Contents

Plan

- Favourable stability properties and selectable order
- Theoretical concepts also used in the numerical solution of PDEs
- Very elegant approach to error control and adaptivity

A Different Philosophy

- In traditional methods, such as Runge-Kutta, on approximates the unknown function $u(t)$ at temporal values t_i
- In the Galerkin method we approximate $u(t)$ by simple functions such as polynomials
- In this way the approximation is defined at all points in time
- Also the error $e(t) = u(t) - U(t)$ is defined at all times
- This allows the use of more sophisticated mathematical methods for their analysis
- The presentation follows chapters 6 and 9 from the book "Computational Differential Equations" by Eriksson, Estep, Hansbo and Johnson

Variational (or Weak) Formulation

- Consider the first-order systems of ODEs in $\mathbb{R}^d$ in explicit form

$$u'(t) - f(t, u(t)) = 0, \qquad t \in (t_0, t_0 + T], \qquad u(t_0) = u_0$$

  to determine the unknown function $u : [t_0, t_0 + T] \to \mathbb{R}^d$
- Given a suitable function $\varphi : [t_0, t_0 + T] \to \mathbb{R}^d$ we may multiply and integrate:

$$\int_{t_0}^{t_0+T} (u'(t) - f(t, u(t))) \cdot \varphi(t)\, dt = 0$$

  where $\cdot$ denotes the Euclidean scalar product.
- Demanding this identity for a sufficiently large class of functions $\varphi \in V_0 = \{v : v(t_0) = 0\}$, we may hope this fixes (uniquely) a function $u \in U = \{w : w(t_0) = u_0\}$
- This function is called a *variational (or weak) solution of the ODE*
- Under suitable conditions the weak and strong solution coincide
- The *function* $R[u]$, $R[u](t) = u'(t) - f(t, u(t))$ is called *residual*

Galerkin Method

- The function space V_0 is infinite-dimensional, e.g. all continuous functions (with zero initial value)
- **Idea**: Replace V_0 and U by *finite-dimensional* counter parts!

Example: Use global polynomials. Let us fix $d = 1$

- Define the following classes of polynomials:

$$\begin{aligned}
\mathcal{P}^q &= \text{polynomials of degree } q \\
\mathcal{P}_0^q &= \{p \in \mathcal{P}^q : p(t_0) = 0\} \\
U^q &= \{p \in \mathcal{P}^q : p = u_0 + v, v \in \mathcal{P}_0^q\} =: u_0 + \mathcal{P}_0^q
\end{aligned}$$

  and note: $\mathcal{P}^q$ is a vector space of dimension $q+1$, $\mathcal{P}_0^q$ is a proper subspace of dimension q, U^q is called an affine space
- Then the global Galerkin method reads: Find $U(t) \in U^q$ such that:

$$\int_{t_0}^{t_0+T} (U'(t) - f(t, U(t)))\varphi(t)\, dt = 0 \qquad \forall \varphi \in \mathcal{P}_0^q$$

Galerkin Method. Example continued

- How to make this method practical?
- Choose a basis representation:

$$\mathcal{P}^q = \text{span}\{1, t-t_0, \ldots, (t-t_0)^q\}, \qquad \mathcal{P}_0^q = \text{span}\{t-t_0, \ldots, (t-t_0)^q\}$$

- Make the ansatz $U(t) = u_0 + \sum_{j=1}^{q} \xi_j (t-t_0)^j$ and insert:

$$\int_{t_0}^{t_0+T} (U'(t) - f(t, U(t)))\varphi(t)\, dt = 0 \quad \forall \varphi \in \mathcal{P}_0^q$$

$$\int_{t_0}^{t_0+T} \left(\sum_{j=1}^{q} \xi_j j (t-t_0)^{j-1} - f\left(t, u_0 + \sum_{j=1}^{q} \xi_j (t-t_0)^j \right)\right)(t-t_0)^i\, dt = 0 \quad 1 \le i \le q$$

$$\sum_{j=1}^{q} \xi_j j \frac{T^{i+j}}{i+j} - \int_{t_0}^{t_0+T} f\left(t, u_0 + \sum_{j=1}^{q} \xi_j (t-t_0)^j \right)(t-t_0)^i\, dt = 0 \quad 1 \le i \le q$$

- Need to solve q coupled nonlinear equations for the coefficients ξ_j

Some Choices

The accuracy of the method can be controlled by

- Increasing the polynomial degree (called *p*-method)
  - Algebraic problem might become ill-conditioned
  - Remedied by choosing an appropriate basis
  - Needs sufficient regularity of the solution of the ODE
- Using piecewise polynomials of degree q (called *h*-method)
  - We will follow this approach below
- Combination of both (called *hp*-method)
- Using of trigonometric polynomials (spectral method)

- *Error control*: What is the error in the computed solution $U(t)$?
- *Adaptivity*: How to choose q and Δt_i to control the error?

Piecewise Polynomial Functions

- As before we treat $d = 1$, extend to arbitrary d by making each component a polynomial
- Choose N time steps as before

$$t_0 < t_1 < t_2 < \ldots < t_{N-1} < t_N = t_0 + T, \qquad \Delta t_i = t_{i+1} - t_i,$$
$$I = (t_0, t_0 + T), \quad I_i = (t_i, t_{i+1}), \quad \mathcal{T}_N = \{I_i : 0 \le i < N\}.$$

- Continuous piecewise polynomials of degree q are

$$V_N^q = \{v \in C^0(\bar{I}) : v|_{I_i} \in \mathcal{P}^q, 0 \le i < N\}$$

- Continuous piecewise polynomials of degree q with zero initial value

$$V_{N,0}^q = \{v \in V_N^q : v(t_0) = 0\} \subset V_N^q$$

- Discontinuous piecewise polynomials:

$$W_N^q = \{v \in L^2(I)) : v|_{I_i} \in \mathcal{P}^q, 0 \le i < N\}$$

- By $v_i = v|_{I_i}$ we denote the piece on interval I_i

cG(q) Method

Find $U(t) \in u_0 + V^q_{N,0}$ such that

$$\int_{t_0}^{t_0+T} (U'(t) - f(t, U(t)))\varphi(t)\, dt = 0 \qquad \forall \varphi \in W_N^{q-1}$$

- Note the use of test functions $\mathcal{P}^{q-1}$ instead of $\mathcal{P}^q_0$ on I_i
- The choice of *discontinuous* test functions is essential, since it allows to solve the problem sequentially! The solution U can be determined as follows
- Consider $I_0 = (t_0, t_1]$, restricting the test functions $\phi_i = 0$, $i > 1$:

$$\text{Find } U_0(t) \in u_0 + \mathcal{P}^q_0\colon \int_{I_0} (U_0'(t) - f(t, U_0(t)))\varphi(t)\, dt = 0 \qquad \forall \varphi \in \mathcal{P}^{q-1}$$

- Considering $I_i = (t_i, t_{i+1}]$, $i > 0$, assume $U_{i-1}(t)$ is available:

$$\text{Find } U_i(t) \in U_{i-1}(t_i) + \mathcal{P}^q_0\colon \int_{I_i} (U_i'(t) - f(t, U_i(t)))\varphi(t)\, dt = 0 \qquad \forall \varphi \in \mathcal{P}^{q-1}$$

- The value at the end of interval I_{i-1} is used as initial value in I_i

dG(q) Method

- Now we approximate U in W_N^q, i.e. U might be discontinuous at t_i
- For $v \in W_N^q$ introduce the notation

$$v_i^+ = v_i(t_i), \quad v_i^- = v_{i-1}(t_i), \quad v_0^- = v_0 \text{ } (v_0 \text{ a given number})$$

and the *jump*

$$[v]_i = v_i^+ - v_i^-, \qquad 0 \leq i < N$$

- Then the dG(q) methods reads: Find $U(t) \in W_N^q$ such that

$$\sum_{i=0}^{N-1} \left\{ \int_{I_i} (U'(t) - f(t, U(t)))\varphi(t)\, dt + [U]_i \varphi_i^+ \right\} = 0 \qquad \forall \varphi \in W_N^q$$

- Note that both, the solution and the test functions, are in W_N^q
- Of course this needs some explanation

dG(q) Method: Sequential Solution

- Without the jump term the solutions in the intervals I_i would be completely independent of each other, with the jump term we get
- In the interval $I_0 = (t_0, t_1]$: Find $U_0(t) \in \mathcal{P}^q$:

$$\int_{I_0} (U_0'(t) - f(t, U_0(t)))\varphi(t)\, dt + (U_0(t_0) - u_0)\varphi_0(t_0) = 0 \qquad \forall \varphi \in \mathcal{P}^q$$

- In interval $I_i = (t_i, t_{i+1}]$, $i > 0$: Find $U_i(t) \in \mathcal{P}^q$:

$$\int_{I_i} (U_i'(t) - f(t, U_i(t)))\varphi(t)\, dt + (U_i(t_i) - U_{i-1}(t_i))\varphi_i(t_i) = 0 \qquad \forall \varphi \in \mathcal{P}^q$$

- Thus, the U_i are determined sequentially by solving $q+1$ nonlinear algebraic equations in each interval
- The method can be extended to the vector-valued case

dG(q) Method: Jump Term Explained (1)

- Consider the simple problem

$$u'(t) = 0 \text{ in } (t_0, t_0 + T], \qquad u(t_0) = u_0$$

  which has the constant solution $u(t) = u_0$
- Using the weak formulation we obtain using integration by parts:

$$\sum_{i=0}^{N-1} \int_{I_i} U_i'(t)\varphi_i(t)\,dt = \sum_{i=0}^{N-1} \left\{ -\int_{I_i} U_i(t)\varphi_i'(t)\,dt + U_i(t_{i+1})\varphi(t_{i+1}) - U_i(t_i)\varphi(t_i) \right\} = 0$$

dG(q) Method: Jump Term Explained (2)

- With a small change we obtain the correct solution:

$$\sum_{i=0}^{N-1} \left\{ -\int_{I_i} U_i(t)\varphi_i'(t)\,dt + U_i(t_{i+1})\varphi(t_{i+1}) - U_{i-1}(t_i)\varphi(t_i) \right\} = 0$$

(Observe that $U_i(t) = U_{i-1}(t_i)$ solves the problem in each interval)

- The change may be expressed as

$$\sum_{i=0}^{N-1} \int_{I_i} U_i'(t)\varphi_i(t)\,dt + U_0(t_0)\varphi_0(t_0) - u_0\varphi_0(t_0) + \sum_{i=1}^{N-1} \left(U_i(t_i)\varphi_i(t_i) - U_{i-1}(t_i)\varphi_i(t_i) \right) = 0$$

- This is exactly the jump term in the formulation

Example: cG(1) Method

- Recall: Find $U_i(t) \in U_{i-1}(t_i) + \mathcal{P}_0^1$:

$$\int_{I_i} (U_i'(t) - f(t, U_i(t)))\varphi(t)\,dt = 0 \qquad \forall \varphi \in \mathcal{P}^0$$

- $\mathcal{P}^0 = \text{span}\{1\}$, for $U_i(t) \in U_{i-1}(t_i) + \mathcal{P}_0^1$ make the Ansatz

$$U_i(t) = \underbrace{y_i^N}_{=U_{i-1}(t_i)} \underbrace{\frac{t_{i+1} - t}{\Delta t_i}}_{\psi_i^0} + y_{i+1}^N \underbrace{\frac{t - t_i}{\Delta t_i}}_{\psi_i^1}$$

- Inserting the Ansatz into the formulation and $\varphi = 1$:

$$\int_{I_i} y_i^N \left(-\frac{1}{\Delta t_i}\right) + y_{i+1}^N \frac{1}{\Delta t_i} - f(t, y_i^N \psi_i^0(t) + y_{i+1}^N \psi_i^1(t))\,dt = 0$$

$$\Leftrightarrow \quad y_{i+1}^N - y_i^N - \int_{I_i} f(t, y_i^N \psi_i^0(t) + y_{i+1}^N \psi_i^1(t))\,dt = 0$$

- Using 2nd order quadrature yields the implicit trapezoidal rule or the implicit midpoint rule

Example: dG(0) Method

- Recall: Find $U(t) \in W_N^0$ such that

$$\sum_{i=0}^{N-1} \left\{ \int_{I_i} (U'(t) - f(t, U(t)))\varphi(t)\, dt + [U]_i \varphi_i^+ \right\} = 0 \qquad \forall \varphi \in W_N^0$$

- Choose the basis and the Ansatz

$$\psi_i(t) = \begin{cases} 1 & t \in I_i \\ 0 & \text{else} \end{cases}, \qquad U(t) = \sum_{i=0}^{N-1} y_{i+1}^N \psi_i(t)$$

  observe that due to $I_i = (t_i, t_{i+1}]$ one may interpret y_{i+1}^N as the value at the end of the time interval I_i

- Inserting in formulation gives

$$-\int_{I_i} f(t, y_{i+1}^N), dt + y_{i+1}^N - y_i^N = 0$$

- Which yields the implicit Euler method upon quadrature:

$$y_{i+1}^N - y_i^N - \Delta t_i f(t, y_{i+1}^N), dt = 0$$

Error Control and Adaptivity

- **Error control**: Stop the computation when

$$J(u - U) \leq TOL$$

where J is some functional of the error $e = u - U$ and TOL is a user given tolerance

- **Adaptivity**: Choose $\mathcal{T}_N$ such that $J(u - U) \leq TOL$ is achieved with N as small as possible
- An example would be $J(e) = e(T)$
- Adaptive time step control for traditional methods tries to estimate the leading order term of the truncation error
- Galerkin methods allow a much more rigorous and flexible approach that achieves error control

Dual Problem in A Posteriori Error Estimation

- We restrict ourselves to the *linear* ODE

$$u'(t) + a(t)u(t) = f(t),\ t \in (0, T], \qquad u(0) = u_0$$

- We will have a glimpse on the *dual weighted residual* (DWR) method [4] to estimate the error

$$e(t) = u(t) - U(t), \qquad \text{in particular } e(T)$$

- DWR is based on a so-called *dual problem* which reads in this case

$$-\varphi'(t) + a(t)\varphi(t) = 0,\ t \in [0, T), \qquad \varphi(T) = e(T)$$

- Note, that this problem runs *backward* in time with the error $e(T)$ given as initial value
- With the change of variables $t(\tilde{t}) = T - \tilde{t}$ and the identification of $\tilde{\varphi}(\tilde{t}) = \varphi(t(\tilde{t})) = \varphi(T - \tilde{t})$ we obtain an equation for $\tilde{\varphi}$:

$$\tilde{\varphi}(\tilde{t}) + a(T - \tilde{t})\tilde{\varphi}(\tilde{t}) = 0,\ t \in (0, T], \qquad \tilde{\varphi}(0) = e(T)$$

[4] Becker, Rannacher; 1996/2001

Error Representation

- Using the dual problem one obtains

$$
\begin{aligned}
0 &= \int_0^T e(t)\underbrace{\left(-\varphi(t) + a(t)\varphi(t)\right)}_{=0}\,dt \\
&= \int_0^T e'(t)\varphi(t) + e(t)a(t)\varphi(t)\,dt - e(T)\underbrace{\varphi(T)}_{=e(T)} + \underbrace{e(0)}_{=0}\varphi(0) \\
\Leftrightarrow \quad e^2(T) &= \int_0^T (e'(t) + a(t)e(t))\varphi(t)\,dt = \int_0^T (e'(t) + a(t)e(t))\varphi(t)\,dt \\
&= \int_0^T (u'(t) - U'(t) + a(t)u(t) - a(t)U(t))\varphi(t)\,dt \\
\Rightarrow \quad \underbrace{e^2(T)}_{\text{error at } T} &= \int_0^T \underbrace{(f(t) - U'(t) - a(t)U(t))}_{\text{residual } R[U]} \quad \underbrace{\varphi(t)}_{\text{solution of dual problem}} \quad dt
\end{aligned}
$$

the last step is due to u being the solution of the ODE $u' + au = f$

- This is an exact representation of the error in terms of the computable residual $R[U]$ and the solution of the dual problem

A Posteriori Error Estimate

- From the exact error representation one may proceed in different ways to produce an error estimate
- One uses Cauchy Schwarz inequality:

$$e^2(T) = \int_0^T R[u](t)\varphi(t)\,dt = \sum_{i=0}^{N-1} \int_{I_i} R[u](t)\varphi(t)\,dt$$
$$\leq \sum_{i=0}^{N-1} \left(\int_{I_i} R^2[u](t)\,dt \right)^{\frac{1}{2}} \left(\int_{I_i} \varphi^2(t)\,dt \right)^{\frac{1}{2}} = \sum_{i=0}^{N-1} \|R[U]\|_{0,I_i} \|\varphi\|_{0,I_i}$$

- This is interpreted as follows:
  - The fully computable term $\|R[U]\|_{0,I_i}$ measures the error contribution in interval I_i
  - The term $\|\varphi\|_{0,I_i}$ gives the weight of this contribution in the final result
  - This explains the name DWR
- The solution of the dual problem can be approximated (including the initial condition $e(T)$) as it determines only *the relative importance* of the residual contribution

Towards an Adaptive Time-stepping Scheme

How to do error control and adaptivity with the formula

$$|e(T)| \leq \left(\sum_{i=0}^{N-1} \|R[U]\|_{0,I_i} \|\varphi\|_{0,I_i} \right)^{\frac{1}{2}} \quad ? \tag{22}$$

1. Choose Δt_0. Compute solutions U_0 and U_1 with mesh size Δt_0 and $\Delta t_1 = \Delta t_0 / 2$. From this estimate the error at final time:

$$\begin{aligned} |\tilde{e}(T)| = |u(T) - U_0(T)| &= |u(T) - U_1(T) + U_1(T) - U_0(T)| \\ &\leq |u(T) - U_1(T)| + |U_1(T) - U_0(T)| \\ &\leq \alpha |u(T) - U_0(T)| + |U_1(T) - U_0(T)| \\ \Leftrightarrow |\tilde{e}(T)| &\leq \frac{1}{1-\alpha} |U_1(T) - U_0(T)| \end{aligned}$$

2. Given an estimate $\tilde{e}(T)$ of $e(T)$ solve the dual problem
3. Compute estimate for $|e(T)|$ using (22), if $|e(T)| \leq TOL$ STOP
4. Halfen the intervals I_i giving the largest error contribution
5. Recompute U on new mesh, recompute estimate $\tilde{e}(T)$, goto 2

Galerkin Orthogonality

- We first need a further result called **Galerkin orthogonality**
- We may define a piecewise continuous solution $u_i(t)$

$$\int_{I_i} (u_i'(t) + a(t)u_i(t))\varphi_i(t)\,dt = \int_{I_i} f(t)\varphi_i(t)\,dt \quad \forall \varphi_i \in V(I_i) \quad u_i(t_i) = u_{i-1}(t_i)$$

- and the discrete solution in, say cG(q)

$$\int_{I_i} (U_i'(t) + a(t)U_i(t))\varphi_i(t)\,dt = \int_{I_i} f(t)\varphi_i(t)\,dt \quad \forall \varphi_i \in \mathcal{P}^{q-1} \quad U_i(t_i) = U_{i-1}(t_i)$$

- Subtracting and summing over all intervals gives the Galerkin orthogonality relation

$$\sum_{i=0}^{N-1} \int_{I_i} (e'(t) + a(t)e(t))\varphi(t)\,dt = \sum_{i=0}^{N-1} \int_{I_i} R[u](t)\varphi(t)\,dt = 0 \quad \forall \varphi \in \mathcal{P}^{q-1}$$

together with $e_i(t_i) = e_{i-1}(t_i)$

Another A Posteriori Error Estimate

- The second approach uses an analytical estimate to avoid the solution of a dual problem
- Again we start from the error relation

$$e^2(T) = \sum_{i=0}^{N-1} \int_{I_i} R[U](t)\varphi(t)\,dt$$

- Using the L^2-projection $\pi\varphi$ of the dual solution to piecewise polynomials we get

$$\begin{aligned} e^2(T) &= \sum_{i=0}^{N-1} \left\{ \int_{I_i} R[U](t)\varphi(t)\,dt - \int_{I_i} R[U](t)\pi\varphi(t)\,dt \right\} \\ &= \sum_{i=0}^{N-1} \int_{I_i} R[U](t)(\varphi(t) - \pi\varphi(t))\,dt \end{aligned}$$

- For the L^2-projection one has the L^1-estimate

$$\int_{I_i} |\varphi(t) - (\pi\varphi)(t)|\,dt \leq \Delta t_i \int_{I_i} |\varphi'(t)|\,dt$$

Another A Posteriori Error Estimate, ctd.

- and with that we may estimate

$$e^2(T) = \sum_{i=0}^{N-1} \left\{ \|R[U]\|_{\infty,I_i} \Delta t_i \int_{I_i} |\varphi'(t)|\, dt \right\} \le \max_{0\le i<N} (\Delta t_i \|R[U]\|_{\infty,I_i}) \int_0^T |\varphi'(t)|\, dt$$

- We have an analytical solution for φ from which one obtains

$$|\varphi(t)| \le |e(T)| \exp(\mathcal{A}T), \qquad \forall 0 \le t \le T, \qquad |a(t)| \le \mathcal{A}$$

- Introduce the stability factor $S(T) = \int_0^T |\varphi'(t)|\, dt / |e(T)|$ and
  - If $|a(t)| \le \mathcal{A}$ then $S(T) \le \exp(\mathcal{A})$
  - If $a(t) \ge 0$ then $S(T) \le 1$
- To obtain the final estimate

$$|e(T)| \le \max_{0\le i<N} (\Delta t_i \|R[U]\|_{\infty,I_i})$$

What Else ?

There are many more important aspects we could not treat in this first part on ODEs:

- More in-depth treatment of adaptive methods
- E.g. embedded Runge-Kutta methods, extrapolation method
- More in-depth treatment of stiff ODEs
- E.g. different definitions of stiffness, Padé-table, rational approximations of the exponential function
- Implicit higher order Runge-Kutta methods, collocation method to derive them
- Linear multistep methods
- Symplectic methods

Contents

Definition of a PDE

Definition 16 (Partial differential equation (PDE))

A partial differential equation (PDE) is an equation (or equation system) involving an unknown function of two or more variables and certain of its partial derivatives.

Example 11

Often, we have x, y, z as spatial independent variables and t as a temporal variable. These are four independent variables.

Laplace equation / Poisson's equation

Formulation 4 (Laplace problem / Poisson problem)

Let Ω be an open set. The **Laplace equation** reads:

$$-\Delta u = 0 \quad \text{in } \Omega.$$

The **Poisson problem** reads:

$$-\Delta u = f \quad \text{in } \Omega.$$

Definition 17

A C^2 function (C^2 means two times continuously differentiable) that satisfies the Laplace equation is called **harmonic** function.

Notation

We frequently use:

$$\frac{\partial u}{\partial x} = \partial_x u$$

and

$$\frac{\partial u}{\partial t} = \partial_t u$$

and

$$\frac{\partial^2 u}{\partial t \partial t} = \partial_t^2 u = \partial_{tt} u$$

and

$$\frac{\partial^2 u}{\partial x \partial y} = \partial_{xy} u$$

Nabla operators

Well-known in physics, it is convenient to work with the **nabla-operator** to define derivative expressions. The gradient of a single-valued function $v : \mathbb{R}^n \to \mathbb{R}$ reads:

$$\nabla v = \begin{pmatrix} \partial_1 v \\ \vdots \\ \partial_n v \end{pmatrix}.$$

The gradient of a vector-valued function $v : \mathbb{R}^n \to \mathbb{R}^m$ is called **Jacobian matrix** and reads:

$$\nabla v = \begin{pmatrix} \partial_1 v_1 & \dots & \partial_n v_1 \\ \vdots & & \vdots \\ \partial_1 v_m & \dots & \partial_n v_m \end{pmatrix}.$$

Nabla operators

The divergence is defined for vector-valued functions $v : \mathbb{R}^n \to \mathbb{R}^n$:

$$\operatorname{div} v := \nabla \cdot v := \nabla \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{k=1}^{n} \partial_k v_k.$$

The divergence for a tensor $\sigma \in \mathbb{R}^{n\times n}$ is defined as:

$$\nabla \cdot \sigma = \Big(\sum_{j=1}^{n} \frac{\partial \sigma_{ij}}{\partial x_j}\Big)_{1\le i\le n}.$$

The trace of a matrix $A \in \mathbb{R}^{n\times n}$ is defined as

$$tr(A) = \sum_{i=1}^{n} a_{ii}.$$

Nabla operators

Definition 18 (Laplace operator)

The Laplace operator of a two-times continuously differentiable scalar-valued function $u : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\Delta u = \sum_{k=1}^{n} \partial_{kk} u.$$

Definition 19

For a vector-valued function $u : \mathbb{R}^n \to \mathbb{R}^m$, we define the Laplace operator component-wise as

$$\Delta u = \Delta \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} \partial_{kk} u_1 \\ \vdots \\ \sum_{k=1}^{n} \partial_{kk} u_m \end{pmatrix}.$$

Physical interpretation / mathematical modeling of the Laplace operator

The physical interpretation is as follows. Let u denote the density of some quantity, for instance concentration or temperature, in equilibrium. If G is any smooth region $G \subset \Omega$, the flux F of the quantity u through the boundary ∂G is zero:

$$\int_{\partial G} F \cdot n \, dx = 0. \tag{23}$$

Here F denotes the flux density and n the outer normal vector. Gauss' divergence theorem yields:

$$\int_{\partial G} F \cdot n \, dx = \int_G \nabla \cdot F \, dx = 0.$$

Since this integral relation holds for arbitrary G, we obtain

$$\nabla \cdot F = 0 \quad \text{in } \Omega. \tag{24}$$

Physical interpretation

- Now we need a second assumption (or better a relation) between the flux and the quantity u. Such relations do often come from material properties and are so-called **constitutive laws.**
- In many situations it is reasonable to assume that the flux F is proportional to the negative gradient $-\nabla u$ of the quantity u. This means that flow goes from regions with a higher concentration to lower concentration regions.
- For instance, the rate at which energy 'flows' (or diffuses) as heat from a warm body to a colder body is a function of the temperature difference. The larger the temperature difference, the larger the diffusion.
- We consequently obtain as further relation:

$$F = -\nabla u.$$

Physical interpretation

- Plugging into the Equation (24) yields:

$$\nabla \cdot F = \nabla \cdot (-\nabla u) = -\nabla \cdot (\nabla u) = -\Delta u = 0.$$

This is the simplest derivation one can make. Adding more knowledge on the underlying material of the body, a material parameter $a > 0$ can be added:

$$\nabla \cdot F = \nabla \cdot (-a\nabla u) = -\nabla \cdot (a\nabla u) = -a\Delta u = 0.$$

And adding a nonconstant and spatially dependent material further yields:

$$\nabla \cdot F = \nabla \cdot (-a(x)\nabla u) = -\nabla \cdot (a(x)\nabla u) = 0.$$

In this last equation, we do not obtain any more the classical Laplace equation but a diffusion equation in divergence form.

Other fields using Poisson's equation

Some important physical laws are related to the Laplace operator (partially taken from L. Evans; Partial Differential Equations, AMS, 2010):

1. Fick's law of chemical diffusion
2. Fourier's law of heat conduction
3. Ohm's law of electrical conduction
4. Small deformations in elasticity (recall the clothesline problem)

Three important linear PDEs

- Poisson problem: $-\Delta u = f$ is elliptic: second order in space and no time dependence.
- Heat equation: $\partial_t u - \Delta u = f$ is parabolic: second order in space and first order in time.
- Wave equation: $\partial_t^2 u - \Delta u = f$ is hyperbolic: second order in space and second order in time.

Elliptic PDEs: prototype Laplacian

Formulation 5

Let $f:\Omega\to\mathbb{R}$ be given. Furthermore, Ω is an open, bounded set of $\mathbb{R}^n$. We seek the unknown function $u:\bar{\Omega}\to\mathbb{R}$ such that

$$Lu = f \quad \text{in } \Omega, \tag{25}$$

$$u = 0 \quad \text{on } \partial\Omega. \tag{26}$$

Here, the linear second-order differential operator is defined by:

$$Lu := -\sum_{i,j=1}^{n} \partial_{x_j}(a_{ij}(x)\partial_{x_i}u) + \sum_{i=1}^{n} b_i(x)u\partial_{x_i} + c(x)u, \quad u = u(x), \tag{27}$$

with the symmetry assumption $a_{ij} = a_{ji}$ and given coefficient functions a_{ij}, b_i, c. Moreover, we assume that A is positive definite (in order words: the eigenvalues are positive).

Elliptic PDEs: prototype Laplacian

Formulation 6

Alternatively we often use the compact notation with derivatives defined in terms of the nabla-operator:

$$Lu := -\nabla \cdot (a\nabla u) + b\nabla u + cu.$$

Finally we notice that the boundary condition (26) is called **homogeneous Dirichlet condition**.

Elliptic PDEs: prototype Laplacian

Theorem 20 (Strong maximum principle for the Laplace problem)
Suppose $u \in C^2(\Omega) \cap C(\bar{\Omega})$ is a harmonic function. Then

$$\max_{\bar{\Omega}} u = \max_{\partial\Omega} u.$$

Moreover, if Ω is connected and there exists a point $y \in \Omega$ in which

$$u(y) = \max_{\bar{\Omega}} u,$$

then u is constant within Ω. The same holds for $-u$, but then for minima.

Remark 12
The maximum principle has a discrete version and it allows to check whether a numerically-computed discrete solution is correct.

Parabolic PDEs: prototype heat equation

Formulation 7

Let $f : \Omega \times I \to \mathbb{R}$ and $u_0 : \Omega \to \mathbb{R}$ be given. We seek the unknown function $u : \bar{\Omega} \times I \to \mathbb{R}$ such that[5]

$$\partial_t u + Lu = f \quad \text{in } \Omega \times I, \tag{28}$$
$$u = 0 \quad \text{on } \partial\Omega \times [0, T], \tag{29}$$
$$u = u_0 \quad \text{on } \Omega \times \{t = 0\}. \tag{30}$$

Here, the linear second-order differential operator is defined by:

$$Lu := -\sum_{i,j=1}^{n} \partial_{x_j}(a_{ij}(x,t)\partial_{x_i} u) + \sum_{i=1}^{n} b_i(x,t)\partial_{x_i} u + c(x,t)u, \quad u = u(x,t)$$

for given (possibly spatial and time-dependent) coefficient functions a_{ij}, b_i, c.

Parabolic PDEs: prototype heat equation

Formulation 8 (Heat equation)

Setting in Formulation 7, $a_{ij} = \delta_{ij}$ and $b_i = 0$ and $c = 0$, we obtain the Laplace operator. Let $f : \Omega \to \mathbb{R}$ be given. Furthermore, Ω is an open, bounded set of $\mathbb{R}^n$. We seek the unknown function $u : \bar{\Omega} \to \mathbb{R}$ such that

$$\partial_t u + Lu = f \quad \text{in } \Omega \times I, \tag{31}$$

$$u = 0 \quad \text{on } \partial\Omega \times [0, T], \tag{32}$$

$$u = u_0 \quad \text{on } \Omega \times \{t = 0\}. \tag{33}$$

Here, the linear second-order differential operator is defined by:

$$Lu := -\nabla \cdot (\nabla u) = -\Delta u.$$

Hyperbolic PDEs: prototype wave equation

Formulation 9

Let $f : \Omega \times I \to \mathbb{R}$ and $u_0, v_0 : \Omega \to \mathbb{R}$ be given. We seek the unknown function $u : \bar{\Omega} \times I \to \mathbb{R}$ such that

$$\partial_t^2 u + Lu = f \quad \text{in } \Omega \times I, \tag{34}$$

$$u = 0 \quad \text{on } \partial\Omega \times [0, T], \tag{35}$$

$$u = u_0 \quad \text{on } \Omega \times \{t = 0\}, \tag{36}$$

$$\partial_t u = v_0 \quad \text{on } \Omega \times \{t = 0\}. \tag{37}$$

In the last line, $\partial_t u = v$ can be identified as the velocity. Furthermore, the linear second-order differential operator is defined by:

$$Lu := -\sum_{i,j=1}^{n} \partial_{x_j}(a_{ij}(x,t)\partial_{x_i} u) + \sum_{i=1}^{n} b_i(x,t)\partial_{x_i} u + c(x,t)u, \quad u = u(x,t)$$

for given (possibly spatial and time-dependent) coefficient functions a_{ij}, b_i, c.

Hyperbolic PDEs: prototype wave equation

Remark 13

The wave equation is often written in terms of a first-order system in which the velocity is introduced and a second-order time derivative is avoided. Then the previous equation reads: Find $u : \bar{\Omega} \times I \to \mathbb{R}$ and $v : \bar{\Omega} \times I \to \mathbb{R}$ such that

$$
\begin{aligned}
\partial_t v + Lu &= f \quad \text{in } \Omega \times I, && (38)\\
\partial_t u &= v \quad \text{in } \Omega \times I, && (39)\\
u &= 0 \quad \text{on } \partial\Omega \times [0, T], && (40)\\
u &= u_0 \quad \text{on } \Omega \times \{t = 0\}, && (41)\\
v &= v_0 \quad \text{on } \Omega \times \{t = 0\}. && (42)
\end{aligned}
$$

Remarks to boundary data

- Dirichlet (or essential) boundary conditions: $u = g_D$ on $\partial\Omega_D$; when $g_D = 0$ we say 'homogeneous' boundary condition.
- Neumann (or natural) boundary conditions: $\partial_n u = g_N$ on $\partial\Omega_N$; when $g_N = 0$ we say 'homogeneous' boundary condition.
- Robin (third type) boundary condition: $au + b\partial_n u = g_R$ on $\partial\Omega_R$; when $g_R = 0$ we say 'homogeneous' boundary condition.

Example temperature in a room

We consider the heat equation: Find $T : \Omega \times I \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t T + (v \cdot \nabla)T - \nabla \cdot (K \nabla T) = f \quad & \text{in } \Omega \times I, \\ T = 18^\circ C \quad & \text{on } \partial_D \Omega \times I, \\ K \nabla T \cdot n = 0 \quad & \text{on } \partial_N \Omega \times I, \\ T(0) = 15^\circ C \quad & \text{in } \Omega \times \{0\}. \end{aligned}$$

The homogeneous Neumann condition means that there is no heat exchange on the respective walls (thus neighboring rooms will have the same room temperature on the respective walls). The nonhomogeneous Dirichlet condition states that there is a given temperature of $18C$, which is constant in time and space (but this condition may be also non-constant in time and space). Possible heaters in the room can be modeled via the right hand side f. The vector $v : \Omega \to \mathbb{R}^3$ denotes a given flow field yielding a convection of the heat, for instance wind. We can assume $v \approx 0$. Then the above equation is reduced to the original heat equation: $\partial_t T - \nabla \cdot (K \nabla T) = f$.

Further classifications (we recall from our ODE studies)

- Order of a differential equation
- Single equations and PDE systems
- Nonlinear problems:
  - Nonlinearity in the PDE
  - The function set is not a vector space yielding a variational inequality
- Coupled problems and coupled PDE systems.

Further classifications: examples

- p-Laplace equation: Find $u : \Omega \to \mathbb{R}$:

$$-\nabla \cdot (|\nabla u|^{p-2} \nabla u)^{p/2} = f \tag{43}$$

Properties: nonlinear (quasilinear), stationary, scalar-valued.

- Find $u : \Omega \to \mathbb{R}$:

$$-\Delta u + u^2 = f \tag{44}$$

Properties: nonlinear (semilinear), stationary, scalar-valued.

- Incompressible, isothermal Navier-Stokes equations: Find $v : \Omega \to \mathbb{R}^n$ and $p : \Omega \to \mathbb{R}$

$$\partial_t + (v \cdot \nabla) v - \frac{1}{Re} \Delta v + \nabla p = f, \quad \nabla \cdot v = 0 \tag{45}$$

with Re being the Reynolds' number. For $Re \to \infty$ we obtain the Euler equations. Properties: nonlinear (semilinear), nonstationary, vector-valued, PDE system.

Further classifications: examples

- A volume-coupled problem: Find $u : \Omega \to \mathbb{R}$ and $\varphi : \Omega \to \mathbb{R}$

$$-\Delta u = f(\varphi), \tag{46}$$
$$|\nabla u|^2 - \Delta\varphi = g(u) \tag{47}$$

  Properties: nonlinear, coupled problem via right hand sides, stationary. Equations become linear when solution variables are fixed in the other equation.

- An interface-coupled problem: Let Ω_1 and Ω_2 with $\Omega_1 \cap \Omega_2 = 0$ and $\bar{\Omega}_1 \cap \bar{\Omega}_2 = \Gamma$ and $\bar{\Omega}_1 \cup \bar{\Omega}_2 = \Omega$. Find $u_1 : \Omega_1 \to \mathbb{R}$ and $u_2 : \Omega_2 \to \mathbb{R}$:

$$-\Delta u_1 = f_1 \quad \text{in } \Omega_1, \tag{48}$$
$$-\Delta u_2 = f_2 \quad \text{in } \Omega_2, \tag{49}$$
$$u_1 = u_2 \quad \text{on } \Gamma, \tag{50}$$
$$\partial_n u_1 = \partial_n u_2 \quad \text{on } \Gamma \tag{51}$$

  Properties: linear, coupled problem via interface conditions, stationary.

Advanced examples (to give an outlook): elasticity

This example is already difficult because a system of nonlinear equations is considered:

Formulation 10

Let $\widehat{\Omega}_s \subset \mathbb{R}^n$, $n = 3$ with the boundary $\partial\widehat{\Omega} := \widehat{\Gamma}_D \cup \widehat{\Gamma}_N$. Furthermore, let $I := (0, T]$ where $T > 0$ is the end time value. The equations for geometrically non-linear elastodynamics in the reference configuration $\widehat{\Omega}$ are given as follows: Find vector-valued displacements $\hat{u}_s := (\hat{u}_s^{(x)}, \hat{u}_s^{(y)}, \hat{u}_s^{(z)}) : \widehat{\Omega}_s \times I \to \mathbb{R}^n$ such that

$$
\begin{aligned}
\hat{\rho}_s \partial_t^2 \hat{u}_s - \widehat{\nabla} \cdot (\widehat{F}\widehat{\Sigma}) &= 0 \quad && \text{in } \widehat{\Omega}_s \times I, \\
\hat{u}_s &= 0 && \text{on } \widehat{\Gamma}_D \times I, \\
\widehat{F}\widehat{\Sigma} \cdot \hat{n}_s &= \hat{h}_s && \text{on } \widehat{\Gamma}_N \times I, \\
\hat{u}_s(0) &= \hat{u}_0 && \text{in } \widehat{\Omega}_s \times \{0\}, \\
\hat{v}_s(0) &= \hat{v}_0 && \text{in } \widehat{\Omega}_s \times \{0\}.
\end{aligned}
$$

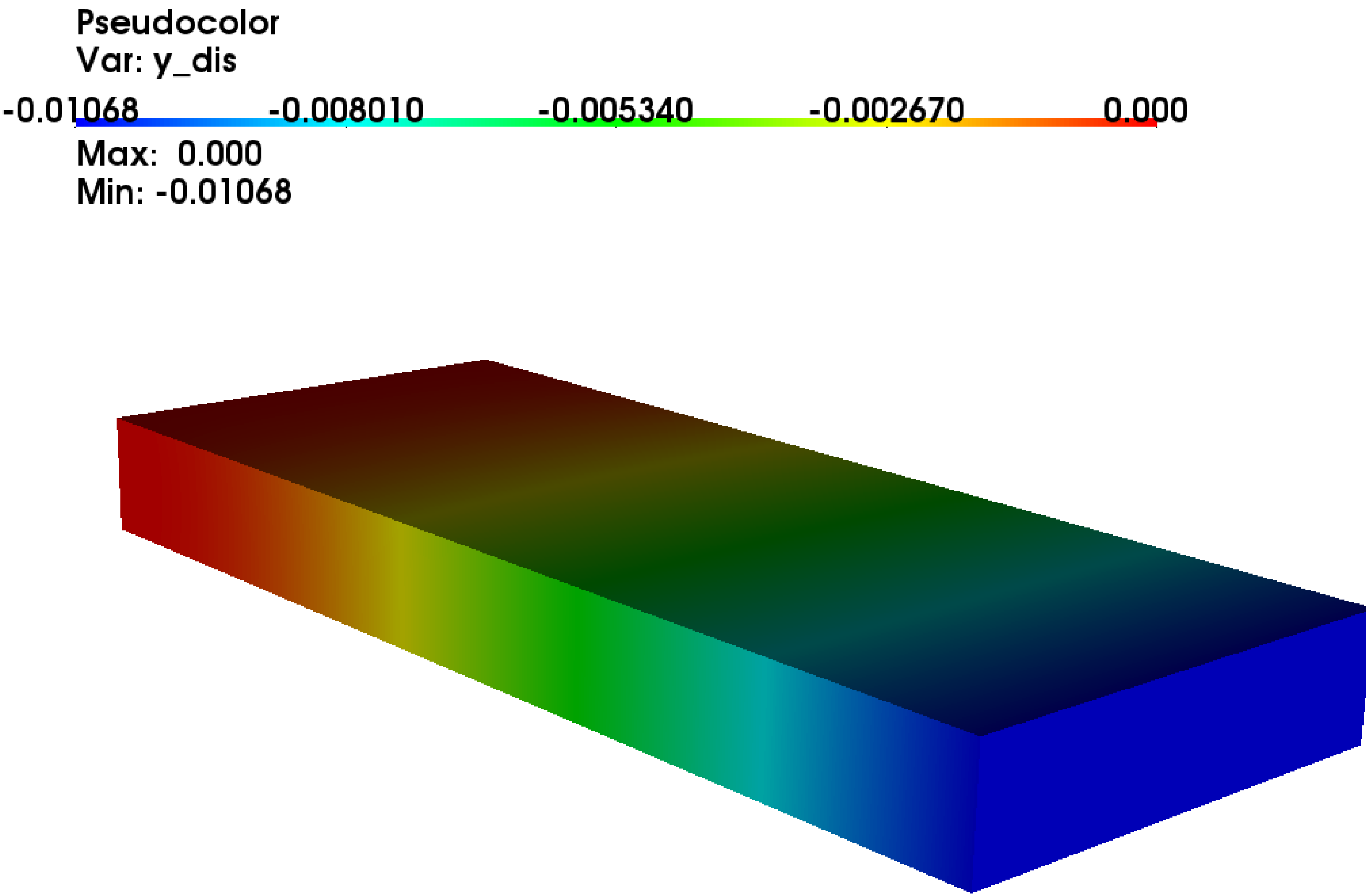
Advanced examples (to give an outlook): elasticity

We deal with two types of boundary conditions: Dirichlet and Neumann conditions. Furthermore, two initial conditions on the displacements and the velocity are required. The constitutive law is given by the geometrically nonlinear tensors (see e.g., Ciarlet 1984):

$$\widehat{\Sigma} = \widehat{\Sigma}_s(\hat{u}_s) = 2\mu\widehat{E} + \lambda tr(\widehat{E})I, \quad \widehat{E} = \frac{1}{2}(\widehat{F}^T\widehat{F} - I). \tag{52}$$

Here, μ and λ are the Lamé coefficients for the solid. The solid density is denoted by $\hat{\rho}_s$ and the solid deformation gradient is $\widehat{F} = \hat{I} + \widehat{\nabla}\hat{u}_s$ where $\hat{I} \in \mathbb{R}^{3\times 3}$ is the identity matrix. Furthermore, $\hat{n}_s$ denotes the normal vector.

Advanced examples (to give an outlook): elasticity



Advanced examples (to give an outlook): incompressible flow - Navier-Stokes equations

Flow equations in general are extremely important and have an incredible amount of possible applications such as for example

- water (fluids),
- blood flow,
- wind,
- weather forecast,
- aerodynamics:

Advanced examples (to give an outlook): incompressible flow - Navier-Stokes equations

Formulation 11

Let $\Omega_f \subset \mathbb{R}^n$, $n = 3$. Furthermore, let the boundary be split into $\partial\Omega_f := \Gamma_{in} \cup \Gamma_{out} \cup \Gamma_D \cup \Gamma_i$. The isothermal, incompressible (non-linear) Navier-Stokes equations read: Find vector-valued velocities $v_f : \Omega_f \times I \to \mathbb{R}^n$ and a scalar-valued pressure $p_f : \Omega_f \times I \to \mathbb{R}$ such that

$$
\begin{aligned}
\rho_f \partial_t v_f + \rho_f v_f \cdot \nabla v_f - \nabla \cdot \sigma_f(v_f, p_f) &= 0 && \text{in } \Omega_f \times I, \\
\nabla \cdot v_f &= 0 && \text{in } \Omega_f \times I, \\
v_f^D &= v_{in} && \text{on } \Gamma_{in} \times I, \\
v_f &= 0 && \text{on } \Gamma_D \times I, \\
-p_f n_f + \rho_f \nu_f \nabla v_f \cdot n_f &= 0 && \text{on } \Gamma_{out} \times I, \\
v_f &= h_f && \text{on } \Gamma_i \times I, \\
v_f(0) &= v_0 && \text{in } \Omega_f \times \{t = 0\},
\end{aligned}
$$

Advanced examples (to give an outlook): incompressible flow - Navier-Stokes equations

Here the (symmetric) Cauchy stress is given by

$$\sigma_f(v_f, p_f) := -p_f I + \rho_f \nu_f (\nabla v_f + \nabla v_f^T),$$

with the density ρ_f and the kinematic viscosity ν_f. The normal vector is denoted by n_f.

Advanced examples (to give an outlook): incompressible flow - Navier-Stokes equations

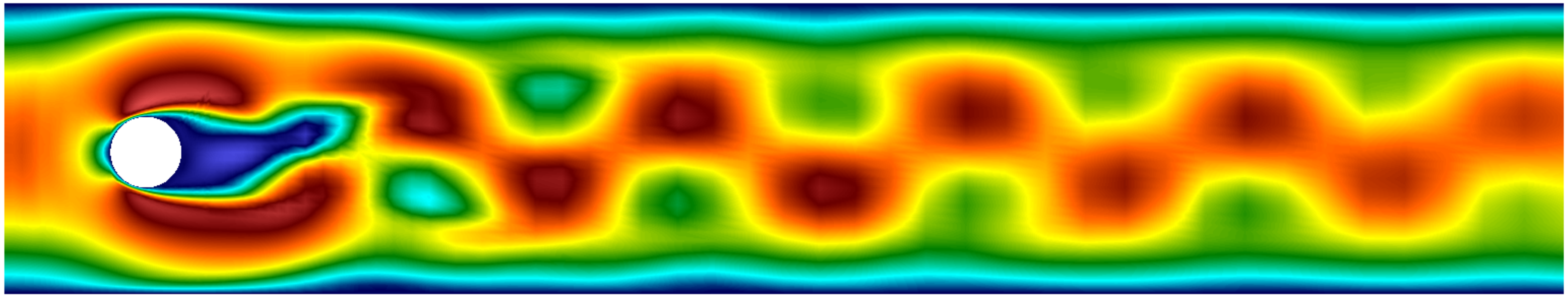


Figure: Prototype example of a fluid mechanics problem (isothermal, incompressible Navier-Stokes equations): the famous Karman vortex street. The setting is based on the benchmark setting Schaefer/Turek et al. 1996 and the code can be found in NonStat Example 1 in DOpElib `www.dopelib.net`.

Summary lecture 06

- Different types of PDEs
- Modeling and physical explanations
- Classifications of the order, linear/nonlinear, PDE systems
- Various further examples

Contents

Recall model problem in 1D: Poisson's problem

$$
\begin{aligned}
-u'' &= f \quad \text{in } \Omega = (0,1), && (53)\\
u(0) &= u(1) = 0. && (54)
\end{aligned}
$$

Equivalent formulations

We first introduce the scalar product

$$(v, w) = \int_0^1 v(x)w(x)\,dx.$$

Furthermore we introduce the linear space

$$V := \{v|\ v \in C[0,1], v' \text{ is piecewise continuous and bounded on } [0,1], v(0) = v(1) = 0\}. \tag{55}$$

We also introduce the linear functional $F : V \to \mathbb{R}$ such that

$$F(v) = \frac{1}{2}(v', v') - (f, v).$$

Equivalent formulations

Definition 21

We deal with three (equivalent) problems:

(D) Find $u \in C^2$ such that $-u'' = f$ with $u(0) = u(1) = 0$;

(M) Find $u \in V$ such that $F(u) \leq F(v)$ for all $v \in V$;

(V) Find $u \in V$ such that $(u', v') = (f, v)$ for all $v \in V$.

- In physics, the quantity $F(v)$ stands for the **total potential energy** of the underlying model.
- Moreover, the first term in $F(v)$ denotes the internal elastic energy and (f, v) the load potential.
- Therefore, formulation (M) corresponds to the fundamental **principle of minimal potential energy** and the variational problem (V) to the **principle of virtual work** (e.g., Ciarlet 1984).
- The proofs of their equivalence will be provided in the following.

Equivalent formulations

Proposition 14

It holds

$$(D) \to (V).$$

Proof.

We multiply $u'' = f$ with an arbitrary function ϕ (a so-called **test function**) from the space V defined in (55). Then we integrate over the interval $\Omega = (0, 1)$ yielding

$$-u'' = f \tag{56}$$

$$\Rightarrow -\int_\Omega u''\phi\,dx = \int_\Omega f\phi\,dx \tag{57}$$

$$\Rightarrow \int_\Omega u'\phi'\,dx - u'(1)\phi(1) + u'(0)\phi(0) = \int_\Omega f\phi\,dx \tag{58}$$

$$\Rightarrow \int_\Omega u'\phi'\,dx = \int_\Omega f\phi\,dx \quad \forall\phi \in V. \tag{59}$$

Equivalent formulations

In the second last term, we used integration by parts.
The boundary terms vanish because $\phi \in V$. This shows that

$$\int_\Omega u' \phi' \, dx = \int_\Omega f \phi \, dx$$

is a solution of (V).

Remark 15
The technique used in this proof is of paramount importance since the integration by parts is THE standard trick in the finite element method.

Equivalent formulations

Proposition 16

It holds

$$(V) \leftrightarrow (M).$$

Proof.

We first assume that u is a solution to (V). Let $\phi \in V$ and set $w = \phi - u$ such that $\phi = u + w$ and $w \in V$. We obtain

$$\begin{aligned} F(\phi) = F(u+w) &= \frac{1}{2}(u' + w', u' + w') - (f, u+w) \\ &= \frac{1}{2}(u', u') - (f, u) + (u', w') - (f, w) + \frac{1}{2}(w', w') \geq F(u) \end{aligned}$$

We use now the fact that (V) holds true, namely

$$(u', w') - (f, w) = 0.$$

Equivalent formulations

And also that $(w', w') \geq 0$. Thus, we have shown that u is a solution to (M). We show now that $(M) \to (V)$ holds true as well. For any $\phi \in V$ and $\varepsilon \in \mathbb{R}$ we have

$$F(u) \leq F(u + \varepsilon\phi),$$

because $u + \varepsilon\phi \in V$. We differentiate with respect to ε and show that (V) is a first order necessary condition to (M) with a minimum at $\varepsilon = 0$. To do so, we define

$$g(\varepsilon) := F(u + \varepsilon\phi) = \frac{1}{2}(u', u') + \varepsilon(u', \phi') + \frac{\varepsilon^2}{2}(\phi', \phi') - (f, u) - \varepsilon(f, \phi).$$

Equivalent formulations

Thus

$$g'(\varepsilon) = (u', \phi') + \varepsilon(\phi', \phi') - (f, \phi).$$

A minimum is obtained for $\varepsilon = 0$. Consequently,

$$g'(0) = 0.$$

In detail:

$$(u', \phi') - (f, \phi) = 0,$$

which is nothing else than the solution of (V).

Equivalent formulations

Proposition 17

It holds

$$(V) \to (D).$$

Proof. We assume that u is a solution to (V), i.e.,

$$(u', \phi') = (f, \phi) \quad \forall \phi \in V.$$

If we assume sufficient regularity of u (in particular $u \in C^2$), then u'' exists and we can integrate backwards. Moreover, we use that $\phi(0) = \phi(1) = 0$ since $\phi \in V$. Then:

$$(-u'' - f, \phi) = 0 \quad \forall \phi \in V.$$

Equivalent formulations

Since we assumed sufficient regularity for u'' and f the difference is continuous. We can now apply the fundamental principle (see Proposition 18):

$$w \in C(\Omega) \quad \Rightarrow \quad \int_\Omega w\phi \, dx = 0 \quad \Rightarrow \quad w \equiv 0.$$

We proof this result later. Before, we obtain

$$(-u'' - f, \phi) = 0 \quad \Rightarrow \quad -u'' - f = 0,$$

which yields the desired expression. Since we know that $(D) \to (V)$ holds true, u has the assumed regularity properties and we have shown the equivalence.

Fundamental lemma of calculus of variations

Proposition 18

Let $\Omega = [a, b]$ be a compact interval and let $w \in C(\Omega)$. Let $\phi \in C^\infty$ with $\phi(a) = \phi(b) = 0$, i.e., $\phi \in C_c^\infty(\Omega)$. If for all ϕ it holds

$$\int_\Omega w(x)\phi(x)\, dx = 0,$$

then, $w \equiv 0$ in Ω.

Proof.

We perform an indirect proof. We suppose that there exist a point $x_0 \in \Omega$ with $w(x_0) \neq 0$. Without loss of generality, we can assume $w(x_0) > 0$. Since w is continuous, there exists a small (open) neighborhood $\omega \subset \Omega$ with $w(x) > 0$ for all $x \in \omega$; otherwise $w \equiv 0$ in $\Omega \setminus \omega$.

Proof continued.

Let ϕ now be a positive test function (recall that ϕ can be arbitrary, specifically positive if we wish) in Ω and thus also in ω. Then:

$$\int_\Omega w(x)\phi(x)\,dx = \int_\omega w(x)\phi(x)\,dx.$$

But this is a contradiction to the hypothesis on w. Thus $w(x) = 0$ for all in $x \in \omega$. Extending this result to all open neighborhoods in Ω we arrive at the final result.

Remark 19

The general form of the proof can be found in P. Ciarlet; 2013: Linear and nonlinear functional analysis with applications.

Derivation of a weak (variational) form

$$-u'' = f \tag{60}$$

$$\Rightarrow -\int_\Omega u''\phi\,dx = \int_\Omega f\phi\,dx \tag{61}$$

$$\Rightarrow \int_\Omega u'\phi'\,dx - \int_{\partial\Omega} \partial_n u\phi\,ds = \int_\Omega f\phi\,dx \tag{62}$$

$$\Rightarrow \int_\Omega u'\phi'\,dx = \int_\Omega f\phi\,dx. \tag{63}$$

To summarize we have:

$$\int_\Omega u'\phi'\,dx = \int_\Omega f\phi\,dx \tag{64}$$

Derivation of a weak (variational) form

A common short-hand notation in mathematics is to use parentheses for L^2 scalar products: $\int_\Omega ab\,dx =: (a, b)$:

$$(u', \phi') = (f, \phi) \tag{65}$$

A mathematically-correct statement is:

Formulation 12
Find $u \in V$ such that

$$(u', \phi') = (f, \phi) \quad \forall \phi \in V. \tag{66}$$

In the following, we introduce some tools from functional analysis that are required to analyze further the variational form.

Hilbert spaces

Definition 22 (Hilbert space)

A complete space endowed with an inner product is called a **Hilbert space**. The norm is defined by

$$\|u\| := \sqrt{(u, u)}.$$

Example 20

The space $\mathbb{R}^n$ from before has a scalar product and is complete, thus a Hilbert space. The space $\{C(\Omega), \|\cdot\|_{L^2}\}$ has a scalar product, but is not complete, and therefore not a Hilbert space. The space $\{C(\Omega), \|\cdot\|_{C(\Omega)}\}$ is complete, but the norm is not induced by a scalar product and is therefore not a Hilbert space, but only a Banach space.

Hilbert spaces: L_2

Definition 23 (The L^2 space in 1D)

Let $\Omega = (a, b)$ be an interval (recall 1D Poisson). The space of square-integrable functions on Ω is defined by

$$L^2(\Omega) = \{v : \int_\Omega v^2\,dx < \infty\}$$

The space L^2 is a Hilbert space equipped with the scalar product

$$(v, w) = \int_\Omega vw\,dx$$

and the induced norm

$$\|v\|_{L^2} := \sqrt{(v, v)}.$$

Hilbert spaces: L_2

Using Cauchy's inequality

$$|(v, w)| \leq \|v\|_{L^2} \|w\|_{L^2},$$

we observe that the scalar product is well-defined when $v, w \in L^2$. A mathematically very correct definition must include in which sense (Riemann or Lebesgue) the integral exists. In general, all L spaces are defined in the sense of the Lebuesgue integral (see for instance books introducing Lebesgue spaces).

Hilbert spaces: H^1

Definition 24 (The H^1 space in 1D)

We define the $H^1(\Omega)$ space with $\Omega = (a, b)$ as

$$H^1(\Omega) = \{v : \; v \text{ and } v' \text{ belong to } L^2\}$$

This space is equipped with the following scalar product:

$$(v, w)_{H^1} = \int_\Omega (vw + v'w')\, dx$$

and the norm

$$\|v\|_{H^1} := \sqrt{(v, v)_{H^1}}.$$

Hilbert spaces: H_0^1

Definition 25 (The H_0^1 space in 1D)

We define the $H_0^1(\Omega)$ space with $\Omega = (a, b)$ as

$$H_0^1(\Omega) = \{v \in H^1(\Omega) : \; v(a) = v(b) = 0\}.$$

The scalar product is the same as for the H^1 space.

Well-posedness: existence, uniqueness and stability

Formulation 13 (Abstract model problem)

Let V be a Hilbert space with norm $||\cdot||_V$. Find $u \in V$ such that

$$a(u,\phi) = l(\phi) \quad \forall \phi \in V.$$

with

$$a(u,\phi) := (u',\phi'),$$
$$l(\phi) := (f,\phi)$$

Well-posedness: existence, uniqueness and stability

Definition 26 (Assumptions)

We suppose:

1. $l(\cdot)$ is a bounded linear form:
$$|l(u)| \leq C\|u\| \quad \text{for all } u \in V.$$
2. $a(\cdot,\cdot)$ is a bilinear form on $V \times V$ and continuous:
$$|a(u,v)| \leq \gamma \|u\|_V \|v\|_V, \quad \gamma > 0, \quad \forall u, v \in V.$$
3. $a(\cdot,\cdot)$ is coercive (or V-elliptic):
$$a(u,u) \geq \alpha \|u\|_V^2, \quad \alpha > 0, \quad \forall u \in V.$$

The Lax-Milgram lemma

Lemma 27 (Lax-Milgram)

Let $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ be a continuous, V-elliptic bilinear form. Then, for each $l \in V^*$ the variational problem

$$a(u,\phi) = l(\phi) \quad \forall \phi \in V$$

has a unique solution $u \in V$. Moreover, we have the stability estimate:

$$\|u\| \leq \frac{1}{\alpha}\|l\|_{V^*}.$$

with

$$\|l\|_{V^*} := \sup_{\varphi \neq 0} \frac{|l(\varphi)|}{||\varphi||_V}.$$

The proof can be found in P. Ciarlet; 2013.

The energy norm

The continuity and coercivity of the bilinear form yield the **energy norm**:

$$\|v\|_a^2 := a(v,v), \quad v \in V.$$

This norm is aquivalent to the V-norm of the space V, i.e.,

$$c\|v\|_V \le \|v\|_a \le C\|v\|_V, \quad \forall v \in V$$

and two positive constants c and C. We can even precisely determine these two constants:

$$\alpha||u||_V^2 \le a(u,u) \le \gamma\|u\|_V^2$$

yielding $c = \sqrt{\alpha}$ and $C = \sqrt{\gamma}$. The corresponding scalar product is defined by

$$(v,w)_a := a(v,w).$$

The energy norm: example

For the Poisson problem, the energy norm reads:

- Given $a(v, v) = (v', v') = \int_\Omega (v'(x))^2\, dx$
- Then:

$$\|v\|_a^2 = \int_\Omega (v'(x))^2\, dx.$$

- The energy norm is the ‘natural’ norm to measure results of Poisson’s problem
- For instance: a computational convergence analysis (see lecture 04), could be done with the energy norm
- Moreover, the energy norm measures indeed the ‘physical’ energy of the given system

Summary of lecture 07

- Equivalent formulations
- Derivation of a weak form from a strong form
- Hilbert spaces
- Well-posedness of linear, stationary PDEs: Lax-Milgram lemma
- Energy norm: natural norm for the Laplace operator

Exercise 3

Let $\alpha \in \mathbb{R}$. We are given the Poisson problem in 1D on the interval $\Omega = (0, 1)$:

$$-\alpha u''(x) = f \quad \text{in } \Omega$$
$$u(0) = u(1) = 0$$

and $\alpha = 1$ and the right hand side $f = -a$ with $a > 0$. The code of this example can be found here:

`https://cloud.ifam.uni-hannover.de/index.php/s/Cwe4ZqwLRMixS3J`

with the password that is known to you.

Remark 21

Please be careful that the above form is only correct when α is constant. The general formulation is

$$-\frac{d}{dx}(\alpha u')$$

which reduces to the above one, when α is constant.

Exercise 3

Please work on the following tasks:

1. Please run the code and observe the results using gnuplot.
   Hint: Please work in the optimized compiling mode
2. We play now with two parameters:
   1. Please vary the **discretization parameter** h and use other parameters. What do you observe?
   2. Vary now the **model parameter** α. What do you observe?
   3. Choose now a different **right hand side** f. What do you observe?
3. We study in this final task the structure of the code. Go into the code and try to understand the different functions and methods that are implemented therein. Please have a specific look into the `assemble_system` method.