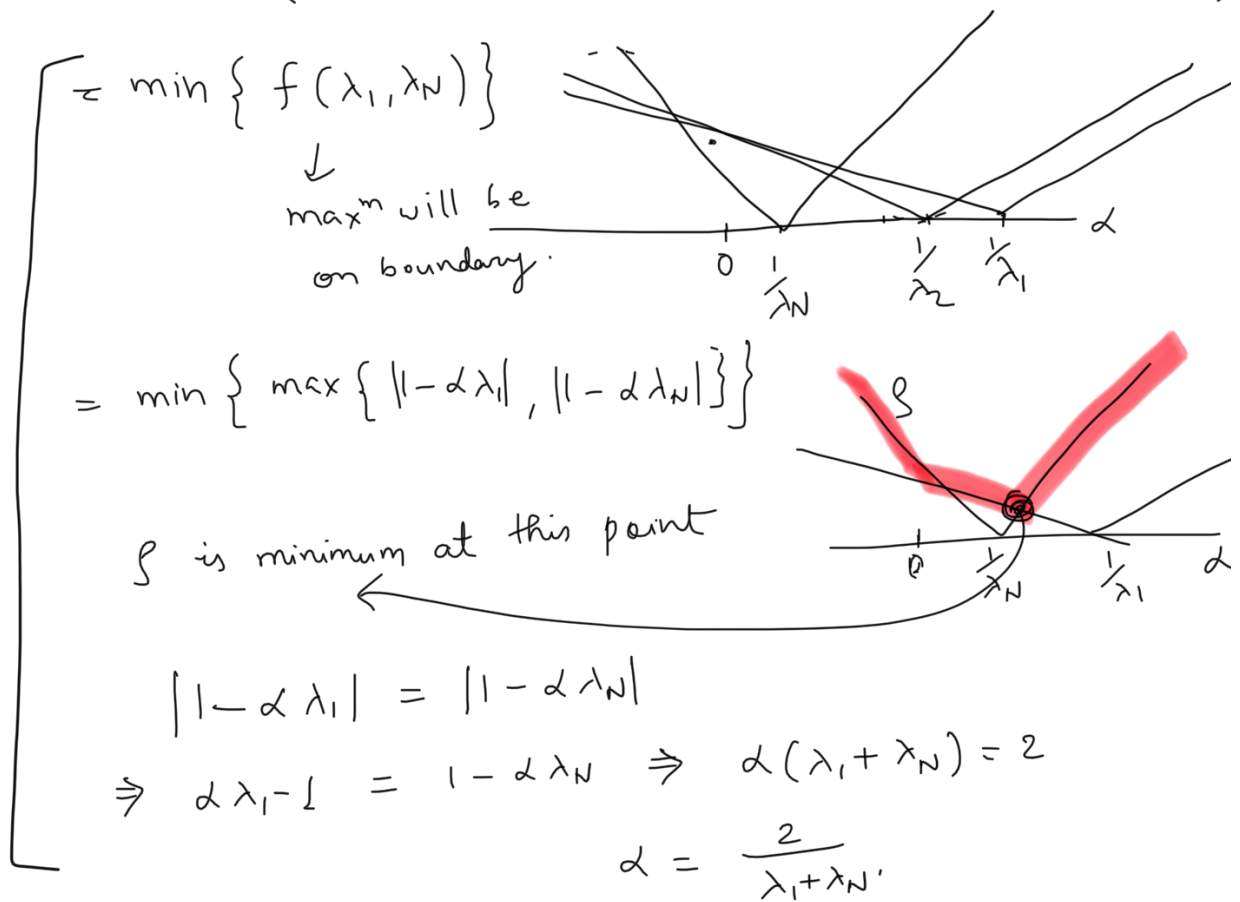
$$\|e_{n+1}\| \leq \rho \|e_n\|, \text{ where } \rho = \max_{1 \leq j \leq N} |1 - \alpha \lambda_j|$$

$$\Rightarrow \|e_n\| \leq \rho^n \|e_0\|$$

(iii) TP: $\rho = \max_{1 \leq j \leq N} |1 - \alpha \lambda_j|$ is minimized by choosing

$$\alpha = \frac{2}{\lambda_1 + \lambda_N}$$

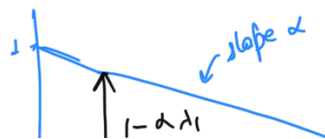
$$\min \rho = \min \left(\max \left\{ |1 - \alpha \lambda_1|, |1 - \alpha \lambda_2|, \dots, |1 - \alpha \lambda_N| \right\} \right)$$



Alternatively,

$$\min \rho = \min \left(\max \left\{ |1 - \alpha \lambda_1|, |1 - \alpha \lambda_2|, \dots, |1 - \alpha \lambda_N| \right\} \right)$$

It should be such that
it balances out on



both extremes. to get
the overall minimum.



$$\Rightarrow 1 - \alpha \lambda_1 = \alpha \lambda_N^{-1}$$

$$\Rightarrow \alpha(\lambda_1 + \lambda_N) = 2 \quad \Rightarrow \quad \boxed{\alpha = \frac{2}{\lambda_1 + \lambda_N}}$$

$$\rho \Big|_{\alpha = \frac{2}{\lambda_1 + \lambda_N}} = \left| 1 - \frac{2 \lambda_N}{\lambda_1 + \lambda_N} \right| = \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1}$$

$$= \frac{k-1}{k+1} < 1 ; \quad k = \frac{\lambda_N}{\lambda_1}$$

(iv) $0 < c \leq \lambda_1 \leq \lambda_N \leq C < \infty$

$$\alpha = \frac{2}{c + C}$$

$$\rho = \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} \leq \frac{(\lambda_N - \lambda_1) + \overset{\nearrow \geq 0}{k}}{(\lambda_N + \lambda_1) + \underbrace{m}_{\leq k}} \leq$$

$$\leq \frac{(\lambda_N - \lambda_1) + (E_N + E_1)}{(\lambda_N + \lambda_1) + (E_N - E_1)}$$

$$\leq \frac{(\lambda_N + E_N) - (\lambda_1 - E_1)}{(\lambda_N + E_N) + (\lambda_1 - E_1)}$$

$$\leq \frac{C - c}{C + c} = \frac{K' - 1}{K' + 1} < 1$$

where $K' = \frac{C}{c}$

(9)

$$q_n = \frac{x_n}{\|x_n\|}$$

$$0 \leq n \leq n^* - 1$$

$$\{q_0, \dots, q_{n-1}\} \text{ orthonormal basis for the Krylov space } \mathcal{K}_n$$

$$\omega_\eta = A p_\eta$$

$$\alpha_n = \frac{x_n^T x_n}{p_n^T \omega_n} = \frac{x_n^T x_n}{p_n^T A p_n}$$

$$x_{n+1} = x_n + \alpha_n p_n$$

$$\begin{aligned} r_{n+1} &= r_n - d_n \omega_n. \\ &= r_n - d_n A p_n \end{aligned}$$

$$\beta_n = \frac{\tau_{n+1}^t \tau_{n+1}}{\tau_n^t \tau_n}$$

$$p_n = \alpha_n + \beta_{n-1} p_{n-1}$$

(i) TS :- $\gamma_1 = \gamma_0 - \alpha_0 A \gamma_0$

$$r_1 = r_0 - \alpha_0 A p_0 \quad ; \quad p_0 = r_0$$

$$= r_0 - \alpha_0 A r_0$$

$$(ii) \quad TS: \quad x_{n+1} = x_n - \alpha_n A x_n + \frac{\alpha_n \beta_{n-1}}{\alpha_{n-1}} (x_n - x_{n-1})$$

$$\begin{aligned} x_{n+1} &= x_n - \alpha_n A p_n \\ &= x_n - \alpha_n A (x_n + \beta_{n-1} p_{n-1}) \\ &= x_n - \alpha_n A x_n + \underbrace{\alpha_n \beta_{n-1} A p_{n-1}}_{= \frac{x_n - x_{n-1}}{\alpha_{n-1}}} \\ &= x_n - \alpha_n A x_n + \frac{\alpha_n \beta_{n-1}}{\alpha_{n-1}} (x_n - x_{n-1}) \end{aligned}$$

$$(iii) \quad \gamma_0 = \frac{1}{\alpha_0} \quad ; \quad \gamma_n = \frac{1}{\alpha_n} + \frac{\beta_{n-1}}{\alpha_{n-1}} \quad (1 \leq n \leq n^*-1)$$

$$\sigma_n \Rightarrow \frac{\sqrt{\beta_n}}{\alpha_n}$$

$$\begin{aligned} A q_0 &= A \frac{r_0}{\|r_0\|} = \frac{1}{\|r_0\|} \left(\frac{r_0 - r_1}{\alpha_0} \right) = \left(\frac{1}{\alpha_0} \right) \frac{r_0}{\|r_0\|} - \left(\frac{1}{\alpha_0} \right) \frac{r_1}{\|r_0\|} \\ &= \gamma_0 q_0 - \frac{1}{\alpha_0} \frac{r_1}{\|r_1\|} \frac{\|r_1\|}{\|r_0\|} = \gamma_0 q_0 - \frac{\sqrt{\beta_0}}{\alpha_0} q_1 \end{aligned}$$

$$Aq_0 = \gamma_0 q_0 - \delta_0 q_1 \quad \left(\beta_b = \frac{||\eta||^2}{||\sigma_b||^2} \right)$$

$$\text{TS: } A q_n = -\delta_{n-1} q_{n-1} + \gamma_n q_n - \delta_n q_{n+1} \quad 1 \leq n \leq n^*-1$$

$$Aq_n = A \frac{x_n}{\|x_n\|} = \frac{A(P_n - \beta_{n-1} P_{n-1})}{\|x_n\|} = \frac{A P_n - \beta_{n-1} A P_{n-1}}{\|x_n\|}$$

$$= \frac{1}{\|\gamma_n\|} \left[\frac{\gamma_n - \gamma_{n+1}}{\alpha_n} - \frac{\beta_{n-1} (\gamma_{n-1} - \gamma_n)}{\alpha_{n-1}} \right]$$

$$= \underbrace{\frac{\gamma_n}{\|\gamma_n\|}}_{q_n} \left(\underbrace{\frac{1}{\alpha_n} + \frac{\beta_{n-1}}{\alpha_{n-1}}}_{=\gamma_n} \right) - \frac{\gamma_{n+1}}{\|\gamma_n\| \alpha_n} - \frac{\beta_{n-1} \gamma_{n-1}}{\|\gamma_n\| \alpha_{n-1}}$$

$$= \gamma_n q_n - \frac{\gamma_{n+1}}{\|\gamma_{n+1}\|} \frac{\|\gamma_{n+1}\|}{\|\gamma_n\|} \frac{1}{\alpha_n} - \beta_{n-1} \frac{\gamma_{n-1}}{\|\gamma_{n-1}\|} \frac{\|\gamma_{n-1}\|}{\|\gamma_n\|} \frac{1}{\alpha_{n-1}}$$

$$\left(\text{using } \beta_n = \frac{\tau_{n+1}^t \tau_{n+1}}{\tau_n^t \tau_n} = \frac{\|\gamma_{n+1}\|^2}{\|\gamma_n\|^2} \right)$$

$$x_1, x_2, \dots, x_n \in \mathbb{R}^n, \quad 1 \leq n \leq n^* - 1$$

$$[Aq_n = \gamma_n v_n - \delta_n v_{n+1} - \delta_{n-1} v_{n-1}] \text{ for } 1 = \dots$$

$$(iv) \quad \begin{matrix} A & Q_n & = & Q_n T_n & - & \delta_{n-1} e_n^T \\ N \times N & N \times n & & N \times n & n \times n & N \times 1 & 1 \times n \end{matrix} \quad q_n \in \mathbb{R}^N$$

$$Q_n = [q_0 \ q_1 \ \dots \ q_{n-1}]$$

$$\text{LHS: } AQ_n = [Aq_0, Aq_1, \dots, Aq_{n-1}]$$

$$= \left[\gamma_0 q_0 - \delta_0 q_1, \left\{ -\delta_{n-1} q_{n-1} + \gamma_n q_n - \delta_n q_{n+1} \right\}_{n=1}^{n-1} \right]$$

$$= [q_0, q_1, \dots, q_{n-1}] \begin{bmatrix} \gamma_0 & -\delta_0 & & & \\ -\delta_0 & \gamma_1 & -\delta_1 & & \\ & -\delta_1 & \gamma_2 & -\delta_2 & \\ & & -\delta_2 & \ddots & \\ & & & -\gamma_{n-2} & -\delta_{n-2} \\ & & & -\delta_{n-2} & \gamma_{n-1} \end{bmatrix} - [0, 0, \dots, \delta_{n-1} q_n]$$

$= [0 \ \dots \ 1] \delta_{n-1}$
 $= e_n^T$
 $e_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$

$$(Aq_{n-1} = -\delta_{n-2} q_{n-2} + \gamma_{n-1} q_{n-1} - \delta_{n-1} q_n)$$

$$= Q_n T_n - \delta_{n-1} q_n e_n^T$$

$$\Rightarrow AQ_n = Q_n T_n - \delta_{n-1} q_n e_n^T$$

$$(v) \quad Q_n^T A Q_n = T_n$$

$$\dots \dots \dots \sigma^T$$

$$A Q_n = Q_n I_n - \alpha_{n-1} e_n$$

$$\underbrace{Q_n^T}_{n \times n} \underbrace{A}_{n \times N} \underbrace{Q_n}_{N \times N} = \underbrace{Q_n^T}_{n \times N} \underbrace{Q_n}_{N \times N} \underbrace{T_n}_{N \times n} - \underbrace{Q_n^T}_{n \times N} \underbrace{\alpha_{n-1}}_{1 \times 1} \underbrace{q_n}_{N \times 1} \underbrace{e_n^T}_{1 \times n}$$

$$= T_n \quad \left[\begin{array}{l} \text{using } Q_n^T Q_n = I \\ \& Q_n^T q_n = 0 \\ \text{as } [q_0 \dots q_{n-1}]^T q_n = \vec{0} \\ \text{orthogonal} \end{array} \right]$$

CS/MTH 714: Homework 2

Abhay Kumar

CODE at https://github.com/abhayk1201/CS714_Hw2/

B

CODE at https://github.com/abhayk1201/CS714_Hw2/hw2_b_interpolant.m

N = 100

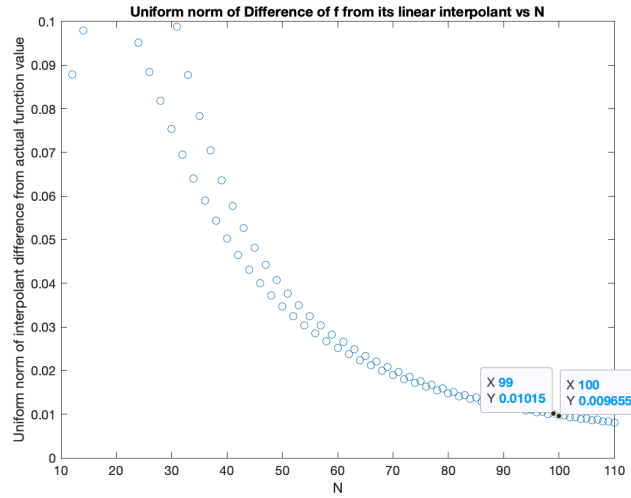


Figure 1: Uniform norm of interpolant difference from actual function value.

Given $f(x) = e^{-400(x-0.5)^2}$

$$\begin{aligned} f'(x) &= -800 e^{-400(x-0.5)^2} (x - 0.5) \\ f''(x) &= -800 e^{-400(x-0.5)^2} [-800(x - 0.5)^2 + 1] \end{aligned} \quad (1)$$

Using Taylor expansion, we have

$$\begin{aligned} f(x+h) &= f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \text{higher order terms} \\ f_{\text{linear interpolant}} &= f(x) + hf'(x) \\ \text{Interpolation Error} &= \frac{h^2}{2} f''(x) + \text{higher order terms} \end{aligned} \quad (2)$$

For $N=100$, $h=0.01$ and numerically, we get Interpolation Error at mid-point between two interpolant points (i.e at $x + h/2$) $= \frac{0.01^2}{2} \times 800 = \frac{(h/2)^2}{2} f''(x) \leq 0.01$ (Max value of $|f''(x)|$ is $\approx 796 < 800$ for $x \in [0, 1]$. (refer Fig-3)

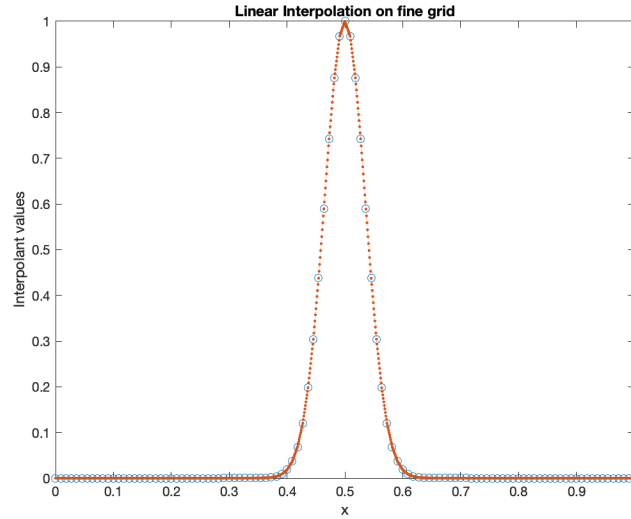


Figure 2: Linear Interpolation on fine grid

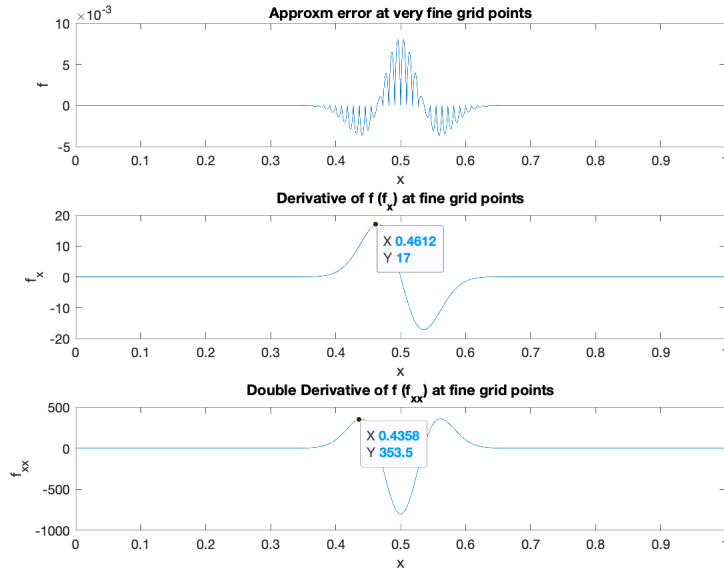


Figure 3: a) Approxm error at very fine grid points; b) Derivative of f (f_x) at fine grid points; c) Double Derivative of f (f_{xx}) at fine grid points

C

CODE at https://github.com/abhayk1201/CS714_Hw2

C(a)

CODE at https://github.com/abhayk1201/CS714_Hw2/hw2_c_a_2D_wave_eq_log_log_error.m

$$u_t t = \Delta u$$

$$\frac{u_{j_1, j_2}^{n+1} + u_{j_1, j_2}^{n-1} - 2u_{j_1, j_2}^n}{\Delta t^2} = \frac{u_{j_1+1, j_2}^n + u_{j_1-1, j_2}^n + u_{j_1, j_2+1}^n + u_{j_1, j_2-1}^n - 4u_{j_1, j_2}^n}{\Delta x^2} \quad (3)$$

on the rectangular domain $[0,1] \times [0,1]$ with homogeneous Dirichlet boundary conditions. ($u(1, :, :) = 0$; $u(:, 1, :) = 0$; $u(N+1, :, :) = 0$; $u(:, N+1, :) = 0$;) I initialize the scheme with the $u(x, y, 0) = 0$ i.e starting with a zero space wave for time $t=0$; and ensure that the $u_t(x, y, 0) = f(x)f(y)$ by setting $u(:, :, 2) = \Delta t \times (\text{transpose}(\exp(-400 \cdot (x - 0.5)^2)) * (\exp(-400 \cdot (y - 0.5)^2)))$

Space interval $[0,1] \times [0,1]$

Space discretization step $\Delta x = \Delta y = \frac{1}{512}$ (finest) and varied for with multiple of 2 to $\frac{1}{32}$

Also, tried with $\Delta x = \Delta y = \frac{1}{1024}$ (finest), but it takes much more time due to higher computation and bigger matrix size

Time discretization step $\Delta t = \frac{\Delta x}{\sqrt{2}} = \frac{1}{512} * \frac{1}{\sqrt{2}}$

Amount of time $T=1$; Number of time steps $= \frac{T}{\Delta t}$

Initial condition $u(x, y, 0) = 0$ and $u_t(x, y, 0) \exp(-400(x - 0.5)^2) * (\exp(-400 * (y - 0.5)^2))$

And I set Δx and Δt s.t the CFL condition is satisfied i.e $\frac{\Delta t^2}{\Delta x^2} \leq \frac{1}{2}$.

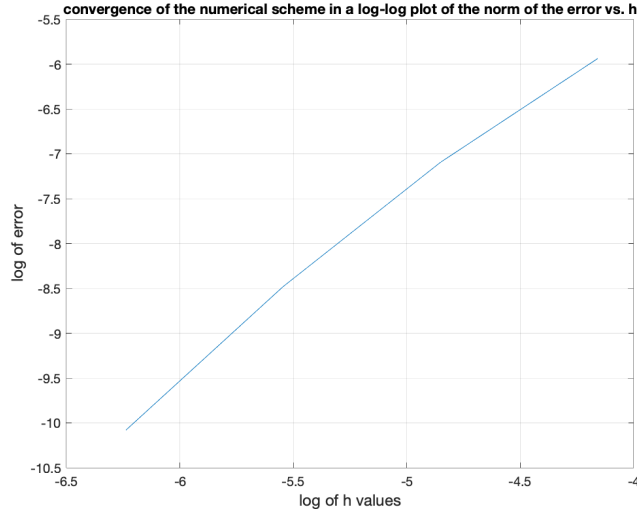


Figure 4: log-log plot of the error vs. the grid spacing, h when finest grid is $\Delta x = \Delta y = \frac{1}{1024}$

Yes, the slope is ≈ 2 and this means that the method is second-order accurate.

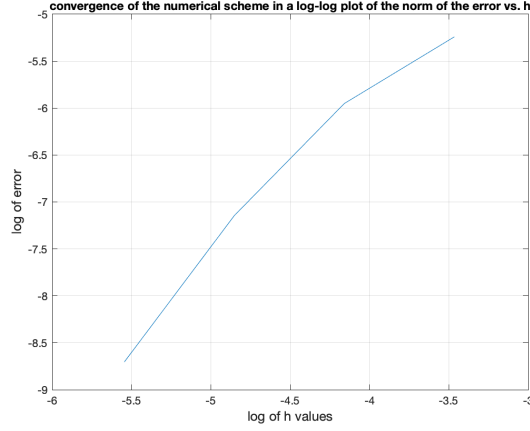


Figure 5: log-log plot of the error vs. the grid spacing, h when finest grid is $\Delta x = \Delta y = \frac{1}{512}$

Yes, the slope is ≈ 2 and this means that the method is second-order accurate.

C(b)

ODE $y''(t) = \lambda y$, and the 3-point rule for y'' as a two-step explicit time integrator. Find the region of stability of this ODE solver in terms of $\lambda(\Delta t)$, and plot it in the complex plane.

$$y''(t) = \lambda y$$

3-point rule for $y'' = \frac{y^{n+1} + y^{n-1} - 2y^n}{\Delta t^2}$;

$$\frac{y^{n+1} + y^{n-1} - 2y^n}{\Delta t^2} = \lambda y^n \quad (4)$$

Let the solution be of the form

$$y_j^n = \rho^n \exp(iKj\Delta y) \quad (5)$$

Also, we have the following relation between ρ^{n+1} and ρ^n ;

$$\begin{aligned} y_j^{n+1} &= \rho^{n+1} \exp(iKj\Delta y) \\ y_j^{n+1} &= \rho^2 y_j^{n-1} \\ y_j^n &= \rho y_j^{n-1} \end{aligned} \quad (6)$$

Using (5) in (4) and taking out y_j^{n-1} as common, we get the following-

$$\begin{aligned} \frac{\rho^2 + 1 - 2\rho}{\Delta t^2} &= \lambda \rho \\ \rho^2 + 1 - 2\rho &= \Delta t^2 \lambda \rho \\ \rho^2 - (2 + \lambda \Delta t^2) \rho + 1 &= 0 \end{aligned} \quad (7)$$

Take $z = \lambda \Delta t^2$, we can simplify (7) as-

$$\rho^2 - (2 + z)\rho + 1 = 0 \quad (8)$$

Roots of above quadratic are-

$$\rho_{1,2} = \frac{(2 + z) \pm \sqrt{(2 + z)^2 - 4}}{2} \quad (9)$$

We need to find out possible values of z such that $|\rho| \leq 1$, i.e.

$$-1 \leq \rho_{1,2} \leq 1$$

$$-1 \leq \left(1 + \frac{z}{2} \pm \sqrt{\left(1 + \frac{z}{2}\right)^2 - 1}\right) \leq 1 \quad (10)$$

I have solved the above equation numerically using matlab, assuming z as complex number of the form $x + iy$

CODE at https://github.com/abhayk1201/CS714_Hw2/hw2_c_b_stability_region_plot.m

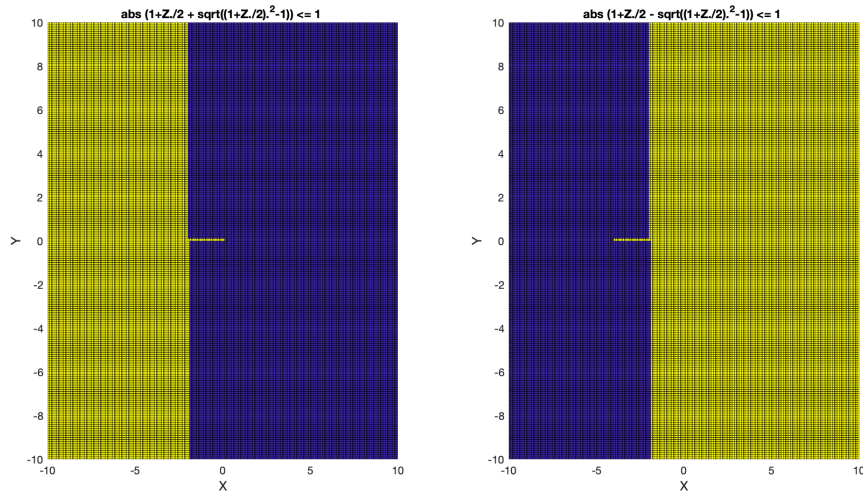


Figure 6: Individual regions for $abs(\rho_1) \leq 1$ or $abs(\rho_2) \leq 1$. Note stability region is the intersection of these two regions. (plotted in next figure)

The stability region will be such that both $abs(\rho_{1,2}) \leq 1$, it will be intersection of above two plots.
i.e. $-4 \leq \lambda \Delta t^2 \leq 0$

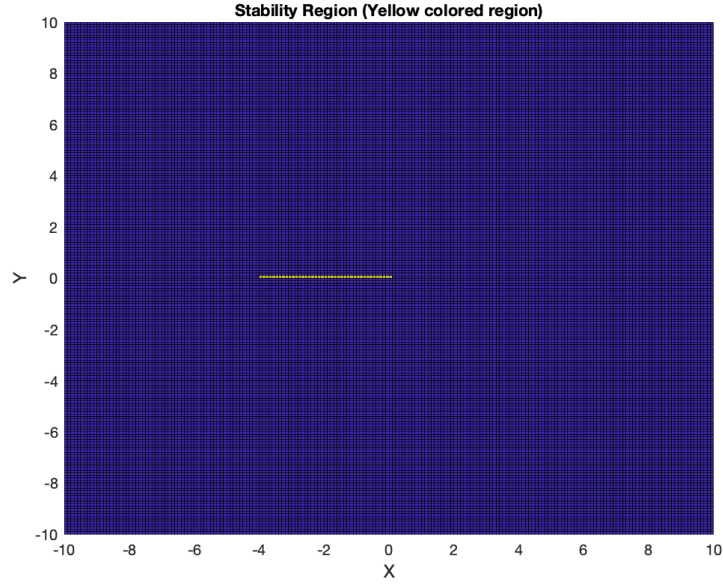


Figure 7: Stability regions (YELLOW color) where both $abs(\rho_{1,2}) \leq 1$.

C(c)

From your answer to (b), and your knowledge of the spectrum of the discrete Laplacian, perform the "method of lines" stability analysis for the method in (a). What CFL condition does this analysis result in?

$$u_{tt} = \Delta u$$

can be written using "method of lines" as following, using Laplacian matrix (\mathbb{L}) in space

$$\frac{U^{n+1} + U^{n-1} - 2U^n}{\Delta t^2} = \mathbb{L}U^n \quad (11)$$

$$\mathbb{L} = A \otimes I_1 + I_2 \otimes A$$

Where,

$$A = \frac{1}{\Delta x^2} \begin{bmatrix} -2 & 1 & \dots & & \\ 1 & -2 & 1 & \dots & \\ & 1 & -2 & 1 & \dots \\ \vdots & & \ddots & \ddots & \ddots \\ & \vdots & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix}$$

The equation (11) is similar to the ODE in part (b) i.e $y''(t) = \lambda y$, here we need to show the same equation as part (b) for all eigen-values of \mathbb{L} .
From the stability region analysis, we get

$$-4 \leq \lambda(\mathbb{L})\Delta t^2 \leq 0 \quad (12)$$

From HW1, we get that $\lambda(\mathbb{L})$ is $(\lambda_i + \lambda_j)$
where

$$\lambda_k = -\frac{4}{\Delta x^2} \sin^2 \left(\frac{k\pi\Delta x}{2} \right)$$

(Similar from HW1: let $\lambda_i \in \sigma(A)$ with corresponding eigenvector x_i , and let $\mu_j \in \sigma(B)$ with corresponding eigenvector y_j . Then $\lambda_i + \mu_j$ is an eigenvalue of $A \otimes B$ with corresponding eigenvector $y_j \otimes x_i$. $(i, j)^{th}$ eigenvalues of $(A \otimes I + I \otimes A)$ is $(\lambda_i + \lambda_j)$ and the corresponding eigenvectors are $x_j \otimes x_i$)

From eq (13), we have: use the fact $0 \leq \sin^2(x) \leq 1$

$$\begin{aligned} -4 &\leq -\frac{4\Delta t^2}{\Delta x^2} \sin^2 \left(\frac{i\pi\Delta x}{2} \right) + -\frac{4\Delta t^2}{\Delta x^2} \sin^2 \left(\frac{j\pi\Delta x}{2} \right) \leq 0 \\ -4 &\leq -\frac{8\Delta t^2}{\Delta x^2} \implies \frac{8\Delta t^2}{\Delta x^2} \leq 4 \implies \frac{\Delta t^2}{\Delta x^2} \leq \frac{1}{2} \end{aligned} \quad (13)$$

We got the following CFL condition-

$$\frac{\Delta t^2}{\Delta x^2} \leq \frac{1}{2} \quad (14)$$

C(d)

Perform the von Neumann stability analysis for the method in (a), and check if the resulting CFL condition agrees with what you found in the previous question.

$$u_{tt} = \Delta u$$

$$\frac{u_{j_1, j_2}^{n+1} + u_{j_1, j_2}^{n-1} - 2u_{j_1, j_2}^n}{\Delta t^2} = \frac{u_{j_1+1, j_2}^n + u_{j_1-1, j_2}^n + u_{j_1, j_2+1}^n + u_{j_1, j_2-1}^n - 4u_{j_1, j_2}^n}{\Delta x^2} \quad (15)$$

Since this is a 2D problem, a plane wave is

$$u_{j_1, j_2}^n = \rho^n \exp(ik_1 j_1 \Delta x) \times \exp(ik_2 j_2 \Delta x) \quad (16)$$

where K_1, K_2 are the wave numbers.

Also, we have the following relation between ρ^{n+1} and ρ^n ;

$$\begin{aligned} \rho^{n+1} &= g(k_1, k_2) \rho^n \\ \rho^n &= g(k_1, k_2) \rho^{n-1} \\ \implies \rho^{n+1} &= g(k_1, k_2)^2 \rho^{n-1} \end{aligned} \quad (17)$$

Substituting (16) in (15); using (17) and taking $\rho^{n-1} \exp(ik_1 j_1 \Delta x) \times \exp(ik_2 j_2 \Delta x)$ as common from both sides, we get:

$$\begin{aligned} \frac{g^2 + 1 - 2g}{\Delta t^2} &= g \times \frac{\exp(-ik_1 \Delta x) + \exp(ik_1 \Delta x) + \exp(-ik_2 \Delta x) + \exp(ik_2 \Delta x) - 4}{\Delta x^2} \\ g^2 + 1 - 2g &= g \times \left(\frac{\Delta t^2}{\Delta x^2} \right) (\exp(-ik_1 \Delta x) + \exp(ik_1 \Delta x) + \exp(-ik_2 \Delta x) + \exp(ik_2 \Delta x) - 4) \end{aligned} \quad (18)$$

simplifying above equation using

$$\begin{aligned} (\exp(-ik_1 \Delta x) + \exp(ik_1 \Delta x) - 2) &= 2 \cos(k_1 \Delta x) - 2 = -4 \sin\left(\frac{k_1 \Delta x}{2}\right)^2 \\ g^2 + 1 - 2g &= g \times \left(\frac{\Delta t^2}{\Delta x^2} \right) \left((-4) \left[\sin\left(\frac{k_1 \Delta x}{2}\right)^2 + \sin\left(\frac{k_2 \Delta x}{2}\right)^2 \right] \right) \end{aligned} \quad (19)$$

Taking $\beta = 1 - 2 \left(\frac{\Delta t^2}{\Delta x^2} \right) \left[\sin\left(\frac{k_1 \Delta x}{2}\right)^2 + \sin\left(\frac{k_2 \Delta x}{2}\right)^2 \right]$, eq:(19) becomes-

$$g^2 - 2\beta g + 1 = 0 \quad (20)$$

Solutions of the above equation for g are-

$$g_{1,2} = \beta \pm \sqrt{(\beta^2 - 1)} \quad (21)$$

Note that if $\beta > 1$ then at least one of absolute value of $g_{1,2}$ is bigger than one. Therefore, we need $\beta < 1$. The solutions will be-

$$g_{1,2} = \beta \pm i\sqrt{(\beta^2 - 1)}$$

$|g|^2 = 1$, i.e., the scheme is conditional stable. The stability condition becomes $-1 \leq \beta \leq 1$

$$-1 \leq 1 - 2 \left(\frac{\Delta t^2}{\Delta x^2} \right) \left[\sin\left(\frac{k_1 \Delta x}{2}\right)^2 + \sin\left(\frac{k_2 \Delta x}{2}\right)^2 \right] \leq 1 \quad (22)$$

using $0 \leq \sin\left(\frac{k_1 \Delta x}{2}\right)^2 \leq 1$

$$-1 \leq 1 - 2 \left(\frac{\Delta t^2}{\Delta x^2} \right) \left[\sin\left(\frac{k_1 \Delta x}{2}\right)^2 + \sin\left(\frac{k_2 \Delta x}{2}\right)^2 \right] \leq 1 \quad (23)$$

$$-1 \leq 1 - 4 \left(\frac{\Delta t^2}{\Delta x^2} \right) \leq 1$$

$$-1 \leq 1 - 4 \left(\frac{\Delta t^2}{\Delta x^2} \right)$$

$$4 \left(\frac{\Delta t^2}{\Delta x^2} \right) \leq 2$$

$$\begin{aligned}\left(\frac{\Delta t^2}{\Delta x^2}\right) &\leq \frac{1}{2} \\ \left(\frac{\Delta t}{\Delta x}\right)^2 &\leq \frac{1}{2}\end{aligned}\tag{24}$$

YES, this resulting CFL condition in eq:(24) matches with the CFL condition from part (c) as in eq:(14).

C(e)

Find the modified equation that corresponds to the numerical method in (a). Solve it via Fourier series, and comment on the physics of the extra terms. Are they dissipative, dispersive, or something else?

$$\frac{u_{j_1,j_2}^{n+1} + u_{j_1,j_2}^{n-1} - 2u_{j_1,j_2}^n}{\Delta t^2} = \frac{u_{j_1+1,j_2}^n + u_{j_1-1,j_2}^n + u_{j_1,j_2+1}^n + u_{j_1,j_2-1}^n - 4u_{j_1,j_2}^n}{\Delta x^2}\tag{25}$$

We can use taylor series expansion to get the following terms:

$$\begin{aligned}u(t + \Delta t) &= u(t) + \Delta t u'(t) + \frac{\Delta t^2}{2} u''(t) + \frac{\Delta t^3}{6} u'''(t) + \frac{\Delta t^4}{24} u''''(t) \\ u(t - \Delta t) &= u(t) - \Delta t u'(t) + \frac{\Delta t^2}{2} u''(t) - \frac{\Delta t^3}{6} u'''(t) + \frac{\Delta t^4}{24} u''''(t) \\ u(x + \Delta x) &= u(x) + \Delta x u'(x) + \frac{\Delta x^2}{2} u''(x) + \frac{\Delta x^3}{6} u'''(x) + \frac{\Delta x^4}{24} u''''(x) \\ u(x - \Delta x) &= u(x) - \Delta x u'(x) + \frac{\Delta x^2}{2} u''(x) - \frac{\Delta x^3}{6} u'''(x) + \frac{\Delta x^4}{24} u''''(x)\end{aligned}\tag{26}$$

$$\begin{aligned}u_{tt} &= \Delta u \\ &= u_{xx} + u_{yy} \\ u_{tttt} &= \delta_t \delta_t (u_{xx} + u_{yy}) \\ &= \delta_t \delta_t (u_{xx}) + \delta_t \delta_t (u_{yy}) \\ &= \delta_x \delta_x (u_{tt}) + \delta_y \delta_y (u_{tt}) \\ &= \delta_x \delta_x (u_{xx} + u_{yy}) + \delta_y \delta_y (u_{xx} + u_{yy}) \\ &= u_{xxxx} + \delta_x \delta_x (u_{yy}) + u_{yyyy} + \delta_y \delta_y (u_{xx}) \\ &= u_{xxxx} + u_{yyyy} + 2u_{xxyy}\end{aligned}\tag{27}$$

using (26) and (27) in (25), we get:

$$\begin{aligned}
\text{LHS of (25)} &= u_{tt} + \frac{\Delta t^2}{12} u_{tttt} + \text{higher order terms} \\
\text{RHS of (25)} &= u_{xx} + \frac{\Delta x^2}{12} u_{xxxx} + u_{yy} + \frac{\Delta x^2}{12} u_{yyyy} + \text{higher order terms} \\
u_{tt} + \frac{\Delta t^2}{12} u_{tttt} &= u_{xx} + \frac{\Delta x^2}{12} u_{xxxx} + u_{yy} + \frac{\Delta x^2}{12} u_{yyyy} \\
u_{tt} - \Delta u &= \frac{\Delta x^2}{12} u_{xxxx} + \frac{\Delta x^2}{12} u_{yyyy} - \frac{\Delta t^2}{12} u_{tttt} \\
u_{tt} - \Delta u &= \frac{\Delta x^2}{12} u_{xxxx} + \frac{\Delta x^2}{12} u_{yyyy} - \frac{\Delta t^2}{12} (u_{xxxx} + u_{yyyy} + 2u_{xxyy})
\end{aligned} \tag{28}$$

Fourier transform and inverse fourier transform in 2-d space is defined as -

$$\begin{aligned}
\hat{u}(\xi_1, \xi_2, t) &= \int_{-\infty}^{\infty} u(x, y, t) e^{-i2\pi(\xi_1 x + \xi_2 y)} dx dy \\
u(x, y, t) &= \int_{-\infty}^{\infty} \hat{u}(\xi_1, \xi_2, t) e^{i2\pi(\xi_1 x + \xi_2 y)} d\xi_1 d\xi_2
\end{aligned} \tag{29}$$

(source: <https://engineering.purdue.edu/bouman/ece637/notes/pdf/CSFT.pdf>)

using properties of fourier transform, we get:

$$\begin{aligned}
\hat{u}_{tt}(\xi_1, \xi_2, t) &= -(2\pi)^2 (\xi_1^2 + \xi_2^2) \hat{u}(\xi_1, \xi_2, t) \\
\hat{u}_{tttt}(\xi_1, \xi_2, t) &= (2\pi)^4 (\xi_1^4 + \xi_2^4 + 2\xi_1^2 \xi_2^2) \hat{u}(\xi_1, \xi_2, t)
\end{aligned} \tag{30}$$

$$\begin{aligned}
\hat{u}(\xi_1, \xi_2, t) &\propto C e^{-(2\pi)^2 (\xi_1^2 + \xi_2^2) t} \hat{u}_o(\xi_1, \xi_2, t) \\
u(x, y, t) &\propto \int_{-\infty}^{\infty} \hat{u}(\xi_1, \xi_2, t) e^{i2\pi(\xi_1 x + \xi_2 y)} d\xi_1 d\xi_2 \\
u(x, y, t) &\propto \int_{-\infty}^{\infty} e^{i2\pi(\xi_1 x + \xi_2 y)} e^{-(2\pi)^2 (\xi_1^2 + \xi_2^2) t} \hat{u}_o(\xi_1, \xi_2, t) d\xi_1 d\xi_2 \\
u(x, y, t) &\propto \int_{-\infty}^{\infty} e^{i2\pi((\xi_1 x + \xi_2 y) + i2\pi(\xi_1^2 + \xi_2^2) t)} \hat{u}_o(\xi_1, \xi_2, t) d\xi_1 d\xi_2
\end{aligned} \tag{31}$$

$$\begin{aligned}
\hat{u}_{tttt}(\xi_1, \xi_2, t) &= (2\pi)^4 (\xi_1^4 + \xi_2^4 + 2\xi_1^2 \xi_2^2) \hat{u}(\xi_1, \xi_2, t) \\
\hat{u}(\xi_1, \xi_2, t) &\propto C e^{((2\pi)^4 (\xi_1^4 + \xi_2^4 + 2\xi_1^2 \xi_2^2) t)} \hat{u}_o(\xi_1, \xi_2, t) \\
u(x, y, t) &\propto \int_{-\infty}^{\infty} \hat{u}(\xi_1, \xi_2, t) e^{i2\pi(\xi_1 x + \xi_2 y)} d\xi_1 d\xi_2 \\
u(x, y, t) &\propto \int_{-\infty}^{\infty} e^{i2\pi(\xi_1 x + \xi_2 y)} e^{((2\pi)^4 (\xi_1^4 + \xi_2^4 + 2\xi_1^2 \xi_2^2) t)} \hat{u}_o(\xi_1, \xi_2, t) d\xi_1 d\xi_2 \\
u(x, y, t) &\propto \int_{-\infty}^{\infty} e^{i2\pi((\xi_1 x + \xi_2 y) - i(2\pi)^3 (\xi_1^2 + \xi_2^2)^2 t)} \hat{u}_o(\xi_1, \xi_2, t) d\xi_1 d\xi_2
\end{aligned} \tag{32}$$

From (24), we see that the 2D wave equation admits wave solution of the form $e^{i2\pi((\xi_1 x + \xi_2 y) + (\xi_1^2 + \xi_2^2)t)}$ for any wave numbers ξ_1, ξ_2 and frequency w that satisfy $w^2 = \xi_1^2 + \xi_2^2$ and curves $w = \text{constant}$ for this dispersion relation will be concentric circles in ξ_1, ξ_2 plane. This are dispersive as different modes have different speeds.

D

(Bonus 20 pts) Formulate and prove one extension of the Lax equivalence theorem to the case of linear ODE with two time derivatives.

Original version Lax Equivalence Theorem for linear ODE with one time derivatives

A consistent method of the form

$$u^{n+1} = B(\Delta t)u^n + b^n(\Delta t) \quad (33)$$

is convergent if and only if it weakly stable.

Weak Stability:

$\forall T > 0, \exists C_T > 0$ such that $\|B(\Delta t)^n\|_2 \leq C_T \forall \Delta t > 0$ s.t. $n\Delta t < T$

Now, consider linear ODE with two time derivatives

$$u''(t) = \lambda u$$

Using FD approach, we get

$$\begin{aligned} u^{n+1} + u^{n-1} - 2u^n &= \lambda \Delta t^2 u^n \\ u^{n+1} &= (2 + \lambda \Delta t^2)u^n - u^{n-1} \end{aligned} \quad (34)$$

Suppose, we define

$$v^{n+1} = \begin{bmatrix} u^{n+1} \\ u^n \end{bmatrix}$$

Then we can write the eq: (34) as following:

$$\begin{aligned} v^{n+1} = \begin{bmatrix} u^{n+1} \\ u^n \end{bmatrix} &= \begin{bmatrix} \lambda\Delta t^2 + 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u^n \\ u^{n-1} \end{bmatrix} \\ v^{n+1} &= \begin{bmatrix} \lambda\Delta t^2 + 2 & -1 \\ 1 & 0 \end{bmatrix} v^n \end{aligned} \quad (35)$$

Note that the recurrence form in eq: (35) is similar to that in eq: (33)., where

$$\hat{B}(\Delta t) = \begin{bmatrix} \lambda\Delta t^2 + 2 & -1 \\ 1 & 0 \end{bmatrix}$$

Considering more generic case, we can have the following recurrence relation-

$$\begin{aligned} u^{n+1} &= \hat{B}_{11}(\Delta t)u^n + \hat{B}_{12}(\Delta t)u^{n-1} + \hat{b}^n(\Delta t) \\ v^{n+1} &= \begin{bmatrix} u^{n+1} \\ u^n \end{bmatrix} = \begin{bmatrix} \hat{B}_{11}(\Delta t) & \hat{B}_{12}(\Delta t) \\ I & 0 \end{bmatrix} \begin{bmatrix} u^n \\ u^{n-1} \end{bmatrix} + \hat{b}^n(\Delta t) \\ v^{n+1} &= \begin{bmatrix} \hat{B}_{11}(\Delta t) & \hat{B}_{12}(\Delta t) \\ I & 0 \end{bmatrix} v^n + \hat{b}^n(\Delta t) \\ \text{where, } \hat{B}(\Delta t) &= \begin{bmatrix} \hat{B}_{11}(\Delta t)u^n & \hat{B}_{12}(\Delta t)u^n \\ I & 0 \end{bmatrix} \\ v^{n+1} &= \hat{B}(\Delta t)v^n + \hat{b}^n(\Delta t) \end{aligned} \quad (36)$$

$$(37)$$

If we have

$$V^n = \begin{bmatrix} v(x_1, t_n) \\ v(x_2, t_n) \\ \vdots \\ v(x_m, t_n) \end{bmatrix}$$

then we get the following form as well;

$$\begin{aligned} V^{n+1} &= B(\Delta t)V^n + b^n(\Delta t) \\ V^{n+1} &= BV^n + b^n \quad (\text{dropping } \Delta \text{ for clarity}) \end{aligned} \quad (38)$$

where, $B(\Delta t) = I \otimes \hat{B}(\Delta t)$ i.e $\hat{B}(\Delta t)$ matrix replicated on the diagonal block. And $b^n(\Delta t)$ has $\hat{b}^n(\Delta t)$ duplicated vertically.

If we apply the numerical method to the exact solution $v(x, t)$, we obtain

$$\tilde{v}^{n+1} = B\tilde{v}^n + b^n + (\Delta t)\tau^n \quad (39)$$

where numerical solution is denoted as

$$\tilde{v}^n = \begin{bmatrix} \tilde{v}(x_1, t_n) \\ \tilde{v}(x_2, t_n) \\ \vdots \\ \tilde{v}(x_m, t_n) \end{bmatrix}$$

Subtracting eq:(39) (numerical solution) from eq:(38), we get gives the difference equation for the global error $E^n = V^n - \tilde{v}^n$ as follows-

$$E^{n+1} = BE^n - (\Delta t)\tau^n \quad (40)$$

After N time steps,

$$\begin{aligned} E^N &= B^N E^0 - (\Delta t) \sum_{n=1}^N B^{N-n} \tau^{n-1} \\ \implies \|E^N\| &\leq \|B^N\| \|E^0\| + (\Delta t) \sum_{n=1}^N \|B^{N-n}\| \|\tau^{n-1}\| \end{aligned} \quad (41)$$

Now, re-define weak stability for the case of linear ODE with two time derivatives.
 $\forall T > 0, \exists C_T > 0$ such that $\|B_{11}(\Delta t)^n\|_2 \leq C_T^1 \quad \forall \Delta t > 0$ s.t. $n\Delta t < T$
 $\forall T > 0, \exists C_T > 0$ such that $\|B_{12}(\Delta t)^n\|_2 \leq C_T^2 \quad \forall \Delta t > 0$ s.t. $n\Delta t < T$
 $\implies \forall T > 0, \exists C_T > 0$ such that we can get $\|B(\Delta t)^n\|_2 \leq C_T \quad \forall \Delta t > 0$ s.t. $n\Delta t < T$, where $C_T \leq C_T^1 C_T^2$

$$\begin{aligned} \|E^N\| &\leq C_T \|E^0\| + (\Delta t) \max_{1 \leq n \leq N} \|\tau^{n-1}\| \sum_{n=1}^N \|B^{N-n}\| \\ \implies \|E^N\| &\leq C_T \|E^0\| + (\Delta t) N C_T \max_{1 \leq n \leq N} \|\tau^{n-1}\| \\ &\text{using } N\Delta t \leq T \text{ and consistency i.e } \|\tau\| \rightarrow 0 \\ \implies \|E^N\| &\rightarrow 0 \text{ as } \Delta t \rightarrow 0 \text{ for } N\Delta t \leq T \end{aligned} \quad (42)$$

proved extension of the Lax equivalence theorem.
