Solution of the exercises

Nicola Perin

February 19, 2025

Contents

1	Lec	ture 1	2
	1.1	Proof that $A(u,v)$ and $F(v)$ in the Poisson problem satisfy the	
		hypotheses of the Lax-Milgram lemma	2
		1.1.1 Continuity of $A(u,v)$	3
		1.1.2 Coercivity of $A(u,v)$	3
		1.1.3 Continuity of $F(v)$	3
	1.2	Proof that if u minimizes the energy functional then is solution	
		of the weak problem	4
2	Lec	ture 2	4
	2.1	Proof of the Strong maximum principle for harmonic functions .	4
	2.2	Proof of the corollary (uniqueness for Poisson)	5
3	Lec	ture 3	6
	3.1	Convergence of the truncation error	6
		3.1.1 Second derivative term	6
		3.1.2 First derivative term	6
		3.1.3 Combine the errors	6
	3.2	Construction of a 3-point scheme: method of undetermined coef-	
		ficients	6
4	Lec	ture 4	7
	4.1	Truncation error bound for 2D Poisson b.v.p	7
	4.2	Derivation of FD scheme for the mixed operator $\delta^x_{2h}\delta^y_{2h}u_{ij}$	8
5	Lec	ture 5	8
	5.1	Proof that $ v _1$ is a full norm for the space H_0^1	8
6	Lec	ture 6	9
	6.1	Proof of the continuity of the bilinear form in Lax-Milgram	9
			9
			9
			0
	6.2	Proof of special case of Ceà's lemma	0

7	Lec	ture 11	11
	7.1	Local truncation error (consistency) of the θ -method	11
		7.1.1 The time part R_{time}	12
		7.1.2 The space part R_{space}	13
	7.2	Sufficient condition for stability of the θ -method	14
	7.3	Fourier (von Neumann) Analysis for the $\theta\text{-method}$ in 1D	14
8	Lec	ture 14	17
	8.1	Wave equation equivalent system	17
	8.2	Conservation of energy for the wave equation	18
9	Lec	ture 15	19
	9.1	Consistency proof for the Lax-Wendroff scheme	19
	9.2	Leap frog scheme for wave equation	20
		9.2.1 First-order system formulation	21
		9.2.2 Discretize system by leap-frog method	21
		9.2.3 Derivation of the leap frog scheme for U	22

1 Lecture 1

1.1 Proof that A(u,v) and F(v) in the Poisson problem satisfy the hypotheses of the Lax-Milgram lemma

We need to show that the bilinear form

$$A(u,v) = \int_0^1 u'(x)v'(x) dx.$$

is continuous:

$$\exists \gamma > 0: |A(u,v)| \leq \gamma \|u\|_V \|v\|_V \quad \forall u,v \in V.$$

and coercive:

$$\exists \alpha_0 > 0 : A(u, u) \ge \alpha_0 ||u||_V^2 \quad \forall u \in V.$$

We also need to show that the linear functional

$$F(v) = \int_0^1 f(x)v(x) dx.$$

is continuous:

$$\exists \beta > 0 : |F(v)| \le \beta ||v||_V \quad \forall v \in V.$$

1.1.1 Continuity of A(u, v)

Recall that the norm on $V = H_0^1(0,1)$ is given by

$$||u||_V = \left(\int_0^1 |u'(x)|^2 dx\right)^{1/2}.$$

Now, consider the absolute value of A(u, v):

$$|A(u,v)| = \left| \int_0^1 u'(x)v'(x) \, dx \right|.$$

Using the Cauchy-Schwarz inequality:

$$|A(u,v)| \le \left(\int_0^1 |u'(x)|^2 dx\right)^{1/2} \left(\int_0^1 |v'(x)|^2 dx\right)^{1/2}.$$

But this is exactly $||u||_V ||v||_V$. Therefore, the bilinear form A(u, v) is continuous with $\gamma = 1$:

$$|A(u,v)| \le ||u||_V ||v||_V.$$

1.1.2 Coercivity of A(u, v)

$$A(u, u) = \int_0^1 |u'(x)|^2 dx = ||u||_V^2.$$

If we take $\alpha = 1$ we get:

$$A(u, u) = ||u||_V^2 \ge \alpha ||u||_V^2.$$

Thus A(u, v) is coercive with $\alpha = 1$.

1.1.3 Continuity of F(v)

Using the Cauchy-Schwarz inequality:

$$|F(v)| = \left| \int_0^1 f(x)v(x) \, dx \right| \le \left(\int_0^1 |f(x)|^2 \, dx \right)^{1/2} \left(\int_0^1 |v(x)|^2 \, dx \right)^{1/2}$$

and recognizing the L^2 -norm:

$$||f||_{L^2(0,1)} = \left(\int_0^1 |f(x)|^2 dx\right)^{1/2}$$

we obtain:

$$|F(v)| \le ||f||_{L^2(0,1)} ||v||_{L^2(0,1)}.$$

Now from the Poincaré inequality we know that for $v \in H_0^1(0,1)$:

$$||v||_{L^2(0,1)} \le C_p ||v'||_{L^2(0,1)} = C_p ||v||_V$$

Thus, F(v) is continuous with $\beta = ||f||_{L^2(0,1)}C_p$:

$$|F(v)| \le ||f||_{L^2(0,1)} C_p ||v||_V.$$

1.2 Proof that if u minimizes the energy functional then is solution of the weak problem

Given the potential energy functional

$$J(v) = \frac{1}{2}A(v,v) - F(v)$$

we want to prove that if a function u minimizes such functional, is then solution of the weak problem A(u, v) = F(v). Let's introduce the auxiliary $u_{\epsilon} = u + \epsilon v$ and compute

$$\frac{d}{d\epsilon} \left[J(u + \epsilon v) \right]_{\epsilon=0} = 0$$

• substitute the expression for J:

$$\frac{d}{d\epsilon} \left[\frac{1}{2} A(u + \epsilon v, u + \epsilon v) - F(u + \epsilon v) \right]_{\epsilon = 0} = 0$$

• expand both A and F:

$$\frac{d}{d\epsilon} \left[\frac{1}{2} A(u, u) + \epsilon A(u, v) + \frac{1}{2} \epsilon^2 A(v, v) - F(u) - \epsilon F(v) \right]_{\epsilon=0} = 0$$

• compute the derivative:

$$[A(u,v) + \epsilon A(v,v) - F(v)]_{\epsilon=0} = 0$$

• evaluating at $\epsilon = 0$ and rearranging we get

$$A(u, v) = F(v)$$

which is the weak formulation we wanted.

2 Lecture 2

2.1 Proof of the Strong maximum principle for harmonic functions

Let Ω be a connected open subset of \mathbb{R}^n and u be a harmonic function, i.e., $\Delta u = 0$, such that $u \in C^2(\Omega) \cap C(\overline{\Omega})$. Then:

- 1. $\max_{\overline{\Omega}} u = \max_{\partial \Omega} u$.
- 2. If $\exists x_0 \in \Omega$ such that $u(x_0) = \max_{\overline{\Omega}} u = M$, then u is constant in Ω .

By contradiction suppose that u attains its maximum at some point $x_0 \in \Omega$, i.e. $u(x_0) = \max_{\overline{\Omega}} u$; now letting $B(x_0, r) \subset \Omega$, it follows that since u is harmonic, by the mean value property its value at a point is equal to the average integral over a sphere of any radius centred at that point:

$$M = u(x_0) = \frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} u(x) dx.$$

For the average to be equal to the maximum value it must be that u(x) = M $\forall x \in B(x_0, r)$.

Now consider the set

$$\{x \in \Omega : u(x) = M\}$$

this set is open because inside the open ball $B(x_0,r)$ where $u(x_0)=M$, all nearby points satisfy also u(y)=M. Since u is continuous, if a sequence of points in $B(x_0,r)$ converges to a point in Ω , that point must also be in $B(x_0,r)$; this implies that the set is also relatively closed in Ω . If a non-empty set S is both open and closed within a connected space U, then it must be the entire space:

- 1. suppose S is not the entire space U; then $U \setminus S$ is non-empty
- 2. since S is open and relatively closed, $U\backslash S$ is both relatively open and relatively closed
- 3. this would imply that U can be divided into two disjoint, non-empty, open subsets, which contradicts the assumption that it is connected

We conclude that the set

$$\{x \in \Omega : u(x) = M\}$$

is the entire space Ω and thus u is constant in Ω .

2.2 Proof of the corollary (uniqueness for Poisson)

Suppose there exist two solutions u and v to the Poisson problem:

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = g & \text{on } \partial \Omega. \end{cases}$$

and

$$\begin{cases} -\Delta v = f & \text{in } \Omega, \\ v = g & \text{on } \partial \Omega. \end{cases}$$

defining w = u - v we get:

$$\begin{cases} -\Delta w = -\Delta u + \Delta v = f - f = 0 & \text{in } \Omega, \\ w = u - v = g - g = 0 & \text{on } \partial \Omega. \end{cases}$$

so w is a harmonic function that is null on $\partial\Omega$ and we can apply the strong maximum principle: since w achieves both its maximum and minimum on the boundary, it must be constant on Ω , thus proving u = v.

3 Lecture 3

3.1 Convergence of the truncation error

Given the finite difference discretization

$$-\frac{a_i}{h^2}(u_{i+1} - 2u_i + u_{i-1}) + \frac{b_i}{2h}(u_{i+1} - u_{i-1}) = f_i$$

we want to show that the truncation error $|T| = |Lu(x_i) - L_h(u(x_i))|$ is bounded by

$$|T| \le \frac{h^2}{12} ||a|| ||u^{(4)}|| + \frac{h^2}{6} ||b|| ||u^{(3)}||$$

3.1.1 Second derivative term

Using Taylor expansion around x_i :

$$u_{i+1} = u(x_i + h) = u(x_i) + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{6}u^{(3)}(x_i) + \frac{h^4}{24}u^{(4)}(x_i) + O(h^5)$$

$$u_{i-1} = u(x_i - h) = u(x_i) - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{6}u^{(3)}(x_i) + \frac{h^4}{24}u^{(4)}(x_i) + O(h^5)$$

Thus
$$u_{i+1} - 2u_i + u_{i-1} = h^2 u''(x_i) + \frac{h^4}{12} u^{(4)}(x_i) + O(h^6)$$

3.1.2 First derivative term

Using the same Taylor expansion around x_i we get:

$$u_{i+1} - u_{i-1} = 2hu'(x_i) + \frac{2h^3}{6}u^{(3)}(x_i) + O(h^5)$$

3.1.3 Combine the errors

Combining the truncation errors for both terms, we have:

$$T_i = -a_i \left(\frac{h^2}{12}u^{(4)}(x_i)\right) + b_i \left(\frac{h^2}{6}u^{(3)}(x_i)\right)$$

Taking the absolute value and using the norms ||a|| and ||a||, we get:

$$|T_i| \le \frac{h^2}{12} ||a|| ||u^{(4)}|| + \frac{h^2}{6} ||b|| ||u^{(3)}||.$$

3.2 Construction of a 3-point scheme: method of undetermined coefficients

We want to find the best combination of values for the coefficients $\alpha, \beta, \gamma \in \mathbb{R}$ that appear in the finite difference operator

$$L = \alpha u_{i+1} + \beta u_i + \gamma u_{i-1}$$

so that the approximation $u''(x_i) = u_i''$ is the best possible. We start by expanding using Taylor series:

$$L = \alpha \left(u_i + h_{i+1}u_i' + \frac{h_{i+1}^2}{2}u_i'' + \frac{h_i^3}{6}u'''(\xi_{i+1}) \right) + \beta u_i + \gamma \left(u_i - h_i u_i' + \frac{h_i^2}{2}u_i'' + \frac{h_i^3}{6}u'''(\xi_i) \right)$$

The terms of order 0 and 1 must cancel out, while the second order term must have coefficient 1, so we get the following conditions on the coefficients:

$$\begin{cases} \alpha + \beta + \gamma = 0 \\ \alpha h_{i+1} - \gamma h_i = 0 \\ \alpha \frac{h_{i+1}^2}{2} + \gamma \frac{h_i^2}{2} = 1 \end{cases}$$

from the second equation we get $\alpha = \frac{h_i}{h_{i+1}} \gamma$; by substituting in the third find $\gamma = \frac{2}{h_i(h_{i+1}+h_i)}$ which gives $\alpha = \frac{2}{h_{i+1}(h_{i+1}+h_i)}$ and $\beta = -\frac{2}{h_{i+1}+h_i}$. The resulting FD scheme is therefore:

$$\begin{cases} u_0 = 0 \\ \frac{2}{h_{i+1}(h_{i+1} + h_i)} u_{i+1} - \frac{2}{h_{i+1} + h_i} u_i + \frac{2}{h_i(h_{i+1} + h_i)} u_{i-1} = f_i & \text{for } i = 1, ..., N - 1 \\ u_N = 0 \end{cases}$$

4 Lecture 4

4.1 Truncation error bound for 2D Poisson b.v.p.

If $u \in C^4(\Omega) \cap C^0(\Omega)$ then the truncation error of the 5-point scheme is bounded by:

$$|T(x)| \le \frac{h^2}{12} (\|u_{xxxx}\|_{C(\overline{\Omega})} + \|u_{yyyy}\|_{C(\overline{\Omega})})$$

To prove this we remember that the truncation error in this case is defined as:

$$T_{ij} = \frac{1}{h^2} \left(u(x_{i+1}, y_j) + u(x_{i-1}, y_j) + u(x_i, y_{j+1}) + u(x_i, y_{j-1}) - 4u(x_i, y_j) \right) - f(x_i, y_j)$$

which we rewrite as:

$$T_{ij} = \frac{u(x_{i+1}, y_j) - 2u(x_i, y_j) + u(x_{i-1}, y_j)}{h^2} + \frac{u(x_i, y_{j+1}) - 2u(x_i, y_j) + u(x_i, y_{j-1})}{h^2} - f(x_i, y_j)$$

using Taylor expansion we can prove that the first two terms are equivalent to:

$$u_{xx}(x_i, y_j) + \frac{h^2}{24}(u_{xxxx}(\xi_1, y_j) + u_{xxxx}(\zeta_1, y_j))$$

and

$$u_{yy}(x_i, y_j) + \frac{h^2}{24}(u_{yyyy}(x_i, \xi_2) + u_{yyyy}(x_i, \zeta_2))$$

for some $\xi_1, \zeta_1 \in [x_{i-1}, x_{i+1}]$ and $\xi_2, \zeta_2 \in [y_{i-1}, y_{i+1}]$ respectively. Thus, combining the result and taking the maximum of the absolute value for each derivative we get the following bound:

$$|T_i| \le \frac{h^2}{12} \left(\max_{(x,y) \in \overline{\Omega}} |u_{xxxx}(x,y)| + \max_{(x,y) \in \overline{\Omega}} |u_{yyyy}(x,y)| \right)$$

which is equivalent to the result we want to prove if we recall the definition of L^{∞} norm.

4.2 Derivation of FD scheme for the mixed operator $\delta^x_{2h}\delta^y_{2h}u_{ij}$

$$(\delta_h^x u)_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2h}, \qquad (\delta_h^y u)_{i,j} = \frac{u_{i,j+1} - u_{i,j-1}}{2h}.$$

Applying δ_h^x first and then δ_h^y :

$$(\delta_h^y \delta_h^x u)_{i,j} = \frac{(\delta_h^x u)_{i,j+1} - (\delta_h^x u)_{i,j-1}}{2h}.$$

Substitute $(\delta_h^x u)_{i,j\pm 1}$ into the above and simplify:

$$(\delta_h^x \delta_h^y u)_{i,j} = \frac{u_{i+1,j+1} - u_{i+1,j-1} - u_{i-1,j+1} + u_{i-1,j-1}}{4h^2}$$

5 Lecture 5

5.1 Proof that $||v||_1$ is a full norm for the space H_0^1

1. Definiteness: show that

$$||v||_1 = 0 \iff v \equiv 0 \text{ (in }\Omega).$$

- (\Longrightarrow) If $\|\nabla v\|_{L^2(\Omega)} = 0$, then $\nabla v = 0$ almost everywhere. Thus v is (almost everywhere) a constant C. Since $v \in H_0^1(\Omega)$, it satisfies v = 0 on $\partial\Omega$. Hence that constant must be C = 0. Therefore $v \equiv 0$ in Ω .
- (\longleftarrow) Clearly if $v \equiv 0$, then $\nabla v \equiv 0$ and thus $\|\nabla v\|_{L^2(\Omega)} = 0$.

Hence

$$||v||_1 = ||\nabla v||_{L^2(\Omega)} = 0 \iff v \equiv 0.$$

2. Homogeneity: $\forall \alpha \in \mathbb{R} \text{ and } \forall v \in H_0^1(\Omega) \text{ show that }$

$$\|\alpha v\|_1 = |\alpha| \|v\|_1.$$

By definition,

$$\|\alpha v\|_1 = \|\nabla(\alpha v)\|_{L^2(\Omega)} = \|\alpha \nabla v\|_{L^2(\Omega)} = |\alpha| \|\nabla v\|_{L^2(\Omega)} = |\alpha| \|v\|_1.$$

3. Triangle inequality: $\forall u, v \in H_0^1(\Omega)$ show that

$$||u+v||_1 \le ||u||_1 + ||v||_1.$$

Indeed,

$$\|u+v\|_1 = \|\nabla(u+v)\|_{L^2(\Omega)} = \|\nabla u+\nabla v\|_{L^2(\Omega)} \le \|\nabla u\|_{L^2(\Omega)} + \|\nabla v\|_{L^2(\Omega)} = \|u\|_1 + \|v\|_1.$$

6 Lecture 6

6.1 Proof of the continuity of the bilinear form in Lax-Milgram

The bilinear form A is continuous if:

$$|A(u,v)| \le ||u|| ||v|| \forall u, v \in V.$$

We study the three terms separately.

6.1.1 First term

Assume A is a bounded, symmetric and positive definite matrix with components $A_{ij}(x)$, such that:

$$\lambda_{min}|\xi|^2 \le A(x)\xi \cdot \xi \le \lambda_{max}|\xi|^2$$

 $\forall \xi \in \mathbb{R}^n$ and $\forall x \in \Omega$, where λ_{min} and λ_{max} are the minimum and maximum eigenvalues of A. Then taking the absolute value the following inequality holds:

$$\left| \int_{\Omega} A \nabla u \nabla v dx \right| \le \lambda_{max} \int_{\Omega} |\nabla u| |\nabla v| dx$$

using the Cauchy-Schwarz inequality:

$$\int_{\Omega} |\nabla u| \, |\nabla v| \, dx \le \left(\int_{\Omega} |\nabla u|^2 \, dx \right)^{\frac{1}{2}} \left(\int_{\Omega} |\nabla v|^2 \, dx \right)^{\frac{1}{2}} = \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)}$$

SO

$$\left| \int_{\Omega} A \nabla u \nabla v dx \right| \leq \lambda_{max} \|\nabla u\|_{L^{2}(\Omega)} \|\nabla v\|_{L^{2}(\Omega)}$$

6.1.2 Second term

Assume b(x) is a bounded vector field. To obtain boundedness it is sufficient to assume that $b(x) \in L^{\infty}(\Omega)$. Using then also Cauchy-Schwarz we get:

$$\left| \int_{\Omega} \left(b \cdot \nabla u \right) v dx \right| \leq \int_{\Omega} \left| b \cdot \nabla u \right| \left| u \right| \leq \|b\|_{L^{\infty}(\Omega)} \int_{\Omega} \left| \nabla u \right| \left| v \right| \leq \|b\|_{L^{\infty}(\Omega)} \|\nabla u\|_{L^{2}(\Omega)} \|v\|_{L^{2}(\Omega)}$$

6.1.3 Third term

Assume c(x) is a bounded scalar field and use Cauchy-Schwarz as before:

$$\left| \int_{\Omega} cuv dx \right| \le \|c\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}$$

Now combining the results:

$$|A(u,v)| \leq \lambda_{max} \|\nabla u\|_{L^{2}(\Omega)} \|\nabla v\|_{L^{2}(\Omega)} + \|b\|_{L^{\infty}(\Omega)} \|\nabla u\|_{L^{2}(\Omega)} \|v\|_{L^{2}(\Omega)} + \|c\|_{L^{\infty}(\Omega)} \|u\|_{L^{2}(\Omega)} \|v\|_{L^{2}(\Omega)}$$

Now recalling the definition of the H^1 norm $||u||_{H^1(\Omega)}$, which is defined as

$$||u||_{H^1(\Omega)} = \left(||u||_{L^2(\Omega)}^2 + ||\nabla u||_{L^2(\Omega)}^2\right)^{\frac{1}{2}}$$

we can bound each term with the H^1 norm of u or v, for example:

$$\lambda_{max} \|\nabla u\|_{L^{2}(\Omega)} \|\nabla v\|_{L^{2}(\Omega)} \le \lambda_{max} \|u\|_{H^{1}(\Omega)} \|v\|_{H^{1}(\Omega)}$$

so we can factor out $||u||_{H^1(\Omega)}$ and $||v||_{H^1(\Omega)}$ from each term and obtain:

$$|A(u,v)| \le C||u||_{H^1(\Omega)}||u||_{H^1(\Omega)}$$

where the constant C is given by:

$$C = \max \left(\lambda_{max}, \|b\|_{L^{\infty}(\Omega)}, \|c\|_{L^{\infty}(\Omega)} \right)$$

which proves the continuity of A(u, v).

6.2 Proof of special case of Ceà's lemma

Let V be a Hilbert space and V_h a finite-dimensional subspace of V. Consider a bilinear form $a:V\times V\to \mathbb{R}$ and a linear functional $F:V\to \mathbb{R}$. Assume a is coercive, continuous and symmetric. Let $u\in V$ be the solution of the variational problem: $a(u,v)=F(v)\ \forall v\in V$, and let $u_h\in V_h$ be the solution of the corresponding finite-dimensional problem $a(u_h,v_h)=F(v_h)\ \forall v_h\in V_h$. Then, we have:

$$||u - u_h||_V \le \sqrt{\frac{\gamma}{\alpha_0}} \min_{v_h \in V_h} ||u - v_h||$$

Proof

adding symmetry to the properties of the form a makes it an inner product on the space V, which allows us to define the norm $||u||_a = \sqrt{a(u,u)}$. Thus,

$$\|u - u_h\|_a^2 = a(u - u_h, u - u_h) = a(u - u_h, u - v_h) \le \|u - u_h\|_a \|u - v_h\|_a \ \forall v \in V_h$$

which leads to

$$||u - u_h||_a \le ||u - v||_a \ \forall v \in V_h$$

Using this result together with continuity and coercivity we can derive:

$$\alpha \|u - u_h\|^2 \le a(u - u_h, u - u_h) = \|u - u_h\|_a^2 \le \|u - v_h\|_a^2 \le \gamma \|u - v_h\|^2 \ \forall v_h \in V_h$$

form which follows:

$$||u - u_h|| \le \sqrt{\frac{\gamma}{\alpha}} ||u - v_h|| \ \forall v_h \in V_h$$

7 Lecture 11

7.1 Local truncation error (consistency) of the θ -method

Prove that the local truncation error of the θ -method for the heat equation $u_t = a u_{xx}$ is given by:

$$T_i^{n+\theta} = \delta_k^t u_i^{n+\theta} - \theta \, a \, \delta_h^2 u_i^{n+1} - (1-\theta) \, a \, \delta_h^2 u_i^n$$

and that:

• For $\theta \neq \frac{1}{2}$, the scheme is **first order** in time:

$$T_i^{n+\theta} = O(k) + O(h^2)$$

• For $\theta = \frac{1}{2}$ (the Crank–Nicolson case), it is **second order** in time:

$$T_i^{n+1/2} = O(k^2) + O(h^2)$$

Proof

Let $k = t_{n+1} - t_n$, $h = x_i - x_{i-1}$, $u_i^n = u(x_i, t_n)$. Our scheme is:

$$\frac{U_i^{\,n+1} - U_i^{\,n}}{k} \; = \; \theta \, a \, \delta_h^2 U_i^{\,n+1} \; + \; (1-\theta) \, a \, \delta_h^2 U_i^{\,n},$$

where δ_h^2 is the centered-difference operator in space:

$$\delta_h^2 v_i = \frac{v_{i+1} - 2v_i + v_{i-1}}{h^2}.$$

The local truncation error is obtained by substituting the true PDE solution u into the discrete scheme and measuring the residual:

- 1. Replace U_i^n by $u_i^n = u(x_i, t_n)$.
- 2. Define

$$T_i^{n+\theta} := \underbrace{\frac{u_i^{n+1} - u_i^n}{k}}_{\delta_k^t u_i^{n+\theta}} - \left[\theta \, a \, \delta_h^2 u_i^{n+1} \, + \, (1-\theta) \, a \, \delta_h^2 u_i^n\right].$$

Because u satisfies $u_t - a u_{xx} = 0$, we can write:

$$\delta_k^t u_i^{n+\theta} - a u_{xx}(t_{n+\theta}, x_i) = 0.$$

Now using Taylor expansion we will show that:

- If $\theta \neq \frac{1}{2}$, then $T_i^{n+\theta} = \mathcal{O}(k) + \mathcal{O}(h^2)$.
- If $\theta=\frac{1}{2}$ (the Crank–Nicolson/trapezoid rule in time), then $T_i^{n+1/2}=\mathcal{O}(k^2)+\mathcal{O}(h^2).$

To see the orders cleanly, we rewrite the truncation error as follows:

$$T_i^{n+\theta} = \left[\delta_k^t u_i^{n+\theta} - u_t(t_{n+\theta}, x_i) \right] - a \left[\theta \, \delta_h^2 u_i^{n+1} + (1-\theta) \, \delta_h^2 u_i^{n} - u_{xx}(t_{n+\theta}, x_i) \right].$$

and name the two main brackets as:

$$T_i^{n+\theta} = R_{\text{time}} - a R_{\text{space}}.$$

7.1.1 The time part R_{time}

$$R_{\text{time}} = \frac{u(x_i, t_{n+1}) - u(x_i, t_n)}{k} - u_t(t_{n+\theta}, x_i)$$

Case (a): $\theta \neq \frac{1}{2}$

We expand $u(t_{n+1}, x_i)$ and $u(t_n, x_i)$ in Taylor series around $t_{n+\theta}$. Let $t_{n+\theta} = t_n + \theta k$. Then

$$t_{n+1} = t_{n+\theta} + (1-\theta) k, \quad t_n = t_{n+\theta} - \theta k.$$

which gives

$$u(t_{n+1}, x_i) = u(t_{n+\theta}, x_i) + (1 - \theta) k u_t(t_{n+\theta}, x_i) + \frac{(1 - \theta)^2 k^2}{2} u_{tt}(\xi_1, x_i),$$

for some $\xi_1 \in (t_{n+\theta}, t_{n+1})$, and similarly

$$u(t_n, x_i) = u(t_{n+\theta}, x_i) - \theta k u_t(t_{n+\theta}, x_i) + \frac{\theta^2 k^2}{2} u_{tt}(\xi_2, x_i).$$

Subtracting and dividing by k:

$$\frac{u_i^{n+1} - u_i^n}{k} = u_t(t_{n+\theta}, x_i) + \frac{\left[(1-\theta)^2 + \theta^2 \right] k}{2} u_{tt}(\xi_3, x_i).$$

Therefore

$$R_{\text{time}} = \frac{u_i^{n+1} - u_i^n}{k} - u_t(t_{n+\theta}, x_i) = \frac{\left[(1-\theta)^2 + \theta^2 \right] k}{2} u_{tt}(\xi_3, x_i) = \mathcal{O}(k).$$

Hence for $\theta \neq \frac{1}{2}$, R_{time} is first-order in k.

Case (b): $\theta = \frac{1}{2}$

In this special case, the method is exactly the trapezoid rule in time (Crank–Nicolson):

$$\frac{u(t_{n+1}, x_i) - u(t_n, x_i)}{k} - u_t \left(\frac{t_n + t_{n+1}}{2}, x_i\right) = \mathcal{O}(k^2).$$

The leading error term (proportional to k) cancels, leaving a remainder of order k^2 . Thus

$$R_{\text{time}} = \mathcal{O}(k^2) \text{ when } \theta = \frac{1}{2}.$$

7.1.2 The space part R_{space}

$$R_{\text{space}} = \theta \, \delta_h^2 u_i^{n+1} + (1-\theta) \, \delta_h^2 u_i^n - u_{xx}(t_{n+\theta}, x_i).$$

In space, recalling the standard second-order finite-difference analysis, we know that:

$$\delta_h^2 u(t, x_i) - u_{xx}(t, x_i) = \mathcal{O}(h^2)$$

Now we must also do a Taylor expansion in time so that $\delta_h^2 u(t_n, x_i)$ can be compared to $\delta_h^2 u(t_{n+\theta}, x_i)$:

$$\delta_h^2 u(t_n, x_i) = \delta_h^2 u(t_{n+\theta}, x_i) - \theta k \frac{\partial}{\partial t} \left[\delta_h^2 u(t_{n+\theta}, x_i) \right] + \mathcal{O}(k^2)$$

$$\delta_h^2 u(t_{n+1}, x_i) = \delta_h^2 u(t_{n+\theta}, x_i) + (1 - \theta)k \frac{\partial}{\partial t} \left[\delta_h^2 u(t_{n+\theta}, x_i) \right] + \mathcal{O}(k^2),$$

which when substituting back cancel out and give:

$$\theta \, \delta_h^2 u(t_{n+1}, \, x_i) + (1-\theta) \, \delta_h^2 u(t_n, \, x_i) \; = \; \delta_h^2 u\big(t_{n+\theta}, \, x_i\big) \; + \; \mathcal{O}(k^2).$$

Now putting everything together:

$$\theta\,\delta_h^2 u_i^{\,n+1} + (1-\theta)\,\delta_h^2 u_i^{\,n} = \delta_h^2 u_i^{\,n+\theta} + \mathcal{O}(k^2),$$

and

$$\delta_h^2 u_i^{n+\theta} - u_{xx}(t_{n+\theta}, x_i) = \mathcal{O}(h^2).$$

SO

$$R_{\rm space} \; = \; \delta_h^2 u_i^{n+\theta} - u_{xx}(t_{n+\theta}, \, x_i) \; + \; \mathcal{O}(k^2) \; = \; \mathcal{O}(h^2) \; + \; \mathcal{O}(k^2)$$

which holds for any value of θ , thus we arrive at the desired result:

Local truncation error =
$$\begin{cases} \mathcal{O}(k+h^2), & \theta \neq \frac{1}{2}, \\ \mathcal{O}(k^2+h^2), & \theta = \frac{1}{2}. \end{cases}$$

7.2 Sufficient condition for stability of the θ -method

We want to prove that the condition:

$$\mu(1-\theta) \le 1/2$$

is sufficient for the stability of the following θ -method scheme:

$$-\mu\theta U_{i+1}^{n+1} + (1+2\mu\theta)U_i^{n+1} - \mu\theta U_{i-1}^{n+1} = -\mu\theta U_{i+1}^n + (1-2\mu(1-\theta))U_i^n - \mu\theta U_{i-1}^n$$

Let's start by bounding the right hand side. By taking the absolute value we get:

$$RHS \le |\mu(1-\theta)||U_{i+1}^n| + |1-2\mu(1-\theta)||U_i^n| + |\mu(1-\theta)||U_{i-1}^n|$$

Now remember that by definition $\mu = \frac{h}{k^2} > 0$ and $0 \le \theta \le 1$, and let $\mu(1-\theta) \le 1/2$ (our hypothetical sufficient condition). This means that the coefficients $\mu(1-\theta)$ and $1-2\mu(1-\theta)$ are both non-negative, so we can remove the absolute value. Let's also define $U_{max}^n = max_i(U_i^n)$ so we can write:

$$RHS \le |U_{max}^n|[\mu(1-\theta) + 1 - 2\mu(1-\theta) + \mu(1-\theta)] = |U_{max}^n| = ||U^n||_{\infty}$$

Now let's bound the left hand side as well:

$$LHS \le -|\mu\theta||U_{i+1}^{n+1}| + |1 + 2\mu\theta||U_{i}^{n+1}| - |\mu\theta||U_{i-1}^{n+1}|$$

by similar consideration we obtain

$$LHS \le |U_{max}^{n+1}|(-\mu\theta + 1 + 2\mu\theta - \mu\theta) = |U_{max}^{n+1}| = ||U^{n+1}||_{\infty}$$

hence $LHS \leq RHS$ translates to $||U^{n+1}||_{\infty} \leq ||U^n||_{\infty} \, \forall n$, which by extension implies $||U^n||_{\infty} \leq ||U^0||_{\infty}$.

7.3 Fourier (von Neumann) Analysis for the θ -method in 1D

We consider the 1D heat equation

$$u_t = a u_{xx}, \quad x \in (0,1), \ t > 0,$$

with homogeneous Dirichlet boundary conditions. We discretize spatially with a uniform grid

$$x_j = j h, \quad j = 0, \dots, M, \quad h = \frac{1}{M},$$

and discretize time with

$$t_n = n k$$
.

Let U_i^n approximate $u(x_j, t_n)$. The θ -method for the heat equation in 1D is

$$\frac{U_j^{n+1} - U_j^n}{k} = \theta \, a \, \delta_h^2 U_j^{n+1} + (1 - \theta) \, a \, \delta_h^2 U_j^n,$$

where

$$\delta_h^2 U_j^n = \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2},$$

for $j=1,\ldots,M-1$, with boundary conditions $U_0^n=U_M^n=0$. Define

$$\mu = \frac{a k}{h^2}.$$

We look for solutions of the form

$$U_i^n = \lambda^n \exp(i \alpha x_i) = \lambda^n \exp(i \alpha j h),$$

where

- j is the spatial index (integer),
- i is the imaginary unit,
- α is a wavenumber

Our goal is to find λ (the amplification factor) in terms of α , μ , and θ .

Now we compute the second difference $\delta_h^2 U_i^n$

$$\delta_h^2 U_j^n = \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2}$$

substitute $U_i^n = \lambda^n \exp(i \alpha j h)$:

$$U_{j+1}^n = \lambda^n \, \exp \bigl(\mathrm{i} \, \alpha \, (j+1) \, h \bigr) = \lambda^n \, \exp(\mathrm{i} \, \alpha \, j \, h) \, \exp(\mathrm{i} \, \alpha \, h)$$

$$U_i^n = \lambda^n \exp(\mathrm{i} \, \alpha \, j \, h)$$

$$U_{j-1}^n = \lambda^n \, \exp\!\!\left(\mathrm{i}\,\alpha\,(j-1)\,h\right) = \lambda^n \, \exp\!\!\left(\mathrm{i}\,\alpha\,j\,h\right) \, \exp\!\left(-\mathrm{i}\,\alpha\,h\right)$$

hence

$$\delta_h^2 U_j^n = \frac{\lambda^n \, \exp(\mathrm{i} \, \alpha \, j \, h)}{h^2} \left[\exp(\mathrm{i} \, \alpha \, h) - 2 + \exp(-\mathrm{i} \, \alpha \, h) \right]$$

but

$$\exp(i \alpha h) + \exp(-i \alpha h) = 2 \cos(\alpha h)$$

so

$$\exp(i \alpha h) - 2 + \exp(-i \alpha h) = -4 \sin^2\left(\frac{\alpha h}{2}\right)$$

therefore,

$$\delta_h^2 U_j^n = -\,\frac{4\,\sin^2(\frac{\alpha\,h}{2})}{h^2}\,\,\lambda^n\,\exp(\mathrm{i}\,\alpha\,j\,h)$$

which when substituted into the θ -method gives:

$$\frac{U_j^{n+1} - U_j^n}{k} = \theta \, a \, \delta_h^2 U_j^{n+1} + (1 - \theta) \, a \, \delta_h^2 U_j^n.$$

Now we look at the time difference (LHS):

$$U_j^{n+1} = \lambda^{n+1} \exp(\mathrm{i}\,\alpha\,j\,h), \quad U_j^n = \lambda^n \exp(\mathrm{i}\,\alpha\,j\,h).$$

$$\frac{U_j^{n+1}-U_j^n}{k} = \frac{\lambda^{n+1}-\lambda^n}{k} \, \exp(\mathrm{i}\,\alpha\,j\,h) = \exp(\mathrm{i}\,\alpha\,j\,h)\,\frac{\lambda^n}{k}\,(\lambda-1)$$

and at the space difference (RHS):

$$\begin{split} \theta \, a \, \delta_h^2 U_j^{n+1} \, + \, (1-\theta) \, a \, \delta_h^2 U_j^n &= \theta \, a \, \left[-\frac{4 \, \sin^2(\frac{\alpha \, h}{2})}{h^2} \right] \lambda^{n+1} \exp(\mathrm{i} \, \alpha \, j \, h) + (1-\theta) \, a \, \left[-\frac{4 \, \sin^2(\frac{\alpha \, h}{2})}{h^2} \right] \lambda^n \exp(\mathrm{i} \, \alpha \, j \, h) \\ &= -\frac{4 \, a \, \sin^2(\frac{\alpha \, h}{2})}{h^2} \, \exp(\mathrm{i} \, \alpha \, j \, h) \left[\theta \, \lambda^{n+1} + (1-\theta) \, \lambda^n \right] \\ &= -\frac{4 \, a \, \sin^2(\frac{\alpha \, h}{2})}{h^2} \, \exp(\mathrm{i} \, \alpha \, j \, h) \, \lambda^n \, \left[\theta \, \lambda + (1-\theta) \right] \end{split}$$

and finally equating them we find that the original:

$$\frac{U_j^{n+1} - U_j^n}{k} = \theta \, a \, \delta_h^2 U_j^{n+1} + (1 - \theta) \, a \, \delta_h^2 U_j^n$$

becomes

$$\exp(\mathrm{i}\,\alpha\,j\,h)\,\frac{\lambda^n}{k}\,(\lambda-1) = -\,\frac{4\,a\,\sin^2(\frac{\alpha\,h}{2})}{h^2}\,\,\exp(\mathrm{i}\,\alpha\,j\,h)\,\lambda^n\,\big[\theta\,\lambda + (1-\theta)\big]$$

and after cancelling $\exp(i \alpha j h) \lambda^n$ we get:

$$\frac{\lambda - 1}{k} = -\frac{4 a \sin^2(\frac{\alpha h}{2})}{h^2} \left[\theta \lambda + (1 - \theta)\right].$$

Now by multiplying both sides by k and remembering the substitution $\mu = \frac{a\,k}{h^2}$ we get:

$$\lambda + 4 \mu \theta \sin^2(\frac{\alpha h}{2}) \lambda = 1 - 4 \mu (1 - \theta) \sin^2(\frac{\alpha h}{2}).$$

from which we obtain the amplification factor λ :

$$\lambda(\alpha) = \frac{1 - 4\mu(1 - \theta)\sin^2(\frac{\alpha h}{2})}{1 + 4\mu\theta\sin^2(\frac{\alpha h}{2})}.$$

The stability condition is then $|\lambda(\alpha)| \leq 1$ for all relevant α .

• If $\theta \ge 1/2$: the denominator exceeds or equals the numerator in absolute value for any $\mu > 0$, so $|\lambda(\alpha)| \le 1$ unconditionally.

• If $0 \le \theta < 1/2$: the worst case is $\sin^2(\frac{\alpha h}{2}) = 1$, then

$$\lambda(\alpha) = \frac{1 - 4\mu(1 - \theta)}{1 + 4\mu\theta}$$

and requiring $|\lambda(\alpha)| \leq 1$ forces

$$\mu \left(1 - 2\theta \right) \le \frac{1}{2}.$$

8 Lecture 14

8.1 Wave equation equivalent system

Consider the classical wave equation:

$$u_{tt} - u_{xx} = 0.$$

Introduce an auxiliary variable v and consider the system:

$$u_t + v_x = 0, (1)$$

$$u_x + v_t = 0. (2)$$

This system is equivalent to the wave equation.

Proof that u satisfies the wave equation:

1. Differentiate equation (1) with respect to t:

$$u_{tt} + v_{xt} = 0.$$

2. Differentiate equation (2) with respect to x:

$$u_{xx} + v_{tx} = 0.$$

3. Since partial derivatives commute (i.e. $v_{xt} = v_{tx}$), subtract the second equation from the first:

$$u_{tt} - u_{xx} = 0.$$

Thus, u satisfies the wave equation.

Proof that v satisfies the wave equation:

1. Differentiate equation (2) with respect to t:

$$u_{xt} + v_{tt} = 0.$$

2. Differentiate equation (1) with respect to x:

$$u_{tx} + v_{xx} = 0.$$

3. Notice that $u_{xt} = u_{tx}$ (same as before), subtract the second equation from the first:

$$v_{tt} - v_{xx} = 0.$$

Thus, v also satisfies the wave equation.

8.2 Conservation of energy for the wave equation

Consider the classical wave equation:

$$u_{tt} - u_{xx} = 0, \quad (t, x) \in (0, T] \times \mathbb{R}.$$

Show that the energy

$$E(t) = \frac{1}{2} \int_{\mathbb{R}} \left[\left(u_t(x,t) \right)^2 + \left(u_x(x,t) \right)^2 \right] dx$$

is conserved, i.e., E(t) is constant in time.

Solution:

1. Multiply the wave equation by $u_t(x,t)$ to obtain:

$$u_t u_{tt} - u_t u_{xx} = 0.$$

2. Integrate the above expression over \mathbb{R} with respect to x:

$$\int_{\mathbb{R}} \left(u_t \, u_{tt} - u_t \, u_{xx} \right) dx = 0.$$

3. Process the time derivative term:

recognize that

$$u_t u_{tt} = \frac{1}{2} \frac{\partial}{\partial t} (u_t^2)$$

thus,

$$\int_{\mathbb{R}} u_t \, u_{tt} \, dx = \frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}} u_t^2 \, dx.$$

4. Process the spatial derivative term:

integrate by parts in x by setting

$$v = u_t$$
 and $dw = u_{xx} dx$

so that $dv = u_{tx} dx$ and $w = u_x$. Then,

$$- \int_{\mathbb{R}} u_t \, u_{xx} \, dx = - \left[u_t \, u_x \right]_{-\infty}^{+\infty} + \int_{\mathbb{R}} u_{tx} \, u_x \, dx.$$

Assuming that u_t and u_x vanish at infinity, the boundary term is zero.

Notice that

$$u_{tx} u_x = \frac{1}{2} \frac{\partial}{\partial t} (u_x^2)$$

hence

$$\int_{\mathbb{R}} u_{tx} \, u_x \, dx = \frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}} u_x^2 \, dx.$$

5. Combine the results:

putting the two parts together, we have

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}}u_t^2\,dx + \frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}}u_x^2\,dx = 0$$

this can be rewritten as:

$$\frac{d}{dt} \left[\frac{1}{2} \int_{\mathbb{R}} \left(u_t^2 + u_x^2 \right) dx \right] = 0$$

therefore, the energy

$$E(t) = \frac{1}{2} \int_{\mathbb{R}} \left(u_t^2 + u_x^2 \right) dx$$

is conserved in time.

9 Lecture 15

9.1 Consistency proof for the Lax-Wendroff scheme

We consider the linear advection PDE:

$$\begin{cases} u_t + a u_x = 0, \\ u(0, x) = u_0(x), \\ u(b, x) = u_b(x), \end{cases}$$

where a is taken as a constant for simplicity. The **Lax–Wendroff scheme** for $u_i^n \approx u(x_i, t_n)$ is written as:

$$\frac{u_i^{n+1} - u_i^n}{k} + a \frac{u_{i+1}^n - u_{i-1}^n}{2h} - \frac{k(a)^2}{2h^2} \left(u_{i+1}^n - 2u_i^n + u_{i-1}^n \right) = 0.$$

We set $\nu := \frac{k}{h}$ (the Courant number). Our goal is to show that the *local truncation error* T_i^n is

$$|T_i^n| \le \frac{k^2}{6} \|u_{ttt}\|_{\infty} + |a| \frac{h^2}{6} \|u_{xxx}\|_{\infty},$$

i.e. it is $\mathcal{O}(k^2 + h^2)$.

Local truncation error

1. Denoting the exact PDE solution at grid points as $U_i^n = u(x_i, t_n)$, substituting into the Lax-Wendroff scheme we get the residual:

$$T_i^n := \frac{U_i^{n+1} - U_i^n}{k} + a \frac{U_{i+1}^n - U_{i-1}^n}{2h} - \frac{k a^2}{2h^2} \left(U_{i+1}^n - 2 U_i^n + U_{i-1}^n \right).$$

2. We expand the time shift:

$$U_i^{n+1} = U_i^n + k U_t^n + \frac{k^2}{2} U_{tt}^n + \frac{k^3}{6} U_{ttt}(x_i, t_n) + \mathcal{O}(k^4).$$

3. and the space shifts:

$$U_{i+1}^{n} = U_{i}^{n} + h U_{x}^{n} + \frac{h^{2}}{2} U_{xx}^{n} + \frac{h^{3}}{6} U_{xxx}(x_{i}, t_{n}) + \mathcal{O}(h^{4}),$$

$$U_{i-1}^{n} = U_{i}^{n} - h U_{x}^{n} + \frac{h^{2}}{2} U_{xx}^{n} - \frac{h^{3}}{6} U_{xxx}(x_{i}, t_{n}) + \mathcal{O}(h^{4}),$$

Substitute into the scheme

1.
$$\frac{U_i^{n+1} - U_i^n}{k} = U_t^n + \frac{k}{2} U_{tt}^n + \frac{k^2}{6} U_{ttt}(x_i, t_n) + \mathcal{O}(k^3)$$

2.
$$a \frac{U_{i+1}^n - U_{i-1}^n}{2h} = a \left(U_x^n + \frac{h^2}{6} U_{xxx}^n \right) + \mathcal{O}(h^3)$$

3.
$$\frac{k a^2}{2} \frac{U_{i+1}^n - U_i^n + U_{i-1}^n}{h^2} = \frac{k a^2}{2} U_{xx} + \mathcal{O}(h^2)$$

Notice that from the PDE we have:

$$U_t = -a U_x, U_{tt} = a^2 U_{xx}, U_{ttt} = -a^3 U_{xxx}$$

so the expression for the truncation error is dominated by $k^2 U_{ttt}$ and $a h^2 U_{xxx}$ terms, hence $|T_i^n|$ is bounded by a combination of $k^2 |U_{ttt}|$ and $|a| h^2 |U_{xxx}|$, plus higher order terms. Consequently,

$$\left|T_i^n\right| \leq \frac{k^2}{6} M_{ttt} + |a| \frac{h^2}{6} M_{xxx} + \cdots,$$

which shows $T_i^n = \mathcal{O}(k^2 + h^2)$.

9.2 Leap frog scheme for wave equation

We consider the model problem

$$u_{tt} - u_{rr} = 0,$$

and recall its standard leap-frog discretization:

$$\frac{U_i^{n+1} - 2U_i^n + U_i^{n-1}}{k^2} = \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{h^2},$$

where k and h denote the time and space step sizes, respectively.

9.2.1 First-order system formulation

To rewrite the second-order wave equation as a first-order system we define the auxiliary variable v such that:

$$u_t + v_x = 0, (1)$$

$$v_t + u_x = 0. (2)$$

Now we show how u solves $u_{tt} - u_{xx} = 0$:

• differentiatiating (1) w.r.t. t we get:

$$u_{tt} = -\frac{\partial}{\partial t}(v_x) = -v_{xt}$$

• differentiatiating (2) w.r.t. x we get:

$$v_{xt} = -u_{xx}$$

• combining, $u_{tt} = -v_{xt} = -[-u_{xx}] = u_{xx}$, hence $u_{tt} - u_{xx} = 0$.

9.2.2 Discretize system by leap-frog method

We now discretize the system (1) and (2) using a leap-frog method with staggered grids. In particular, we choose:

• the solution u is approximated at integer grid points in space and time:

$$U_i^n \approx u(x_i, t^n)$$

with $x_i = i h$ and $t^n = n k$;

ullet the auxiliary variable v is approximated at half-grid points in both space and time:

$$V_{i+\frac{1}{2}}^{n+\frac{1}{2}} \approx v\left(x_{i+\frac{1}{2}}, t^{n+\frac{1}{2}}\right),$$

where $x_{i+\frac{1}{2}}=\left(i+\frac{1}{2}\right)h$ and $t^{n+\frac{1}{2}}=\left(n+\frac{1}{2}\right)k.$

We then have:

$$\frac{U_i^{n+1} - U_i^n}{k} + \frac{V_{i+\frac{1}{2}}^{n+\frac{1}{2}} - V_{i-\frac{1}{2}}^{n+\frac{1}{2}}}{h} = 0,$$
(3)

$$\frac{V_{i+\frac{1}{2}}^{n+\frac{1}{2}} - V_{i+\frac{1}{2}}^{n-\frac{1}{2}}}{k} + \frac{U_{i+1}^n - U_i^n}{h} = 0.$$
 (4)

9.2.3 Derivation of the leap frog scheme for U

The goal now is to eliminate the variable V from (3) and (4) to derive a single scheme for U:

• from (3) we have:

$$\frac{U_i^{n+1} - U_i^n}{k} = -\frac{V_{i+\frac{1}{2}}^{n+\frac{1}{2}} - V_{i-\frac{1}{2}}^{n+\frac{1}{2}}}{h},$$

so that

$$U_i^{n+1} = U_i^n - \frac{k}{h} \left(V_{i+\frac{1}{2}}^{n+\frac{1}{2}} - V_{i-\frac{1}{2}}^{n+\frac{1}{2}} \right)$$
 (5)

• from (4) we have:

$$\frac{V_{i+\frac{1}{2}}^{n+\frac{1}{2}} - V_{i+\frac{1}{2}}^{n-\frac{1}{2}}}{k} = -\frac{U_{i+1}^n - U_i^n}{h},$$

so that

$$V_{i+\frac{1}{2}}^{n+\frac{1}{2}} = V_{i+\frac{1}{2}}^{n-\frac{1}{2}} - \frac{k}{h} \left(U_{i+1}^n - U_i^n \right) \tag{6}$$

• similarly, for the point $i - \frac{1}{2}$ we have:

$$V_{i-\frac{1}{2}}^{n+\frac{1}{2}} = V_{i-\frac{1}{2}}^{n-\frac{1}{2}} - \frac{k}{h} \left(U_i^n - U_{i-1}^n \right) \tag{7}$$

• now subtract (7) from (6):

$$V_{i+\frac{1}{2}}^{n+\frac{1}{2}} - V_{i-\frac{1}{2}}^{n+\frac{1}{2}} = \left(V_{i+\frac{1}{2}}^{n-\frac{1}{2}} - V_{i-\frac{1}{2}}^{n-\frac{1}{2}}\right) - \frac{k}{h} \left[(U_{i+1}^n - U_i^n) - (U_i^n - U_{i-1}^n) \right]$$

$$= \left(V_{i+\frac{1}{2}}^{n-\frac{1}{2}} - V_{i-\frac{1}{2}}^{n-\frac{1}{2}}\right) - \frac{k}{h} \left(U_{i+1}^n - 2U_i^n + U_{i-1}^n\right)$$
(8)

• Now we want to express $V_{i+\frac{1}{2}}^{n-\frac{1}{2}} - V_{i-\frac{1}{2}}^{n-\frac{1}{2}}$ in terms of U. Notice that the same discretization (3) applied at the previous time level gives:

$$\frac{U_i^n - U_i^{n-1}}{k} + \frac{V_{i+\frac{1}{2}}^{n-\frac{1}{2}} - V_{i-\frac{1}{2}}^{n-\frac{1}{2}}}{h} = 0$$

so that

$$V_{i+\frac{1}{2}}^{n-\frac{1}{2}} - V_{i-\frac{1}{2}}^{n-\frac{1}{2}} = -\frac{h}{k} \left(U_i^n - U_i^{n-1} \right) \tag{9}$$

• using (9) in (8) yields

$$V_{i+\frac{1}{2}}^{n+\frac{1}{2}} - V_{i-\frac{1}{2}}^{n+\frac{1}{2}} = -\frac{h}{k} \left(U_i^n - U_i^{n-1} \right) - \frac{k}{h} \left(U_{i+1}^n - 2U_i^n + U_{i-1}^n \right) \tag{10}$$

• recall from (5):

$$U_i^{n+1} = U_i^n - \frac{k}{h} \left(V_{i+\frac{1}{2}}^{n+\frac{1}{2}} - V_{i-\frac{1}{2}}^{n+\frac{1}{2}} \right).$$

• substitute (10) into this expression and rearrange:

$$\begin{split} U_i^{n+1} &= U_i^n - \frac{k}{h} \left[-\frac{h}{k} \left(U_i^n - U_i^{n-1} \right) - \frac{k}{h} \left(U_{i+1}^n - 2U_i^n + U_{i-1}^n \right) \right] \\ &= U_i^n + \left(U_i^n - U_i^{n-1} \right) + \frac{k^2}{h^2} \left(U_{i+1}^n - 2U_i^n + U_{i-1}^n \right) \\ &= 2U_i^n - U_i^{n-1} + \frac{k^2}{h^2} \left(U_{i+1}^n - 2U_i^n + U_{i-1}^n \right). \end{split}$$

• Finally, rewriting this in the standard leap-frog form yields:

$$\frac{U_i^{n+1} - 2U_i^n + U_i^{n-1}}{k^2} = \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{h^2}.$$