6 Numerical Solution of Parabolic Equations

6.1 The One-Dimensional Model Problem

We consider the following initial boundary value problem (IBVP) modelling heat flow in a thin rod, i.e., in one space dimension:

$$u_t = \kappa u_{xx}, \qquad x \in (0,1),\ t > 0, \tag{6.1a}$$

$$u(0,t) = g_0(t), \qquad t > 0, \tag{6.1b}$$

$$u(1,t) = g_1(t), \qquad t > 0, \tag{6.1c}$$

$$u(x,0) = u_0(x), \qquad x \in [0,1] \tag{6.1d}$$

with given (constant) heat conductivity κ (which we set to one in the following) as well as (possibly time-dependent) Dirichlet boundary values g_0, g_1 and initial data u_0.

- Steady-state version:

$$u_{xx} = 0, \qquad u(0) = g_0, \quad u(1) = g_1.$$

- Analogous IBVP in 2D and 3D:

$$u_t = \Delta u \quad + \quad \text{initial and boundary data.}$$

- Related: linear ordinary differential equation

$$\boldsymbol{u}_t = A\boldsymbol{u}, \quad \boldsymbol{u} : t \mapsto \boldsymbol{u}(t) \in \mathbb{R}^n.$$

  Here linear differential operator ∂_{xx} in place of matrix A.

Series solution by separation of variables: in special cases an analytic representation of the (exact) solution may be constructed using the technique of separation of variables. This is helpful for checking numerical approximations and provides important insight into the structure of the solution.

Inserting the special trial solution $u(x,t) = f(x)g(t)$ into the PDE $u_t = u_{xx}$ results in

$$fg' = f''g, \qquad \text{i.e.} \qquad \frac{g'}{g} = \frac{f''}{f} = \text{const.} =: -k^2.$$

For each value of k we obtain a solution

$$u_k(x,t) = e^{-k^2 t} \sin(kx)$$

of $u_t = u_{xx}$. The boundary condition $u(0,t) = u(1,t) = 0$ constrains k to the discrete values

$$k = k_m := m\pi, \qquad m \in \mathbb{N}.$$

Due to the linearity and homogeneity of $u_t = u_{xx}$ any linear combination of these solutions is also a solution. If we succeed in finding coefficients a_m in such a way that

$$u_0(x) = \sum_{m=1}^{\infty} a_m \sin(m\pi x),$$

then the series

$$u(x,t) := \sum_{m=1}^{\infty} a_m e^{-m^2\pi^2 t} \sin(m\pi x)$$

solves the complete IBVP (6.1).

Since the functions $\{\sin(m\pi x)\}_{m=1}^{\infty}$ form a complete orthogonal system of the function space $L^2(0,1)$, this is possible for all $u_0 \in L^2(0,1)$.

The coefficients are given by

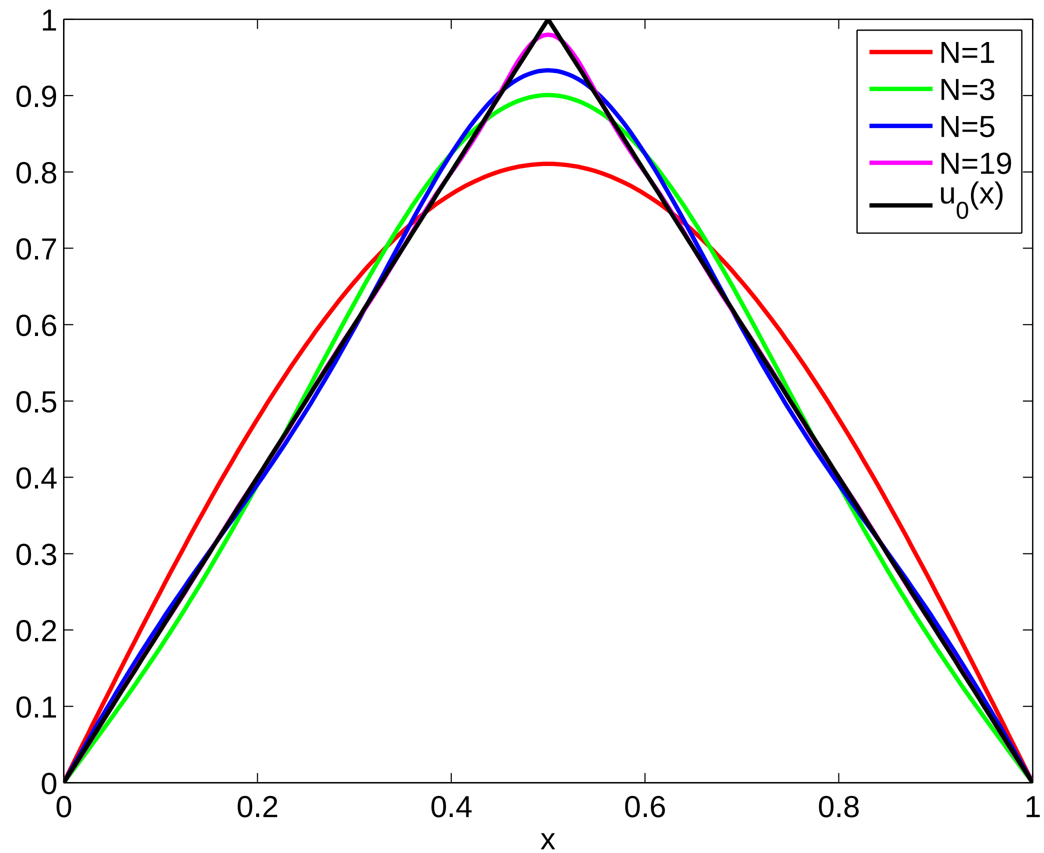
$$a_m = 2\int_0^1 u_0(x) \sin(m\pi x)\, dx.$$

Example:

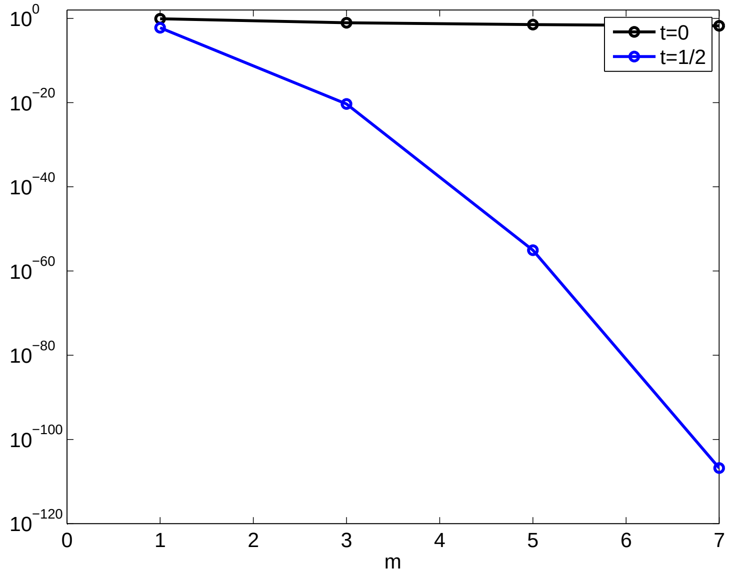
$$u_0(x) = 1 - 2|x - \tfrac{1}{2}|.$$

Here the coefficients are

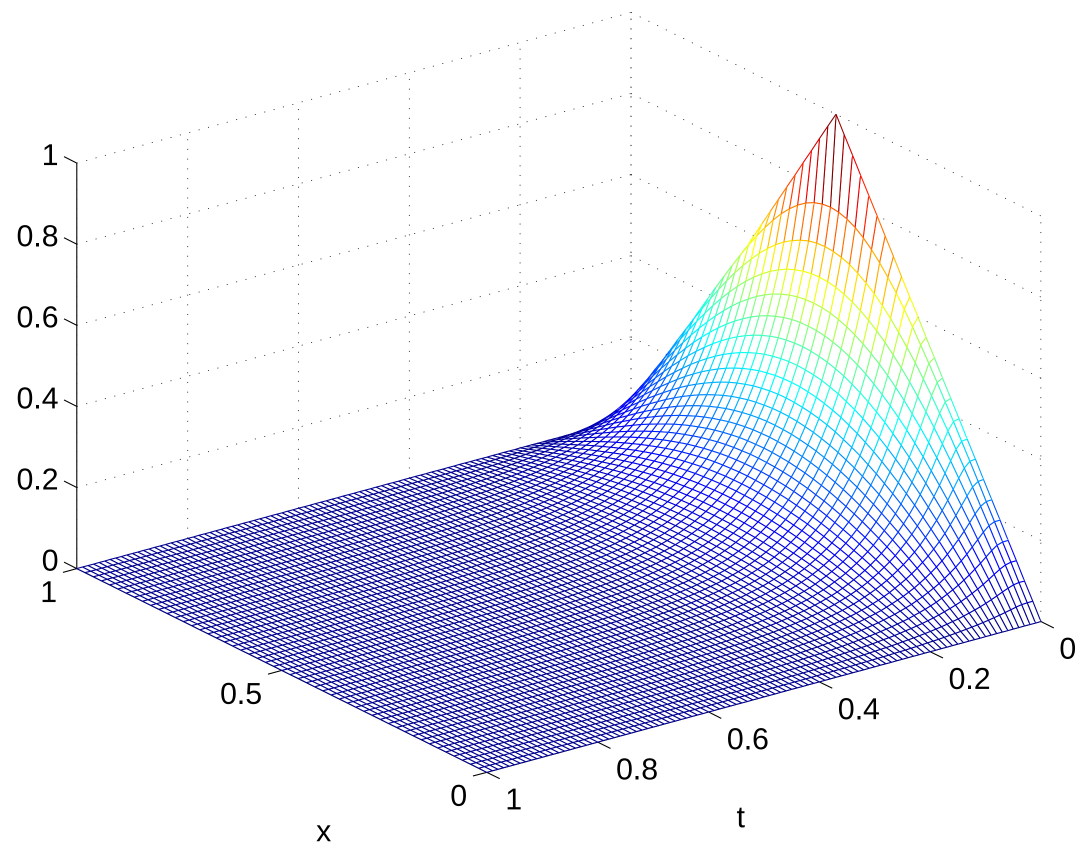
$$a_m = \frac{8}{m^2\pi^2} \sin \frac{m\pi}{2}, \qquad m \in \mathbb{N}.$$



Partial sums $\sum_{m=1}^{N} a_m \sin(m\pi x)$ of the Fourier series of u_0.



Decay of first 4 nonzero Fourier coefficients of u_0.



Fourier series solution of IBVP (6.1) with $\kappa = 1$ and $u_0(x) = 1 - 2|x - \frac{1}{2}|$ in domain $(x,t) \in [0,1] \times [0,1]$.

We first discretize the IBVP (6.1) in the spatial variable x only, leaving time t continuous. To this end we proceed as in the elliptic case and introduce the grid points

$$0 = x_0 < x_1 < \cdots < x_J < x_{J+1} = 1$$

using a fixed grid spacing $\Delta x = 1/(J+1)$ and approximate

$$u_{xx}|_{x=x_j} \approx [A_{\Delta x}\boldsymbol{u}]_j, \qquad j = 1, 2, \dots, J,$$

with

$$A_{\Delta x} = \frac{1}{\Delta x^2}\,\mathrm{tridiag}(1, -2, 1). \tag{6.2}$$

If $\boldsymbol{u} = \boldsymbol{u}(t)$ denotes the vector with components

$$u_j(t) \approx u(x_j, t), \quad t > 0, \quad j = 1, 2, \dots, J,$$

then (6.1) is transformed into the semi-discrete system of ODEs

$$\boldsymbol{u}'(t) = A_{\Delta x}\boldsymbol{u}(t) + \boldsymbol{g}(t), \tag{6.3a}$$

$$\boldsymbol{u}(0) = \boldsymbol{u}_0, \tag{6.3b}$$

with $[\boldsymbol{u}_0]_j = u_0(x_j)$, $j = 1, 2, \dots, J$ as well as

$$\boldsymbol{g}(t) = 1/(\Delta x^2)[g_0(t), 0, \dots, 0, g_1(t)]^\top \in \mathbb{R}^J.$$

We can now solve (6.3) with known numerical methods for solving ODEs. Introducing the fixed time step $\Delta t > 0$, we set

$$U_j^n \approx [\boldsymbol{u}(t_n)]_j \approx u(x_j, t_n), \qquad t_n = n\Delta t.$$

The approximation of the solution of a time-dependent PDE as a system of ODEs along the "lines" $\{(x_j, t) : t > 0\}$ is known as the method of lines.

Applying the explicit Euler method to (6.3) (setting $g_0(t) = g_1(t) = 0$ for now) leads to

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x^2}\left(U_{j-1}^n - 2U_j^n + U_{j+1}^n\right), \qquad 1 \le j \le J, \quad n = 0, 1, 2 \ldots .$$

This corresponds to the finite difference approximation

$$\underbrace{\frac{U_j^{n+1} - U_j^n}{\Delta t}}_{\approx u_t} = \underbrace{\frac{U_{j-1}^n - 2U_j^n + U_{j+1}^n}{\Delta x^2}}_{\approx u_{xx}} \tag{6.4}$$

of the differential equation (6.1a).

In matrix notation:

$$\boldsymbol{U}^{n+1} = (I + \Delta t\, A_{\Delta x})\,\boldsymbol{U}^n, \qquad n = 0, 1, 2, \ldots ,$$

with $\boldsymbol{U}^n = [U_1^n, U_2^n, \ldots, U_J^n]^\top$.

We define the local discretisation error of the difference scheme (6.4) to be the residual obtained on inserting the exact solution into the difference scheme:

$$d(x,t) := \frac{u(x,t+\Delta t) - u(x,t)}{\Delta t} - \frac{u(x-\Delta x,t) - 2u(x,t) + u(x+\Delta x,t)}{\Delta x^2}.$$

Using Taylor expansions in (x,t) one easily obtains:

$$d^{\mathrm{eE}}(x,t) = \left(\frac{\Delta t}{2} - \frac{\Delta x^2}{12}\right) u_{xxxx} + O(\Delta t^2) + O(\Delta x^4). \tag{6.5}$$

Terminology: the explicit Euler method for solving the heat equation is consistent of first order in time and of second order in space.

Besides the asymptotic statement (6.5) we also require upper bounds for d. Truncating the Taylor expansions with a remainder term, we obtain in time

$$u(x,t+\Delta t) = (u + \Delta t\, u_t)\,|_{(x,t)} + \frac{\Delta t^2}{2} u_{tt}(x,\tau), \quad \tau \in (t, t+\Delta t).$$

Proceeding analogously in x yields

$$d^{\mathrm{eE}}(x,t) = \frac{\Delta t}{2} u_{tt}(x,\tau) - \frac{\Delta x^2}{12} u_{xxxx}(\xi,t), \qquad \xi \in (x-\Delta x, x+\Delta x),$$

and we obtain, setting $\mu := \frac{\Delta t}{\Delta x^2}$,

$$|d^{\mathrm{eE}}(x,t)| \le \frac{\Delta t}{2} M_{tt} - \frac{\Delta x^2}{12} M_{xxxx} = \frac{\Delta t}{2}\left(M_{tt} + \frac{1}{6\mu} M_{xxxx}\right), \qquad (6.6)$$

assuming $|u_{tt}| \le M_{tt}$ und $|u_{xxxx}| \le M_{xxxx}$ on $[0,1] \times [0,T]$.

Applying the implicit Euler method to (6.3) one obtains, in place of (6.4), the implicit difference scheme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \frac{U_{j-1}^{n+1} - 2U_j^{n+1} + U_{j+1}^{n+1}}{\Delta x^2}. \tag{6.7}$$

The calculation of $\boldsymbol{u}^{n+1}$ from $\boldsymbol{u}^n$ is seen to require the solution of a linear system of equations with coefficient matrix $I - \Delta t A_{\Delta x}$.

Here Taylor expansion results in a local discretization error of

$$d^{\mathrm{iE}}(x,t) = -\left(\frac{\Delta t}{2} + \frac{\Delta x^2}{12}\right) u_{xxxx} + O(\Delta t^2) + O(\Delta x^4).$$

Applying instead the trapezoidal rule, which for an ODE $\boldsymbol{y}'(t) = \boldsymbol{f}(t, \boldsymbol{y}(t))$ is given by

$$\boldsymbol{y}^{n+1} = \boldsymbol{y}^n + \frac{\Delta t}{2}\left[\boldsymbol{f}(t_n, \boldsymbol{y}^n) + \boldsymbol{f}(t_{n+1}, \boldsymbol{y}^{n+1})\right]$$

yields another implicit scheme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \frac{1}{2}\left(\frac{U_{j-1}^n - 2U_j^n + U_{j+1}^n}{\Delta x^2} + \frac{U_{j-1}^{n+1} - 2U_j^{n+1} + U_{j+1}^{n+1}}{\Delta x^2}\right),$$

which in this context is known as the Crank-Nicolson scheme[a]. This method is also implicit, requiring in each time step the solution of a linear system of equations with the coefficient matrix $I - \frac{\Delta t}{2} A_{\Delta x}$.

For Crank-Nicolson (CN) there holds

$$d^{\mathsf{CN}}(x,t) = O(\Delta t^2) + O(\Delta x^2).$$

[a] J. Crank and P. Nicolson (1947)

6.2 Convergence

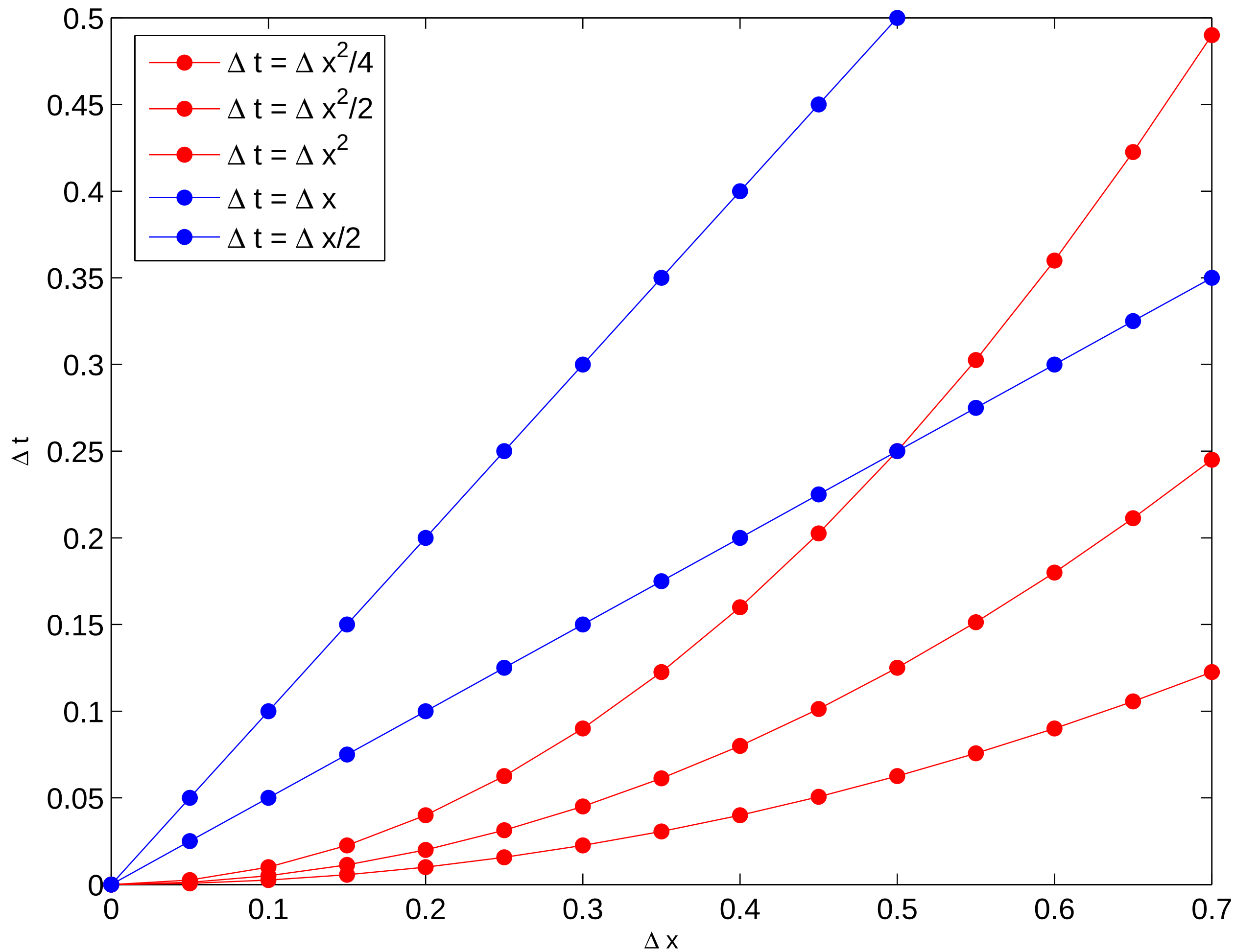
All three methods considered so far are consistent, i.e., at all points (x,t) of the domain we have $d(x,t) \to 0$ as $\Delta x \to 0$ and $\Delta t \to 0$.

To analyze their convergence, we proceed as in the case of numerical methods for ODEs and consider a finite time interval $t \in [0,T]$, $T > 0$ as well as a sequence of grids with grid spacings $\Delta x \to 0$, $\Delta t \to 0$ and determine whether at every fixed grid point (x_j,t_n) also the global error $u(x_j,t_n) - U_j^n$ tends to zero uniformly.

A sequence of grid spacings $\{(\Delta x)_k, (\Delta t)_k\}$ can approach the point $(0,0)$ in different ways. The following figure shows different “refinement curves” in the $(\Delta x, \Delta t)$-plane.

We will see that the explicit Euler method converges only if the refinement satisfies

$$\mu := \frac{\Delta t}{\Delta x^2} \le \frac{1}{2}.$$



Different refinement curves $\Delta t, \Delta x \to 0$:
red: $\Delta t / \Delta x^2 =$ const, blue: $\Delta t / \Delta x =$ const.

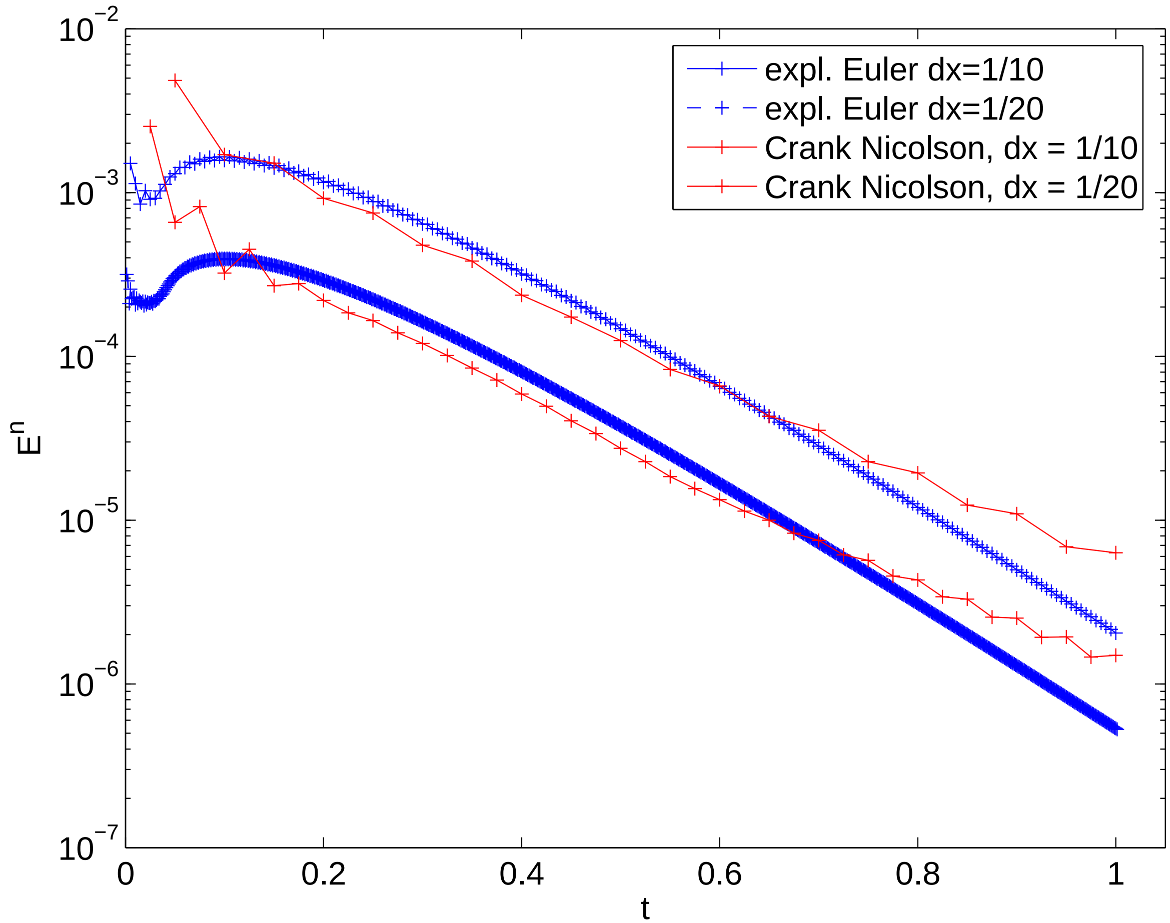
Theorem 6.1 *If a sequence of grids satisfies*

$$\mu_k = \frac{(\Delta t)_k}{(\Delta x)_k^2} \leq \frac{1}{2} \qquad \textit{for all } k \textit{ sufficiently large},$$

and if for the corresponding sequences $\{j_k\}$ *and* $\{n_k\}$ *there holds*

$$n_k(\Delta t)_k \to t \in [0, T], \qquad j_k(\Delta x)_k \to x \in [0, 1],$$

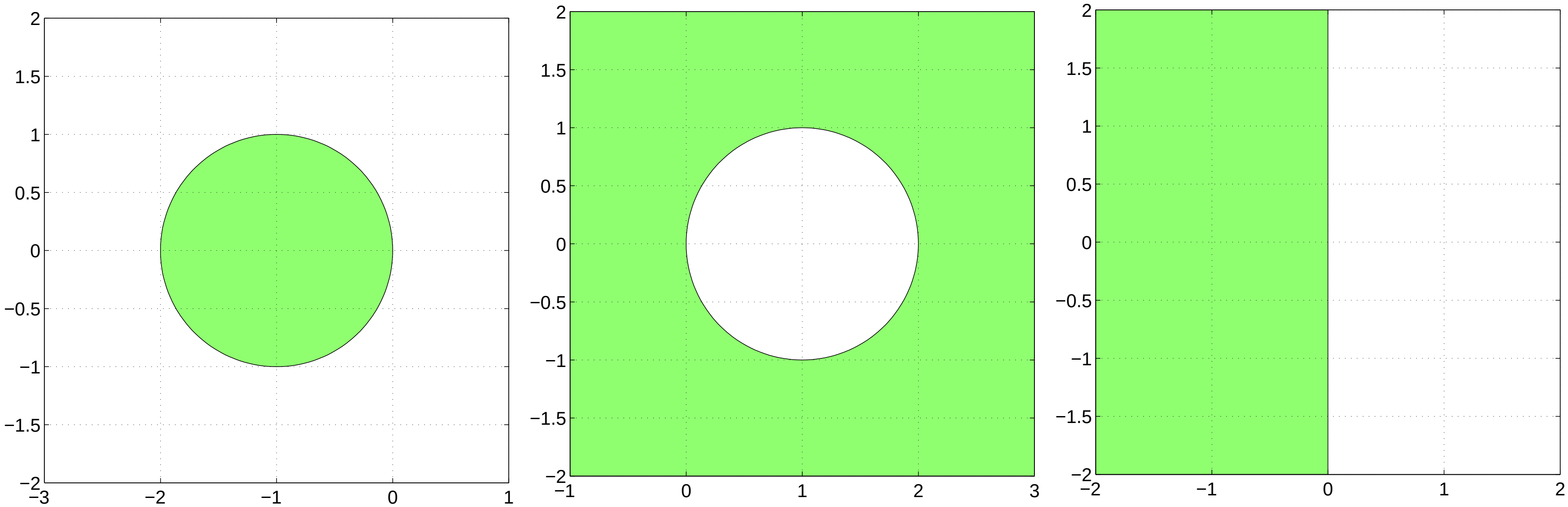
then, under the assumption that $|u_{xxxx}| \leq M_{xxxx}$ *uniformly in* $[0, 1] \times [0, T]$*, the approximations generated by the explicit Euler method* (6.4) *converge to the exact solution* $u(x, t)$ *as* $k \to \infty$ *uniformly on* $[0, 1] \times [0, T]$*.*



Error $E(t_n) = \max_j |u_j^n - U_j^n|$ of numerical solution of model problem(6.1) with $g_0 = g_1 = 0$ and $u_0(x) = x(1-x)$ obtained with explicit Euler method with $\Delta t = \Delta x^2/2$ and the Crank-Nicolson method with $\Delta t = \Delta x/2$.

6.3 Stability

In the method of lines, an ODE solver is applied to the semidiscrete system (6.3). Recalling the regions of absolute stability of the explicit and implicit Euler and trapezoidal methods, we are led to investigate the behavior of the eigenvalues of the matrix $A_{\Delta x}$ in (6.2) as $\Delta x \to 0$.



explicit Euler $R(\widehat{\lambda}) = 1 + \widehat{\lambda}$

implicit Euler $R(\widehat{\lambda}) = \frac{1}{1-\widehat{\lambda}}$

trapezoidal $R(\widehat{\lambda}) = \frac{1+\widehat{\lambda}/2}{1-\widehat{\lambda}/2}$

The eigenvalues of $A_{\Delta x}$ (cf. (5.7)) are given by

$$\lambda_j = -\frac{4}{\Delta x^2}\sin^2\frac{j\pi\Delta x}{2}, \qquad j = 1, 2 \dots, J.$$

The eigenvalue of largest magnitude is

$$\lambda_J = -\frac{4}{\Delta x^2}\sin^2\left(\frac{\pi}{2}(1-\Delta x)\right) = \frac{-4}{\Delta x^2} + \pi^2 + O\left(\Delta x^2\right).$$

Therefore, in order to satisfy

$$|R(\Delta t\,\lambda_j)| \le 1 \qquad \text{for all } j = 1, 2, \dots, J$$

in the explicit Euler method, it is necessary that

$$\frac{\Delta t}{\Delta x^2} \le \frac{1}{2}. \tag{6.8}$$

Both implicit Euler and the trapezoidal rule are absolutely stable methods. Therefore, as all eigenvalues of $A_{\Delta x}$ lie in the left half plane, the requirement of absolute stability places no constraints on the time step Δt.
Since the CN scheme (trapezoidal rule) is consisten of order 2 in both x and t, this suggests a time step $\Delta t \approx \Delta x$.

Absolute stability describes the behavior of a numerical approximation $\boldsymbol{u}^n$ at time t_n of the solution of an ODE as $n \to \infty$. The appropriate stability concept for the convergence analysis of numerical methods for solving IVPs for PDEs was developed by Lax[a] and Richtmyer[b].

All three methods considered so far can be written in the form[c]

$$\boldsymbol{U}^{n+1} = B(\Delta t)\,\boldsymbol{U}^n + \boldsymbol{g}^n(\Delta t), \qquad n = 0, 1, \dots \tag{6.9}$$

with different matrices $B(\Delta t)$ given in each case by

$$B^{\mathsf{eE}}(\Delta t) = I + \Delta t\, A_{\Delta x}, \tag{6.10a}$$

$$B^{\mathsf{iE}}(\Delta t) = (I - \Delta t\, A_{\Delta x})^{-1}, \tag{6.10b}$$

$$B^{\mathsf{CN}}(\Delta t) = \left(I - \frac{\Delta t}{2} A_{\Delta x}\right)^{-1} \left(I + \frac{\Delta t}{2} A_{\Delta x}\right). \tag{6.10c}$$

[a] Peter D. Lax ($*$ 1906)
[b] Robert Davis Richtmyer (1910–2003)
[c] We assume that Δt is chosen as a fixed given function of Δx.

A linear finite difference scheme of the form (6.9) is called Lax-Richtmyer stable, if for any fixed stopping time T there exists a constant $K_T > 0$ such that

$$\|B(\Delta t)^n\| \leq K_T, \qquad \forall \Delta t > 0 \text{ and } n \in \mathbb{N}, n\Delta t \leq T.$$

Theorem 6.2 (Lax-Richtmyer equivalence theorem) *A consistent linear finite difference scheme* (6.9) *is convergent if, and only if, it is Lax-Richtmyer stable.*

Examples: For the explicit Euler method applied to (6.1) the resulting matrix (6.10a) is normal, implying $\|B(\Delta t)\|_2 = \rho(B(\Delta t))$. If condition (6.8) is satisfied, we have $\|B(\Delta t)\|_2 \leq 1$, implying that the explicit Euler method is Lax-Richtmyer stable and therefore convergent.

The implicit Euler and Crank-Nicolson schemes are also Lax-Richtmyer stable but without constraints on the mesh ratio or time step.

Remark 6.3 *In all three schemes considered so far we were able to show the stronger statement* $\|B(\Delta t)\| \leq 1$*, which is sometimes called strong stability. However, strong stability is not necessary for Lax-Richtmyer stability.*

More precisely, for Lax-Richtmyer stability it is necessary that there exist a constant α *such that*

$$\|B(\Delta t)\| \leq 1 + \alpha \Delta t \tag{6.11}$$

for all sufficiently small time steps Δt*, which is evident from the inequality*

$$\|B(\Delta t)^n\| \leq (1 + \alpha \Delta t)^n \leq e^{\alpha T}, \qquad \forall\, n\Delta t \leq T.$$

6.4 Von-Neumann Analysis

A formal procedure based on Fourier analysis for determining stability bounds was introduced by von Neumann [a].

In general, von Neumann analysis only yields necessary conditions, which are also sufficient only in special cases, among these linear PDEs with constant coefficients, unbounded spatial domain (Cauchy problem) or periodic boundary conditions, uniform grid.

Fundamental fact: in the continuous case ($x \in \mathbb{R}$) the functions $e^{i\xi x}, \xi \in \mathbb{R}$, are eigenfunctions of the differential operator ∂_x, i.e.,

$$\partial_x e^{i\xi x} = i\xi\, e^{i\xi x},$$

and thus are also eigenfunctions of any linear differential operator with constant coefficients.

[a] JOHN VON NEUMANN (1903–1957)

On the unbounded spatial grid $\{x_j = j\Delta x\}_{j\in\mathbb{Z}}$ the grid function

$$v_j := e^{i\xi x_j} = e^{i\xi j\Delta x}, \qquad j \in \mathbb{Z},$$

is an eigenfunction of any difference operator with constant coefficients.

Example: For the central difference operator δ_0 defined as $(\delta_0 v)_j := (v_{j+1} - v_{j-1})/(2\Delta x)$ there holds

$$\delta_0 v_j := (\delta_0 v)_j = \frac{e^{i\xi(j+1)\Delta x} - e^{i\xi(j-1)\Delta x}}{2\Delta x} = \frac{e^{i\xi\Delta x} - e^{-i\xi\Delta x}}{2\Delta x} e^{i\xi j\Delta x}$$
$$= \frac{i}{\Delta x}\sin(\xi\Delta x)\, v_j$$

The power series expansion

$$\frac{i}{\Delta x}\sin(\xi\Delta x) = \frac{i}{\Delta x}\left(\xi\Delta x - \frac{(\xi\Delta x)^3}{6} + O\big((\xi\Delta x)^5\big)\right) = i\xi + O(\Delta x^2\xi^3)$$

shows that the eigenvalue of δ_0 approximates the corresponding eigenvalue of ∂_x.

Any grid function on $\{x_j = j\Delta x : j \in \mathbb{Z}\}$ can be represented as the superposition

$$v_j = \frac{1}{2\pi} \int_{-\pi/\Delta x}^{\pi/\Delta x} \widehat{v}(\xi) e^{i\xi j \Delta x} \, d\xi, \qquad j \in \mathbb{Z}$$

of “grid waves” $e^{i\xi j\Delta x}$, $\xi \in [\frac{-\pi}{\Delta x}, \frac{\pi}{\Delta x}]$ with Fourier coefficients $\widehat{v}(\xi)$ given by

$$\widehat{v}(\xi) = \Delta x \sum_{j=-\infty}^{\infty} v_j e^{-i\xi j \Delta x}, \qquad \xi \in \left[\frac{-\pi}{\Delta x}, \frac{\pi}{\Delta x}\right].$$

If the magnitude of grid functions and their Fourier coefficients are measured by

$$\|v\| = \left(\Delta x \sum_{j=-\infty}^{\infty} |v_j|^2\right)^{1/2} \quad \text{and} \quad \|\widehat{v}\| = \left(\int_{-\pi/\Delta x}^{\pi/\Delta x} |\widehat{v}(\xi)|^2 \, d\xi\right)^{1/2},$$

then Parseval’s identity states that

$$\|\widehat{v}\| = \sqrt{2\pi} \|v\|. \tag{6.12}$$

The bound $\|B\| \le 1 + \alpha\Delta t$ requires verifying

$$\|Bu\| \le (1 + \alpha\Delta t)\|u\| \quad \forall u.$$

According to (6.12), an equivalent condition is

$$\|\widehat{Bu}\| \le (1 + \alpha\Delta t)\|\widehat{u}\| \quad \forall \widehat{u}, \tag{6.13}$$

which is easier to show because the Fourier coefficients are decoupled.

For a one-step method one obtains

$$\widehat{u}^{n+1}(\xi) = g(\xi)\, \widehat{u}^{n}(\xi)$$

with a so-called amplification factor $g(\xi)$. If we are able to show that

$$|g(\xi)| \le 1 + \alpha\Delta t.$$

with a constant α which is independent of ξ, then this implies (6.13).

Examples:

(a) For the explicit Euler scheme applied to (6.3) we obtain the amplification factor

$$g(\xi) = 1 - \frac{4\Delta t}{\Delta x^2} \sin^2 \frac{\xi \Delta x}{2}.$$

(b) For Crank-Nicolson,

$$g(\xi) = \frac{1 + z(\xi)}{1 - z(\xi)}, \qquad z(\xi) = \frac{\Delta t}{\Delta x^2} \Big(-1 + \cos(\xi \Delta x) \Big) \le 0 \quad \forall \xi.$$