Numerical Methods for Differential Equations

Chapter 5: Partial differential equations – elliptic and parabolic

Gustaf Söderlind and Carmen Arévalo

Numerical Analysis, Lund University

Textbooks: *A First Course in the Numerical Analysis of Differential Equations*, by Arieh Iserles
and *Introduction to Mathematical Modelling with Differential Equations*, by Lennart Edsberg

1. Brief overview of PDE problems

Classification: Three basic types, four prototype equations

- *Elliptic*

$$\Delta u = 0 \qquad \text{+ BC}$$

- *Parabolic*

$$u_t = \Delta u \qquad \text{+ BC \& IC}$$

- *Hyperbolic*

$$\begin{aligned} u_{tt} &= \Delta u \qquad &\text{+ BC \& IC} \\ u_t + a(u)u_x &= 0 \qquad &\text{+ BC \& IC} \end{aligned}$$

Classification of PDEs

Linear PDE with two independent variables

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} + L(u_x, u_y, u, x, y) = 0$$

with L *linear in* u_x, u_y, u. Study

$$\delta := \det \begin{pmatrix} A & B \\ B & C \end{pmatrix} = AC - B^2$$

Elliptic $\delta > 0$ **Parabolic** $\delta = 0$ **Hyperbolic** $\delta < 0$

Standard PDEs. Prototypical problems

$\delta > 0$	Elliptic PDE	*Laplace equation*
	$u_{xx} + u_{yy} = 0$	$A = C = 1;\ \ B = 0$
$\delta = 0$	Parabolic PDE	*Diffusion equation*
	$u_t = u_{xx}$	$A = 1;\ \ B = C = 0$
$\delta < 0$	Hyperbolic PDEs	*Wave equation*
	$u_{tt} = u_{xx}$	$A = 1;\ B = 0;\ \ C = -1$

PDE method types

FDM *Finite difference methods*

FEM *Finite element methods*

FVM *Finite volume methods*

BEM *Boundary element methods*

We will mostly study FDM to cover basic theory

Industrial relevance: FEM

PDE methods for elliptic problems

Simple geometry	FDM or Fourier methods
Complex geometry	FEM
Special problems	FVM or BEM

Large sparse systems

Combine with iterative solvers such as multigrid methods

PDE methods for parabolic problems

Simple geometry	FDM or Fourier methods
Complex geometry	FEM

Stiffness

Always use A–stable time-stepping methods

Need Newton-type solvers for large sparse systems

PDE methods for hyperbolic problems

FDM, FVM. Sometimes FEM

Shocks

Solutions may be discontinuous – example: “sonic boom”

Turbulence

Multiscale phenomena

Hyperbolic problems have several complications and many highly specialized techniques are often needed

2. Elliptic problems. FDM

Laplacian $$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

Laplace equation $\Delta u = 0$
with boundary conditions $u = u_0(x, y, z) \quad x, y, z \in \partial\Omega$

Poisson equation $-\Delta u = f$
with boundary conditions $u = u_0(x, y, z) \quad x, y, z \in \partial\Omega$

Other boundary conditions also of interest ($\partial\Omega$ = bdry of Ω)

Elliptic problems. Some applications

- **Equilibrium problems**

  Structural analysis (strength of materials)

  Heat distribution

- **Potential problems**

  Potential flow (inviscid, subsonic flow)

  Electromagnetics (fields, radiation)

- **Eigenvalue problems**

  Acoustics

  Microphysics

Poisson equation – an elliptic model problem

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y)$$

Computational domain $\Omega = [0, 1] \times [0, 1]$ (unit square)
Dirichlet conditions $u(x, y) = 0$ on boundary

Uniform grid $\{x_i, y_j\}_{i,j=1}^{N,M}$ with equidistant mesh widths
$\Delta x = 1/(N + 1)$ and $\Delta y = 1/(M + 1)$

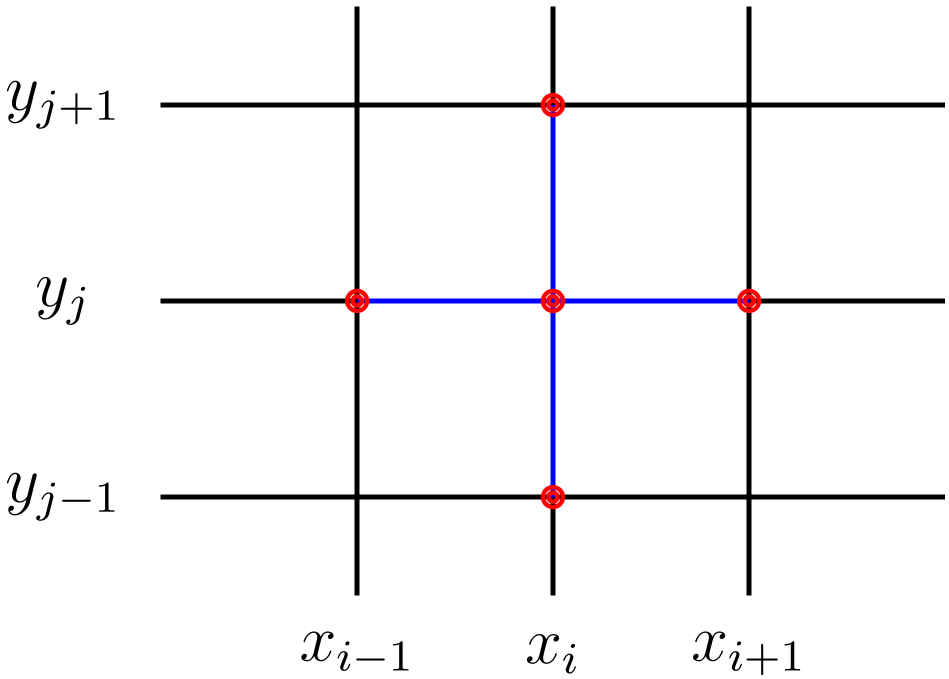
Discretization Finite differences with $u_{i,j} \approx u(x_i, y_j)$

$$\frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\Delta x^2} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{\Delta y^2} = f(x_i, y_j)$$

Equidistant mesh $\Delta x = \Delta y$

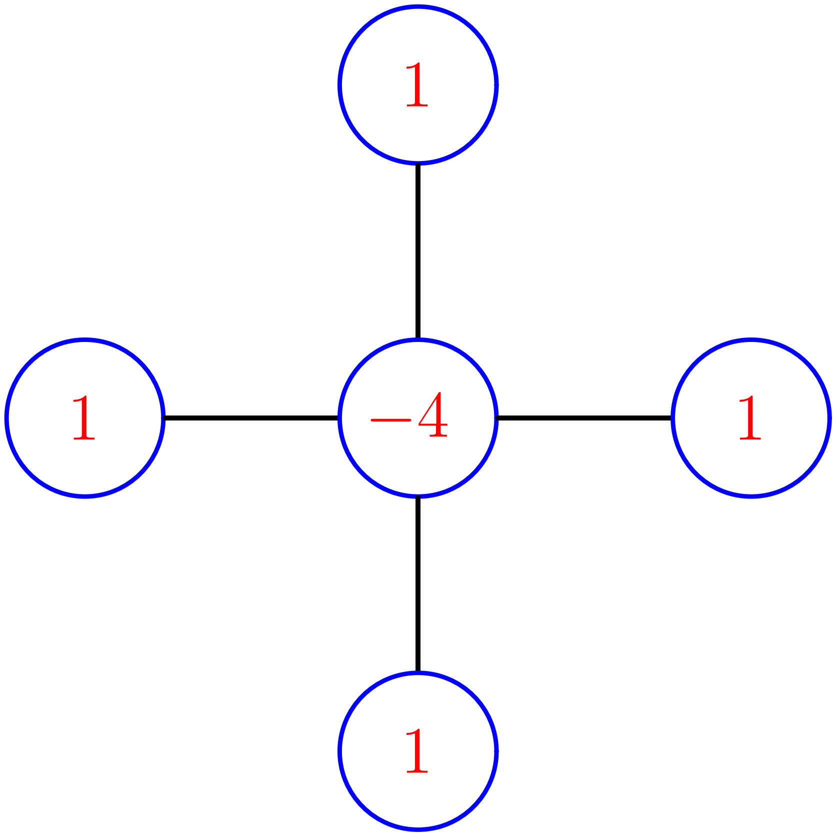
$$\frac{u_{i-1,j} + u_{i,j-1} - 4u_{i,j} + u_{i,j+1} + u_{i+1,j}}{\Delta x^2} = f(x_i, y_j)$$

Participating approximations and mesh points



Computational “stencil” for $\Delta x = \Delta y$

$$\frac{u_{i-1,j} + u_{i,j-1} - 4u_{i,j} + u_{i,j+1} + u_{i+1,j}}{\Delta x^2} = f(x_i, y_j)$$



“Five-point operator”

The FDM linear system of equations

Lexicographic ordering of unknowns $\Rightarrow$ *partitioned system*

$$\frac{1}{\Delta x^2}\begin{pmatrix} T & I & 0 & \dots & & \\ I & T & I & & & \\ & I & T & I & & \\ & & & & & \\ & & & & \ddots & I \\ & \dots & & 0 & I & T \end{pmatrix}\begin{pmatrix} u_{\cdot,1} \\ u_{\cdot,2} \\ u_{\cdot,3} \\ \vdots \\ \\ u_{\cdot,N} \end{pmatrix} = \begin{pmatrix} f(x_\cdot, y_1) \\ f(x_\cdot, y_2) \\ f(x_\cdot, y_3) \\ \vdots \\ \\ f(x_\cdot, y_N) \end{pmatrix}$$

with Toeplitz matrix $T = \mathrm{tridiag}(1 \quad -4 \quad 1)$

The system is $N^2 \times N^2$, hence *large and very sparse*

3. Elliptic problems. FEM

Finite Element Method

PDE	$Lu = 0$		
Ansatz	$u = \sum c_i \varphi_i$	$\Rightarrow$	$Lu = \sum c_i L\varphi_i$
Requirement	$\langle \varphi_i, Lu \rangle = 0$	gives coefficients $\{c_i\}$	

FEM is a *least squares approximation*, fitting a linear combination of basis functions $\{\varphi_i\}$ to the differential equation using *orthogonality*

Simplest case Piecewise linear basis functions

Strong and weak forms

Strong form

$$-\Delta u = f \,; \qquad u = 0 \text{ on } \partial\Omega$$

Take v with $v = 0$ on $\partial\Omega$

$$\int_\Omega -\Delta u \cdot v = \int_\Omega f \cdot v$$

Integrate by parts to get weak form

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega f \cdot v$$

Strong and weak forms. 1D case

Recall integration by parts in 1D

$$\int_0^1 -u''v = [-u'v]_0^1 + \int_0^1 u'v'$$

or in terms of an inner product

$$-\langle u'', v\rangle = \langle u', v'\rangle$$

Generalization to 2D, 3D uses vector calculus

Weak form of $-\Delta u = f$

Define *inner product*

$$\langle v, u \rangle = \int_\Omega v u \, \mathrm{d}\Omega$$

and *energy norm* (note scalar product!)

$$a(v, u) = \int_\Omega \nabla v \cdot \nabla u \, \mathrm{d}\Omega$$

to get the *weak form* of $-\Delta u = f$ as

$$a(v, u) = \langle v, f \rangle$$

Galerkin method (Finite Element Method)

1. *Basis functions* $\{\varphi_i\}$
2. *Approximate* $u = \sum c_j \varphi_j$
3. *Determine* c_j *from* $\sum c_j \int \nabla \varphi_i \cdot \nabla \varphi_j = \int f \varphi_i$

The c_j are determined by the *linear system*

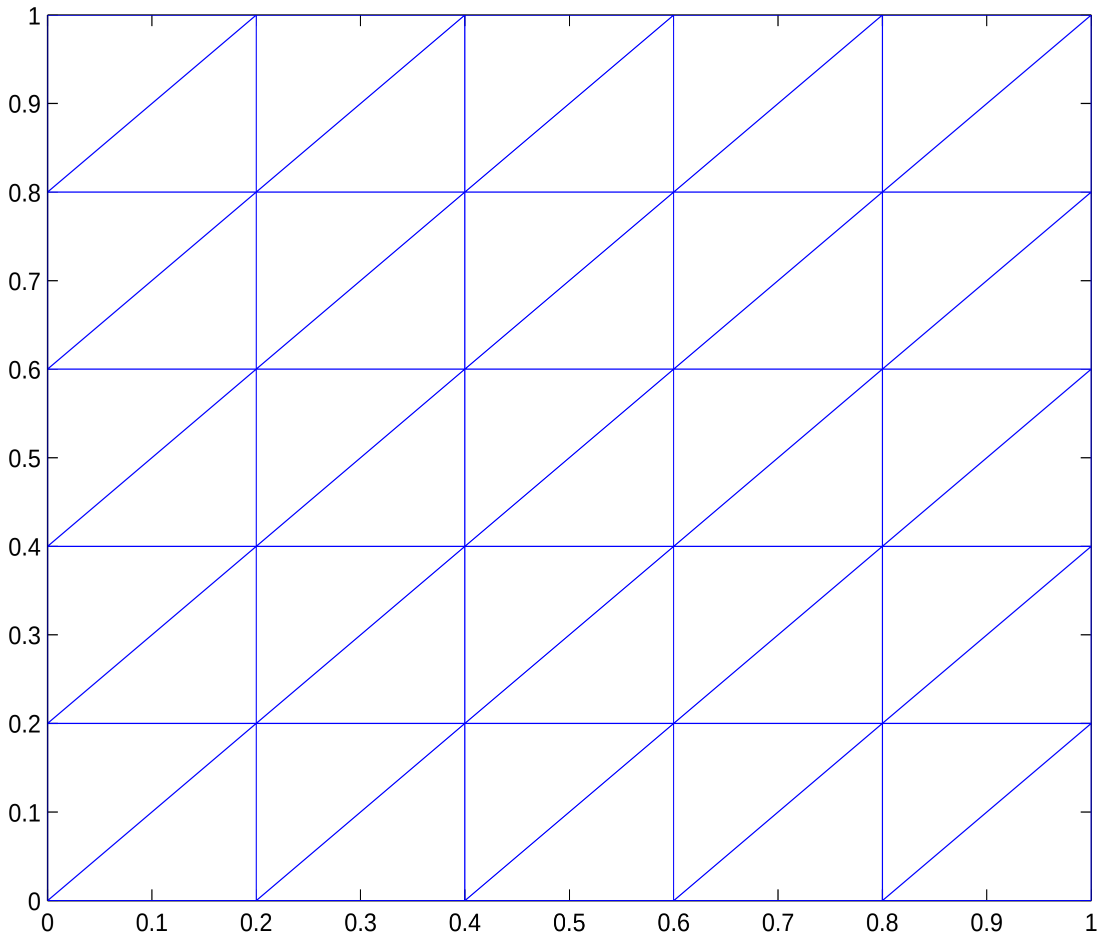
$$Kc = F$$

The matrix K is called *stiffness matrix*

Stiffness matrix elements $k_{ij} = \int \nabla \varphi_i \cdot \nabla \varphi_j = a(\varphi_i, \varphi_j)$
Right-hand side $F_i = \int \varphi_i f = \langle \varphi_i, f \rangle$

The FEM mesh. Domain triangulation

Piecewise linear basis $\{\varphi_j\}$ require domain *triangulation*



4. Parabolic problems

The prototypical equation is the

Diffusion equation $u_t = \Delta u$

Nonlinear diffusion

$$u_t = \mathrm{div}\,(k(u)\mathrm{grad}u)$$

Boundary and initial conditions are needed

Parabolic problems. Some applications

- **Diffusive processes**

  Heat conduction $u_t = d \cdot u_{xx}$

- **Chemical reactions**

  Reaction–diffusion $u_t = d \cdot u_{xx} + f(u)$

  Convection–diffusion $u_t = u_x + \frac{1}{\mathrm{Pe}} u_{xx}$

Irreversibility $u_t = -\Delta u$ *is not well-posed!*

Diffusion – a parabolic model problem

Equation	$u_t = u_{xx}$
Initial values	$u(x,0) = g(x)$
Boundary values	$u(0,t) = u(1,t) = 0$

Separation of variables $u(x,t) := X(x)T(t) \Rightarrow$

$$u_t = X\dot{T}, \quad u_{xx} = X''T \quad \Rightarrow \quad \frac{\dot{T}}{T} = \frac{X''}{X} =: \lambda$$

$$T = C\mathrm{e}^{\lambda t} \qquad X = A\sin\sqrt{-\lambda}\,x + B\cos\sqrt{-\lambda}\,x$$

Parabolic model problem...

Boundary values

$X(0) = X(1) = 0 \;\Rightarrow\; \lambda_k = -(k\pi)^2$, therefore

$$X_k(x) = \sqrt{2}\sin k\pi x \qquad T_k(t) = \mathrm{e}^{-(k\pi)^2 t}$$

Initial values

Fourier expansion $g(x) = \sum_1^\infty c_k \sqrt{2}\sin k\pi x \;\Rightarrow$

$$u(x,t) = \sqrt{2}\sum_{k=1}^{\infty} c_k \mathrm{e}^{-(k\pi)^2 t} \sin k\pi x$$

5. Method of lines (MOL) discretization

In $u_t = u_{xx}$, discretize $\partial^2/\partial x^2$ by

$$u_{xx} \approx \frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2}$$

System of ODEs (semidiscretization) $\dot{u} = T_{\Delta x} u$

$$\dot{u} = \frac{1}{\Delta x^2} \begin{pmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & & \ddots & \\ & & 1 & -2 \end{pmatrix} u$$

Full FDM discretization

Note that $u_i(t) \approx u(x,t)$ along the *line* $x = x_i$ *in* (x,t) *plane*

Use time-stepping to solve the IVP; Explicit Euler with $u_{i,j} \approx u(x_i, t_j)$ implies

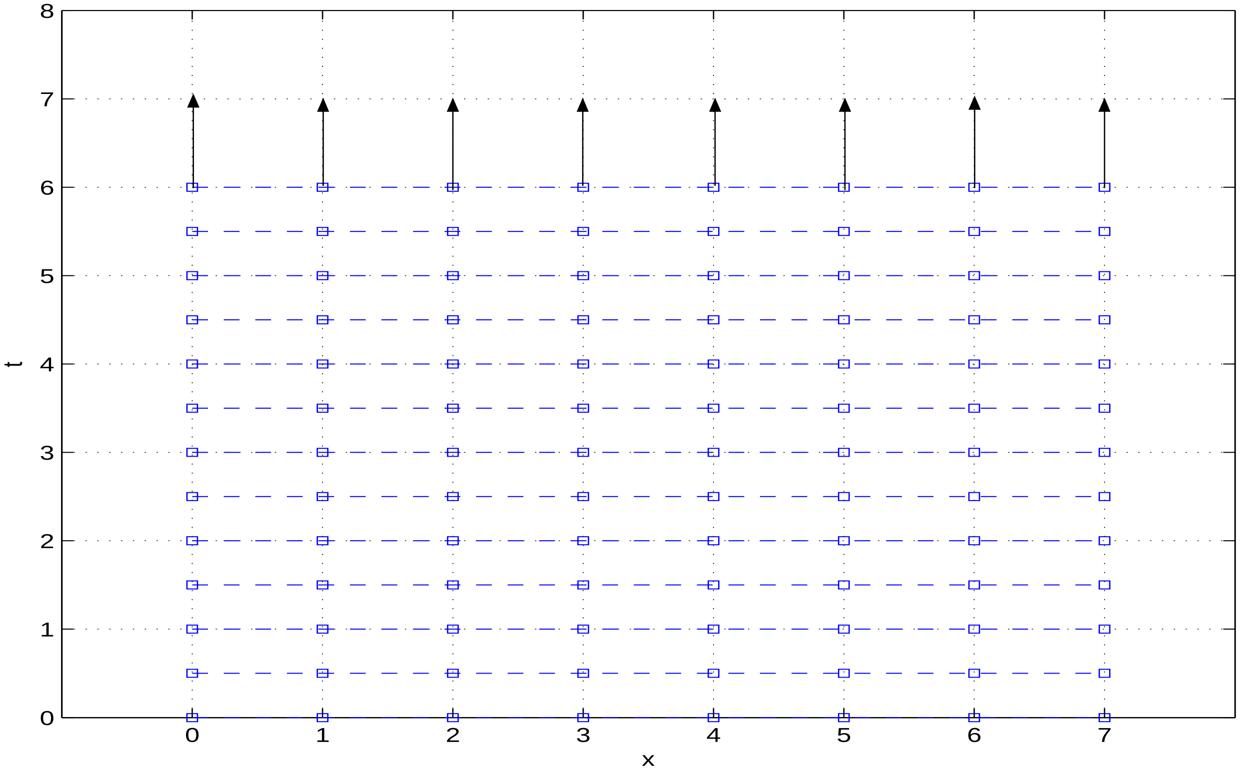
$$\frac{u_{i,j+1} - u_{i,j}}{\Delta t} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{\Delta x^2}$$

With the *Courant number* $\mu = \Delta t / \Delta x^2$ we obtain recursion

$$u_{i,j+1} = u_{i,j} + \mu \cdot (u_{i-1,j} - 2u_{i,j} + u_{i+1,j})$$

Method of lines. The grid

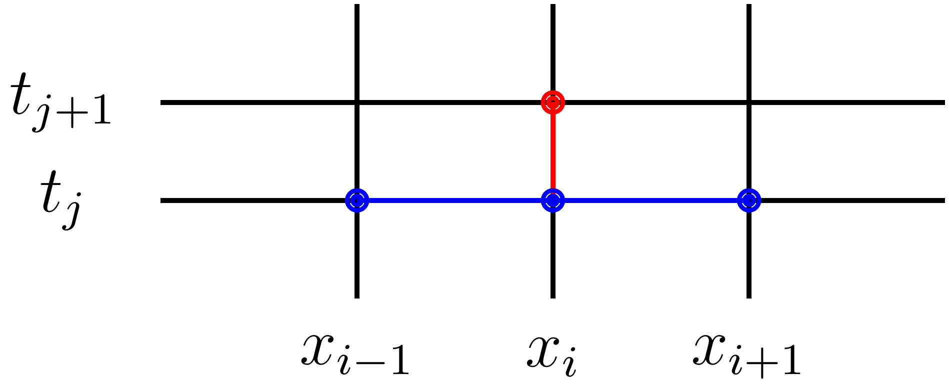
Rectangular grid $\{(i \cdot \Delta x, j \cdot \Delta t),\ i = 0 : N+1,\ j \geq 0\}$ with $\Delta x = 1/(N+1)$

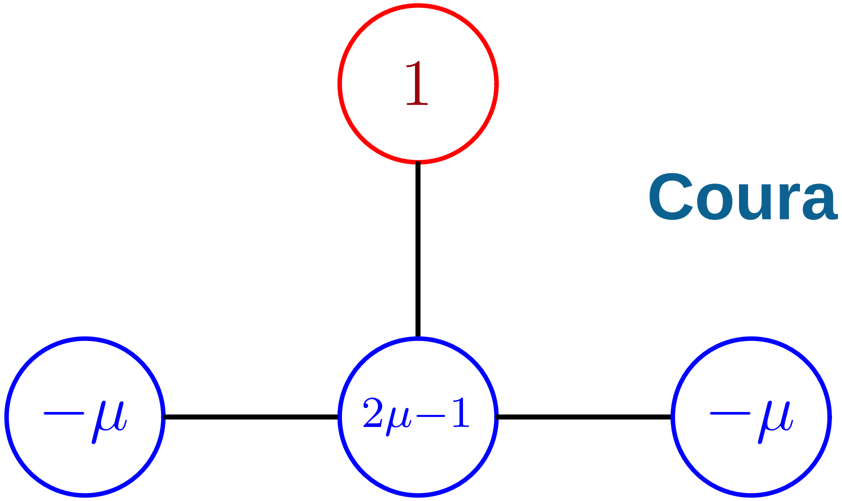


$$u_{i,j} \approx u(i \cdot \Delta x, j \cdot \Delta t)$$

Method of lines. Computational stencil

Explicit Euler time stepping. Participating grid points





Courant number $\mu = \Delta t / \Delta x^2$

Method of lines. Stability and the CFL condition

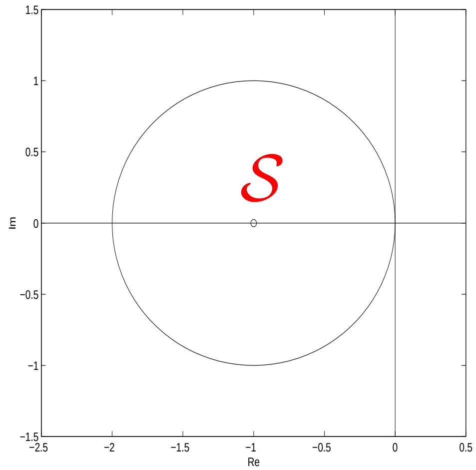
Explicit Euler with $\Delta x = 1/(N+1)$ implies recursion

$$u_{\cdot,j+1} = u_{\cdot,j} + \Delta t \cdot T_{\Delta x} u_{\cdot,j}$$

Recall $\lambda_k[T_{\Delta x}] = -4(N+1)^2 \sin^2 \frac{k\pi}{2(N+1)}$ for $k = 1 : N$

Stability *requires* $\Delta t \cdot \lambda_k \in \mathcal{S}$ *for all eigenvalues*

$$\Delta t \cdot \lambda_k \in \left[-\frac{4\Delta t}{\Delta x^2}, -\pi^2 \Delta t \right]$$



The CFL condition

For stability we need $4\Delta t/\Delta x^2 \leq 2$

CFL condition (Courant, Friedrichs, Lewy 1928)

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1}{2}$$

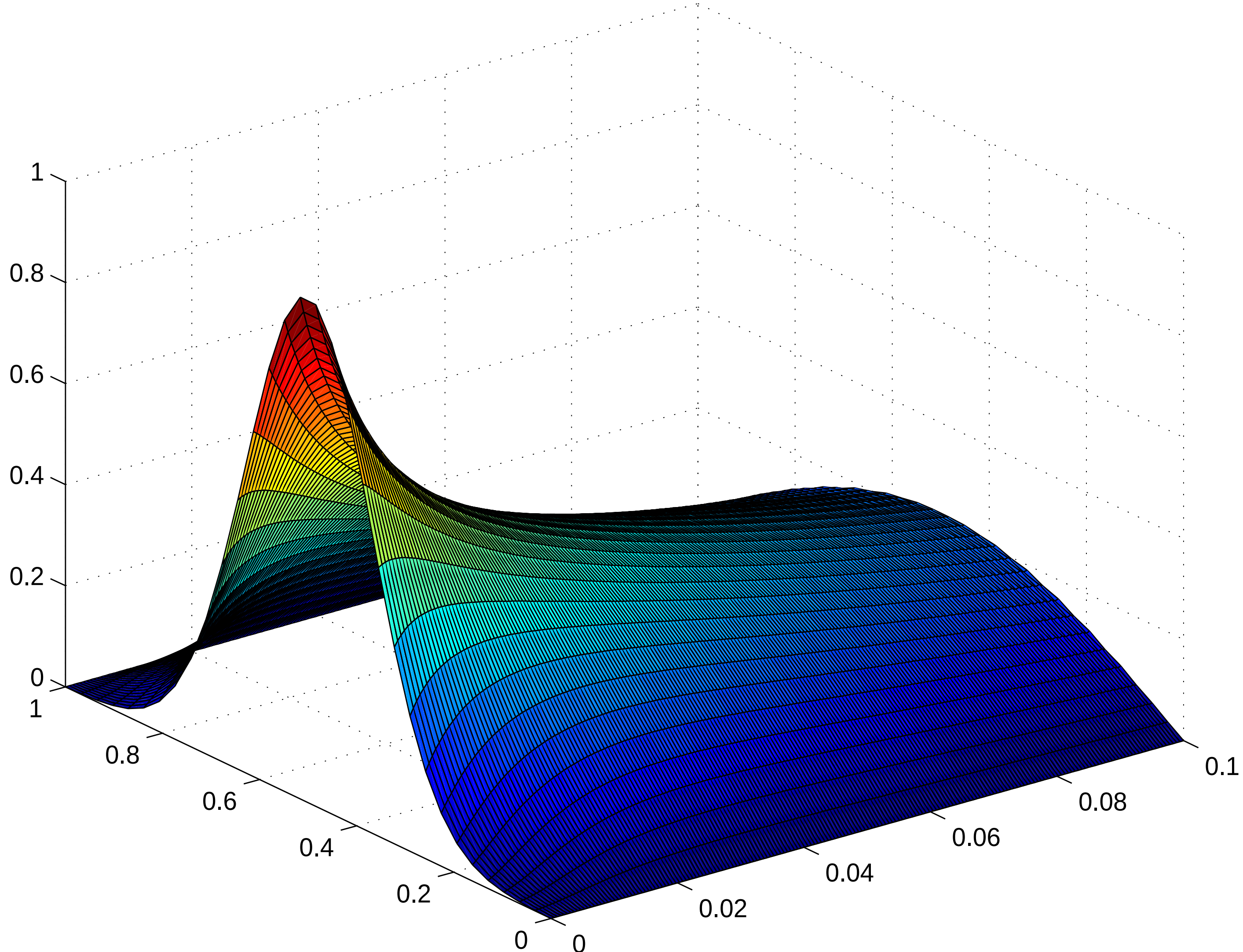
The CFL condition is a severe restriction on time step Δt

Stiffness The CFL condition can be avoided by using A-stable methods, e.g. Trapezoidal Rule or Implicit Euler

Experimental stability investigation

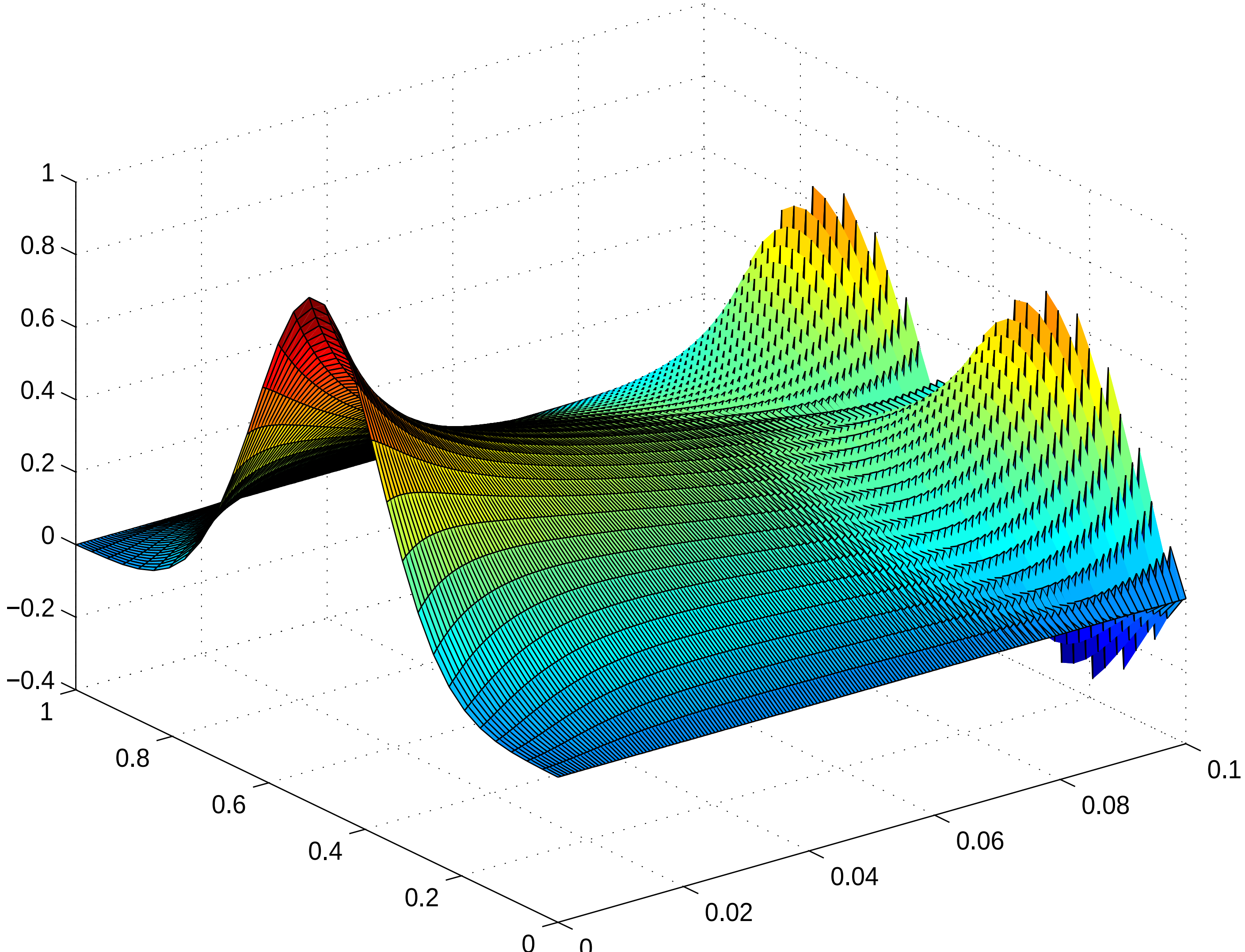
$N = 30$ internal pts in $[0, 1]$, $M = 187$ time steps on $[0, 0.1]$

Stable solution at CFL $= .514$



Violating the CFL condition. Instability

$N = 30$ internal pts in $[0, 1]$, $M = 184$ time steps on $[0, 0.1]$
Unstable solution at CFL $= .522$



Crank–Nicolson method (1947)

Crank–Nicolson method $\Leftrightarrow$ Trapezoidal Rule for PDEs

The trapezoidal rule is

- implicit $\Rightarrow$ *more work/step*
- A–stable $\Rightarrow$ *no restriction on* Δt

Theorem *Crank–Nicolson is unconditionally stable*

There is *no CFL condition* on the time-step Δt

Crank–Nicolson method...

Courant number $\mu = \Delta t / \Delta x^2 \quad \Rightarrow$ recursion

$$(I - \frac{\mu}{2} T) u_{\cdot, j+1} = (I + \frac{\mu}{2} T) u_{\cdot, j}$$

with Toeplitz matrix $T = \mathrm{tridiag}(1 \quad -2 \quad 1)$

Tridiagonal structure $\Rightarrow$ low complexity

Refactorize only if Courant number $\mu = \Delta t / \Delta x^2$ changes

6. Error analysis. Convergence

MOL with explicit Euler for $u_t = u_{xx}$

Global error $e_{i,j} = u_{i,j} - u(x_i, t_j)$

Local error Insert exact solution to get

$$\frac{u(x_i, t_{j+1}) - u(x_i, t_j)}{\Delta t} =$$
$$= \frac{u(x_{i-1}, t_j) - 2u(x_i, t_j) + u(x_{i+1}, t_j)}{\Delta x^2} - l_{i,j}$$

Expand *local error* $l_{i,j}$ in Taylor series

Local error

Taylor expansion

$$-l_{i,j} = u_t - u_{xx} + \frac{\Delta t}{2} u_{tt} + \frac{\Delta x^2}{12} u_{xxxx} + O(\Delta t^2, \Delta x^4)$$

Therefore

$$-l_{i,j} = \frac{\Delta t}{2} u_{tt} + \frac{\Delta x^2}{12} u_{xxxx} + O(\Delta t^2, \Delta x^4)$$

The Lax Principle

Conclusion

Consistency	$l_{i,j} \to 0$	as $\Delta t, \Delta x \to 0$
Stability	CFL condition	$\Delta t/\Delta x^2 \leq 1/2$
Convergence	$e_{i,j} \to 0$	as $\Delta t, \Delta x \to 0$

Theorem (Lax Principle)

Consistency + *Stability* $\Rightarrow$ *Convergence*

Note Choice of norm is very important

The order of the method

With local error

$$-l_{i,j} = \frac{\Delta t}{2} u_{tt} + \frac{\Delta x^2}{12} u_{xxxx} = O(\Delta t, \Delta x^2)$$

and stability in terms of CFL condition $\mu = \Delta t / \Delta x^2 \leq 1/2$
we have global error $e_{i,j} = O(\Delta t, \Delta x^2)$

For fixed μ we have $\Delta t \sim \Delta x^2$ and it follows that

Global error $e_{i,j} = O(\Delta t, \Delta x^2) = O(\Delta x^2) \ \Rightarrow$

Theorem *The order of convergence is* $p = 2$

Order of PDE discretizations

Discretization process from PDE to SD to FD

$$u_t = u_{xx} \to \dot{v}_{\Delta x} = P_{\Delta x} v_{\Delta x} + h_{\Delta x}(t) \to u_\mu^{j+1} = A_\mu u_\mu^j + k_\mu^j$$

Suppose

- order of SD scheme for spatial variables is p_1
- order of ODE time discretization is p_2

Theorem *If* $\Delta t = \mu \Delta x^2$ *the FD order of convergence is* $p = \min\{p_1, 2p_2\}$

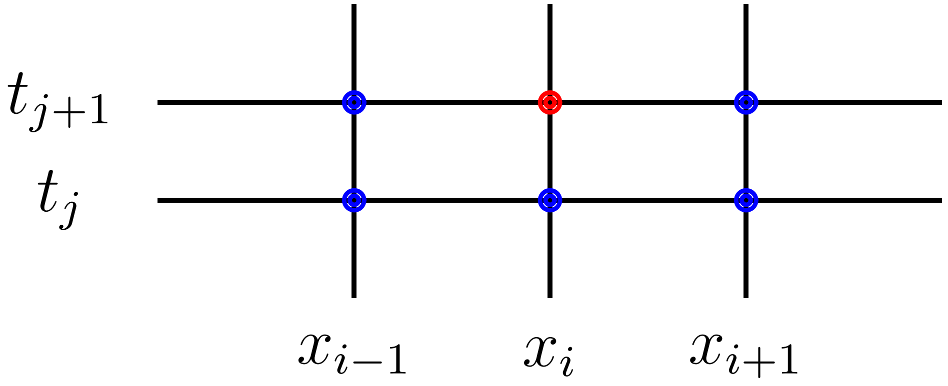
Example. The Crank–Nicolson method

From Trapezoidal rule $y_{n+1} = y_n + \frac{h}{2}[f(y_n) + f(y_{n+1})]$

$$u_i^{j+1} = u_i^j + \frac{\Delta t}{2\Delta x^2}[(u_{i-1}^j - 2u_i^j + u_{i+1}^j) + (u_{i-1}^{j+1} - 2u_i^{j+1} + u_{i+1}^{j+1})]$$

$$-\frac{\mu}{2}u_{i-1}^{j+1} + (1+\mu)u_i^{j+1} - \frac{\mu}{2}u_{i+1}^{j+1} = \frac{\mu}{2}u_{i-1}^j + (1-\mu)u_i^j + \frac{\mu}{2}u_{i+1}^j$$

Same order $p = \min\{2, 4\} = 2$ as with Explicit Euler



Crank–Nicolson. Stability

$A_\mu = (I - \frac{\mu}{2}T)^{-1}(I + \frac{\mu}{2}T)$ with usual Toeplitz matrix T

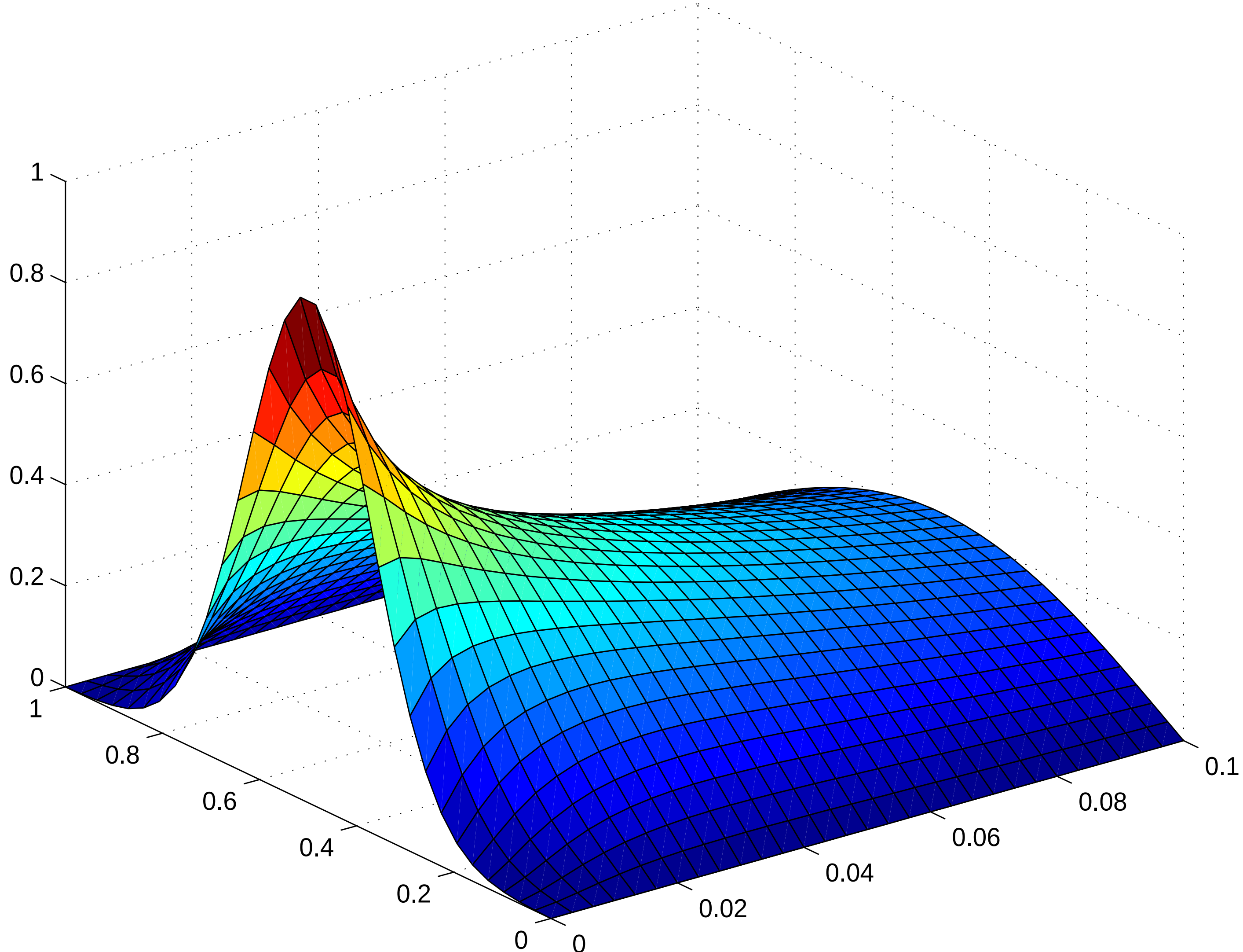
Theorem *The eigenvalues are* $\lambda[A_\mu] = \dfrac{1 + \frac{\mu}{2}\lambda[T]}{1 - \frac{\mu}{2}\lambda[T]}$

Note $\lambda[T] \in (-4, 0) \Rightarrow -1 < \lambda[A_\mu] < 1$ *This implies that there is no CFL stability condition on the Courant ratio* μ*!* *The method is stable for all* $\Delta t > 0$

Theorem *Crank–Nicolson is unconditionally stable*

Experimental stability investigation

$N = 30$ internal pts in $[0, 1]$, $M = 30$ time steps on $[0, 0.1]$
Stable solution at CFL $= 3.2$



7. Parabolic problems. FEM

Consider diffusion problem in strong form $u_t - u_{xx} = 0$ with Dirichlet boundary conditions

Multiply by test function v and integrate by parts

$$\int_0^1 v u_t \, \mathrm{d}x + \int_0^1 v' u' \, \mathrm{d}x = 0$$

In terms of inner product and energy norm –

Weak form $\langle v, u_t \rangle + a(v, u) = 0$ for all v with $v(0) = v(1) = 0$

Galerkin method (Finite Element Method)

1. *Basis functions* $\{\varphi_i\}$
2. *Approximate* $u(t,x) = \sum c_j(t)\varphi_j(x)$
3. *Determine* c_j *from* $\langle \varphi_i, u_t\rangle + a(\varphi_i, u) = 0$

Note $\langle \varphi_i, u_t\rangle = \sum \dot{c}_j \langle \varphi_i, \varphi_j\rangle$ and $a(\varphi_i, u) = \sum c_j \langle \varphi_i', \varphi_j'\rangle$

We get an initial value problem

$$B_{\Delta x}\dot{c} + K_{\Delta x}c = 0$$

for the determination of the coefficients $c_j(t)$ with $c(0)$ determined by the initial condition

Galerkin method

Simplest case
Piecewise linear basis functions on equidistant grid

Stiffness matrix elements $k_{ij} = \langle \varphi_i', \varphi_j' \rangle$

$$K_{\Delta x} = \frac{1}{\Delta x} \operatorname{tridiag}(-1 \quad 2 \quad -1)$$

and *mass matrix* elements $b_{ij} = \langle \varphi_i, \varphi_j \rangle$

$$B_{\Delta x} = \frac{\Delta x}{6} \operatorname{tridiag}(1 \quad 4 \quad 1)$$

Simplest Galerkin FEM...

Note that in the initial value problem

$$B_{\Delta x}\dot{c} + K_{\Delta x}c = 0$$

the matrix $B_{\Delta x}$ *is tridiagonal* $\Rightarrow$ *no advantage from explicit time stepping methods*

Explicit Euler

$$B_{\Delta x}(c_{n+1} - c_n) = -\Delta t \cdot K_{\Delta x}c_n$$

requires the solution of a tridiagonal system on every step

Simplest Galerkin FEM...

As the system is *stiff*, consider implicit A-stable method

$$B_{\Delta x}(c_{n+1} - c_n) = -\frac{\Delta t}{2} \cdot K_{\Delta x}(c_n + c_{n+1})$$

and solve tridiagonal system

$$(B_{\Delta x} + \frac{\Delta t}{2} K_{\Delta x}) c_{n+1} = (B_{\Delta x} - \frac{\Delta t}{2} K_{\Delta x}) c_n$$

on every step

Trapezoidal rule has *same cost, but better stability*

8. Well-posedness

Linear partial differential equation

$$\begin{aligned} u_t &= \mathcal{L}u + f, \quad 0 \le x \le 1, \quad t \ge 0, \quad u(x,0) = h(x), \\ u(0,t) &= \phi_0(t), \quad u(1,t) = \phi_1(t) \end{aligned}$$

Suppose $\begin{cases} w_t = \mathcal{L}w + f, & w(x,0) = h(x) \\ v_t = \mathcal{L}v + f, & v(x,0) = h(x) + g(x) \end{cases}$

Subtract to get *homogeneous Dirichlet problem*

$$u_t = \mathcal{L}u, \quad u(x,0) = g(x), \quad \phi_0(t) \equiv 0, \quad \phi_1(t) \equiv 0$$

Well-posedness. Time evolution

Suppose *time evolution* $u(x,t) = \mathcal{E}(t)g(x)$

Definition *The equation is well-posed if for every* $t^* > 0$ *there is a constant* $0 < C(t^*) < \infty$ *such that* $\|\mathcal{E}(t)\| \leq C(t^*)$ *for all* $0 \leq t \leq t^*$

Theorem *A well-posed equation has a solution that*

- *depends continuously on the initial value (the "data")*
- *is uniformly bounded in any compact interval*

$u_t = u_{xx}$ is well posed

Fourier series expansion $g(x) = \sqrt{2}\sum_1^\infty c_k \sin k\pi x$ implies

$$u(x,t) = \sqrt{2}\sum_{k=1}^{\infty} c_k \mathrm{e}^{-(k\pi)^2 t} \sin k\pi x$$

$$\begin{aligned}
\|\mathcal{E}(t)g\|_2^2 &= \int_0^1 |u(x,t)|^2 \,\mathrm{d}x \\
&= 2\sum_{k=1}^{\infty}\sum_{j=1}^{\infty} c_k c_j \mathrm{e}^{-(k^2+j^2)\pi^2 t} \int_0^1 \sin k\pi x \sin j\pi x \,\mathrm{d}x \\
&= \sum_{k=1}^{\infty} c_k^2 \mathrm{e}^{-2(k\pi)^2 t} \le \sum_{k=1}^{\infty} c_k^2 = \|g\|_2^2
\end{aligned}$$

Hence $\|\mathcal{E}(t)\|_2 \le 1$ for every $t \ge 0$