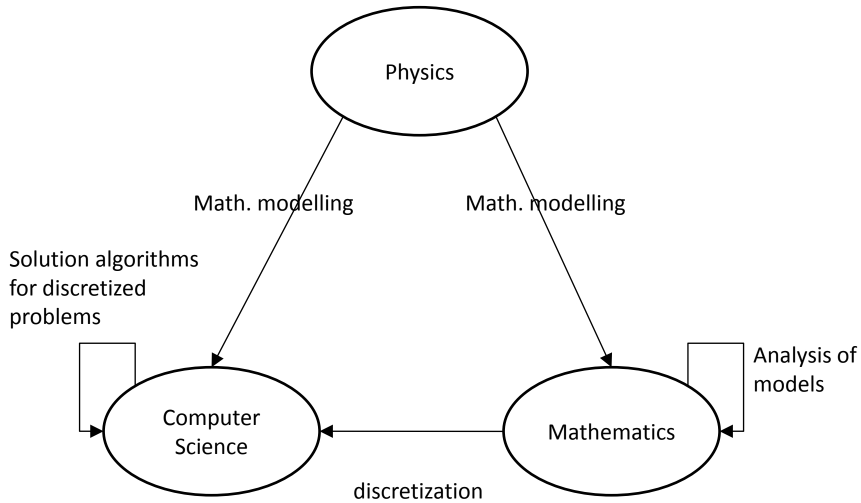
1 Introduction

Mathematical models describing physical phenomena are often written in the language of *(partial) differential equations*.



An initial value problem for an *ordinary differential equation* (ODE):

$$\begin{aligned} y''(t) &= f(t, y(t), y'(t)), \quad t > t_0 \\ y(t_0) &= y_0 \\ y'(t_0) &= y_1 \end{aligned}$$

A boundary value problem for an ordinary differential equation:

$$\begin{aligned} y''(t) &= f(t, y(t), y'(t)), \quad t_0 < t < t_1 \\ y(t_0) &= y_0 \\ y(t_1) &= y_1 \end{aligned}$$

General second order *partial differential equation* (PDE) of two variables:

$$\begin{aligned} &F(u, u_x, u_y, u_{xy}, u_{xx}, u_{yy}) = 0, \quad (x, y) \in \Omega \subset \mathbb{R}^2 \\ &u_x := \frac{\partial u}{\partial x}, \quad u_{xx} := \frac{\partial^2 u}{\partial x^2}, \quad u_{xy} := \frac{\partial^2 u}{\partial x \partial y}. \end{aligned}$$

To have a unique solution, PDE obviously needs some *extra conditions* (boundary condition on $\partial\Omega$ and/or initial condition)!
Some classical partial differential equations:

- Poisson equation

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

  (if $f \equiv 0$ then it is called Laplace's equation). Usually we write $\Delta u := u_{xx} + u_{yy}$ for short, i.e. the Poisson equation reads $-\Delta u = f$.

- Heat equation

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = f(x,t)$$

- Wave equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = f(x,t)$$

Abstract form

$$L(u) = f.$$

PDE is said to be *homogeneous* if $f = 0$.
PDE is said to be *linear* if

$$L(\alpha u + \beta v) = \alpha L(u) + \beta L(v) \quad \forall \alpha, \beta \in \mathbb{R}.$$

A PDE can have variable coefficients or it can be even nonlinear.

Example 1.1.

$$\begin{aligned} a u_x + b u_y &= 0 \quad \text{constant coefficients} \\ y u_x + x u_y &= 0 \quad \text{variable coefficients} \\ u u_x + x u_y &= 0 \quad \text{nonlinear + variable coeff.} \end{aligned}$$

Example 1.2. Laplace equation in unit disk $\Omega = B(0,1)$:

$$\begin{cases} \Delta u = 0 & \text{in } \Omega \\ u = g & \text{on } \partial\Omega \end{cases}$$

Solution u can be expressed as

$$u(x) = \frac{1}{2\pi} \int_{\partial\Omega} g(z) \frac{1 - |x|^2}{|z - x|^2} dS_z, \quad x \in \Omega.$$

Example 1.3. Poisson equation in $\Omega =]0,\pi[\times]0,\pi[$:

$$\begin{cases} \Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

Solution u can be expressed as

$$u(x_1, x_2) = -\sum_{\ell=1}^{\infty} \sum_{k=1}^{\infty} c_{\ell,k} \sin(\ell x_1) \sin(k x_2),$$

where

$$c_{\ell,k} = \frac{\pi^2}{4(\ell^2 + k^2)} \int_0^{\pi} \int_0^{\pi} f(x) \sin(\ell x_1) \sin(k x_2)\, dx_1\, dx_2.$$

From above examples one immediately sees that even for the simplest PDEs the solution can not generally be given in a simple analytical expression.
Therefore, for practical modeling tasks umerical solution methods are preferred. These include:

- finite difference methods
- weighted residual methods (finite element method FEM, finite volume method,...)

$$A(u) = f \implies \langle A(u) - f, w \rangle = 0$$

- integral equation methods

2 Finite difference method for elliptic boundary value problems

2.1 One-dimensional case

Consider the following second order boundary value problem:

$$\begin{cases} -u''(x) + q(x)u(x) = f(x), & a < x < b \\ u(a) = \alpha, \quad u(b) = \beta, \end{cases} \tag{1}$$

where $q, f \in C([a,b])$ and $q \geq 0$.
One can show that there exists a unique solution $u \in C^1([a,b])$ to (1).
Let us subdivide the interval $[a,b]$ in subintervals as follows:

$$a = x_0 < x_1 < \ldots < x_{n+1} = b, \quad x_j = a + jh, \quad h = \frac{b-a}{n+1}. \tag{2}$$

Let us suppose that $u \in C^4([a,b])$. Then it follows from Taylor's expansion that

$$\tau_i := u''(x_i) - \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} = -\frac{h^2}{12}u^{(4)}(x_i + \theta_i h), \quad |\theta_i| < 1. \tag{3}$$

Thus the "nodal values" $u(x_i)$ of the exact solution satisfy the system of equations

$$\begin{cases} u(x_0) = \alpha, \\ \dfrac{2u(x_i) - u(x_{i-1}) - u(x_{i+1})}{h^2} + q(x_i)u(x_i) = f(x_i) + \tau_i, \quad i = 1, ..., n, \\ u(x_{n+1}) = \beta. \end{cases} \tag{4}$$

Neglecting truncation errors τ_i the nodal values of the approximate solution u_h can be solved from the linear system of equations

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f}. \tag{5}$$

Here we denote by $\boldsymbol{u} = [u_1\ u_2\ \ldots\ u_n]^{\mathrm{T}}$, $u_i := u_h(x_i)$. The $n \times n$ matrix $\boldsymbol{A}$ is defined as follows:

$$\boldsymbol{A} = \frac{1}{h^2}\begin{bmatrix} 2 + q(x_1)h^2 & -1 & & & \\ -1 & 2 + q(x_2)h^2 & -1 & & \\ & & \ddots & & \\ & & & -1 & 2 + q(x_n)h^2 \end{bmatrix}.$$

And, finally the right-hand side vector is defined by $\boldsymbol{f} = [f_1 + \frac{\alpha}{h^2}\ f_2\ \ldots\ f_n + \frac{\beta}{h^2}]^{\mathrm{T}}$, $f_i := f(x_i)$, $i = 1, ..., n$.

2.2 On the convergence of the finite difference method

Let us next study the relation of the solutions to the original problem (1) and the matrix problem (5).
Let us first consider the situation in more abstract level. Let we have a continuous (linear) differential equation:

$$L(u) = f \tag{6}$$

and a discrete algebraic problem arising froma a finite difference discretization

$$L_h(u_h) = f_h. \tag{7}$$

Then, the question rises: On what condition (formally) $\|u - u_h\| \to 0,\ h \to 0$?

Definition 2.1. An approximate scheme is *consistent* if for *fixed* u it holds

$$\begin{cases} L_h(u) \to L(u), & h \to 0 \\ f_h \to f, & h \to 0. \end{cases} \tag{8}$$

Definition 2.2. An approximate scheme is *stable* if the inverse of the discrete operator is uniformly bounded, i.e.

$$\|L_h^{-1}\| \le C \quad \forall h > 0. \tag{9}$$

Theorem 2.1. *A scheme if consistent and stable $\implies$ the scheme is convergent.*

Proof.

$$\begin{aligned} \|u - u_h\| = \|u - L_h^{-1}(L(u) - f) - u_h\| &= \|L_h^{-1}L_h(u) - L_h^{-1}(u) + L_h^{-1}f - L_h^{-1}f_h\| \\ &\le \|L_h^{-1}\|\|L_h(u) - L(u)\| + \|L_h^{-1}\|\|f - f_h\| \\ &\underset{\text{stability}}{\le} C\|L_h(u) - L(u)\| + C\|f - f_h\| \underset{\text{consistency}}{\longrightarrow} 0. \end{aligned}$$

The last inequality follows from the stability and the limit from the consistency. □

Theorem 2.2. *If the solution of boundary value problem* (1) *satisfies* $|u^{(4)}(x)| \le M\ \forall x \in [a, b]$, *then*

$$\|L_h(u) - L(u)\| := \max_{1\le i\le n} |(L_h u)(x_i) - (Lu)(x_i)| \to 0, \quad h \to 0. \tag{10}$$

Thus the scheme (4) *is consistent.*

Proof. is evident. □

Theorem 2.3. *If* $q \ge 0$ *then* $\boldsymbol{A}$ *is positive definite and* $0 \le \boldsymbol{A}^{-1} \le \tilde{\boldsymbol{A}}^{-1}$, *where*

$$\tilde{\boldsymbol{A}} = \frac{1}{h^2}\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & & \ddots & \\ & & -1 & 2 \end{bmatrix}.$$

Theorem 2.4. *If the solution of boundary value problem* (1) *satisfies* $|u^{(4)}(x)| \le M\ \forall x \in [a, b]$, *then*

$$|u(x_i) - u_i| \le \frac{Mh^2}{96}(b - a)^2. \tag{11}$$

Proof. Let us denote

$$\|u - u_h\| := \max_{1\le i\le n} |u(x_i) - u_h(x_i)| = \|\boldsymbol{u}^* - \boldsymbol{u}\|_\infty,$$

where

$$\begin{aligned} \boldsymbol{u}^* &= [u(x_1)\ u(x_2)\ \ldots\ u(x_n)]^{\mathrm{T}} \\ \boldsymbol{u} &= [u_h(x_1)\ u_h(x_2)\ \ldots\ u_h(x_n)]^{\mathrm{T}} =: [u_1\ u_2\ \ldots\ u_n]^{\mathrm{T}}. \end{aligned}$$

As $\|f - f_h\| = 0$ it follows that

$$\|\boldsymbol{u}^* - \boldsymbol{u}\|_\infty \le \|A^{-1}\|_\infty \|\tau\|_\infty. \tag{12}$$

Let us show that $\|A^{-1}\|_\infty \le C \ \forall h > 0$:

$$\|\boldsymbol{A}^{-1}\|_\infty = \max_{\|y\|_\infty = 1} \|\boldsymbol{A}^{-1}\boldsymbol{y}\|_\infty = \max_{\|y\|_\infty=1} \max_i \left| \sum_j (\boldsymbol{A}^{-1})_{ij} y_j \right|$$

$$\left| \sum_j (\boldsymbol{A}^{-1})_{ij} y_j \right| \le \sum_j (\boldsymbol{A}^{-1})_{ij} |y_j| \le \sum_j (\boldsymbol{A}^{-1})_{ij} e_j,$$

where $\boldsymbol{e} = [1\ 1\ \dots\ 1]^{\mathrm{T}}$. Thus $\|\boldsymbol{A}^{-1}\|_\infty \le \|A^{-1}\boldsymbol{e}\|_\infty$.
Consider problem

$$\begin{cases} -z''(x) = 1, & a < x < b \\ \qquad z(a) = z(b) = 0. \end{cases}$$

Its exact solution is $z(x) = \frac{1}{2}(x-a)(b-x)$ and the finite difference scheme (4) is exact too, i.e. $\tau_i = 0,\ i = 1, \dots, n$, and the nodal values satisfy the equation

$$\tilde{\boldsymbol{A}}\boldsymbol{z}^* = \boldsymbol{e}.$$

Thus,

$$\begin{aligned} \|\boldsymbol{A}^{-1}\|_\infty &\le \|\tilde{\boldsymbol{A}}^{-1}\|_\infty \le \|\tilde{\boldsymbol{A}}^{-1}\boldsymbol{e}\|_\infty = \|\boldsymbol{z}^*\|_\infty \\ &= \max_{1 \le i \le n} \left|\frac{1}{2}(x_i - a)(b - x_i)\right| \le \max_{a \le x \le b} \frac{1}{2}(x-a)(b-x) = \frac{1}{8}(b-a)^2 =: C. \end{aligned}$$

As $\|\tau\|_\infty \le \frac{h^2}{12}M$ the inequality (11) follows.

□

2.3 Finite difference method for Poisson problem in a rectangle

Consider the two-dimensional Poisson equation in a rectangle:

$$\begin{cases} -\Delta u(x,y) = f(x,y), & \text{on } \Omega :=]a,b[\times]c,d[, \\ \quad u(x,y) = g(x,y), & \text{on } \partial\Omega. \end{cases} \tag{13}$$

We form $(n+2) \times (m+2)$ grid (see picture below) on Ω, defined by the points

$$(ih, jh),\ i = 0, \dots, n+1,\ j = 0, \dots, m+1, \qquad h_1 = \frac{b-a}{n+1}, \qquad h_2 = \frac{d-c}{m+1}$$

```
o---o---o---o---o---o---o
|   |   |   |   |   |   |
o---o---o---o---o---o---o
|   |   |   |   |   |   |
o---o---o---o---o---o---o
|   |   |   |   |   |   |
o---o---o---o---o---o---o
|   |   |   |   |   |   | h2
o---o---o---o---o---o--o
 h1
```

We denote the nodal values of the problemdata f, g and the approximate solution u_h by

$$u_{ij} := u_h(a + ih_1, c + jh_2), \quad f_{ij} := f(a + ih_1, c + jh_2), \quad g_{ij} := g(a + ih_1, c + jh_2).$$

The central difference approximation (5-point formula) is the generalization of the one-dimensional finite difference method and reads:

$$-\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_1^2} - \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_2^2} = f_{i,j}, \quad -i = 1, ..., n, \quad j = 1, ..., m. \tag{14}$$

Taking into account known boundary values one gets e.g.

$$\left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) u_{11} - \frac{1}{h_1^2} u_{21} - \frac{1}{h_2^2} u_{12} = f_{11} + \frac{1}{h_1^2} g_{01} + \frac{1}{h_2^2} g_{10}.$$

In the simplest case $g \equiv 0,\ n = m,\ h_1 = h_2 =: h$, the system of equations (14) takes the form

$$4u_{i,j} - u_{i+1,j} - u_{i-1,j} - u_{i,j+1} - u_{i,j-1} = h^2 f_{i,j}, \quad i, j = 1, ..., n. \tag{15}$$

If we use the "natural" numbering of internal nodes, e.g.

```
+   +   +   +   +   +
+  13  14  15  16   +
+   9  10  11  12   +
+   5   6   7   8   +
+   1   2   3   4   +
+   +   +   +   +   +
```

the resulting coefficient matrix is pentadiagonal and can be in *block tridiagonal form*

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{T} & -\boldsymbol{I} & \dots & & & \\ -\boldsymbol{I} & \boldsymbol{T} & -\boldsymbol{I} & & & \\ & -\boldsymbol{I} & \boldsymbol{T} & -\boldsymbol{I} & & \\ & & & \ddots & & \\ & & & -\boldsymbol{I} & \boldsymbol{T} & -\boldsymbol{I} \\ & & & & -\boldsymbol{I} & \boldsymbol{T} \end{bmatrix} \in \mathbb{R}^{n^2 \times n^2}, \tag{16}$$

where

$$\boldsymbol{T} = \begin{bmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & & \ddots & & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad \boldsymbol{I} = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

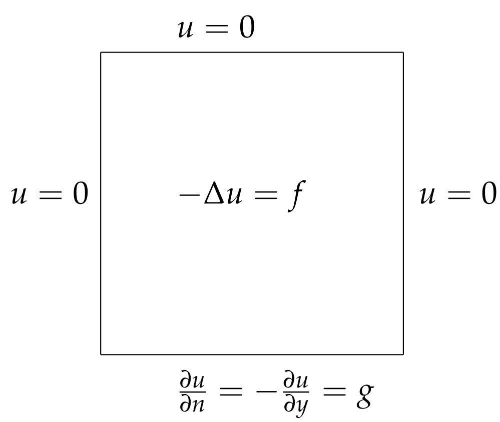
Matrix $\boldsymbol{A}$ is symmetric and positive definite band matrix.

Remark 2.1. Iterative methods for the numerical solution of equation (15) do not require explicit construction of the coefficient matrix $\boldsymbol{A}$. For example the Gauss–Seidel method can be implemented as

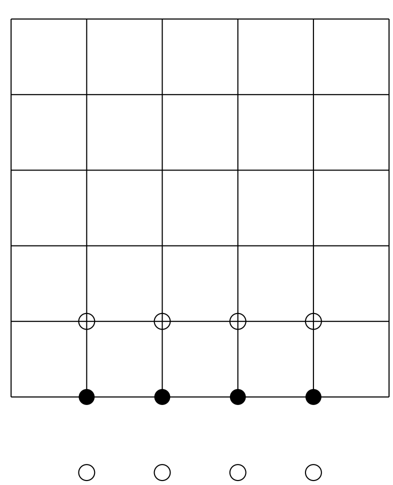
```
u=zeros(n+2);
for iter=1:itmax
    for j=2:n+1
        for i=2:n+1
            u(i,j)=0.25*(h*h*f(i,j)+u(i+1,j)+u(i-1,j) ...
                  +u(i,j+1)+u(i,j-1));
        end
    end
end
```

Neumann boundary value condition

Let us assume that $\Omega =]0,1[\times]0,1[$ and $h_1 = h_2 = h$. Consider the following mixed boundary value problem for the Poisson equation:



As the central difference approximation inside Ω is $\mathcal{O}(h^2)$ accurate, the same should hold also on the boundary. For central difference approximation of $\frac{\partial u}{\partial y}$ we introduce a set of "ghost nodes" outside of Ω.



Then the discrete analogue of $-\frac{\partial u}{\partial y} = g$ reads

$$\begin{cases} -\dfrac{u_{11} - u_{1,-1}}{2h} = g_{10} & \iff \quad u_{1,-1} = u_{1}1 + 2hg_{10} \\ -\dfrac{u_{21} - u_{2,-1}}{2h} = g_{20} & \iff \quad u_{2,-1} = u_{21} + 2hg_{20} \\ \qquad\qquad\vdots & \\ -\dfrac{u_{n1} - u_{n,-1}}{2h} = g_{n0} & \iff \quad u_{n,-1} = u_{n1} + 2hg_{n0} \end{cases}$$

We eliminate the ghost values (those with negative indices) by assuming that the Poisson equation is satisfied also on the boundary. For example

$$4u_{10} - \underbrace{u_{00}}_{=0} - u_{20} - u_{1,-1} - u_{11} = h^2 f_{10}$$

$$4u_{10} - u_{20} - u_{11} - u_{11} = h^2 f_{10} + 2hg_{10}.$$

Thus, in general we have:

$$4u_{i,0} - u_{i-1,0} - u_{i+1,0} - 2u_{i,1} = h^2 f_{i,0} + 2hg_{i,0}, \quad i = 1, ..., n.$$

In matrix form the previous mixed boundary value problem reads

$$\begin{bmatrix} \boldsymbol{T} & -2\boldsymbol{I} & & & \\ -\boldsymbol{I} & \boldsymbol{T} & -\boldsymbol{I} & & \\ & \ddots & \ddots & \ddots & \\ & & -\boldsymbol{I} & \boldsymbol{T} & -\boldsymbol{I} \\ & & & -\boldsymbol{I} & \boldsymbol{T} \end{bmatrix} \left[\begin{array}{c} u_{10} \\ u_{20} \\ \vdots \\ u_{n0} \\ \hline u_{11} \\ u_{21} \\ \vdots \\ \hline \vdots \\ \hline u_{1n} \\ \vdots \\ u_{nn} \end{array}\right] = \left[\begin{array}{c} h^2 f_{10} + 2hg_{10} \\ h^2 f_{20} + 2hg_{20} \\ \vdots \\ h^2 f_{n0} + 2hg_{n0} \\ \hline h^2 f_{11} \\ h^2 f_{21} \\ \vdots \\ \hline \vdots \\ \hline h^2 f_{1n} \\ \vdots \\ h^2 f_{nn} \end{array}\right] \tag{17}$$