Scientific Computing WS 2018/2019

Lecture 14

Jürgen Fuhrmann

juergen.fuhrmann@wias-berlin.de

Partial Differential Equations

Differential operators: notations

Given: domain $\Omega \subset \mathbb{R}^d$.

- Dot product: for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{d} x_i y_i$
- Bounded domain $\Omega \subset \mathbb{R}^d$, with piecewise smooth boundary
- Scalar function $u : \Omega \to \mathbb{R}$
- Vector function $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix} : \Omega \to \mathbb{R}^d$
- Write $\partial_i u = \frac{\partial u}{\partial x_i}$
- For a multiindex $\alpha = (\alpha_1 \dots \alpha_d)$, let
  - $|\alpha| = \alpha_1 + \cdots + \alpha_d$
  - $\partial^\alpha u = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$

Basic Differential operators

- Gradient of scalar function $u : \Omega \to \mathbb{R}$

$$\mathrm{grad} = \nabla = \begin{pmatrix} \partial_1 \\ \vdots \\ \partial_d \end{pmatrix} : \; u \mapsto \nabla u = \begin{pmatrix} \partial_1 u \\ \vdots \\ \partial_d u \end{pmatrix}$$

- Divergence of vector function $\mathbf{v} = \Omega \to \mathbb{R}^d$

$$\mathrm{div} = \nabla \cdot : \; \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix} \mapsto \nabla \cdot \mathbf{v} = \partial_1 v_1 + \cdots + \partial_d v_d$$

- Laplace operator of scalar function $u : \Omega \to \mathbb{R}$

$$\Delta = \mathrm{div} \cdot \mathrm{grad} = \nabla \cdot \nabla : \; u \mapsto \Delta u = \partial_{11} u + \cdots + \partial_{dd} u$$

Lipschitz domains

Definition:

- Let $D \subset \mathbb{R}^n$. A function $f : D \to \mathbb{R}^m$ is called *Lipschitz continuous* if there exists $c > 0$ such that $\|f(x) - f(y)\| \leq c\|x - y\|$
- A hypersurface in $\mathbb{R}^n$ is a *graph* if for some k it can be represented as

$$x_k = f(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$$

  defined on some domain $D \subset \mathbb{R}^{n-1}$
- A domain $\Omega \subset \mathbb{R}^n$ is a *Lipschitz domain* if for all $x \in \partial\Omega$, there exists a neigborhood of x on $\partial\Omega$ which can be represented as the graph of a Lipschitz continuous function.

Corollaries

- Boundaries of Lipschitz domains are continuous
- Boundaries of Lipschitz domains have no cusps (e.g. the graph of $y = \sqrt{|x|}$ has a cusp at $x = 0$)
- Polygonal domains are Lipschitz
- Standard PDE calculus happens in Lipschitz domains

Divergence theorem (Gauss' theorem)

Theorem: Let Ω be a bounded Lipschitz domain and $\mathbf{v} : \Omega \to \mathbb{R}^d$ be a continuously differentiable vector function. Let $\mathbf{n}$ be the outward normal to Ω. Then,

$$\int_\Omega \nabla \cdot \mathbf{v} \, d\mathbf{x} = \int_{\partial\Omega} \mathbf{v} \cdot \mathbf{n} \, ds$$

□

Species balance over an REV

- Let $u(\mathbf{x}, t) : \Omega \times [0, T] \to \mathbb{R}$ be the local amount of some species.
- Assume *representative elementary volume (REV)* $\omega \subset \Omega$
- Subinterval in time $(t_0, t1) \subset (0, T)$
- $-\delta \nabla u \cdot \mathbf{n}$ describes the flux of these species trough $\partial\omega$, where δ is some transfer coefficient
- Let $f(\mathbf{x}, t)$ be some local source of species. Then the flux through the boundary is balanced by the change of the amount of species in ω and the source strength:

$$0 = \int_\omega (u(\mathbf{x}, t_1) - u(\mathbf{x}, t_0))\, d\mathbf{x} - \int_{t_0}^{t_1} \int_{\partial\omega} \delta \nabla u \cdot \mathbf{n}\, ds\, dt - \int_{t_0}^{t_1} \int_\omega f(\mathbf{x}, t)\, ds$$

- Using Gauss' theorem, rewrite this as

$$0 = \int_{t_0}^{t_1} \int_\omega \partial_t u(\mathbf{x}, t)\, d\mathbf{x}\, dt - \int_{t_0}^{t_1} \int_\omega \nabla \cdot (\delta \nabla u)\, d\mathbf{x}\, dt - \int_{t_0}^{t_1} \int_\omega f(\mathbf{x}, t)\, ds$$

- True for all $\omega \subset \Omega$, $(t_0, t1) \subset (0, T) \Rightarrow$ parabolic second order PDE

$$\partial_t u(x, t) - \nabla \cdot (\delta \nabla u(x, t)) = f(x, t)$$

Second order PDEs

$$\partial_t u(x,t) - \nabla \cdot (\delta \nabla u(x,t)) = f(x,t)$$

For solvability we need additional conditions:

- Initial condition in the time dependent case: $u(x,0) = u_0(x)$
- Boundary conditions: behavior of solution on $\partial\Omega$?

Second order parabolic PDEs

- Heat conduction:
  u: temperature
  $\delta = \lambda$: heat conduction coefficient
  f: heat source
  flux$=-\lambda\nabla u$: "Fourier law"
- Diffusion of molecules in a given medium
  u: concentration
  $\delta = D$
  diffusion coefficient
  f: species source
  flux$=-D\nabla u$: "Fick's law"

Second order elliptic PDEs

Stationary case: $\partial_t u = 0 \Rightarrow$ second order *elliptic* PDE

$$-\nabla \cdot (\delta \nabla u(x)) = f(x)$$

- Stationary heat conduction, stationary diffusion
- Incompressible flow in saturated porous media: u: pressure
  $\delta = k$: permeability, flux$=-k\nabla u$: "Darcy's law"
- Electrical conduction: u: electric potential
  $\delta = \sigma$: electric conductivity
  flux$=-\sigma\nabla u \equiv$ current density: "Ohms's law"
- Poisson equation (electrostatics in a constant magnetic field):
  u: electrostatic potential, ∇u: electric field,
  $\delta = \varepsilon$: dielectric permittivity, f: charge density

Second order PDEs: boundary conditions

- Combine PDE in the interior with boundary conditions on variable u and/or or normal flux $\delta \nabla u \cdot \mathbf{n}$
- Assume $\partial\Omega = \cup_{i=1}^{N_\Gamma} \Gamma_i$ is the union of a finite number of non-intersecting subsets Γ_i which are locally Lipschitz.
- On each Γ_i, specify one of
  - Dirichlet ("first kind"): let $g_i : \Gamma_i \to \mathbb{R}$ (homogeneous for $g_i = 0$)

$$u(x) = u_{\Gamma_i}(x) \quad \text{for } x \in \Gamma_i$$

  - Neumann ("second kind"): Let $g_i : \Gamma_i \to \mathbb{R}$ (homogeneus for $g_i = 0$)

$$\delta \nabla u(x) \cdot \mathbf{n} = g_i(x) \quad \text{for } x \in \Gamma_i$$

  - Robin ("third kind"): let $\alpha_i, g_i : \Gamma_i \to \mathbb{R}$

$$\delta \nabla u(x) \cdot \mathbf{n} + \alpha_i(x)\,(u(x) - g_i(x)) = 0 \quad \text{for } x \in \Gamma_i$$

- Boundary functions may be time dependent.

The Dirichlet penalty method

- We will see later that the implementation of Dirichlet boundary may be connected with certain technical difficulties
- The Dirichlet penalty method provides a simple way to avoid these difficulties: let $\varepsilon > 0$:

$$\delta \nabla u_\varepsilon(x) \cdot \mathbf{n} + \frac{1}{\varepsilon} \left(u_\varepsilon(x) - g_i(x) \right) = 0 \quad \text{for } x \in \Gamma_i$$

- It is conceivable that for $\varepsilon \to 0$, u_ε converges to u with

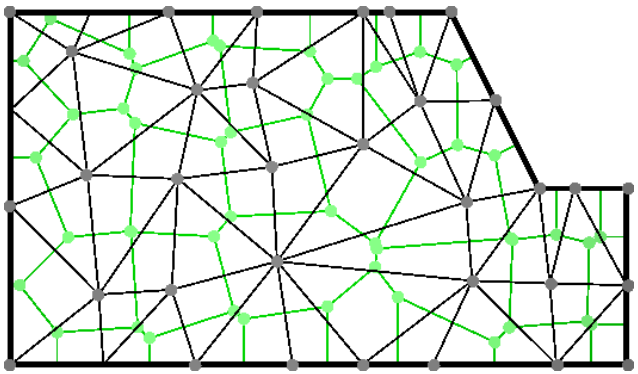
$$u(x) = g_i(x) \quad \text{for } x \in \Gamma_i$$

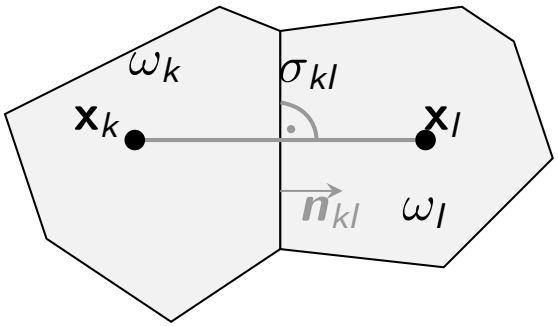
- Proper way of handling the parameter leads to exact fulfillment of Dirichlet boundary condition in the floating point precision

Constructing control volumes I

- Assume Ω is a polygon
- Subdivide the domain Ω into a finite number of **control volumes** : $\bar{\Omega} = \bigcup_{k \in \mathcal{N}} \bar{\omega}_k$ such that
  - ω_k are open (not containing their boundary) convex domains
  - $\omega_k \cap \omega_l = \emptyset$ if $\omega_k \neq \omega_l$
  - $\sigma_{kl} = \bar{\omega}_k \cap \bar{\omega}_l$ are either empty, points or straight lines
  - we will write $|\sigma_{kl}|$ for the length
  - if $|\sigma_{kl}| > 0$ we say that ω_k, ω_l are neighbours
  - neighbours of ω_k: $\mathcal{N}_k = \{l \in \mathcal{N} : |\sigma_{kl}| > 0\}$
- To each control volume ω_k assign a **collocation point**: $\mathbf{x}_k \in \bar{\omega}_k$ such that
  - **admissibility condition**:
    if $l \in \mathcal{N}_k$ then the line $\mathbf{x}_k\mathbf{x}_l$ is orthogonal to σ_{kl}
  - **placement of boundary unknowns**:
    if ω_k is situated at the boundary, i.e. for $|\partial\omega_k \cap \partial\Omega| > 0$, then $\mathbf{x}_k \in \partial\Omega$, and $\partial\omega_k \cap \partial\Omega = \cup_{i=1}^{N_\Gamma} \gamma_{i,k}$ (where $\gamma_{i,k} = \emptyset$ is possible).

Constructing control volumes II





We know how to construct such a partitioning:

- obtain a boundary conforming Delaunay triangulation with vertices x_k
- construct restricted Voronoi cells ω_k with $x_k \in \omega_k$
- Delaunay triangulation gives connected neigborhood graph of Voronoi cells
- Admissibility condition fulfilled in a natural way
- Boundary placement of triangle nodes

Discretization ansatz for Robin boundary value problem

Given constants $\kappa > 0$, $\alpha_i \geq 0$ $(i = 1 \ldots N_\Gamma)$

$$-\nabla \cdot \kappa \nabla u = f \text{ in } \Omega$$
$$\kappa \nabla u \cdot \mathbf{n} + \alpha_i (u - g_i) = 0 \text{ on } \Gamma_i \ (i = 1 \ldots N_\Gamma) \qquad (*)$$

- Given control volume ω_k, $k \in \mathcal{N}$, integrate

$$\begin{aligned}
0 &= \int_{\omega_k} (-\nabla \cdot \kappa \nabla u - f)\, d\omega \\
&= -\int_{\partial \omega_k} \kappa \nabla u \cdot \mathbf{n}_k d\gamma - \int_{\omega_k} f d\omega \qquad \text{(Gauss)} \\
&= -\sum_{l \in \mathcal{N}_k} \int_{\sigma_{kl}} \kappa \nabla u \cdot \mathbf{n}_{kl} d\gamma - \sum_{i=1}^{N_\Gamma} \int_{\gamma_{ik}} \kappa \nabla u \cdot \mathbf{n} d\gamma - \int_{\omega_k} f d\omega \\
&\approx \sum_{L \in \mathcal{N}_k} \underbrace{\kappa \frac{|\sigma_{kl}|}{h_{kl}} (u_k - u_l)}_{\nabla u \cdot \mathbf{n} \approx \frac{u_l - u_k}{h_{kl}}} + \sum_{i=1}^{N_\Gamma} \underbrace{|\gamma_{i,k}| \alpha_i (u_k - g_{i,k})}_{\text{bound. cond. } (*)} - \underbrace{|\omega_k| f_k}_{\text{quadrature}}
\end{aligned}$$

- Here, $u_k = u(\mathbf{x}_k)$, $g_{i,k} = g_i(\mathbf{x}_k)$, $f_k = f(\mathbf{x}_k)$

Properties of discretization matrix

- $N = |\mathcal{N}|$ equations (one for each control volume ω_k)
- $N = |\mathcal{N}|$ unknowns (one for each collocation point $x_k \in \omega_k$)
- weighted connected edge graph of triangulation $\equiv N \times N$ irreducible sparse discretization matrix $A = (a_{kl})$:

$$a_{kl} = \begin{cases} \sum_{l' \in \mathcal{N}_k} \kappa \frac{|\sigma_{kl'}|}{h_{kl'}} + \sum_{i=1}^{N_\Gamma} |\gamma_{i,k}| \alpha_i, & l = k \\ -\kappa \frac{\sigma_{kl}}{h_{kl}}, & l \in \mathcal{N}_k \\ 0, & else \end{cases}$$

- A is irreducibly diagonally dominant if at least for one i, $|\gamma_{i,k}|\alpha_i > 0$
- Main diagonal entries are positive, off diagonal entries are non-positive
- $\Rightarrow$ A has the M-property.
- A is symmetric $\Rightarrow$ A is positive definite

Matrix assembly: summary

- Sufficient to keep list of triangles, boundary segments – they typically come out of the mesh generator
- Be able to calculate triangular contributions to form factors: $|\omega_k \cap \tau|$, $|\sigma_{kl} \cap \tau|$ – we need only the numbers, and not the construction of the geometrical objects
- $O(N)$ operation, one loop over triangles, one loop over boundary elements

Interpretation of results

- One solution value per control volume ω_k allocated to the collocation point $x_k \Rightarrow$ piecewise constant function on collection of control volumes
- But: x_k are at the same time nodes of the corresponding Delaunay mesh $\Rightarrow$ representation as piecewise linear function on triangles

Convergence

- With finer discretization meshes we expect closer approximation to solution of continuous problem
- Current convergence theory for finite volume methods uses rather deep results from functional analysis: sequences of piecewise constant functions converging to differentiable solutions
- On structured grids, investigations based on Taylor expansions are possible
- How to find a simple ansatz for a convergence theory ?
  - Formulation of PDE in some space V
  - Approximation by elements of a subspace V_h

Problems with "strong formulation"

Writing the PDE with divergence and gradient assumes smoothness of coefficients and at least second derivatives for the solution.

- δ may not be continuous – what is then $\nabla \cdot (\delta \nabla u)$?
- Approximation of solution u e.g. by piecewise linear functions what does ∇u mean ?
- Spaces of twice, and even once continuously differentiable functions is not well suited:
  - Favorable approximation functions (e.g. piecewise linear ones) are not contained
  - Though they can be equipped with norms ($\Rightarrow$ Banach spaces) they have no scalar product $\Rightarrow$ no Hilbert spaces
  - Not complete: Cauchy sequences of functions may not converge to elements in these spaces

Cauchy sequences of functions

- Let Ω be a Lipschitz domain, let V be a metric space of functions $f : \Omega \to \mathbb{R}$
- Regard sequences of functions $f_n = \{f_n\}_{n=1}^{\infty} \subset V$
- A *Cauchy sequence* is a sequence f_n of functions where the norm of the difference between two elements can be made arbitrarily small by increasing the element numbers:

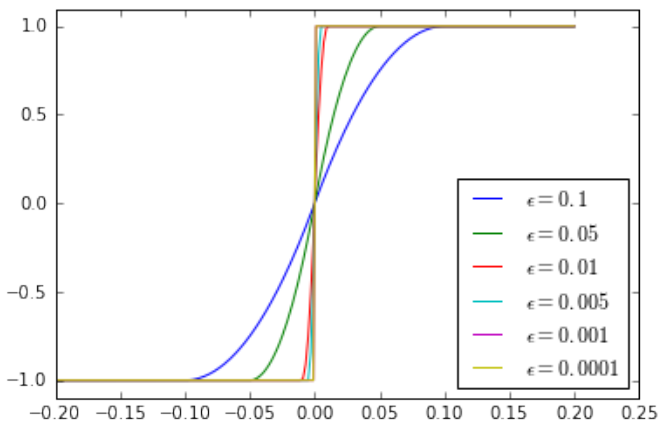
$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} : \forall m, n > n_0, ||f_n - f_m|| < \varepsilon$$

- All convergent sequences of functions are Cauchy sequences
- A metric space V is *complete* if all Cauchy sequences f_n of its elements have a limit $f = \lim_{n \to \infty} f_n \in V$ within this space

Completion of a metric space

- Let V be a metric space. Its completion is the space $\bar{V}$ consisting of all elements of V and all possible limits of Cauchy sequences of elements of V.
- This procedure allows to carry over definitions which are applicable only to elements of V to more general ones
- Example: step function

$$f_\epsilon(x) = \begin{cases} 1, & x \geq \epsilon \\ -(\frac{x-\epsilon}{\epsilon})^2 + 1, & 0 \leq x < \epsilon \\ (\frac{x+\epsilon}{\epsilon})^2 - 1, & -\epsilon \leq x < 0 \\ -1, & x < -\epsilon \end{cases} \xrightarrow{\epsilon \to 0} f(x) = \begin{cases} 1, & x \geq 0 \\ -1, & \text{else} \end{cases}$$



Riemann integral $\to$ Lebesgue integral

- Let Ω be a Lipschitz domain, let $C_c(\Omega)$ be the set of continuous functions $f : \Omega \to \mathbb{R}$ with compact support. ($\Rightarrow$ they vanish on $\partial\Omega$)
- For these functions, the Riemann integral $\int_\Omega f(x)d\mathbf{x}$ is well defined, and $\|f\|_{L^1} := \int_\Omega |f(x)|d\mathbf{x}$ provides a norm, and induces a metric.
- Let $L^1(\Omega)$ be the completion of $C_c(\Omega)$ with respect to the metric defined by the norm $\|\cdot\|_{L^1}$. That means that $L^1(\Omega)$ consists of all elements of $C_c(\Omega)$, and of all limites of Cauchy sequences of elements of $C_c(\Omega)$. Such functions are called *measurable*.
- For any measurable $f = \lim_{n\to\infty} f_n \in L^1(\Omega)$ with $f_n \in C_c(\Omega)$, define the *Lebesque integral*

$$\int_\Omega f(x)\, d\mathbf{x} := \lim_{n\to\infty} \int_\Omega f_n(x)\, d\mathbf{x}$$

as the limit of a sequence of Riemann integrals of continuous functions

Examples for Lebesgue integrable (measurable) functions

- Bounded functions which are continuous except in a finite number of points
- Step functions
- Equality of L^1 functions is elusive as they are not necessarily continuous: best what we can say is that they are equal "almost everywhere".
- In particular, L^1 functions whose values differ in a finite number of points are equal almost everywhere.

Spaces of integrable functions

- For $1 \leq p \leq \infty$, let $L^p(\Omega)$ be the space of measureable functions such that

$$\int_\Omega |f(x)|^p d\mathbf{x} < \infty$$

equipped with the norm

$$||f||_p = \left(\int_\Omega |f(x)|^p d\mathbf{x} \right)^{\frac{1}{p}}$$

- These spaces are *Banach spaces*, i.e. complete, normed vector spaces.
- The space $L^2(\Omega)$ is a *Hilbert space*, i.e. a Banach space equipped with a scalar product $(\cdot, \cdot)$ whose norm is induced by that scalar product, i.e. $||u|| = \sqrt{(u, u)}$. The scalar product in L^2 is

$$(f, g) = \int_\Omega f(x) g(x) d\mathbf{x}.$$

Green's theorem for smooth functions

Theorem Let $u, v \in C^1(\overline{\Omega})$ (continuously differentiable). Then for $\mathbf{n} = (n_1 \dots n_d)$ being the outward normal to Ω,

$$\int_\Omega u \partial_i v \, d\mathbf{x} = \int_{\partial\Omega} u v n_i \, ds - \int_\Omega v \partial_i u \, d\mathbf{x}$$

Corollaries

- Let $\mathbf{u} = (u_1 \dots u_d)$. Then

$$\int_\Omega (\sum_{i=1}^d u_i \partial_i v) \, d\mathbf{x} = \int_{\partial\Omega} v \sum_{i=1}^d (u_i n_i) \, ds - \int_\Omega v \sum_{i=1}^d (\partial_i u_i) \, d\mathbf{x}$$

$$\int_\Omega \mathbf{u} \cdot \nabla v \, d\mathbf{x} = \int_{\partial\Omega} v \mathbf{u} \cdot \mathbf{n} \, ds - \int_\Omega v \nabla \cdot \mathbf{u} \, d\mathbf{x}$$

- If $v = 0$ on $\partial\Omega$:

$$\int_\Omega u \partial_i v \, d\mathbf{x} = -\int_\Omega v \partial_i u \, d\mathbf{x}$$

$$\int_\Omega \mathbf{u} \cdot \nabla v \, d\mathbf{x} = -\int_\Omega v \nabla \cdot \mathbf{u} \, d\mathbf{x}$$

Weak derivative

- Let $L^1_{loc}(\Omega)$ be the set of functions which are Lebesgue integrable on every compact subset $K \subset \Omega$. Let $C_0^\infty(\Omega)$ be the set of functions infinitely differentiable with zero values on the boundary.
- For $u \in L^1_{loc}(\Omega)$ we *define* $\partial_i u$ by

$$\int_\Omega v \partial_i u d\mathbf{x} = -\int_\Omega u \partial_i v d\mathbf{x} \quad \forall v \in C_0^\infty(\Omega)$$

and $\partial^\alpha u$ by

$$\int_\Omega v \partial^\alpha u d\mathbf{x} = (-1)^{|\alpha|} \int_\Omega u \partial_i v d\mathbf{x} \quad \forall v \in C_0^\infty(\Omega)$$

if these integrals exist.

- For smooth functions, weak derivatives coincide with with the usual derivative

Sobolev spaces

- For $k \geq 0$ and $1 \leq p < \infty$, the *Sobolev space* $W^{k,p}(\Omega)$ is the space functions where all up to the k-th derivatives are in L^p:

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) : \partial^\alpha u \in L^p(\Omega)\ \forall |\alpha| \leq k\}$$

  with then norm

$$||u||_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha| \leq k} ||\partial^\alpha u||^p_{L^p(\Omega)} \right)^{\frac{1}{p}}$$

- Alternatively, they can be defined as the completion of C^∞ in the norm $||u||_{W^{k,p}(\Omega)}$
- $W_0^{k,p}(\Omega)$ is the completion of C_0^∞ in the norm $||u||_{W^{k,p}(\Omega)}$
- The Sobolev spaces are Banach spaces.

Sobolev spaces of square integrable functions

- $H^k(\Omega) = W^{k,2}(\Omega)$ with the scalar product

$$(u, v)_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \int_\Omega \partial^\alpha u \partial^\alpha v \, d\mathbf{x}$$

  is a Hilbert space.
- $H_0^k(\Omega) = W_0^{k,2}(\Omega)$ with the scalar product

$$(u, v)_{H_0^k(\Omega)} = \sum_{|\alpha| = k} \int_\Omega \partial^\alpha u \partial^\alpha v \, d\mathbf{x}$$

  is a Hilbert space as well.
- For this course the most important:
  - $L^2(\Omega)$, scalar product $(u, v)_{L^2(\Omega)} = (u, v)_{0,\Omega} = \int_\Omega uv \, d\mathbf{x}$
  - $H^1(\Omega)$, scalar product $(u, v)_{H^1(\Omega)} = (u, v)_{1,\Omega} = \int_\Omega (uv + \nabla u \cdot \nabla v) \, d\mathbf{x}$
  - $H_0^1(\Omega)$, scalar product $(u, v)_{H_0^1(\Omega)} = \int_\Omega (\nabla u \cdot \nabla v) \, d\mathbf{x}$
- Inequalities:

$$|(u, v)|^2 \leq (u, u)(v, v) \quad \text{Cauchy-Schwarz}$$
$$||u + v|| \leq ||u|| + ||v|| \quad \text{Triangle inequality}$$

A trace theorem

The notion of function values on the boundary initially is only well defined for continouos functions. So we need an extension of this notion to functions from Sobolev spaces.

Theorem: Let Ω be a bounded Lipschitz domain. Then there exists a bounded linear mapping

$$\mathrm{tr} : H^1(\Omega) \to L^2(\partial\Omega)$$

such that

(i) $\exists c > 0$ such that $\|\mathrm{tr}\, u\|_{0,\partial\Omega} \leq c\|u\|_{1,\Omega}$
(ii) $\forall u \in C^1(\bar{\Omega})$, $\mathrm{tr}\, u = u|_{\partial\Omega}$

□

Derivation of weak formulation

- Sobolev space theory provides a convenient framework to formulate existence, uniqueness and approximations of solutions of PDEs.
- Stationary heat conduction equation with homogeneous Dirichlet boundary conditions:

$$-\nabla \cdot \lambda \nabla u(x) = f(x) \text{ in } \Omega$$
$$u = 0 \text{ on } \partial\Omega$$

Multiply and integrate with an arbitrary *test function* $v \in C_0^\infty(\Omega)$ and apply Green's theorem using $v = 0$ on $\partial\Omega$

$$-\int_\Omega (\nabla \cdot \lambda \nabla u) v \, d\mathbf{x} = \int_\Omega f v \, d\mathbf{x}$$
$$\int_\Omega \lambda \nabla u \nabla v \, d\mathbf{x} = \int_\Omega f v \, d\mathbf{x}$$

Weak formulation of homogeneous Dirichlet problem

- Search $u \in H_0^1(\Omega)$ (here, $\operatorname{tr} u = 0$) such that

$$\int_\Omega \lambda \nabla u \nabla v \, d\mathbf{x} = \int_\Omega fv \, d\mathbf{x} \; \forall v \in H_0^1(\Omega)$$

- Then,

$$a(u, v) := \int_\Omega \lambda \nabla u \nabla v \, d\mathbf{x}$$

  is a self-adjoint bilinear form defined on the Hilbert space $H_0^1(\Omega)$.
- It is bounded due to Cauchy-Schwarz:

$$|a(u, v)| = |\lambda| \cdot | \int_\Omega \nabla u \nabla v \, d\mathbf{x}| \leq ||u||_{H_0^1(\Omega)} \cdot ||v||_{H_0^1(\Omega)}$$

- $f(v) = \int_\Omega fv \, d\mathbf{x}$ is a linear functional on $H_0^1(\Omega)$. For Hilbert spaces V the dual space V' (the space of linear functionals) can be identified with the space itself.

The Lax-Milgram lemma

Theorem: Let V be a Hilbert space. Let $a : V \times V \to \mathbb{R}$ be a self-adjoint bilinear form, and f a linear functional on V. Assume a is coercive, i.e.

$$\exists \alpha > 0 : \forall u \in V, a(u,u) \geq \alpha ||u||_V^2.$$

Then the problem: find $u \in V$ such that

$$a(u,v) = f(v) \; \forall v \in V$$

admits one and only one solution with an a priori estimate

$$||u||_V \leq \frac{1}{\alpha} ||f||_{V'}$$

□

Coercivity of weak formulation

Theorem: Assume $\lambda > 0$. Then the weak formulation of the heat conduction problem: search $u \in H_0^1(\Omega)$ such that

$$\int_\Omega \lambda \nabla u \nabla v \, d\mathbf{x} = \int_\Omega f v \, d\mathbf{x} \; \forall v \in H_0^1(\Omega)$$

has an unique solution.

Proof: $a(u, v)$ is cocercive:

$$a(u, v) = \int_\Omega \lambda \nabla u \nabla u \, d\mathbf{x} = \lambda ||u||^2_{H_0^1(\Omega)}$$

□

Weak formulation of inhomogeneous Dirichlet problem

$$-\nabla \cdot \lambda \nabla u = f \text{ in } \Omega$$
$$u = g \text{ on } \partial\Omega$$

- If g is smooth enough, there exists a *lifting* $u_g \in H^1(\Omega)$ such that $u_g|_{\partial\Omega} = g$. Then, we can re-formulate:

$$-\nabla \cdot \lambda \nabla (u - u_g) = f + \nabla \cdot \lambda \nabla u_g \text{ in } \Omega$$
$$u - u_g = 0 \text{ on } \partial\Omega$$

- Search $u \in H^1(\Omega)$ such that

$$u = u_g + \phi$$
$$\int_\Omega \lambda \nabla \phi \nabla v \, d\mathbf{x} = \int_\Omega f v \, d\mathbf{x} + \int_\Omega \lambda \nabla u_g \nabla v \quad \forall v \in H_0^1(\Omega)$$

Here, necessarily, $\phi \in H_0^1(\Omega)$ and we can apply the theory for the homogeneous Dirichlet problem.

Weak formulation of Robin problem

$$-\nabla \cdot \lambda \nabla u = f \text{ in } \Omega$$
$$\lambda \nabla u \cdot \mathbf{n} + \alpha(u - g) = 0 \text{ on } \partial\Omega$$

- Multiply and integrate with an arbitrary *test function* from $C_c^\infty(\Omega)$:

$$-\int_\Omega (\nabla \cdot \lambda \nabla u) v \, d\mathbf{x} = \int_\Omega f v \, d\mathbf{x}$$
$$\int_\Omega \lambda \nabla u \nabla v \, d\mathbf{x} + \int_{\partial\Omega} (\lambda \nabla u \cdot \mathbf{n}) v ds = \int_\Omega f v \, d\mathbf{x}$$
$$\int_\Omega \lambda \nabla u \nabla v \, d\mathbf{x} + \int_{\partial\Omega} \alpha u v \, ds = \int_\Omega f v \, d\mathbf{x} + \int_{\partial\Omega} \alpha g v \, ds$$

Weak formulation of Robin problem II

- Let

$$a^R(u,v) := \int_\Omega \lambda \nabla u \nabla v \, d\mathbf{x} + \int_{\partial\Omega} \alpha u v \, ds$$
$$f^R(v) := \int_\Omega f v \, d\mathbf{x} + \int_{\partial\Omega} \alpha g v \, ds$$

- Search $u \in H^1(\Omega)$ such that

$$a^R(u,v) = f^R(v) \; \forall v \in H^1(\Omega)$$

- If $\lambda > 0$ and $\alpha > 0$ then $a^R(u,v)$ is cocercive.

Neumann boundary conditions

- Homogeneous Neumann:

$$\lambda \nabla u \cdot \mathbf{n} = 0 \text{ on } \partial\Omega$$

- Inhomogeneous Neumann:

$$\lambda \nabla u \cdot \mathbf{n} = g \text{ on } \partial\Omega$$

- Weak formulation: Search $u \in H^1(\Omega)$ such that

$$\int_\Omega \nabla u \nabla v \, d\mathbf{x} = \int_{\partial\Omega} gv \, ds \; \forall v \in H^1(\Omega)$$

Not coercive due to the fact that we can add an arbitrary constant to u and $a(u, u)$ stays the same!

Further discussion on boundary conditions

- Mixed boundary conditions:
  One can have differerent boundary conditions on different parts of the boundary. In particular, if Dirichlet or Robin boundary conditions are applied on at least a part of the boundary of measure larger than zero, the binlinear form becomes coercive.
- Natural boundary conditions: Robin, Neumann
  These are imposed in a "natural" way in the weak formulation
- Essential boundary conditions: Dirichlet
  Explicitely imposed on the function space
- Coefficients $\lambda, \alpha \ldots$ can be functions from Sobolev spaces as long as they do not change integrability of terms in the bilinear forms

The Dirichlet penalty method

- Robin problem: search $u_\alpha \in H^1(\Omega)$ such that

$$\int_\Omega \lambda \nabla u_\alpha \nabla v \, d\mathbf{x} + \int_{\partial\Omega} \alpha u_\alpha v \, ds = \int_\Omega f v \, d\mathbf{x} + \int_{\partial\Omega} \alpha g v \, ds \quad \forall v \in H^1(\Omega)$$

- Dirichlet problem: search $u \in H^1(\Omega)$ such that

$$u = u_g + \phi \quad \text{where } u_g|_{\partial\Omega} = g$$

$$\int_\Omega \lambda \nabla \phi \nabla v \, d\mathbf{x} = \int_\Omega f v \, d\mathbf{x} + \int_\Omega \lambda \nabla u_g \nabla v \, d\mathbf{x} \;\; \forall v \in H_0^1(\Omega)$$

- Penalty limit:

$$\lim_{\alpha \to \infty} u_\alpha = u$$

- Formally, the convergence rate is quite low
- Implementing Dirichlet boundary conditions directly leads to a number of technical problems
- Implementing the penalty method is technically much simpler
- Proper way of handling the parameter leads to exact fulfillment of Dirichlet boundary condition in the floating point precision

The Finite Element Method

The Galerkin method I

- Weak formulations "live" in Hilbert spaces which essentially are infinite dimensional
- For computer representations we need finite dimensional approximations
- The Galerkin method and its modifications provide a general scheme for the derivation of finite dimensional appoximations
- Finite dimensional subspaces of Hilbert spaces are the spans of a set of basis functions, and are Hilbert spaces as well $\Rightarrow$ e.g. the Lax-Milgram lemma is valid there as well

The Galerkin method II

- Let V be a Hilbert space. Let $a : V \times V \to \mathbb{R}$ be a self-adjoint bilinear form, and f a linear functional on V. Assume a is coercive with coercivity constant α, and continuity constant γ.
- Continuous problem: search $u \in V$ such that

$$a(u, v) = f(v) \; \forall v \in V$$

- Let $V_h \subset V$ be a finite dimensional subspace of V
- "Discrete" problem $\equiv$ Galerkin approximation:
  Search $u_h \in V_h$ such that

$$a(u_h, v_h) = f(v_h) \; \forall v_h \in V_h$$

By Lax-Milgram, this problem has a unique solution as well.

Céa's lemma

- What is the connection between u and u_h ?
- Let $v_h \in V_h$ be arbitrary. Then

$$\begin{aligned}
\alpha||u - u_h||^2 &\leq a(u - u_h, u - u_h) \quad \text{(Coercivity)} \\
&= a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h) \\
&= a(u - u_h, u - v_h) \quad \text{(Galerkin Orthogonality)} \\
&\leq \gamma||u - u_h|| \cdot ||u - v_h|| \quad \text{(Boundedness)}
\end{aligned}$$

- As a result

$$||u - u_h|| \leq \frac{\gamma}{\alpha} \inf_{v_h \in V_h} ||u - v_h||$$

- Up to a constant, the error of the Galerkin approximation is the error of the best approximation of the solution in the subspace V_h.

From the Galerkin method to the matrix equation

- Let $\phi_1 \dots \phi_n$ be a set of basis functions of V_h.
- Then, we have the representation $u_h = \sum_{j=1}^{n} u_j \phi_j$
- In order to search $u_h \in V_h$ such that

$$a(u_h, v_h) = f(v_h) \; \forall v_h \in V_h$$

it is actually sufficient to require

$$a(u_h, \phi_i) = f(\phi_i) \; (i = 1 \dots n)$$

$$a\left(\sum_{j=1}^{n} u_j \phi_j, \phi_i\right) = f(\phi_i) \; (i = 1 \dots n)$$

$$\sum_{j=1}^{n} a(\phi_j, \phi_i) u_j = f(\phi_i) \; (i = 1 \dots n)$$

$$AU = F$$

with $A = (a_{ij})$, $a_{ij} = a(\phi_i, \phi_j)$, $F = (f_i)$, $f_i = F(\phi_i)$, $U = (u_i)$.

- Matrix dimension is $n \times n$. Matrix sparsity ?

Obtaining a finite dimensional subspace

- Let $\Omega = (a, b) \subset \mathbb{R}^1$
- Let $a(u, v) = \int_\Omega \lambda(x) \nabla u \nabla v d\mathbf{x}$.
- Analysis I provides a finite dimensional subspace: the space of sin/cos functions up to a certain frequency $\Rightarrow$ *spectral method*
- Ansatz functions have global support $\Rightarrow$ full $n \times n$ matrix
- OTOH: rather fast convergence for smooth data
- Generalization to higher dimensions possible
- Big problem in irregular domains: we need the eigenfunction basis of some operator...
- Spectral methods are successful in cases where one has regular geometry structures and smooth/constant coefficients – e.g. "Spectral Einstein Code"

The finite element idea

- Choose basis functions with local support. In this case, the matrix becomes sparse, as only integrals of basis function pairs with overlapping support contribute to the matrix.
- Linear finite elements in $\Omega = (a, b) \subset \mathbb{R}^1$:
- Partition $a = x_1 \leq x_2 \leq \cdots \leq x_n = b$
- Basis functions (for $i = 1 \ldots n$)

$$\phi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & i > 1, x \in (x_{i-1}, x_i) \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & i < n, x \in (x_i, x_{i+1}) \\ 0, & \text{else} \end{cases}$$

- Any function $u_h \in V_h = \operatorname{span}\{\phi_1 \ldots \phi_n\}$ is piecewise linear, and the coefficients in the representation $u_h = \sum_{i=1}^{n} u_i \phi_i$ are the values $u_h(x_i)$.
- Fortunately, we are working with a weak formulation, and weak derivatives are well defined !

1D matrix elements for heat equation

- Assume ($\lambda = 1$, $x_{i+1} - x_i = h$)

$$a(u, v) = \int_a^b \nabla u \nabla v \, dx + \alpha u(a)v(a) + \alpha u(b)v(b)$$

- The integrals are nonzero for $i = j, i + 1 = j, i - 1 = j$
- Let $j = i + 1$

$$\int_\Omega \nabla \phi_i \nabla \phi_j d\mathbf{x} = \int_{x_i}^{x_{i+1}} \nabla \phi_i \nabla \phi_j d\mathbf{x} = -\int_{x_i}^{x_{i+1}} \frac{1}{h^2} d\mathbf{x} = -\frac{1}{h} d\mathbf{x}$$

- Similarly, for $j = i + 1$, $\int_\Omega \nabla \phi_i \nabla \phi_j d\mathbf{x} = -\frac{1}{h}$
- For $1 < i < N$:

$$\int_\Omega \nabla \phi_i \nabla \phi_i d\mathbf{x} = \int_{x_{i-1}}^{x_{i+1}} \nabla \phi_i \nabla \phi_i d\mathbf{x} = \int_{x_{i-1}}^{x_{i+1}} \frac{1}{h^2} d\mathbf{x} = \frac{2}{h}$$

- For $i = 1$ or $i = N$, $a(\phi_i, \phi_i) = \frac{1}{h} + \alpha$

1D matrix elements II

Adding the boundary integrals yields

$$A = \begin{pmatrix} \alpha + \frac{1}{h} & -\frac{1}{h} & & & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & & \\ & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\ & \ddots & \ddots & \ddots & \ddots & \\ & & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\ & & & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} \\ & & & & -\frac{1}{h} & \frac{1}{h} + \alpha \end{pmatrix}$$

. . . the same matrix as for the finite difference and finite volume methods