MA4255 Numerical Methods in Differential Equations

Chapter 8: FD approximation of parabolic problems

8.1 The heat equation

Parabolic PDEs

Parabolic PDEs: For some given open set $\Omega \subseteq \mathbb{R}^n$, seek a function $u = u(x,t) = u(x_1, \ldots, x_n, t)$ s.t.

$$\partial_t u = -\mathscr{L}_x u, \qquad (x,t) \in \Omega \times (0,\infty), \tag{1}$$

where $\mathscr{L}_x$ is an elliptic differential operator acting on the x-variable, e.g.,

$$\mathscr{L}_x u = -\mathrm{div}_x(A \nabla_x u) + b \cdot \nabla_x u + c\, u$$

with A satisfying the uniform ellipticity condition. We call x the **space variable** and t the **time variable**.

- If $\Omega = \mathbb{R}^n$, the PDE is considered together with an **initial condition (i.c.)**

$$u(x,0) = u_0(x) \quad \text{for} \;\; x \in \Omega, \tag{2}$$

where $u_0 : \Omega \to \mathbb{R}$ is a given function, called an **initial datum**. The PDE (1) together with the i.c. (2) is called an **initial-value problem (IVP)**.

- If $\Omega \subset \mathbb{R}^n$ is a bounded open set, the PDE is considered together with an i.c. (2) and a **boundary condition (b.c.)**

$$u = g \quad \text{on } \partial\Omega \times (0,\infty), \tag{3}$$

where $g : \partial\Omega \times (0,\infty) \to \mathbb{R}$ is a given function. The PDE (1) together with the i.c. (2) and the b.c. (3) is called an **initial-boundary-value problem (IBVP)**.

The heat equation

The **heat equation** (or **diffusion equation**) is the parabolic PDE

$$\partial_t u = -\mathscr{L}_x u \quad \text{in } \Omega \times (0, \infty), \qquad \text{where} \quad \mathscr{L}_x u := -\Delta_x u.$$

We will simply write Δ instead of Δ_x, but keep in mind that it only acts on the space variable $x = (x_1, \ldots, x_n)$, i.e., $\Delta u(x,t) = \sum_{i=1}^n \partial^2_{x_i x_i} u(x,t)$.

First, we focus on the **IVP for the heat equation in one space dimension** ($n = 1$, $\Omega = \mathbb{R}$): Seek a fct $u = u(x,t) : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ s.t.

$$\begin{aligned} \partial_t u(x,t) &= \partial^2_{xx} u(x,t), \qquad && (x,t) \in \mathbb{R} \times (0, \infty), \\ u(x,0) &= u_0(x), && x \in \mathbb{R}, \end{aligned}$$

where $u_0 : \mathbb{R} \to \mathbb{R}$ is some given initial datum.

Solving the IVP for the heat eqn in one space dimension

Let us discuss how to find the true soln to the IVP

$$\begin{aligned} \partial_t u(x,t) &= \partial_{xx}^2 u(x,t), \qquad &(x,t) \in \mathbb{R} \times (0,\infty), \\ u(x,0) &= u_0(x), \qquad &x \in \mathbb{R}. \end{aligned}$$

Key tool: **Fourier transform (FT)** of a fct $v : \mathbb{R} \to \mathbb{C}$ is given by

$$[\mathscr{F}v](\xi) := \hat{v}(\xi) := \int_{-\infty}^{\infty} v(x)\, e^{-ix\xi}\, \mathrm{d}x, \qquad \xi \in \mathbb{R}.$$

We can recover a fct v from its Fourier transform $\hat{v}$ using the **inverse Fourier transform (IFT)**:

$$v(x) = [\mathscr{F}^{-1}\hat{v}](x) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{v}(\xi) e^{ix\xi}\, \mathrm{d}\xi, \qquad x \in \mathbb{R}.$$

We shall assume henceforth that the functions under consideration are sufficiently smooth and that they decay to 0 as $x \to \pm\infty$ sufficiently quickly in order to ensure that our formal manipulations make sense.

Problem:

$$\partial_t u(x,t) = \partial^2_{xx} u(x,t), \qquad (x,t) \in \mathbb{R} \times (0,\infty),$$
$$u(x,0) = u_0(x), \qquad x \in \mathbb{R}.$$

- FT: $[\mathscr{F} v](\xi) := \hat{v}(\xi) := \int_{-\infty}^{\infty} v(x)\, e^{-ix\xi}\, \mathrm{d}x$ for $\xi \in \mathbb{R}$.
- IFT: $v(x) = [\mathscr{F}^{-1}\hat{v}](x) := \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{v}(\xi) e^{ix\xi}\, \mathrm{d}\xi$ for $x \in \mathbb{R}$.

Let $\hat{u}(\xi,t) := \int_{-\infty}^{\infty} u(x,t)\, e^{-ix\xi}\, \mathrm{d}x$ FT of u w.r.t. x-variable. Then,

$$\partial_t \hat{u}(\xi,t) = \frac{\partial}{\partial t}\int_{-\infty}^{\infty} u(x,t)\, e^{-ix\xi}\, \mathrm{d}x = \int_{-\infty}^{\infty} \partial_t u(x,t)\, e^{-ix\xi}\, \mathrm{d}x$$
$$= \int_{-\infty}^{\infty} \partial^2_{xx} u(x,t)\, e^{-ix\xi}\, \mathrm{d}x = (-i\xi)^2 \int_{-\infty}^{\infty} u(x,t)\, e^{-ix\xi}\, \mathrm{d}x = -\,\xi^2 \hat{u}(\xi,t).$$

We see that $y_\xi(t) := \hat{u}(\xi,t)$ satisfies

$$y'_\xi(t) = -\xi^2 y_\xi(t), \qquad y_\xi(0) = \hat{u}_0(\xi).$$

Thus, $\hat{u}(\xi,t) = y_\xi(t) = e^{-t\xi^2}\hat{u}_0(\xi)$. Recover u via IFT:

$$u(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-t\xi^2}\hat{u}_0(\xi) e^{ix\xi}\, \mathrm{d}\xi = \frac{1}{\sqrt{4\pi t}}\int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4t}} u_0(y)\, \mathrm{d}y.$$

(The last equality is not trivial: use defn of $\hat{u}_0(\xi)$, then interchange order of integration, then do some calculation)

$\Longrightarrow$ We have found that the true solution to the IVP for the heat equation in one space dimension, i.e., to the problem

$$\begin{aligned} \partial_t u(x,t) &= \partial_{xx}^2 u(x,t), \qquad && (x,t) \in \mathbb{R} \times (0,\infty), \\ u(x,0) &= u_0(x), && x \in \mathbb{R}, \end{aligned}$$

is given by

$$u(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{4t}} u_0(y)\,\mathrm{d}y = \int_{-\infty}^{\infty} w(x-y,t)\,u_0(y)\,\mathrm{d}y,$$

where w is the so-called **heat kernel** defined as

$$w : \mathbb{R} \times (0,\infty) \to \mathbb{R}, \qquad w(x,t) := \frac{1}{\sqrt{4\pi t}}\, e^{-\frac{x^2}{4t}}.$$

Rk: $w(x,t) > 0 \ \forall (x,t) \in \mathbb{R} \times (0,\infty)$ and $\int_{-\infty}^{\infty} w(x,t)\,\mathrm{d}x = 1 \ \forall t \in (0,\infty)$.
$\Longrightarrow$ If u_0 is a bounded continuous function, then

$$\sup_{x\in\mathbb{R}} |u(x,t)| \le \sup_{x\in\mathbb{R}} |u_0(x)| \qquad \forall t \in (0,\infty).$$

In other words, the 'largest' and 'smallest' values of $x \mapsto u(x,t)$ at $t > 0$ cannot exceed those of u_0.
Next: Derive similar bound in L^2-norm, i.e., $\|u(\cdot,t)\|_{L^2(\mathbb{R})} \le \|u_0\|_{L^2(\mathbb{R})}$.

Stability bound $\|u(\cdot,t)\|_{L^2(\mathbb{R})} \le \|u_0\|_{L^2(\mathbb{R})}$ for heat eqn

For $\Omega \subseteq \mathbb{R}^n$ open and $v : \Omega \to \mathbb{C}$, the $L^2(\Omega)$-norm of v is defined as

$$\|v\|_{L^2(\Omega)} := \left(\int_\Omega |v(x)|^2 \, \mathrm{d}x \right)^{\frac{1}{2}} = \left(\int_\Omega v(x)\overline{v(x)} \, \mathrm{d}x \right)^{\frac{1}{2}}.$$

We write $v \in L^2(\Omega)$ iff $\|v\|_{L^2(\Omega)} < \infty$.

Lemma (Parseval's identity)

Let $v \in L^2(\mathbb{R})$. Then, $\hat{v} \in L^2(\mathbb{R})$ and we have $\|\hat{v}\|^2_{L^2(\mathbb{R})} = 2\pi\|v\|^2_{L^2(\mathbb{R})}$.

For soln of heat eqn $\partial_t u = \partial^2_{xx} u$ in $\mathbb{R} \times (0,\infty)$ with $u(\cdot,0) = u_0$, we have

$$\begin{aligned} 2\pi \int_{-\infty}^{\infty} |u(x,t)|^2 \mathrm{d}x = \int_{-\infty}^{\infty} |\hat{u}(\xi,t)|^2 \mathrm{d}\xi &= \int_{-\infty}^{\infty} |e^{-t\xi^2}\hat{u}_0(\xi)|^2 \mathrm{d}\xi \\ &\le \int_{-\infty}^{\infty} |\hat{u}_0(\xi)|^2 \mathrm{d}\xi = 2\pi \int_{-\infty}^{\infty} |u_0(x)|^2 \mathrm{d}x \end{aligned}$$

for any $t > 0$. Therefore, we have the stability bound

$$\|u(\cdot,t)\|_{L^2(\mathbb{R})} \le \|u_0\|_{L^2(\mathbb{R})} \qquad \forall t \in (0,\infty).$$

Proof of Parseval's identity

Lemma (Parseval's identity)

Let $v \in L^2(\mathbb{R})$. *Then,* $\hat{v} \in L^2(\mathbb{R})$ *and we have* $\|\hat{v}\|_{L^2(\mathbb{R})}^2 = 2\pi \|v\|_{L^2(\mathbb{R})}^2$.

Proof: We have that

$$\begin{aligned}
\|\hat{v}\|_{L^2(\mathbb{R})}^2 = \int_{-\infty}^{\infty} \hat{v}(\xi)\overline{\hat{v}(\xi)}\,\mathrm{d}\xi &= \int_{-\infty}^{\infty} \overline{\hat{v}(\xi)} \int_{-\infty}^{\infty} v(x)\, e^{-ix\xi}\,\mathrm{d}x\,\mathrm{d}\xi \\
&= \int_{-\infty}^{\infty} v(x) \int_{-\infty}^{\infty} \overline{\hat{v}(\xi)}\, e^{-ix\xi}\,\mathrm{d}\xi\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} v(x)\hat{w}(x)\,\mathrm{d}x,
\end{aligned}$$

where $w(s) := \overline{\hat{v}(s)}$. We compute

$$\hat{w}(x) = \int_{-\infty}^{\infty} w(s)\, e^{-ixs}\,\mathrm{d}s = 2\pi \overline{\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{v}(s)\, e^{ixs}\,\mathrm{d}s} = 2\pi \overline{[\mathscr{F}^{-1}\hat{v}](x)} = 2\pi \overline{v(x)}$$

for any $x \in \mathbb{R}$. Thus, $\|\hat{v}\|_{L^2(\mathbb{R})}^2 = 2\pi \int_{-\infty}^{\infty} v(x)\overline{v(x)}\,\mathrm{d}x = 2\pi \|v\|_{L^2(\mathbb{R})}^2$. $\square$

Stability of soln to heat eqn w.r.t. perturb. in initial datum

Recall that we have shown the stability bound

$$\|u(\cdot,t)\|_{L^2(\mathbb{R})} \leq \|u_0\|_{L^2(\mathbb{R})} \qquad \forall t \in (0,\infty)$$

for the soln u to the IVP $\partial_t u = \partial^2_{xx} u$ in $\mathbb{R} \times (0,\infty)$ with $u(\cdot,0) = u_0$.
This implies stability of the soln w.r.t. perturbation in initial datum:

Let $u, \tilde{u}$ be solutions to

$$\begin{aligned} \partial_t u &= \partial^2_{xx} u \quad \text{in } \mathbb{R} \times (0,\infty), \qquad u(\cdot,0) = u_0, \\ \partial_t \tilde{u} &= \partial^2_{xx} \tilde{u} \quad \text{in } \mathbb{R} \times (0,\infty), \qquad \tilde{u}(\cdot,0) = \tilde{u}_0, \end{aligned}$$

where $u_0, \tilde{u}_0 \in L^2(\mathbb{R})$ are given initial data. Then, $w := u - \tilde{u}$ solves

$$\partial_t w = \partial^2_{xx} w \quad \text{in } \mathbb{R} \times (0,\infty), \qquad w(\cdot,0) = u_0 - \tilde{u}_0.$$

By the stability bound, we have that

$$\|u(\cdot,t) - \tilde{u}(\cdot,t)\|_{L^2(\mathbb{R})} = \|w(\cdot,t)\|_{L^2(\mathbb{R})} \leq \|u_0 - \tilde{u}_0\|_{L^2(\mathbb{R})} \qquad \forall t \in (0,\infty).$$

$\Longrightarrow$ small perturbations in u_0 in the $L^2(\mathbb{R})$-norm result in small perturbations in corresponding soln $u(\cdot,t)$ in the $L^2(\mathbb{R})$-norm for all $t > 0$.
$\Longrightarrow$ Important property which we try to mimic with FD approximation

8.2 FD approximation of the heat equation

Explicit Euler scheme for heat eqn (IVP)

Goal: approximate the soln u to the IVP

$$\begin{aligned} \partial_t u(x,t) &= \partial^2_{xx} u(x,t), \qquad && (x,t) \in \mathbb{R} \times (0,\infty), \\ u(x,0) &= u_0(x), && x \in \mathbb{R}. \end{aligned}$$

Computational domain: $\mathbb{R} \times [0,T]$, where $T > 0$ is a given final time.

Step 1: Define the mesh: Choose $\Delta x > 0$, $M \in \mathbb{N}$, and set $\Delta t := \frac{T}{M}$. Writing $x_j := j\Delta x$ and $t_m := m\Delta t$, we take the mesh

$$\{(x_j, t_m) \mid j \in \mathbb{Z}, m \in \{0, \dots, M\}\} \subset \mathbb{R} \times [0,T].$$

Step 2: Approximate derivatives appearing in the PDE at the mesh pts:

$$\partial_t u(x_j,t_m) \approx \frac{u(x_j,t_{m+1}) - u(x_j,t_m)}{\Delta t}, \quad \partial^2_{xx} u(x_j,t_m) \approx \frac{u(x_{j+1},t_m) - 2u(x_j,t_m) + u(x_{j-1},t_m)}{(\Delta x)^2}.$$

$\Longrightarrow$ This gives the **explicit Euler scheme**:

$$\begin{aligned} \frac{U_j^{m+1} - U_j^m}{\Delta t} &= \frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{(\Delta x)^2}, \quad && j \in \mathbb{Z}, \quad m \in \{0, \dots, M-1\}, \\ U_j^0 &= u_0(x_j), && j \in \mathbb{Z}. \end{aligned}$$

(The value U_j^m is our approximation to $u(x_j, t_m)$.)

The explicit Euler scheme

$$\begin{aligned} \frac{U_j^{m+1} - U_j^m}{\Delta t} &= \frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{(\Delta x)^2}, \quad && j \in \mathbb{Z}, \quad m \in \{0, \dots, M-1\}, \\ U_j^0 &= u_0(x_j), && j \in \mathbb{Z} \end{aligned}$$

can equivalently be written as

$$\begin{aligned} U_j^{m+1} &= U_j^m + \mu(U_{j+1}^m - 2U_j^m + U_{j-1}^m), \quad && j \in \mathbb{Z}, \quad m \in \{0, \dots, M-1\}, \\ U_j^0 &= u_0(x_j), && j \in \mathbb{Z}, \end{aligned}$$

where $\mu > 0$ is the so-called **CFL number** (Courant–Friedrichs–Lewy)

$$\mu := \frac{\Delta t}{(\Delta x)^2}.$$

$\Longrightarrow$ The values U_j^{m+1} for time level $m+1$ can be explicitly calculated, for all $j \in \mathbb{Z}$, from the values U_{j+1}^m, U_j^m, U_{j-1}^m from time level m.

Implicit Euler scheme for heat eqn (IVP)

Goal: approximate the soln u to the IVP

$$\begin{aligned} \partial_t u(x,t) &= \partial^2_{xx} u(x,t), \qquad && (x,t) \in \mathbb{R} \times (0,\infty), \\ u(x,0) &= u_0(x), && x \in \mathbb{R}. \end{aligned}$$

Computational domain: $\mathbb{R} \times [0,T]$, where $T > 0$ is a given final time.

Step 1: Define the mesh: Choose $\Delta x > 0$, $M \in \mathbb{N}$, and set $\Delta t := \frac{T}{M}$. Writing $x_j := j\Delta x$ and $t_m := m\Delta t$, we take the mesh

$$\{(x_j, t_m) \mid j \in \mathbb{Z}, m \in \{0,\ldots,M\}\} \subset \mathbb{R} \times [0,T].$$

Step 2: Approximate derivatives appearing in the PDE at the mesh pts:

$$\partial_t u(x_j,t_m) \approx \frac{u(x_j,t_m) - u(x_j,t_{m-1})}{\Delta t}, \quad \partial^2_{xx} u(x_j,t_m) \approx \frac{u(x_{j+1},t_m) - 2u(x_j,t_m) + u(x_{j-1},t_m)}{(\Delta x)^2}.$$

$\Longrightarrow$ This gives the **implicit Euler scheme**:

$$\begin{aligned} \frac{U_j^m - U_j^{m-1}}{\Delta t} &= \frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{(\Delta x)^2}, \quad && j \in \mathbb{Z}, \quad m \in \{1,\ldots,M\}, \\ U_j^0 &= u_0(x_j), && j \in \mathbb{Z}. \end{aligned}$$

(The value U_j^m is our approximation to $u(x_j,t_m)$.)

θ-scheme for heat eqn (IVP)

The explicit and implicit Euler schemes are special cases of a more general one-parameter family of numerical methods for heat eqn, called θ-**scheme**:

$$\begin{aligned} \frac{U_j^{m+1}-U_j^m}{\Delta t} &= (1-\theta)\,\frac{U_{j+1}^m-2U_j^m+U_{j-1}^m}{(\Delta x)^2}+\theta\,\frac{U_{j+1}^{m+1}-2U_j^{m+1}+U_{j-1}^{m+1}}{(\Delta x)^2}, && j\in\mathbb{Z},\ m\in\{0,\dots,M-1\},\\ U_j^0 &= u_0(x_j), && j\in\mathbb{Z}, \end{aligned}$$

where $\theta \in [0,1]$ is a parameter.

Important special cases:

- $\theta = 0$: Explicit Euler scheme
- $\theta = \frac{1}{2}$: **Crank–Nicolson scheme**
- $\theta = 1$: Implicit Euler scheme

Consistency error of the θ-scheme: For $j \in \mathbb{Z}$, $m \in \{0, \dots, M-1\}$,

$$T_j^m := \frac{u_j^{m+1}-u_j^m}{\Delta t} - (1-\theta)\,\frac{u_{j+1}^m-2u_j^m+u_{j-1}^m}{(\Delta x)^2} - \theta\,\frac{u_{j+1}^{m+1}-2u_j^{m+1}+u_{j-1}^{m+1}}{(\Delta x)^2},$$

where we write $u_j^m := u(x_j, t_m)$ with u being the true solution.

Accuracy of the θ-scheme

Let us expand the consistency error

$$T_j^m := \underbrace{\frac{u_j^{m+1} - u_j^m}{\Delta t}}_{=:A_j^m} - (1-\theta)\underbrace{\frac{u_{j+1}^m - 2u_j^m + u_{j-1}^m}{(\Delta x)^2}}_{=:B_j^m} - \theta\underbrace{\frac{u_{j+1}^{m+1} - 2u_j^{m+1} + u_{j-1}^{m+1}}{(\Delta x)^2}}_{=:C_j^m},$$

using Taylor, around the point $(x_j, t_{m+\frac{1}{2}}) := (x_j, t_m + \frac{\Delta t}{2})$.

1) Taylor the term A_j^m: We have that

$$\begin{aligned}(\Delta t)A_j^m &= u\left(x_j, t_{m+\frac{1}{2}} + \frac{\Delta t}{2}\right) - u\left(x_j, t_{m+\frac{1}{2}} - \frac{\Delta t}{2}\right) \\ &= (\Delta t)u_t(x_j, t_{m+\frac{1}{2}}) + \frac{(\Delta t)^3}{24}u_{ttt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta t)^5).\end{aligned}$$

Therefore, we have that

$$A_j^m = u_t(x_j, t_{m+\frac{1}{2}}) + \frac{(\Delta t)^2}{24}u_{ttt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta t)^4).$$

2) Taylor the term B_j^m: We have that

$$
\begin{aligned}
B_j^m &= \frac{u_{j+1}^m - 2u_j^m + u_{j-1}^m}{(\Delta x)^2} \\
&= \frac{u(x_{j+1}, t_{m+\frac{1}{2}} - \frac{\Delta t}{2}) - 2u(x_j, t_{m+\frac{1}{2}} - \frac{\Delta t}{2}) + u(x_{j-1}, t_{m+\frac{1}{2}} - \frac{\Delta t}{2})}{(\Delta x)^2} \\
&= \frac{u(x_j + \Delta x, t_{m+\frac{1}{2}}) - 2u(x_j, t_{m+\frac{1}{2}}) + u(x_j - \Delta x, t_{m+\frac{1}{2}})}{(\Delta x)^2} \\
&\quad - \frac{\Delta t}{2} \frac{u_t(x_j + \Delta x, t_{m+\frac{1}{2}}) - 2u_t(x_j, t_{m+\frac{1}{2}}) + u_t(x_j - \Delta x, t_{m+\frac{1}{2}})}{(\Delta x)^2} \\
&\quad + \frac{(\Delta t)^2}{8} \frac{u_{tt}(x_j + \Delta x, t_{m+\frac{1}{2}}) - 2u_{tt}(x_j, t_{m+\frac{1}{2}}) + u_{tt}(x_j - \Delta x, t_{m+\frac{1}{2}})}{(\Delta x)^2} \\
&\quad + \cdots \\
&= \left[u_{xx}(x_j, t_{m+\frac{1}{2}}) + \frac{(\Delta x)^2}{12} u_{xxxx}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4) \right] \\
&\quad - \frac{\Delta t}{2} \left[u_{xxt}(x_j, t_{m+\frac{1}{2}}) + \frac{(\Delta x)^2}{12} u_{xxxxt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4) \right] \\
&\quad + \frac{(\Delta t)^2}{8} \left[u_{xxtt}(x_j, t_{m+\frac{1}{2}}) + \frac{(\Delta x)^2}{12} u_{xxxxtt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4) \right] \\
&\quad + \mathcal{O}((\Delta t)^3).
\end{aligned}
$$

3) Taylor the term C_j^m: We have that

$$
\begin{aligned}
C_j^m &= \frac{u_{j+1}^{m+1} - 2u_j^{m+1} + u_{j-1}^{m+1}}{(\Delta x)^2} \\
&= \frac{u(x_{j+1}, t_{m+\frac{1}{2}} + \frac{\Delta t}{2}) - 2u(x_j, t_{m+\frac{1}{2}} + \frac{\Delta t}{2}) + u(x_{j-1}, t_{m+\frac{1}{2}} + \frac{\Delta t}{2})}{(\Delta x)^2} \\
&= \frac{u(x_j + \Delta x, t_{m+\frac{1}{2}}) - 2u(x_j, t_{m+\frac{1}{2}}) + u(x_j - \Delta x, t_{m+\frac{1}{2}})}{(\Delta x)^2} \\
&\quad + \frac{\Delta t}{2} \frac{u_t(x_j + \Delta x, t_{m+\frac{1}{2}}) - 2u_t(x_j, t_{m+\frac{1}{2}}) + u_t(x_j - \Delta x, t_{m+\frac{1}{2}})}{(\Delta x)^2} \\
&\quad + \frac{(\Delta t)^2}{8} \frac{u_{tt}(x_j + \Delta x, t_{m+\frac{1}{2}}) - 2u_{tt}(x_j, t_{m+\frac{1}{2}}) + u_{tt}(x_j - \Delta x, t_{m+\frac{1}{2}})}{(\Delta x)^2} \\
&\quad + \cdots \\
&= \left[u_{xx}(x_j, t_{m+\frac{1}{2}}) + \frac{(\Delta x)^2}{12} u_{xxxx}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4) \right] \\
&\quad + \frac{\Delta t}{2} \left[u_{xxt}(x_j, t_{m+\frac{1}{2}}) + \frac{(\Delta x)^2}{12} u_{xxxxt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4) \right] \\
&\quad + \frac{(\Delta t)^2}{8} \left[u_{xxtt}(x_j, t_{m+\frac{1}{2}}) + \frac{(\Delta x)^2}{12} u_{xxxxtt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4) \right] \\
&\quad + \mathcal{O}((\Delta t)^3).
\end{aligned}
$$

4) Altogether: We find that

$$
\begin{aligned}
T_j^m &= A_j^m - (1-\theta)B_j^m - \theta C_j^m \\
&= \left[u_t(x_j, t_{m+\frac{1}{2}}) - u_{xx}(x_j, t_{m+\frac{1}{2}}) - \frac{(\Delta x)^2}{12} u_{xxxx}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4)\right] \\
&\quad + (2\theta - 1)\frac{\Delta t}{2}\left[-u_{xxt}(x_j, t_{m+\frac{1}{2}}) - \frac{(\Delta x)^2}{12} u_{xxxxt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4)\right] \\
&\quad + \frac{(\Delta t)^2}{8}\left[\frac{1}{3}u_{ttt}(x_j, t_{m+\frac{1}{2}}) - u_{xxtt}(x_j, t_{m+\frac{1}{2}}) - \frac{(\Delta x)^2}{12} u_{xxxxtt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4)\right] \\
&\quad + \mathcal{O}((\Delta t)^3) \\
&= \left[-\frac{(\Delta x)^2}{12} u_{tt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4)\right] \\
&\quad + (2\theta - 1)\frac{\Delta t}{2}\left[-u_{tt}(x_j, t_{m+\frac{1}{2}}) - \frac{(\Delta x)^2}{12} u_{ttt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4)\right] \\
&\quad + \frac{(\Delta t)^2}{8}\left[-\frac{2}{3}u_{ttt}(x_j, t_{m+\frac{1}{2}}) - \frac{(\Delta x)^2}{12} u_{tttt}(x_j, t_{m+\frac{1}{2}}) + \mathcal{O}((\Delta x)^4)\right] \\
&\quad + \mathcal{O}((\Delta t)^3),
\end{aligned}
$$

where we have used that $u_t = u_{xx}$ (and hence, also $u_{tt} = u_{xxxx}, \dots$).

We conclude that

$$
T_j^m = \begin{cases} \mathcal{O}((\Delta x)^2 + (\Delta t)^2) & \text{if } \theta = \frac{1}{2}, \\ \mathcal{O}((\Delta x)^2 + \Delta t) & \text{if } \theta \neq \frac{1}{2}. \end{cases}
$$

Fully-discrete vs spatially semi-discrete approximation

- Numerical methods such as the θ-scheme are called **fully-discrete approximations** (we discretize both spatial and time derivatives).

- Alternative: approximate only the spatial partial derivative in the heat eqn, resulting in the following IVP for a system of ODEs:

$$\begin{aligned} \frac{\mathrm{d}U_j(t)}{\mathrm{d}t} &= \frac{U_{j+1}(t) - 2U_j(t) + U_{j-1}(t)}{(\Delta x)^2}, \qquad && j \in \mathbb{Z}, \\ U_j(0) &= u_0(x_j), && j \in \mathbb{Z}. \end{aligned}$$

Here, the function U_j is an approximation to $t \mapsto u(x_j, t)$. This is called a **spatially semi-discrete approximation**, because no discretization with respect to the time variable has taken place.

Rk: Because no discretization in time was performed in the first place, this approach is usually referred to as the **method of lines**.

8.3 Practical stability of FD schemes

Practical stability of FD schemes

Recall: the true soln to the IVP for heat eqn satisfies the stability bound

$$\|u(\cdot,t)\|_{L^2(\mathbb{R})} \leq \|u_0\|_{L^2(\mathbb{R})} \qquad \forall t \in (0,\infty).$$

In order to be able to replicate this stability property at the discrete level, we require an appropriate notion of stability.

Definition (Practical stability of FD schemes)

We say that a FD scheme for the IVP for the heat eqn is **practically stable (in the ℓ^2 norm)** iff for the values $\{U_j^m\}_{j\in\mathbb{Z},m\in\{0,\ldots,M-1\}}$ obtained from the FD scheme there holds

$$\|U^m\|_{\ell^2} \leq \|U^0\|_{\ell^2} \qquad \forall m \in \{1,\ldots,M\},$$

where $U^m := (\ldots, U_{-2}^m, U_{-1}^m, U_0^m, U_1^m, U_2^m, \ldots)$ and

$$\|U^m\|_{\ell^2} := \sqrt{\Delta x \sum_{j=-\infty}^{\infty} |U_j^m|^2}.$$

Key tool for stability analysis: The semidiscrete FT

Definition (SFT and ISFT)

(i) The **semidiscrete Fourier transform (SFT)** of a function U defined on the infinite mesh with mesh-points $x_j = j\Delta x$, $j \in \mathbb{Z}$, is defined by

$$\hat{U}(k) := \Delta x \sum_{j=-\infty}^{\infty} U_j \, e^{-ikx_j}, \qquad k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right],$$

where U_j denotes the value of U at the mesh point x_j.

(ii) For a fct $\hat{U} : [-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}] \to \mathbb{C}$, its **inverse semidiscrete Fourier transform (ISFT)** is the function U defined on the infinite mesh with mesh-points $x_j = j\Delta x$, $j \in \mathbb{Z}$, with the value of U at the mesh point x_j given by

$$U_j := \frac{1}{2\pi} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \hat{U}(k) \, e^{ikx_j} \, \mathrm{d}k, \qquad j \in \mathbb{Z}.$$

Discrete Parseval's identity

Let us recall Parseval's identity: If $u \in L^2(\mathbb{R})$, then its FT $\hat{u} \in L^2(\mathbb{R})$ and we have $\|\hat{u}\|^2_{L^2(\mathbb{R})} = 2\pi\|u\|^2_{L^2(\mathbb{R})}$.

We have a discrete analogue of this result for a mesh fct U and its SFT $\hat{U}$:

Lemma (Discrete Parseval's identity)

Let U be a function defined on the infinite mesh with mesh-points $x_j = j\Delta x$, $j \in \mathbb{Z}$, and let $\hat{U}$ be its SFT. If $\|U\|_{\ell^2} < \infty$, then $\hat{U} \in L^2((-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}))$ and there holds

$$\|\hat{U}\|^2_{L^2((-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}))} = 2\pi\|U\|^2_{\ell^2}.$$

(Recall: $\|U\|^2_{\ell^2} := \Delta x \sum_{j=-\infty}^{\infty} |U_j|^2$ and $\|\hat{U}\|^2_{L^2((-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}))} := \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} |\hat{U}(k)|^2 \, \mathrm{d}k$)

Proof: Exercise. □

Example: Stability analysis of the explicit Euler scheme

Explicit Euler scheme for IVP for heat eqn:

$$\begin{aligned} \frac{U_j^{m+1} - U_j^m}{\Delta t} &= \frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{(\Delta x)^2}, \quad && j \in \mathbb{Z}, \quad m \in \{0, \dots, M-1\}, \\ U_j^0 &= u_0(x_j), && j \in \mathbb{Z}. \end{aligned}$$

By inserting $U_j^m = \frac{1}{2\pi} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{ikj\Delta x} \hat{U}^m(k)\, \mathrm{d}k$, we deduce that

$$\begin{aligned} \frac{1}{2\pi}\int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{ikj\Delta x}\, \frac{\hat{U}^{m+1}(k) - \hat{U}^m(k)}{\Delta t}\, \mathrm{d}k &= \frac{1}{2\pi}\int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \frac{e^{ik(j+1)\Delta x} - 2e^{ikj\Delta x} + e^{ik(j-1)\Delta x}}{(\Delta x)^2} \hat{U}^m(k)\, \mathrm{d}k \\ &= \frac{1}{2\pi}\int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{ikj\Delta x}\, \frac{e^{ik\Delta x} - 2 + e^{-ik\Delta x}}{(\Delta x)^2}\, \hat{U}^m(k)\, \mathrm{d}k. \end{aligned}$$

$\Longrightarrow$ the integrands are identically equal (by injectivity of SFT/ISFT).
Thus,

$$\frac{\hat{U}^{m+1}(k) - \hat{U}^m(k)}{\Delta t} = \frac{e^{ik\Delta x} - 2 + e^{-ik\Delta x}}{(\Delta x)^2}\, \hat{U}^m(k) \qquad \forall k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right].$$

$\Longrightarrow$ We have that

$$\hat{U}^{m+1}(k) = \hat{U}^m(k) + \mu(e^{ik\Delta x} - 2 + e^{-ik\Delta x})\hat{U}^m(k) \qquad \forall k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right],$$

where $\mu := \frac{\Delta t}{(\Delta x)^2}$ is the CFL number. Equivalently,

$$\hat{U}^{m+1}(k) = \lambda(k)\hat{U}^m(k), \qquad \lambda(k) := 1 + \mu(e^{ik\Delta x} - 2 + e^{-ik\Delta x}).$$

We call the function $\lambda = \lambda(k)$ the **amplification factor**. Let us define $\Lambda := \max_{k\in[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}]} |\lambda(k)|$. Then, by discrete Parseval's identity,

$$2\pi\|U^{m+1}\|_{\ell^2}^2 = \|\hat{U}^{m+1}\|_{L^2((-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}))}^2 = \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} |\lambda(k)\hat{U}^m(k)|^2 \,\mathrm{d}k$$

$$\leq \Lambda^2 \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} |\hat{U}^m(k)|^2 \,\mathrm{d}k = \Lambda^2 \|\hat{U}^m\|_{L^2((-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}))}^2 = 2\pi\Lambda^2\|U^m\|_{\ell^2}^2,$$

i.e., we find that

$$\|U^m\|_{\ell^2} \leq \Lambda\|U^{m-1}\|_{\ell^2} \leq \cdots \leq \Lambda^m\|U^0\|_{\ell^2} \qquad \forall m \in \{1, \ldots, M\}.$$

$\Longrightarrow$ For practical stability we demand that $\Lambda \leq 1$.

$\Longrightarrow$ For practical stability, we demand that (note $\mu = \frac{\Delta t}{(\Delta x)^2} > 0$)

$$\max_{k\in[-\frac{\pi}{\Delta x},\frac{\pi}{\Delta x}]} |1 + \mu(e^{ik\Delta x} - 2 + e^{-ik\Delta x})| \le 1$$

$$\Longleftrightarrow \max_{k\in[-\frac{\pi}{\Delta x},\frac{\pi}{\Delta x}]} |1 + 2\mu(\cos(k\Delta x) - 1)| \le 1$$

$$\Longleftrightarrow \max_{k\in[-\frac{\pi}{\Delta x},\frac{\pi}{\Delta x}]} \left|1 - 4\mu \sin^2\left(\frac{k\Delta x}{2}\right)\right| \le 1$$

$$\Longleftrightarrow \forall k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right] : \quad -1 \le 1 - 4\mu \sin^2\left(\frac{k\Delta x}{2}\right) \le 1$$

$$\Longleftrightarrow \forall k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right] : \quad \sin^2\left(\frac{k\Delta x}{2}\right) \le \frac{1}{2\mu}$$

$$\Longleftrightarrow \max_{k\in[-\frac{\pi}{\Delta x},\frac{\pi}{\Delta x}]} \sin^2\left(\frac{k\Delta x}{2}\right) = 1 \le \frac{1}{2\mu}$$

$$\Longleftrightarrow \mu \le \frac{1}{2}.$$

Hence, the explicit Euler scheme is **conditionally practically stable**, with the condition for stability being that $\mu = \frac{\Delta t}{(\Delta x)^2} \le \frac{1}{2}$.

$\Longrightarrow$ We must choose $\Delta x, \Delta t$ s.t. $\Delta t \le \frac{1}{2}(\Delta x)^2$ to have practical stability.

Example: Stability analysis of the implicit Euler scheme

Implicit Euler scheme for IVP for heat eqn:

$$\frac{U_j^{m+1} - U_j^m}{\Delta t} = \frac{U_{j+1}^{m+1} - 2U_j^{m+1} + U_{j-1}^{m+1}}{(\Delta x)^2}, \quad j \in \mathbb{Z}, \quad m \in \{0, \dots, M-1\},$$
$$U_j^0 = u_0(x_j), \qquad j \in \mathbb{Z}.$$

By inserting $U_j^m = \frac{1}{2\pi} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{ikj\Delta x} \hat{U}^m(k)\, \mathrm{d}k$, we deduce that

$$\frac{1}{2\pi}\int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{ikj\Delta x} \frac{\hat{U}^{m+1}(k) - \hat{U}^m(k)}{\Delta t}\, \mathrm{d}k = \frac{1}{2\pi}\int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \frac{e^{ik(j+1)\Delta x} - 2e^{ikj\Delta x} + e^{ik(j-1)\Delta x}}{(\Delta x)^2} \hat{U}^{m+1}(k) \mathrm{d}k$$
$$= \frac{1}{2\pi}\int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} e^{ikj\Delta x} \frac{e^{ik\Delta x} - 2 + e^{-ik\Delta x}}{(\Delta x)^2} \hat{U}^{m+1}(k)\, \mathrm{d}k.$$

$\Longrightarrow$ the integrands are identically equal (by injectivity of SFT/ISFT).
Thus,

$$\frac{\hat{U}^{m+1}(k) - \hat{U}^m(k)}{\Delta t} = \frac{e^{ik\Delta x} - 2 + e^{-ik\Delta x}}{(\Delta x)^2} \hat{U}^{m+1}(k) \qquad \forall k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right].$$

$\Longrightarrow$ We have that

$$\hat{U}^{m+1}(k) = \hat{U}^m(k) + \mu(e^{ik\Delta x} - 2 + e^{-ik\Delta x})\hat{U}^{m+1}(k) \quad \forall k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right],$$

where $\mu := \frac{\Delta t}{(\Delta x)^2}$ is the CFL number. Equivalently,

$$\hat{U}^{m+1}(k) = \lambda(k)\hat{U}^m(k), \qquad \lambda(k) := \frac{1}{1 - \mu(e^{ik\Delta x} - 2 + e^{-ik\Delta x})}.$$

Note that we have for the amplification factor that

$$\begin{aligned}
\max_{k\in[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}]} |\lambda(k)| &= \max_{k\in[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}]} \left| \frac{1}{1 + 4\mu \sin^2\left(\frac{k\Delta x}{2}\right)} \right| \\
&= \max_{k\in[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}]} \frac{1}{1 + 4\mu \sin^2\left(\frac{k\Delta x}{2}\right)} \\
&\leq 1
\end{aligned}$$

for any $\mu > 0$.

$\Longrightarrow$ The implicit Euler scheme is **unconditionally practically stable**, meaning that $\|U^m\|_{\ell^2} \leq \|U^0\|_{\ell^2}$ $\forall m \in \{1, \dots, M\}$ holds without any restrictions on Δx and Δt.

Stability analysis of the θ-scheme

θ-scheme for IVP for heat eqn:

$$\frac{U_j^{m+1}-U_j^m}{\Delta t} = (1-\theta)\,\frac{U_{j+1}^m - 2U_j^m + U_{j-1}^m}{(\Delta x)^2} + \theta\,\frac{U_{j+1}^{m+1} - 2U_j^{m+1} + U_{j-1}^{m+1}}{(\Delta x)^2}, \; j \in \mathbb{Z},\; m \in \{0,\dots,M-1\},$$

$$U_j^0 = u_0(x_j), \qquad j \in \mathbb{Z}.$$

On the problem sheets, you are going to discover the following:

- If $\theta \in [0, \frac{1}{2})$, then the θ-scheme is conditionally practically stable, with the stability condition being that $\mu = \frac{\Delta t}{(\Delta x)^2} \leq \frac{1}{2(1-2\theta)}$.
- If $\theta \in [\frac{1}{2}, 1]$, then the θ-scheme is unconditionally practically stable.

Rk: In particular, the Crank–Nicolson scheme ($\theta = \frac{1}{2}$) is unconditionally practically stable.

8.4 Von Neumann stability

Von Neumann stability

Let us introduce a less demanding notion of stability:

Definition (von Neumann stability)

We say that a FD scheme for the IVP for the heat eqn on $\mathbb{R} \times [0, T]$ is **von Neumann stable** in the ℓ^2-norm, if $\exists$ a constant $C = C(T) > 0$ s.t.

$$\|U^m\|_{\ell^2} \le C \|U^0\|_{\ell^2} \qquad \forall m \in \left\{1, \ldots, M = \frac{T}{\Delta t}\right\}.$$

(Recall: $\|U^m\|_{\ell^2}^2 := \Delta x \sum_{j=-\infty}^{\infty} |U_j^m|^2$.)

Rk: Practical stability implies von Neumann stability with $C = 1$.

Rk: When C depends on T, then typically $C(T) \to \infty$ as $T \to \infty$.

A simple way for verifying von Neumann stability

Lemma (Verifying von Neumann stability in practice)

Suppose that the SFT of the soln $\{U_j^m\}_{j\in\mathbb{Z}, m\in\{0,\dots,M=\frac{T}{\Delta t}\}}$ *of a FD scheme for the IVP for the heat equation satisfies* $\hat{U}^{m+1} = \lambda \hat{U}^m$ *with some amplification factor* $\lambda = \lambda(k)$*, and suppose* $\exists$ *a constant* $C_0 \geq 0$ *s.t.*

$$|\lambda(k)| \leq 1 + C_0 \Delta t \qquad \forall\, k \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right].$$

Then, the scheme is von Neumann stable.
In particular, if $C_0 = 0$*, then the scheme is practically stable.*

Proof: Set $\Lambda := \max_{k\in[-\frac{\pi}{\Delta x},\frac{\pi}{\Delta x}]} |\lambda(k)|$. We have seen before that

$$\|U^{m+1}\|_{\ell^2} \leq \Lambda \|U^m\|_{\ell^2} \qquad \forall m \in \{0,\dots,M-1\}.$$

Hence, using that $\Lambda \leq 1 + C_0\Delta t$, we find for any $m \in \{1,\dots,M\}$ that

$$\|U^m\|_{\ell^2} \leq (1+C_0\Delta t)\|U^{m-1}\|_{\ell^2} \leq \cdots \leq (1+C_0\Delta t)^m \|U^0\|_{\ell^2} \leq e^{mC_0\Delta t}\|U^0\|_{\ell^2}.$$

As $m\Delta t \leq T\ \forall m \in \{1,\dots,M\}$, we have v.N. stab. with $C := e^{C_0 T}$. □

8.5 Initial-boundary-value problems for parabolic problems

The Dirichlet IBVP for the heat equation

For fixed $a, b \in \mathbb{R}$ with $a < b$, and $T > 0$, we consider the heat equation

$$\partial_t u(x,t) = \partial_{xx}^2 u(x,t) \quad \text{for } (x,t) \in (a,b) \times (0,T]$$

subject to the initial condition

$$u(x,0) = u_0(x) \quad \text{for } x \in [a,b]$$

and the Dirichlet boundary condition

$$\begin{aligned} u(a,t) &= A(t) \quad \text{for } t \in (0,T], \\ u(b,t) &= B(t) \quad \text{for } t \in (0,T]. \end{aligned}$$

Here, $u_0 : [a,b] \to \mathbb{R}$ and $A, B : [0,T] \to \mathbb{R}$ are given. We assume that the b.c. is compatible with the i.c., that is, $A(0) = u_0(a)$, $B(0) = u_0(b)$.

θ-scheme for the Dirichlet IBVP for the heat eqn

Recall the IBVP:

$$\partial_t u(x,t) = \partial^2_{xx} u(x,t) \quad \text{for } (x,t) \in (a,b) \times (0,T],$$
$$u(x,0) = u_0(x) \quad \text{for } x \in [a,b],$$
$$u(a,t) = A(t) \quad \text{for } t \in (0,T], \quad u(b,t) = B(t) \quad \text{for } t \in (0,T].$$

Mesh: For $J, M \in \mathbb{N}$ fixed, let $\Delta x := \frac{b-a}{J}$ and $\Delta t := \frac{T}{M}$. Take the mesh

$$\{(x_j, t_m) := (a + j\Delta x, m\Delta t) \,|\, j \in \{0, \dots, J\}, m \in \{0, \dots, M\}\}.$$

The FD scheme: The θ-scheme for the IBVP is the following:

$$\frac{U_j^{m+1} - U_j^m}{\Delta t} = (1-\theta)\frac{\delta^2 U_j^m}{(\Delta x)^2} + \theta \frac{\delta^2 U_j^{m+1}}{(\Delta x)^2}, \quad j \in \{1, \dots, J-1\}, m \in \{0, \dots, M-1\}$$
$$U_j^0 = u_0(x_j), \qquad j \in \{0, \dots, J\},$$
$$U_0^{m+1} = A(t_{m+1}), \quad U_J^{m+1} = B(t_{m+1}), \qquad m \in \{0, \dots, M-1\}$$

where $\theta \in [0,1]$ is a parameter. We have written

$$\delta^2 U_j^m := U_{j+1}^m - 2U_j^m + U_{j-1}^m, \qquad \delta^2 U_j^{m+1} := U_{j+1}^{m+1} - 2U_j^{m+1} + U_{j-1}^{m+1}.$$

θ-scheme as a linear system (time level $m \to m+1$)

With $\mu := \frac{\Delta t}{(\Delta x)^2}$, the θ-scheme can be written as

$$
\begin{aligned}
U_j^{m+1} - \theta\mu\delta^2 U_j^{m+1} &= U_j^m + (1-\theta)\mu\delta^2 U_j^m, \ j \in \{1,\dots,J-1\}, m \in \{0,\dots,M-1\},\\
U_j^0 &= u_0(x_j), \qquad j \in \{0,\dots,J\},\\
U_0^{m+1} &= A(t_{m+1}), \ \ U_J^{m+1} = B(t_{m+1}), \qquad m \in \{0,\dots,M-1\}.
\end{aligned}
$$

Let $I := I_{J-1}$ be the identity matrix in $\mathbb{R}^{(J-1)\times(J-1)}$, and let

$$
\mathscr{A} := \begin{bmatrix} -2 & 1 & & & \mathbf{0} \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ \mathbf{0} & & & 1 & -2 \end{bmatrix} \in \mathbb{R}^{(J-1)\times(J-1)}.
$$

Writing $\mathbf{U}^m := (U_1^m, \dots, U_{J-1}^m)^{\mathrm{T}}$, $\mathbf{F}^m := (A(t_m), 0, \dots, 0, B(t_m))^{\mathrm{T}} \in \mathbb{R}^{J-1}$, the θ-scheme can be written as

$$
\begin{aligned}
(I - \theta\mu\mathscr{A})\mathbf{U}^{m+1} &= (I + (1-\theta)\mu\mathscr{A})\mathbf{U}^m + \theta\mu\mathbf{F}^{m+1} + (1-\theta)\mu\mathbf{F}^m, \ \ m \in \{0,\dots,M-1\}\\
\mathbf{U}^0 &= (u_0(x_1), \dots, u_0(x_{J-1}))^{\mathrm{T}},
\end{aligned}
$$

and $U_0^{m+1} = A(t_{m+1})$, $U_J^{m+1} = B(t_{m+1})$ for $m \in \{0,\dots,M-1\}$.

$\Longrightarrow$ Obtain $\mathbf{U}^{m+1}$ from $\mathbf{U}^m$ by solving linear system with matrix $I - \theta\mu\mathscr{A}$.

Discrete maximum/minimum principle for the θ-scheme

Theorem (Discrete maximum/minimum principle for the θ-scheme)

Consider the θ-scheme for the Dirichlet IBVP for the heat eqn, with $\theta \in [0,1]$. Suppose that

$$(1-\theta)\mu \leq \frac{1}{2}, \quad \text{where } \mu := \frac{\Delta t}{(\Delta x)^2}.$$

Then, for the numerical approximations $\{U_j^m\}_{j\in\{0,\ldots,J\};m\in\{0,\ldots,M\}}$ we have

$$\min\{U_{\min}^0, U_0^{\min}, U_J^{\min}\} \leq U_j^m \leq \max\{U_{\max}^0, U_0^{\max}, U_J^{\max}\}$$

for any $j \in \{0,\ldots,J\}$ and $m \in \{0,\ldots,M\}$, where

$$U_{\min}^0 := \min\{U_0^0, U_1^0, \ldots, U_J^0\}, \qquad U_{\max}^0 := \max\{U_0^0, U_1^0, \ldots, U_J^0\},$$
$$U_0^{\min} := \min\{U_0^0, U_0^1, \ldots, U_0^M\}, \qquad U_0^{\max} := \max\{U_0^0, U_0^1, \ldots, U_0^M\},$$
$$U_J^{\min} := \min\{U_J^0, U_J^1, \ldots, U_J^M\}, \qquad U_J^{\max} := \max\{U_J^0, U_J^1, \ldots, U_J^M\}.$$

Proof of the discrete maximum principle for the θ-scheme

We prove $U_j^m \leq \max\{U_{\max}^0, U_0^{\max}, U_J^{\max}\}\ \forall j, m$. (The other inequality is proved similarly.) We rewrite the θ-scheme as

$$(1+2\theta\mu)\,U_j^{m+1} = \theta\mu\left(U_{j+1}^{m+1} + U_{j-1}^{m+1}\right) + (1-\theta)\mu\left(U_{j+1}^m + U_{j-1}^m\right) + (1-2(1-\theta)\mu)U_j^m.$$

By hypothesis, $\theta\mu \geq 0$, $(1-\theta)\mu \geq 0$, and $1 - 2(1-\theta)\mu \geq 0$. Suppose U attains its maximum value at (x_{j_0}, t_{m_0+1}) for some $j_0 \in \{1, \ldots, J-1\}$, $m_0 \in \{0, \ldots, M-1\}$.

We define $U^\star := \max\{U_{j_0+1}^{m_0+1}, U_{j_0-1}^{m_0+1}, U_{j_0+1}^{m_0}, U_{j_0-1}^{m_0}, U_{j_0}^{m_0}\}$. Then,

$$\begin{aligned}(1+2\theta\mu)\,U_{j_0}^{m_0+1} &\leq 2\theta\mu U^\star + 2(1-\theta)\mu U^\star + (1-2(1-\theta)\mu)\,U^\star \\ &= (1+2\theta\mu)\,U^\star.\end{aligned}$$

$\Longrightarrow U_{j_0}^{m_0+1} \leq U^\star$. Note that also $U^\star \leq U_{j_0}^{m_0+1}$ and hence, $U_{j_0}^{m_0+1} = U^\star$.
$\Longrightarrow$ The maximum value is also attained at each of the points neighbouring (x_{j_0}, t_{m_0+1}) present in the scheme.

The same argument applies to these neighbouring points, and we can repeat this process until the bdry at $x = a$ or $x = b$ or at $t = 0$ is reached. The maximum is therefore attained at a boundary point.

Rk: A classical solution u to the Dirichlet IBVP for the heat eqn attains its maximum and minimum value on the **parabolic boundary**

$$\begin{aligned}\Gamma_T &:= \{t=0\} \cup \{x=a\} \cup \{x=b\} \\ &:= ([a,b]\times\{0\}) \cup (\{a\}\times[0,T]) \cup (\{b\}\times[0,T]),\end{aligned}$$

i.e., there holds the following **maximum**/**minimum** **principle**:

$$\max_{[a,b]\times[0,T]} u = \max_{\Gamma_T} u,$$

$$\min_{[a,b]\times[0,T]} u = \min_{\Gamma_T} u.$$

We have just proved that our numerical approximation obtained from the θ-scheme satisfies a discrete analogue to this result.

Remark on discrete MP and practical stability of θ-scheme

Recall:

- Condition for discrete maximum/minimum principle:

$$\mu(1-\theta) \leq \frac{1}{2}. \tag{4}$$

- Condition for practical stability when $\theta \in [0, \frac{1}{2})$:

$$\mu(1-2\theta) \leq \frac{1}{2}. \tag{5}$$

- When $\theta \in [\frac{1}{2}, 1]$, unconditionally practically stable.

Some comments:

- When $\theta = 0$ (explicit Euler), then (4) $\Leftrightarrow$ (5); both requiring $\mu \leq \frac{1}{2}$.
- When $\theta \in (0, \frac{1}{2})$, condition (4) is more demanding than (5).
- Crank–Nicolson ($\theta = \frac{1}{2}$) only satisfies the discrete MP when $\mu \leq 1$.
- For $\theta \in [\frac{1}{2}, 1]$, the θ-scheme only satisfies the discrete MP unconditionally when $\theta = 1$ (implicit Euler scheme).

Convergence analysis of the θ-scheme

Suppose $\mu(1-\theta) \leq \frac{1}{2}$. We begin by rewriting the scheme as

$$(1+2\theta\mu)\,U_j^{m+1} = \theta\mu\left(U_{j+1}^{m+1} + U_{j-1}^{m+1}\right) + (1-\theta)\mu\left(U_{j+1}^{m} + U_{j-1}^{m}\right) + (1-2(1-\theta)\mu)U_j^m.$$

Recall that the **consistency error** for the θ-scheme is

$$T_j^m := \frac{u_j^{m+1} - u_j^m}{\Delta t} - (1-\theta)\,\frac{u_{j+1}^m - 2u_j^m + u_{j-1}^m}{(\Delta x)^2} - \theta\,\frac{u_{j+1}^{m+1} - 2u_j^{m+1} + u_{j-1}^{m+1}}{(\Delta x)^2},$$

where $u_j^m := u(x_j, t_m)$, and therefore

$$(1+2\theta\mu)\,u_j^{m+1} = \theta\mu\left(u_{j+1}^{m+1} + u_{j-1}^{m+1}\right) + (1-\theta)\mu\left(u_{j+1}^{m} + u_{j-1}^{m}\right) + (1-2(1-\theta)\mu)u_j^m + (\Delta t)T_j^m.$$

Define the **global error**

$$e_j^m := u(x_j, t_m) - U_j^m.$$

Note $e_0^{m+1} = e_J^{m+1} = e_j^0 = 0\ \forall j \in \{0, \ldots, J\}, m \in \{0, \ldots, M-1\}$, and

$$(1+2\theta\mu)\,e_j^{m+1} = \theta\mu\left(e_{j+1}^{m+1} + e_{j-1}^{m+1}\right) + (1-\theta)\mu\left(e_{j+1}^{m} + e_{j-1}^{m}\right) + (1-2(1-\theta)\mu)e_j^m + (\Delta t)T_j^m$$

$$\leq 2\theta\mu E^{m+1} + 2(1-\theta)\mu E^m + (1-2(1-\theta)\mu)\,E^m + (\Delta t)T^m = 2\theta\mu E^{m+1} + E^m + (\Delta t)T^m,$$

where $E^m := \max\{|e_0^m|, \ldots, |e_J^m|\}$ and $T^m := \max\{|T_0^m|, \ldots, |T_J^m|\}$.

$\Longrightarrow$ We find that

$$(1+2\theta\mu)E^{m+1} \leq 2\theta\mu E^{m+1} + E^m + (\Delta t)T^m \quad \forall m \in \{0,\dots,M-1\}.$$

Hence, $E^{m+1} \leq E^m + (\Delta t)T^m \;\forall m \in \{0,\dots,M-1\}$. As $E^0 = 0$, we have

$$\begin{aligned}
E^m &\leq E^{m-1} + (\Delta t)T^{m-1} \\
&\leq E^{m-2} + (\Delta t)T^{m-2} + (\Delta t)T^{m-1} \\
&\;\vdots \\
&\leq (\Delta t)\left(T^0 + T^1 + \cdots + T^{m-1}\right) \leq m(\Delta t) \max_{i\in\{0,\dots,m-1\}} T^i \leq T \max_{i\in\{0,\dots,M-1\}} T^i
\end{aligned}$$

for any $m \in \{1,\dots,M\}$. It follows that

$$\begin{aligned}
\max_{m\in\{0,\dots,M\}} \max_{j\in\{0,\dots,J\}} |e_j^m| &\leq T \max_{i\in\{0,\dots,M-1\}} T^i \\
&= \begin{cases} \mathcal{O}\left((\Delta x)^2 + (\Delta t)^2\right) & \text{if } \theta = 1/2, \\ \mathcal{O}\left((\Delta x)^2 + \Delta t\right) & \text{if } \theta \neq 1/2, \end{cases}
\end{aligned}$$

where we have used our results of the expansion of the consistency error.

8.6 FD approximation of parabolic equations in two space-dimensions

The Dirichlet IBVP in two space-dimensions

On an open rectangle $\Omega := (a, b) \times (c, d)$ in $\mathbb{R}^2$, we consider the heat eqn

$$\partial_t u(x, y, t) = \partial^2_{xx} u(x, y, t) + \partial^2_{yy} u(x, y, t), \qquad (x, y) \in \Omega, \ \ t \in (0, T],$$

subject to the initial condition

$$u(x, y, 0) = u_0(x, y), \qquad (x, y) \in \overline{\Omega},$$

and the Dirichlet boundary condition

$$u(x, y, t) = B(x, y, t), \qquad (x, y) \in \partial\Omega, \ \ t \in (0, T],$$

where $u_0 : \overline{\Omega} \to \mathbb{R}$ and $B : \partial\Omega \times [0, T] \to \mathbb{R}$ are given fcts. We assume the b.c. is compatible with the i.c., i.e., $B(x, y, 0) = u_0(x, y) \ \forall (x, y) \in \partial\Omega$.

Explicit Euler scheme

Let

$$\delta_x^2 U_{i,j} := U_{i+1,j} - 2U_{i,j} + U_{i-1,j}, \qquad \delta_y^2 U_{i,j} := U_{i,j+1} - 2U_{i,j} + U_{i,j-1}.$$

For $J_x, J_y, M \in \mathbb{N}$ fixed, let $\Delta x := \frac{b-a}{J_x}$, $\Delta y := \frac{d-c}{J_y}$, $\Delta t := \frac{T}{M}$, and define

$$\begin{aligned} x_i &:= a + i\Delta x, & i &\in \{0, \ldots, J_x\}, \\ y_j &:= c + j\Delta y, & j &\in \{0, \ldots, J_y\}, \\ t_m &:= m\Delta t, & m &\in \{0, \ldots, M\}. \end{aligned}$$

The explicit Euler FD approximation is

$$\frac{U_{i,j}^{m+1} - U_{i,j}^m}{\Delta t} = \frac{\delta_x^2 U_{i,j}^m}{(\Delta x)^2} + \frac{\delta_y^2 U_{i,j}^m}{(\Delta y)^2},$$

for $i \in \{1, \ldots, J_x - 1\}$, $j \in \{1, \ldots, J_y - 1\}$, $m \in \{0, \ldots, M-1\}$, with i.c.

$$U_{i,j}^0 = u_0(x_i, y_j), \quad i \in \{0, \ldots, J_x\}, \ \ j \in \{0, \ldots, J_y\},$$

and b.c.

$$U_{i,j}^{m+1} := B(x_i, y_j, t_{m+1}), \ \text{ when } (x_i, y_j) \in \partial\Omega, \ m \in \{0, \ldots, M-1\}.$$

Implicit Euler scheme

Let

$$\delta_x^2 U_{i,j} := U_{i+1,j} - 2U_{i,j} + U_{i-1,j}, \qquad \delta_y^2 U_{i,j} := U_{i,j+1} - 2U_{i,j} + U_{i,j-1}.$$

For $J_x, J_y, M \in \mathbb{N}$ fixed, let $\Delta x := \frac{b-a}{J_x}$, $\Delta y := \frac{d-c}{J_y}$, $\Delta t := \frac{T}{M}$, and define

$$\begin{aligned} x_i &:= a + i\Delta x, && i \in \{0, \dots, J_x\}, \\ y_j &:= c + j\Delta y, && j \in \{0, \dots, J_y\}, \\ t_m &:= m\Delta t, && m \in \{0, \dots, M\}. \end{aligned}$$

The implicit Euler FD approximation is

$$\frac{U_{i,j}^{m+1} - U_{i,j}^{m}}{\Delta t} = \frac{\delta_x^2 U_{i,j}^{m+1}}{(\Delta x)^2} + \frac{\delta_y^2 U_{i,j}^{m+1}}{(\Delta y)^2}$$

for $i \in \{1, \dots, J_x - 1\}$, $j \in \{1, \dots, J_y - 1\}$, $m \in \{0, \dots, M-1\}$, with i.c.

$$U_{i,j}^0 = u_0(x_i, y_j), \quad i \in \{0, \dots, J_x\}, \;\; j \in \{0, \dots, J_y\},$$

and b.c.

$$U_{i,j}^{m+1} := B(x_i, y_j, t_{m+1}), \;\; \text{when } (x_i, y_j) \in \partial\Omega, \; m \in \{0, \dots, M-1\}.$$

θ-scheme

Let

$$\delta_x^2 U_{i,j} := U_{i+1,j} - 2U_{i,j} + U_{i-1,j}, \qquad \delta_y^2 U_{i,j} := U_{i,j+1} - 2U_{i,j} + U_{i,j-1}.$$

For $J_x, J_y, M \in \mathbb{N}$ fixed, let $\Delta x := \frac{b-a}{J_x}$, $\Delta y := \frac{d-c}{J_y}$, $\Delta t := \frac{T}{M}$, and define

$$\begin{aligned} x_i &:= a + i\Delta x, \qquad && i \in \{0, \dots, J_x\}, \\ y_j &:= c + j\Delta y, && j \in \{0, \dots, J_y\}, \\ t_m &:= m\Delta t, && m \in \{0, \dots, M\}. \end{aligned}$$

The θ-scheme, $\theta \in [0,1]$, is

$$\frac{U_{i,j}^{m+1} - U_{i,j}^m}{\Delta t} = (1-\theta)\left(\frac{\delta_x^2 U_{i,j}^m}{(\Delta x)^2} + \frac{\delta_y^2 U_{i,j}^m}{(\Delta y)^2}\right) + \theta\left(\frac{\delta_x^2 U_{i,j}^{m+1}}{(\Delta x)^2} + \frac{\delta_y^2 U_{i,j}^{m+1}}{(\Delta y)^2}\right),$$

for $i \in \{1, \dots, J_x - 1\}$, $j \in \{1, \dots, J_y - 1\}$, $m \in \{0, \dots, M-1\}$, with i.c.

$$U_{i,j}^0 = u_0(x_i, y_j), \quad i \in \{0, \dots, J_x\}, \ \ j \in \{0, \dots, J_y\},$$

and b.c.

$$U_{i,j}^{m+1} := B(x_i, y_j, t_{m+1}), \ \text{ when } (x_i, y_j) \in \partial\Omega, \ m \in \{0, \dots, M-1\}.$$

Practical stability of the θ-scheme

The practical stability of the θ-scheme (in the absence of a b.c. now, i.e., for the pure IVP rather than the IBVP) is assessed by inserting

$$U_{i,j}^{m} = \frac{1}{(2\pi)^2} \int_{-\frac{\pi}{\Delta x}}^{\frac{\pi}{\Delta x}} \int_{-\frac{\pi}{\Delta y}}^{\frac{\pi}{\Delta y}} e^{\imath(k_x i\Delta x + k_y j\Delta y)} \hat{U}^{m}(k_x, k_y)\, \mathrm{d}k_y\, \mathrm{d}k_x.$$

(Here, $\imath$ denotes the complex number, and i the index from $U_{i,j}^{m}$). Writing $\mu_x := \frac{\Delta t}{(\Delta x)^2}$ and $\mu_y := \frac{\Delta t}{(\Delta y)^2}$, we find that

$$\hat{U}^{m+1}(k_x, k_y) = \lambda(k_x, k_y)\hat{U}^{m}(k_x, k_y) \quad \forall (k_x, k_y) \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right] \times \left[-\frac{\pi}{\Delta y}, \frac{\pi}{\Delta y}\right]$$

where the amplification factor $\lambda = \lambda(k_x, k_y)$ is given by

$$\lambda(k_x, k_y) := \frac{1 - 4(1-\theta)\left[\mu_x \sin^2\left(\frac{k_x \Delta x}{2}\right) + \mu_y \sin^2\left(\frac{k_y \Delta y}{2}\right)\right]}{1 + 4\theta\left[\mu_x \sin^2\left(\frac{k_x \Delta x}{2}\right) + \mu_y \sin^2\left(\frac{k_y \Delta y}{2}\right)\right]}$$

for $(k_x, k_y) \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right] \times \left[-\frac{\pi}{\Delta y}, \frac{\pi}{\Delta y}\right]$.

Recall from previous slide:

$$\lambda(k_x, k_y) := \frac{1 - 4(1-\theta)\left[\mu_x \sin^2\left(\frac{k_x \Delta x}{2}\right) + \mu_y \sin^2\left(\frac{k_y \Delta y}{2}\right)\right]}{1 + 4\theta\left[\mu_x \sin^2\left(\frac{k_x \Delta x}{2}\right) + \mu_y \sin^2\left(\frac{k_y \Delta y}{2}\right)\right]}.$$

For practical stability, we require that

$$\max_{(k_x, k_y) \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right] \times \left[-\frac{\pi}{\Delta y}, \frac{\pi}{\Delta y}\right]} |\lambda(k_x, k_y)| \leq 1.$$

Note that $\lambda(k_x, k_y) \leq 1$ without any restriction on μ_x, μ_y. Hence, the scheme is practically stable iff

$$\lambda(k_x, k_y) \geq -1$$

for all $(k_x, k_y) \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right] \times \left[-\frac{\pi}{\Delta y}, \frac{\pi}{\Delta y}\right]$, which holds iff

$$(1-2\theta)\left[\mu_x \sin^2\left(\frac{k_x \Delta x}{2}\right) + \mu_y \sin^2\left(\frac{k_y \Delta y}{2}\right)\right] \leq \frac{1}{2}$$

for all $(k_x, k_y) \in \left[-\frac{\pi}{\Delta x}, \frac{\pi}{\Delta x}\right] \times \left[-\frac{\pi}{\Delta y}, \frac{\pi}{\Delta y}\right]$, i.e., iff

$$(1-2\theta)(\mu_x + \mu_y) \leq \frac{1}{2}.$$

We have obtained that the condition for practical stability is

$$(1-2\theta)(\mu_x+\mu_y) \leq \frac{1}{2}.$$

$\Longrightarrow$ In particular, if $\theta \in [\frac{1}{2}, 1]$, then unconditionally practically stable.
If $\theta \in [0, \frac{1}{2})$, then only conditionally practically stable.

For example, implicit Euler ($\theta = 1$) and Crank–Nicolson ($\theta = \frac{1}{2}$) are unconditionally practically stable, while explicit Euler ($\theta = 0$) is only conditionally practically stable with stability condition $\mu_x + \mu_y \leq \frac{1}{2}$.

Discrete maximum/minimum principle for the θ-scheme

Theorem (Discrete maximum/minimum principle for the θ-scheme)

Consider the θ-scheme for the Dirichlet IBVP for the heat eqn in two space-dimensions, with $\theta \in [0, 1]$. Suppose that

$$(1-\theta)(\mu_x + \mu_y) \leq \frac{1}{2}, \quad \textit{where } \mu_x := \frac{\Delta t}{(\Delta x)^2},\ \mu_y := \frac{\Delta t}{(\Delta y)^2}.$$

Then, for the numerical approximations $\{U_{i,j}^m\}_{i \in \{0,\ldots,J_x\}; j \in \{0,\ldots,J_y\}; m \in \{0,\ldots,M\}}$ we have that

$$\min\{U_{\min}^0, U_\partial^{\min}\} \leq U_{i,j}^m \leq \max\{U_{\max}^0, U_\partial^{\max}\}$$

for any $i \in \{0, \ldots, J_x\}$, $j \in \{0, \ldots, J_y\}$, $m \in \{0, \ldots, M\}$, where

$$\begin{aligned}
U_{\min}^0 &:= \min\{U_{i,j}^0 \,|\, i \in \{0, \ldots, J_x\}, j \in \{0, \ldots, J_y\}\},\\
U_\partial^{\min} &:= \min\{U_{i,j}^m \,|\, (x_i, y_j) \in \partial\Omega, m \in \{0, \ldots, M\}\},\\
U_{\max}^0 &:= \max\{U_{i,j}^0 \,|\, i \in \{0, \ldots, J_x\}, j \in \{0, \ldots, J_y\}\},\\
U_\partial^{\max} &:= \max\{U_{i,j}^m \,|\, (x_i, y_j) \in \partial\Omega, m \in \{0, \ldots, M\}\}.
\end{aligned}$$

Convergence analysis of the θ-scheme

Suppose $(1-\theta)(\mu_x+\mu_y) \le \frac{1}{2}$. We begin by rewriting the scheme as

$$\begin{aligned}(1+2\theta(\mu_x+\mu_y))U_{i,j}^{m+1} &= (1-2(1-\theta)(\mu_x+\mu_y))U_{i,j}^{m} \\ &\quad + (1-\theta)\mu_x(U_{i+1,j}^m + U_{i-1,j}^m) + (1-\theta)\mu_y(U_{i,j+1}^m + U_{i,j-1}^m) \\ &\quad + \theta\mu_x(U_{i+1,j}^{m+1} + U_{i-1,j}^{m+1}) + \theta\mu_y(U_{i,j+1}^{m+1} + U_{i,j-1}^{m+1}).\end{aligned}$$

We have the **consistency error**

$$\begin{aligned}T_{i,j}^m &:= \frac{u_{i,j}^{m+1}-u_{i,j}^m}{\Delta t} - (1-\theta)\left(\frac{\delta_x^2 u_{i,j}^m}{(\Delta x)^2} + \frac{\delta_y^2 u_{i,j}^m}{(\Delta y)^2}\right) - \theta\left(\frac{\delta_x^2 u_{i,j}^{m+1}}{(\Delta x)^2} + \frac{\delta_y^2 u_{i,j}^{m+1}}{(\Delta y)^2}\right) \\ &= \begin{cases} \mathcal{O}\left((\Delta x)^2 + (\Delta y)^2 + (\Delta t)^2\right) & \text{if } \theta = 1/2, \\ \mathcal{O}\left((\Delta x)^2 + (\Delta y)^2 + \Delta t\right) & \text{if } \theta \neq 1/2. \end{cases}\end{aligned}$$

where $u_{i,j}^m := u(x_i, y_j, t_m)$. Observe that

$$\begin{aligned}(1+2\theta(\mu_x+\mu_y))u_{i,j}^{m+1} &= (1-2(1-\theta)(\mu_x+\mu_y))u_{i,j}^{m} \\ &\quad + (1-\theta)\mu_x(u_{i+1,j}^m + u_{i-1,j}^m) + (1-\theta)\mu_y(u_{i,j+1}^m + u_{i,j-1}^m) \\ &\quad + \theta\mu_x(u_{i+1,j}^{m+1} + u_{i-1,j}^{m+1}) + \theta\mu_y(u_{i,j+1}^{m+1} + u_{i,j-1}^{m+1}) \\ &\quad + (\Delta t)T_{i,j}^m.\end{aligned}$$

Define the **global error**

$$e_{i,j}^m := u(x_i, y_j, t_m) - U_{i,j}^m.$$

Note $e_{i,j}^0 = 0 \ \forall i,j$, and $e_{i,j}^m = 0 \ \forall (x_i, y_j) \in \partial\Omega$, $m \in \{1, \ldots, M\}$. We have

$$\begin{aligned}(1 + 2\theta(\mu_x + \mu_y))e_{i,j}^{m+1} &= (1 - 2(1-\theta)(\mu_x + \mu_y))e_{i,j}^m \\ &+(1-\theta)\mu_x(e_{i+1,j}^m + e_{i-1,j}^m) + (1-\theta)\mu_y(e_{i,j+1}^m + e_{i,j-1}^m) \\ &+\theta\mu_x(e_{i+1,j}^{m+1} + e_{i-1,j}^{m+1}) + \theta\mu_y(e_{i,j+1}^{m+1} + e_{i,j-1}^{m+1}) \\ &+(\Delta t)T_{i,j}^m \\ &\le 2\theta(\mu_x + \mu_y)E^{m+1} + E^m + (\Delta t)T^m,\end{aligned}$$

where $E^m := \max_{i,j} |e_{i,j}^m|$ and $T^m := \max_{i,j} |T_{i,j}^m|$.

$\Longrightarrow$ We find that for any $m \in \{0, \ldots, M-1\}$ there holds

$$(1 + 2\theta(\mu_x + \mu_y))E^{m+1} \le 2\theta(\mu_x + \mu_y)E^{m+1} + E^m + (\Delta t)T^m.$$

Hence,

$$E^{m+1} \le E^m + (\Delta t)T^m \qquad \forall m \in \{0, \ldots, M-1\}.$$

As $E^0 = 0$, we have

$$
\begin{aligned}
E^m &\le E^{m-1} + (\Delta t) T^{m-1} \\
&\le E^{m-2} + (\Delta t) T^{m-2} + (\Delta t) T^{m-1} \\
&\vdots \\
&\le (\Delta t) \left(T^0 + T^1 + \cdots + T^{m-1} \right) \le m(\Delta t) \max_{l \in \{0,\dots,m-1\}} T^l \le T \max_{l \in \{0,\dots,M-1\}} T^l
\end{aligned}
$$

for any $m \in \{1, \dots, M\}$. It follows that

$$
\begin{aligned}
\max_{m \in \{0,\dots,M\}} \max_{i \in \{0,\dots,J_x\}, j \in \{0,\dots,J_y\}} |e_{i,j}^m| &\le T \max_{l \in \{0,\dots,M-1\}} T^l \\
&= \begin{cases} \mathcal{O}\left((\Delta x)^2 + (\Delta y)^2 + (\Delta t)^2 \right) & \text{if } \theta = 1/2, \\ \mathcal{O}\left((\Delta x)^2 + (\Delta y)^2 + \Delta t \right) & \text{if } \theta \neq 1/2. \end{cases}
\end{aligned}
$$

End of “Chapter 8: FD approximation of parabolic problems”.

End of MA4255 (AY 2022/23).

Thank you for your attention! :-)