3 Finite difference solution of the one-dimensional heat equation

Let us consider the heat equation with Dirichlet boundary conditions:

$$\begin{cases} \dfrac{\partial u}{\partial t} - \beta^2 \dfrac{\partial^2 u}{\partial x^2} = f, \qquad 0 < x < \ell, \quad t > 0 \\ u(0,t) = u(\ell,t) = 0 \\ u(x,0) = u_0(x). \end{cases} \tag{18}$$

As we have now also partial derivative in time, we need an initial condition at $t = 0$.

Example 3.1. Let $\beta = 1,\ f \equiv 0,\ \ell = 1$, and $u_0(x) = \sin(\pi x)$, then the exact solution of (18) is

$$u(x,t) = \exp(-\pi^2 t) \sin \pi x.$$

3.1 Spatial discretization

Let's make first spatial discretization of (18). Set

$$0 = x_0 < x_1 < \ldots < x_n < x_{n+1} = \ell, \qquad x_{i+1} - x_i = h$$

and replace spatial derivatives with difference approximations:

$$\frac{\partial u(x_i,t)}{\partial t} - \beta^2 \frac{u(x_{i+1},t) - 2u(x_i,t) + u(x_{i-1},t)}{h^2} = f(x_i,t), \qquad i = 1,...,n. \tag{19}$$

Thus we have replaced the original PDE by a system of ordinary (linear) differential equations

$$\begin{cases} \dfrac{\partial \boldsymbol{u}(t)}{\partial t} + \boldsymbol{A}\boldsymbol{u}(t) = \boldsymbol{f}(t) \\ \boldsymbol{u}(0) = \boldsymbol{u}^{(0)} \end{cases} \tag{20}$$

As matrix $\boldsymbol{A}$ is symmetric its eigenvalues are real and its eigenvectors form a basis in $\mathbb{R}^n$. If $f \equiv 0$, then the solution of ODE system

$$\frac{\partial u(t)}{\partial t} + \boldsymbol{A}\boldsymbol{u}(t) = 0 \tag{21}$$

can be represented as

$$\boldsymbol{u}(t) = \sum_{i=1}^{n} c_i^{(0)} \exp(-\lambda_i t) \boldsymbol{v}^{(i)},$$

where $(\lambda_i, \boldsymbol{v}^{(i)}),\ i = 1,...,n$ are the eigenvalue/eigenvector pairs of $\boldsymbol{A}$, and

$$\sum_{i=1}^{n} c_i^{(0)} \boldsymbol{v}^{(i)} = \boldsymbol{u}^{(0)}.$$

Theorem 3.1. *Consider a symmetric real tridiagonal matrix*

$$\boldsymbol{A} = \begin{bmatrix} a & b & & & \\ b & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & b & a & b \\ & & & b & a \end{bmatrix}$$

Then, its eigenvalues are

$$\lambda_j = a + 2b\cos\left(\frac{j\pi}{n+1}\right), \quad j = 1, ..., n \tag{22}$$

and eigenvectors $\boldsymbol{v}^{(j)} \in \mathbb{R}^n$:

$$v_i^{(j)} = \sin\left(\frac{ij\pi}{n+1}\right), \quad i = 1, ..., n. \tag{23}$$

3.2 Temporal discretization

The simplest discretization method is the forward difference (Euler) method:

$$\frac{\boldsymbol{u}(t+\Delta t) - \boldsymbol{u}(t)}{\Delta t} + \boldsymbol{A}\boldsymbol{u}(t) = 0$$

Denote $t_k := k\Delta t$ and

$$\boldsymbol{u}^{(k)} := \begin{bmatrix} u_1(t_k) \\ \vdots \\ u_n(t_k) \end{bmatrix} =: \begin{bmatrix} u_{1,k} \\ \vdots \\ u_{n,k} \end{bmatrix}.$$

Then Euler's method can be written as

$$\boldsymbol{u}^{(k+1)} = \boldsymbol{u}^{(k)} - \Delta t \cdot \boldsymbol{A}\boldsymbol{u}^{(k)} = \underbrace{(\boldsymbol{I} - \Delta t\boldsymbol{A})}_{=:\boldsymbol{B}}\boldsymbol{u}^{(k)} = \boldsymbol{B}\boldsymbol{u}^{(k)} = \boldsymbol{B}^2\boldsymbol{u}^{(k-1)} = ... = (\boldsymbol{B})^k\boldsymbol{u}^{(0)}. \tag{24}$$

Consistency: The error of difference approximations is

$$\mathcal{O}(h^2) + \mathcal{O}(\Delta t) = \mathcal{O}(h^2 + \Delta t).$$

To get a convergent scheme the stability of the scheme is required, i.e. $\|\boldsymbol{u}^{(k)}\| \leq C,\ k \to \infty$. A necessary condition for stability is $\rho(\boldsymbol{B}) \leq 1$, where $\rho(\boldsymbol{B}) = \max\{|\lambda| \,:\, \lambda \text{is an eigenvalue of } \boldsymbol{B}$. Matrix $\boldsymbol{B}$ is tridiagonal, with the following diagonal and co-diagonal entries (cf. Theorem 3.1)

$$a = 1 - r, \quad b = r, \quad r := \frac{\Delta t \beta^2}{h^2}.$$

Thus the eigenvalues of $\boldsymbol{B}$ are

$$\lambda_j = 1 - 2r + 2r\cos\left(\frac{j\pi}{n+1}\right), \quad j = 1, ..., n.$$

Using properties of trigonometric functions we get

$$\lambda_j = 1 - \underbrace{4r\overbrace{\sin^2\frac{j\pi}{2(n+1)}}^{\leq 1}}_{\geq 0}.$$

Thus $\lambda_j \leq 1$ for all $r > 0$. On the other hand $\lambda_j \geq -1$ if $1 - 4r \geq -1$ i.e.

$$\Delta t \leq \frac{1}{2}h^2/\beta^2. \tag{25}$$

The concrete meaning of equation (25) is that if we refine the grid in spatial direction by $h \to \frac{1}{2}h$, then at the same time we must refine the temporal discretization by $\Delta t \to \frac{1}{4}\Delta t$.

3.3 θ methods

A family of time discretization schemes can be defined as

$$\frac{\boldsymbol{u}^{(k+1)} - \boldsymbol{u}^{(k)}}{\Delta t} + \theta \boldsymbol{A}\boldsymbol{u}^{(k+1)} + (1-\theta)\boldsymbol{A}\boldsymbol{u}^{(k)} = 0, \quad 0 \le \theta \le 1. \tag{26}$$

Example 3.2. Most common choices of *theta* are

- $\theta = 0$ (Euler method) local accuracy $\mathcal{O}(h^2 + \Delta t)$
- $\theta = 1$ (implicit Euler method) local accuracy $\mathcal{O}(h^2 + \Delta t)$
- $\theta = \frac{1}{2}$ (Crank–Nicolson method) local accuracy $\mathcal{O}(h^2 + (\Delta t)^2)$

Rewriting (26) in matrix form results in

$$\boldsymbol{B}^{(1)}\boldsymbol{u}^{(k+1)} = \boldsymbol{B}^{(2)}\boldsymbol{u}^{(k)}, \tag{27}$$

where

$$\boldsymbol{B}^{(1)} = \boldsymbol{I} + \theta\Delta t\boldsymbol{A}, \qquad \boldsymbol{B}^{(2)} = \boldsymbol{I} - (1-\theta)\Delta t\boldsymbol{A}.$$

Let us now analyze the stability of the scheme. The eigenvalues are

$$\lambda_j(\boldsymbol{B}^{(1)}) = 1 + 2r\theta\left(1 - \cos\frac{j\pi}{n+1}\right), \quad j = 1, ..., n$$
$$\lambda_j(\boldsymbol{B}^{(2)}) = 1 - 2r(1-\theta)\left(1 - \cos\frac{j\pi}{n+1}\right), \quad j = 1, ..., n.$$

From (27) it follows that

$$\boldsymbol{u}^{(k+1)} = (\boldsymbol{B}^{(1)})^{-1}\boldsymbol{B}^{(2)}\boldsymbol{u}^{(k)} =: \boldsymbol{B}\boldsymbol{u}^{(k)}.$$

As $\boldsymbol{B}^{(1)}$ and $\boldsymbol{B}^{(2)}$ have the same eigenvectors, we have

$$\lambda_j(\boldsymbol{B}) = \frac{\lambda_j(\boldsymbol{B}^{(2)})}{\lambda_j(\boldsymbol{B}^{(1)})} = \frac{1 - 2r(1-\theta)\left(1 - \cos\frac{j\pi}{n+1}\right)}{1 + 2r\theta\left(1 - \cos\frac{j\pi}{n+1}\right)}.$$

In order to have a stable θ-scheme, it must hold $|\lambda_j(\boldsymbol{B})| \le 1$, i.e.

$$r(1 - \cos\frac{j\pi}{n+1})(1 - 2\theta) \le 1. \tag{28}$$

If $\theta \ge \frac{1}{2}$ then (28) holds for all $r > 0$. If $\theta < \frac{1}{2}$, then $r \le \frac{1}{2(1-2\theta)}$

3.3.1 Von Neumann stability analysis

Under some assumtions (not presented here) the stability analysis can also be done using *von Neumann* method. Let us assume that the discrete solution at a grid point is of the form

$$u_{j,k} = \xi^k e^{\mathrm{i}j\varphi} \qquad (\mathrm{i}^2 = -1,\ \xi \neq 0). \tag{29}$$

Then for a single grid point, the θ-method reads

$$\begin{aligned} -\theta r u_{j+1,k+1} + (1+2\theta r)u_{j,k+1} - \theta r u_{j-1,k+1} \\ = (1-\theta)r u_{j+1,k} + (1-2(1-\theta)r)u_{j,k} + (1-\theta)r u_{j-1,k}. \end{aligned} \tag{30}$$

Inserting (29) into (30) we obtain

$$\begin{aligned} (1+2\theta r)\xi^{k+1}e^{\mathrm{i}j\varphi} - \theta r\xi^{k+1}e^{\mathrm{i}(j+1)\varphi} - \theta r\xi^{k+1}e^{\mathrm{i}(j-1)\varphi} \\ = (1-2(1-\theta)r)\xi^k e^{\mathrm{i}j\varphi} + (1-\theta)r\xi^k e^{\mathrm{i}(j+1)\varphi} + (1-\theta)r\xi^k e^{\mathrm{i}(j-1)\varphi}. \end{aligned} \tag{31}$$

Dividign (31) by $\xi^k e^{\mathrm{i}j\varphi}$ we obtain

$$(1+2\theta r)\xi - \theta r\xi e^{\mathrm{i}\varphi} - \theta r\xi e^{-\mathrm{i}\varphi} = (1-2(1-\theta)r) + (1-\theta)re^{\mathrm{i}\varphi} + (1-\theta)re^{-\mathrm{i}\varphi}.$$

As $\cos\varphi = \frac{1}{2}(e^{\mathrm{i}\varphi} + e^{-\mathrm{i}\varphi})$ we finally get

$$\xi = \frac{1-2(1-\theta)r(1-\cos\varphi)}{1+2\theta r(1-\cos\varphi)}.$$

As stability requires $|\xi| \leq 1$, then we see that no restriction on stepsize is required for implicit Euler and Crank–Nicholson schemes. For classical Euler method $r \leq \frac{1}{2}$ must hold.

3.4 Alternate Direction Implicit (ADI) method

Peaceman & Rachford, 1955.
ADI is an *operator splitting method*. First "split" the 2D heat equation

$$\frac{\partial u}{\partial t} - \beta^2\frac{\partial^2 u}{\partial x^2} = \beta^2\frac{\partial^2 u}{\partial y^2},$$

then solve the resulting 1D heat equation (the left hand side).
The time step $t_k \to t_{k+1}$ is divided into two substeps $t_k \to t_{k+\frac{1}{2}}$ and $t_{k+\frac{1}{2}} \to t_{k+1}$.

$$\begin{cases} \dfrac{u_{i,j}^{(k+\frac{1}{2})} - u_{i,j}^{(k)}}{\Delta t/2} - \beta^2\dfrac{u_{i+1,j}^{(k+\frac{1}{2})} - 2u_{i,j}^{(k+\frac{1}{2})} + u_{i-1,j}^{(k+\frac{1}{2})}}{h_1^2} - \beta^2\dfrac{u_{i,j+1}^{(k)} - 2u_{i,j}^{(k)} + u_{i,j-1}^{(k)}}{h_2^2} = 0 \\ \dfrac{u_{i,j}^{(k+1)} - u_{i,j}^{(k+\frac{1}{2})}}{\Delta t/2} - \beta^2\dfrac{u_{i+1,j}^{(k+\frac{1}{2})} - 2u_{i,j}^{(k+\frac{1}{2})} + u_{i-1,j}^{(k+\frac{1}{2})}}{h_2^2} - \beta^2\dfrac{u_{i,j+1}^{(k+1)} - 2u_{i,j}^{(k+1)} + u_{i,j-1}^{(k+1)}}{h_2^2} = 0 \end{cases} \tag{32}$$

In the fist substep partial derivative with respect to x is handled implcitly, and in the second substep with respect to y.
Practical implementation (unit square, $h_1 = h_2$):

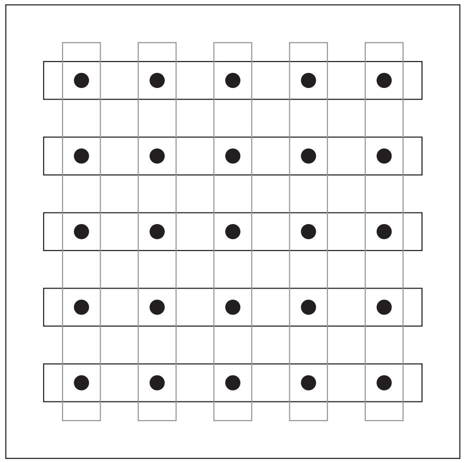

Figure 1:

Let

$$B^{(1)} = \begin{bmatrix} 1+2r & -r & & & \\ -r & 1+2r & -r & & \\ & \ddots & \ddots & \ddots & \\ & & & -r & 1+2r \end{bmatrix}, \quad B^{(2)} = \begin{bmatrix} 1-2r & r & & & \\ r & 1-2r & r & & \\ & \ddots & \ddots & \ddots & \\ & & & r & 1-2r \end{bmatrix},$$

where $r = \Delta t \beta^2 / 2h^2$. Then one step can be implemented as follows (see also Figure 1):

1. Solve unknowns by rows from the tridiagonal system

$$B^{(1)} u^{(k+\frac{1}{2})} = B^{(2)} u^{(k)}$$

2. Solve unknowns by columns from the tridiagonal system

$$B^{(1)} u^{(k+1)} = B^{(2)} u^{(k+\frac{1}{2})}.$$

ADI method was very important breakthrough in the 1950's. Namely, it can be proved that ADI method is stable for all time step sizes. The accuracy of the method is $\mathcal{O}(h^2 + (\Delta t)^2)$. Moreover, the computational cost per timestep is only $\mathcal{O}(N)$, where $N = n_x n_y$.