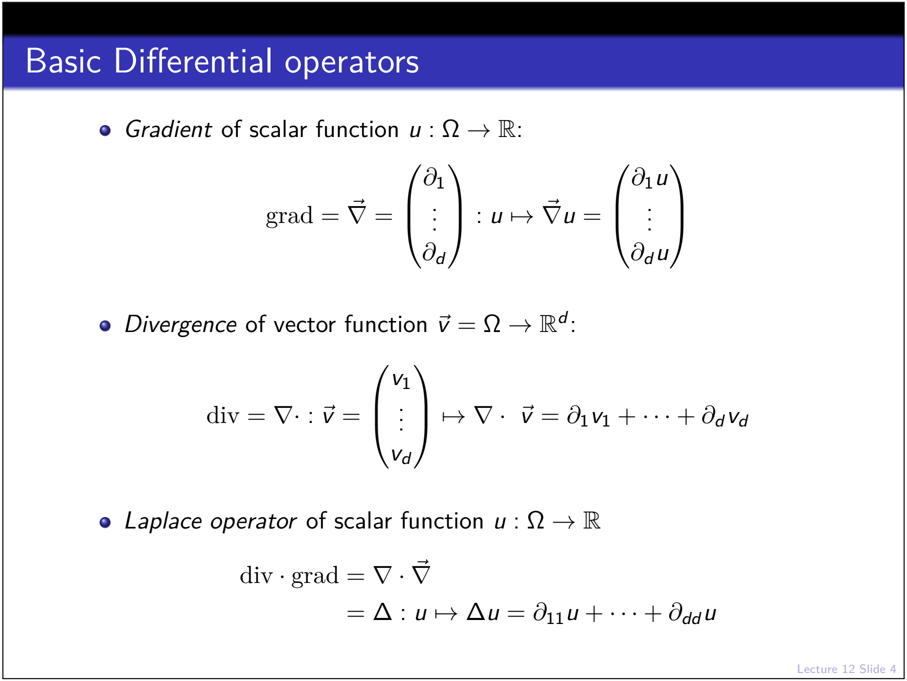
Scientific Computing WS 2019/2020

Lecture 14

Jürgen Fuhrmann

juergen.fuhrmann@wias-berlin.de

Recap: PDEs and Sobolev spaces



Basic Differential operators

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

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

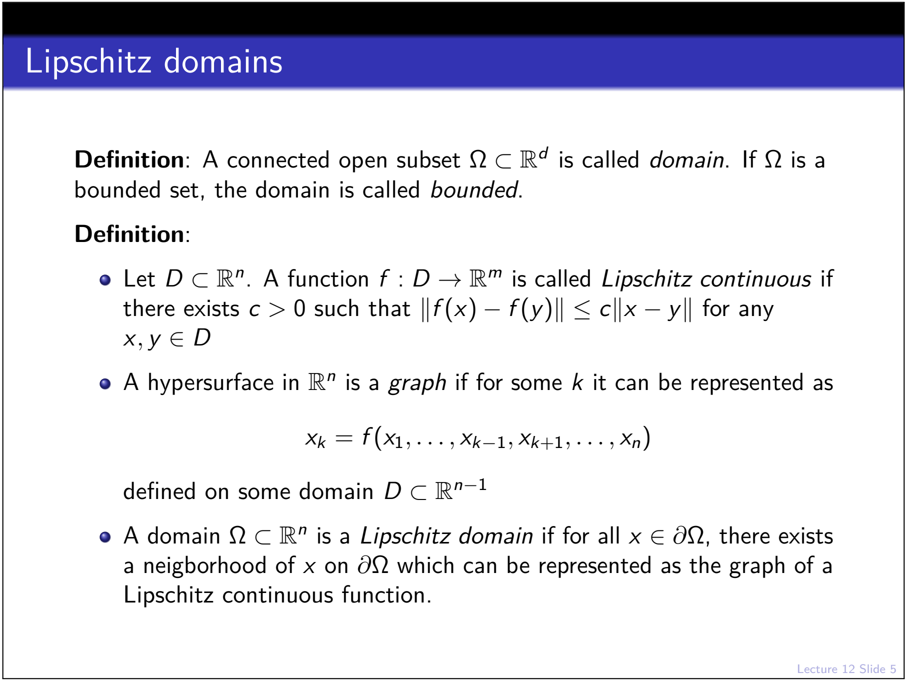
- *Divergence* of vector function $\vec{v} = \Omega \to \mathbb{R}^d$:

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

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

$$\begin{aligned} \mathrm{div} \cdot \mathrm{grad} &= \nabla \cdot \vec{\nabla} \\ &= \Delta : u \mapsto \Delta u = \partial_{11} u + \cdots + \partial_{dd} u \end{aligned}$$

Lecture 12 Slide 4



Lipschitz domains

Definition: A connected open subset $\Omega \subset \mathbb{R}^d$ is called *domain*. If Ω is a bounded set, the domain is called *bounded*.

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 $\|f(x) - f(y)\| \leq c\|x - y\|$ for any $x, y \in D$
- A hypersurface in $\mathbb{R}^n$ is a *graph* if for some k it can be represented as

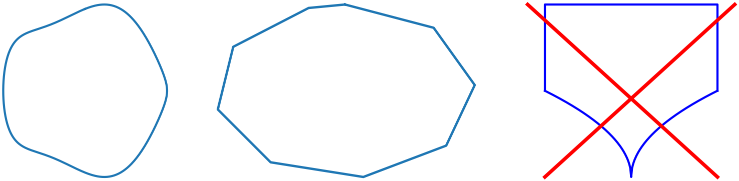
  $$x_k = f(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$$

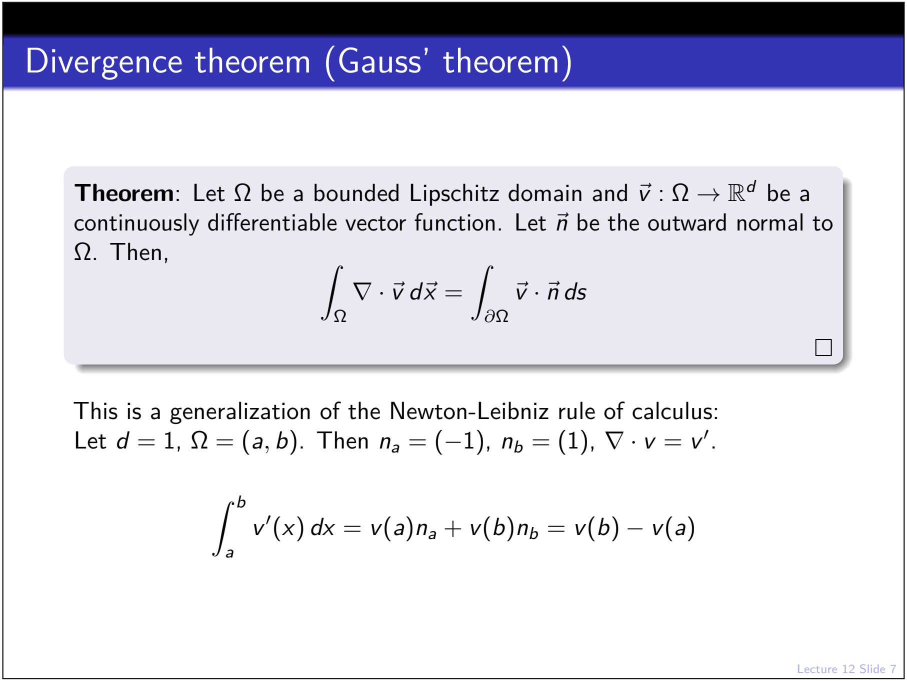
  defined on some domain $D \subset \mathbb{R}^{n-1}$
- A domain $\Omega \subset \mathbb{R}^n$ is a *Lipschitz domain* if for all $x \in \partial\Omega$, there exists a neigborhood of x on $\partial\Omega$ which can be represented as the graph of a Lipschitz continuous function.

Lipschitz domains II

Standard PDE calculus happens in Lipschitz domains

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





Divergence theorem (Gauss' theorem)

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

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

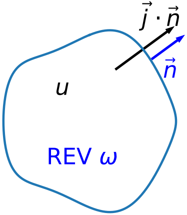
□

This is a generalization of the Newton-Leibniz rule of calculus:
Let $d = 1$, $\Omega = (a, b)$. Then $n_a = (-1)$, $n_b = (1)$, $\nabla \cdot v = v'$.

$$\int_a^b v'(x)\, dx = v(a)n_a + v(b)n_b = v(b) - v(a)$$

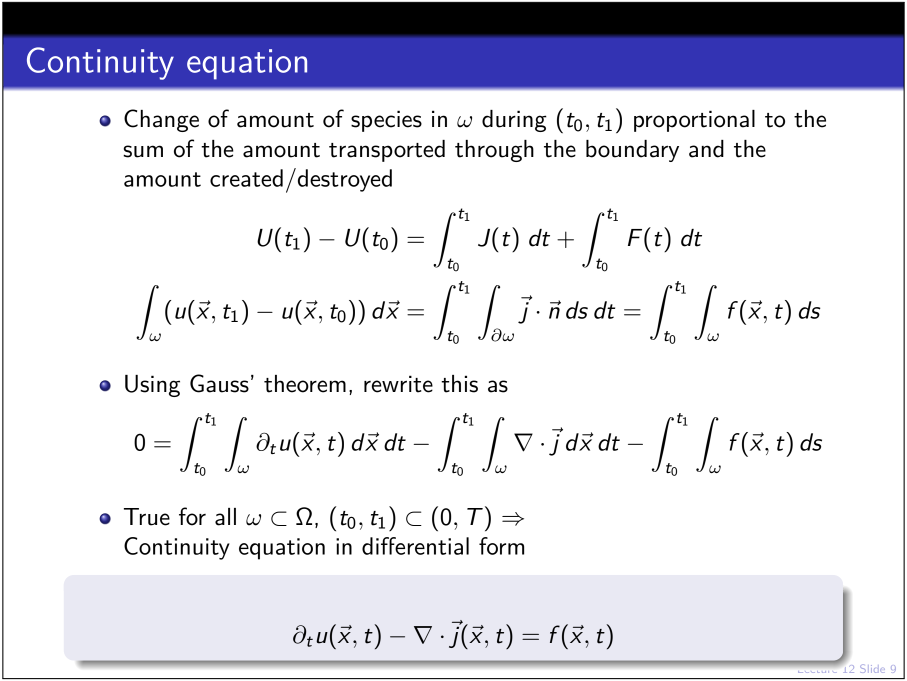
Species flux through boundary of an REV

- $u(\vec{x}, t) : \Omega \times [0, T] \to \mathbb{R}$: time dependent local amount of species
- $f(\vec{x}, t) : \Omega \times [0, T] \to \mathbb{R}$: species sources/sinks
- $\vec{j}(\vec{x}, t)$: vector field of the species flux



- $\omega \subset \Omega$: *representative elementary volume (REV)*
- $(t_0, t_1) \subset (0, T)$: subset of the time interval

- $J(t) = \int_{\partial\omega} \vec{j}(\vec{x}, t) \cdot \vec{n}\, \mathrm{ds}$: flux of species trough $\partial\omega$ at moment t
- $U(t) = \int_\omega u(\vec{x}, t)\, d\vec{x}$: amount of species in ω at moment t
- $F(t) = \int_\omega f(\vec{x}, t)\, d\vec{x}$: rate of creation/destruction at moment t



Continuity equation

- Change of amount of species in ω during (t_0, t_1) proportional to the sum of the amount transported through the boundary and the amount created/destroyed

$$U(t_1) - U(t_0) = \int_{t_0}^{t_1} J(t)\, dt + \int_{t_0}^{t_1} F(t)\, dt$$

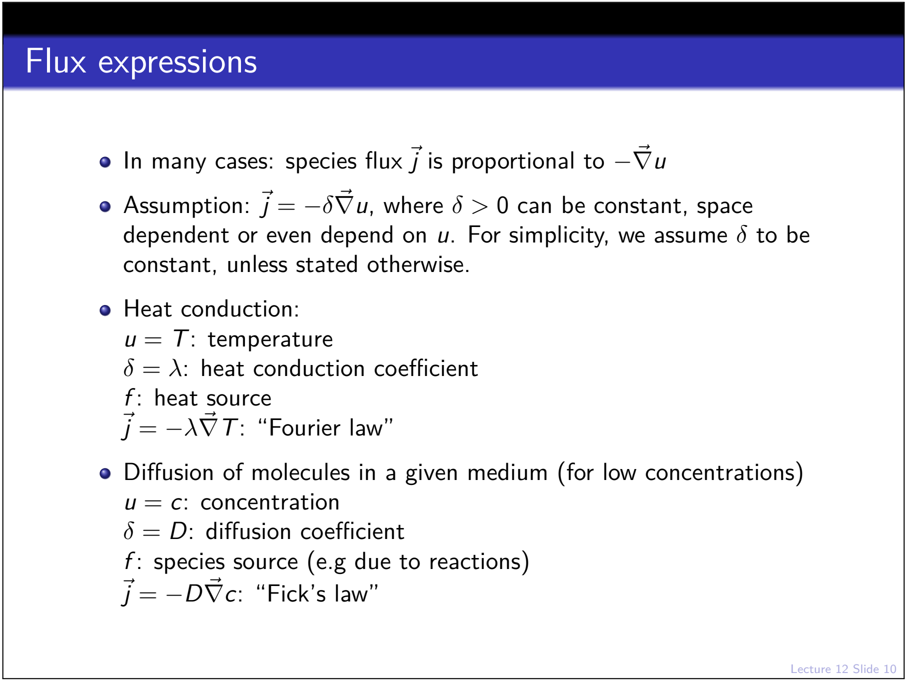
$$\int_\omega (u(\vec{x}, t_1) - u(\vec{x}, t_0))\, d\vec{x} = \int_{t_0}^{t_1} \int_{\partial\omega} \vec{j} \cdot \vec{n}\, ds\, dt = \int_{t_0}^{t_1} \int_\omega f(\vec{x}, t)\, ds$$

- Using Gauss' theorem, rewrite this as

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

- True for all $\omega \subset \Omega$, $(t_0, t_1) \subset (0, T) \Rightarrow$
  Continuity equation in differential form

$$\partial_t u(\vec{x}, t) - \nabla \cdot \vec{j}(\vec{x}, t) = f(\vec{x}, t)$$



Flux expressions

- In many cases: species flux $\vec{j}$ is proportional to $-\vec{\nabla} u$
- Assumption: $\vec{j} = -\delta\vec{\nabla} u$, where $\delta > 0$ can be constant, space dependent or even depend on u. For simplicity, we assume δ to be constant, unless stated otherwise.
- Heat conduction:
  $u = T$: temperature
  $\delta = \lambda$: heat conduction coefficient
  f: heat source
  $\vec{j} = -\lambda\vec{\nabla} T$: “Fourier law”
- Diffusion of molecules in a given medium (for low concentrations)
  $u = c$: concentration
  $\delta = D$: diffusion coefficient
  f: species source (e.g due to reactions)
  $\vec{j} = -D\vec{\nabla} c$: “Fick’s law”

More flux expressions

- Flow in a saturated porous medium:
  $u = p$: pressure
  $\delta = k$: permeability
  $\vec{j} = -k\vec{\nabla}p$: "Darcy's law"
- Electrical conduction:
  $u = \varphi$: electric potential
  $\delta = \sigma$: electric conductivity
  $\vec{j} = -\sigma\vec{\nabla}\varphi \equiv$ current density: "Ohms's law"
- Electrostatics in a constant magnetic field:
  $u = \varphi$: electric potential
  $\delta = \varepsilon$: dielectric permittivity
  $\vec{E} = \vec{\nabla}\phi$: electric field
  $\vec{j} = \vec{D} = \varepsilon\vec{E} = \varepsilon\vec{\nabla}\varphi$: electric displacement field: "Gauss's Law"
  $f = \rho$: charge density

Second order partial differential equaions (PDEs)

Combine continuity equation with flux expression:

- Transient problem:

Parabolic PDE:

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

- Stationary case: $\partial_t u = 0 \Rightarrow$

Elliptic PDE

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

- For solvability we need additional conditions:
  - Initial condition in the time dependent case: $u(\vec{x},0) = u_0(\vec{x})$
  - Boundary conditions: behavior of solution on $\partial\Omega$

Second order PDEs: boundary conditions

- 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
  - Fixed solution at boundary ⇒ *Dirichlet* ("first kind") BC: let $g_i : \Gamma_i \to \mathbb{R}$ (*homogeneous* for $g_i = 0$)

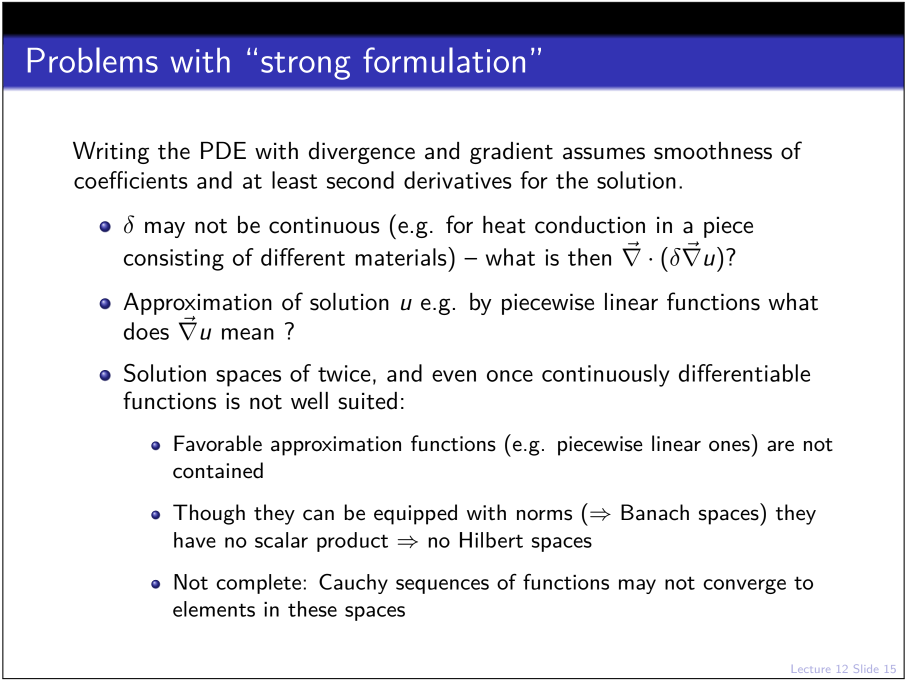
    $$u(\vec{x}, t) = g_i(\vec{x}, t) \quad \text{for } \vec{x} \in \Gamma_i$$

  - Fixed boundary flux ⇒ *Neumann* ("second kind") BC: Let $g_i : \Gamma_i \to \mathbb{R}$ (*homogeneus* for $g_i = 0$)

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

  - Boundary flux proportional to solution ⇒ *Robin* ("third kind") BC: let $\alpha_i > 0, g_i : \Gamma_i \to \mathbb{R}$

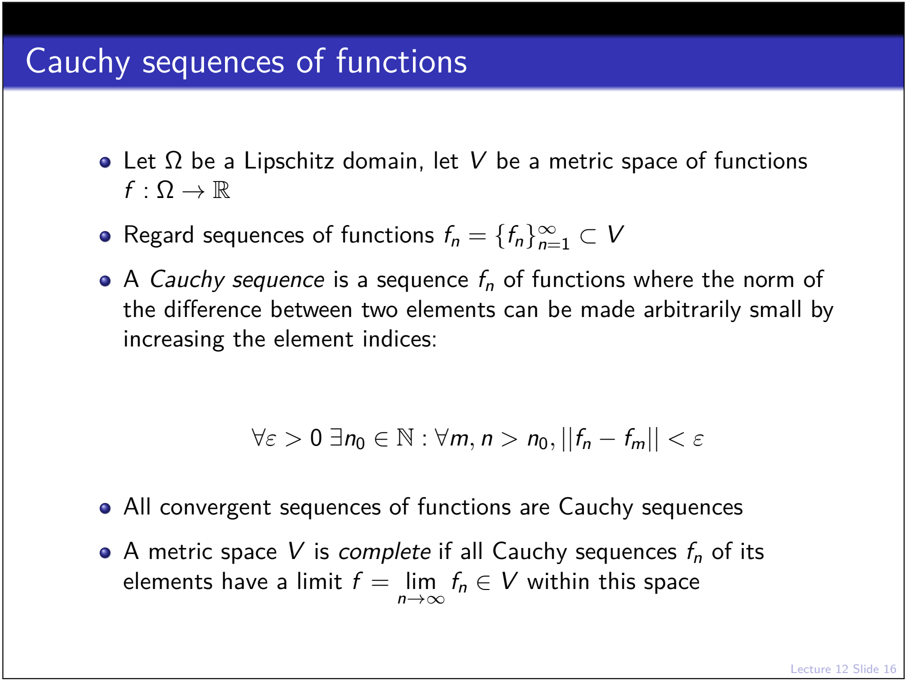
    $$\delta\vec{\nabla} u(\vec{x}, t) \cdot \vec{n} + \alpha_i(\vec{x}, t) u(\vec{x}, t) = g_i(\vec{x}, t) \quad \text{for } \vec{x} \in \Gamma_i$$



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 (e.g. for heat conduction in a piece consisting of different materials) – what is then $\vec{\nabla} \cdot (\delta \vec{\nabla} u)$?
- Approximation of solution u e.g. by piecewise linear functions what does $\vec{\nabla} u$ mean ?
- Solution 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 indices:

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

- 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

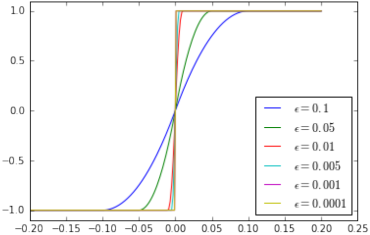
- Let V be a metric space. Its completion is the space $\bar{V}$ consisting of all elements of V and all possible limits of Cauchy sequences of elements of V.
- This procedure allows to carry over definitions which are applicable only to elements of V to more general ones
- This process depends on the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

Example: construction of real numbers $\bar{V} = \mathbb{R}$ from rational numbers $V = \mathbb{Q}$ via Cauchy sequences: every real number is an equivalence class of Cauchy sequences with the same limit.

Completion in function spaces

- Example: step function

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



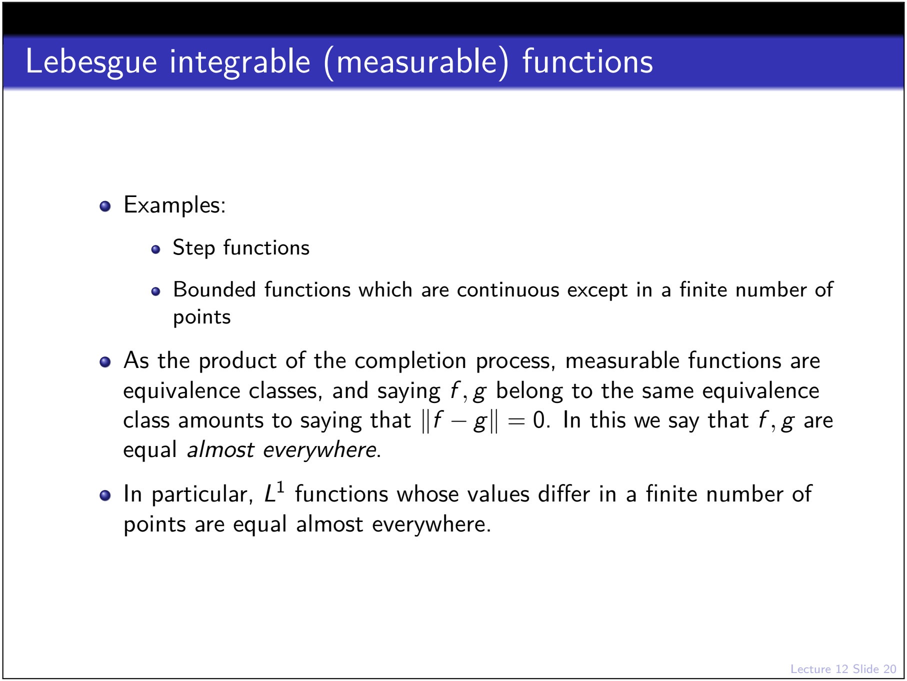
- The discontinuous function $\theta(x)$ is the limit of a sequence of continuously differentiable functions θ_ϵ.

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 $\int_\Omega f(\vec{x})\, d\vec{x}$ is well defined, and $\|f\|_{L^1} := \int_\Omega |f(\vec{x})| d\vec{x}$ 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 limits 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*

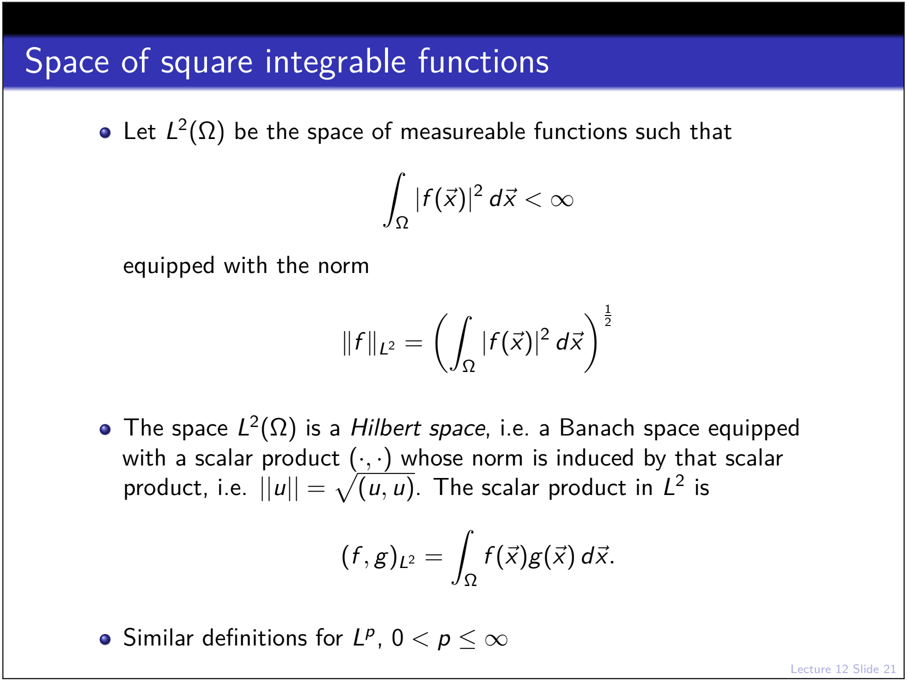
$$\int_\Omega f(\vec{x})\, d\vec{x} := \lim_{n\to\infty} \int_\Omega f_n(\vec{x})\, d\vec{x}$$

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



Lebesgue integrable (measurable) functions

- Examples:
  - Step functions
  - Bounded functions which are continuous except in a finite number of points
- As the product of the completion process, measurable functions are equivalence classes, and saying f, g belong to the same equivalence class amounts to saying that $\|f - g\| = 0$. In this we say that f, g are equal *almost everywhere*.
- In particular, L^1 functions whose values differ in a finite number of points are equal almost everywhere.



Space of square integrable functions

- Let $L^2(\Omega)$ be the space of measureable functions such that

$$\int_\Omega |f(\vec{x})|^2 \, d\vec{x} < \infty$$

equipped with the norm

$$\|f\|_{L^2} = \left(\int_\Omega |f(\vec{x})|^2 \, d\vec{x} \right)^{\frac{1}{2}}$$

- 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

$$(f, g)_{L^2} = \int_\Omega f(\vec{x}) g(\vec{x}) \, d\vec{x}.$$

- Similar definitions for L^p, $0 < p \leq \infty$

Green's theorem for smooth functions

Theorem Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain and $u, v \in C^1(\overline{\Omega})$ (continuously differentiable). Let $\vec{n} = (n_1 \dots n_d)$ being the outward normal $\partial\Omega$. Then Ω,

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

□

This is a generalization of the integration by parts rule of calculus: Let $d = 1$, $\Omega = (a, b)$. Then $n_a = (-1)$, $n_b = 1$, $\partial_i(\cdot) = (\cdot)'$.

$$\int_a^b uv'(x) \, dx = n_a u(a)v(a) + n_b u(b)v(b) - \int_a^b u'v \, dx$$

$$= uv\Big|_a^b - \int_a^b u'v \, dx$$

Corollaries of Green's theorem

- Let $\vec{u} = (u_1 \dots u_d) : \Omega \to \mathbb{R}^d$ and $v : \Omega \to \mathbb{R}$. Then

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

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

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

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

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

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

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

  and $\partial^\alpha u$ by

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

  if these integrals exist.
- For smooth functions, weak derivatives coincide with with the usual derivative

Sobolev spaces of square integrable functios

- For $k \geq 0$ the *Sobolev space* $H^k(\Omega)$ is the space functions where all up to the k-th derivatives are in L^2:

$$H^k(\Omega) = \{u \in L^2(\Omega) : \partial^\alpha u \in L^2(\Omega)\ \forall |\alpha| \leq k\}$$

  with then norm

$$||u||_{H^k(\Omega)} = \left(\sum_{|\alpha| \leq k} ||\partial^\alpha u||^2_{L^2(\Omega)} \right)^{\frac{1}{2}}$$

- Alternatively, H^k is the completion of C^∞ in the norm $||u||_{H^k(\Omega)}$
- $H^k_0(\Omega)$ is the completion of C^∞_0 in the norm $||u||_{H^k(\Omega)}$
- These Sobolev spaces are Banach spaces.
- Similar definitions exist for $p \neq 2$

Important function spaces

- $H^k(\Omega)$ is a Hilbert space with the scalar product

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

- $H_0^k(\Omega)$ is a Hilbert space with the scalar product

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

Hilbert space structure

- For this course the most important:
  - $L^2(\Omega)$, scalar product $(u,v)_{L^2(\Omega)} = (u,v)_{0,\Omega} = \int_\Omega uv\ d\vec{x}$
  - $H^1(\Omega)$, scalar product $(u,v)_{H^1(\Omega)} = (u,v)_{1,\Omega} = \int_\Omega (uv + \vec{\nabla} u \cdot \vec{\nabla} v)\ d\vec{x}$
  - $H^1_0(\Omega)$, scalar product $(u,v)_{H^1_0(\Omega)} = \int_\Omega (\vec{\nabla} u \cdot \vec{\nabla} v)\ d\vec{x}$
- All of them are metric spaces with a scalar product and we have in each of them

$$
\begin{aligned}
\left|(u,v)\right|^2 &\leq (u,u)(v,v) && \text{Cauchy-Schwarz}\\
||u+v|| &\leq ||u|| + ||v|| && \text{Triangle inequality}
\end{aligned}
$$

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) $\forall u \in C^1(\bar{\Omega})$, $\mathrm{tr}\, u = u|_{\partial\Omega}$

□

Corollary: If $u \in H^1_0(\Omega)$ then $\mathrm{tr}\, u = 0$.

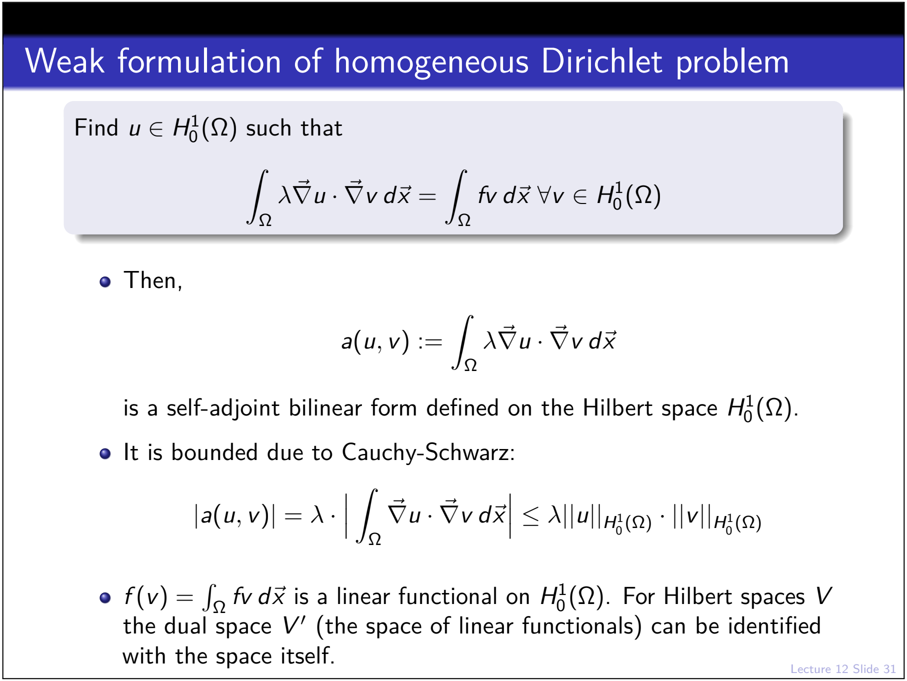
Derivation of weak formulation

- Sobolev space theory provides a convenient framework to formulate existence, uniqueness and approximations of solutions of PDEs.
- Let us first consider the stationary heat conduction equation with homogeneous Dirichlet boundary conditions and constant heat conduction coefficient $\lambda > 0$:

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

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

$$\begin{aligned} & -\int_\Omega (\nabla \cdot \lambda \vec{\nabla} u) v \, d\vec{x} = \int_\Omega f v \, d\vec{x} \\ \Rightarrow \quad & \int_\Omega \lambda \vec{\nabla} u \cdot \vec{\nabla} v \, d\vec{x} = \int_\Omega f v \, d\vec{x} \end{aligned}$$



Weak formulation of homogeneous Dirichlet problem

Find $u \in H^1_0(\Omega)$ such that

$$\int_\Omega \lambda \vec{\nabla} u \cdot \vec{\nabla} v \, d\vec{x} = \int_\Omega fv \, d\vec{x} \; \forall v \in H^1_0(\Omega)$$

- Then,

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

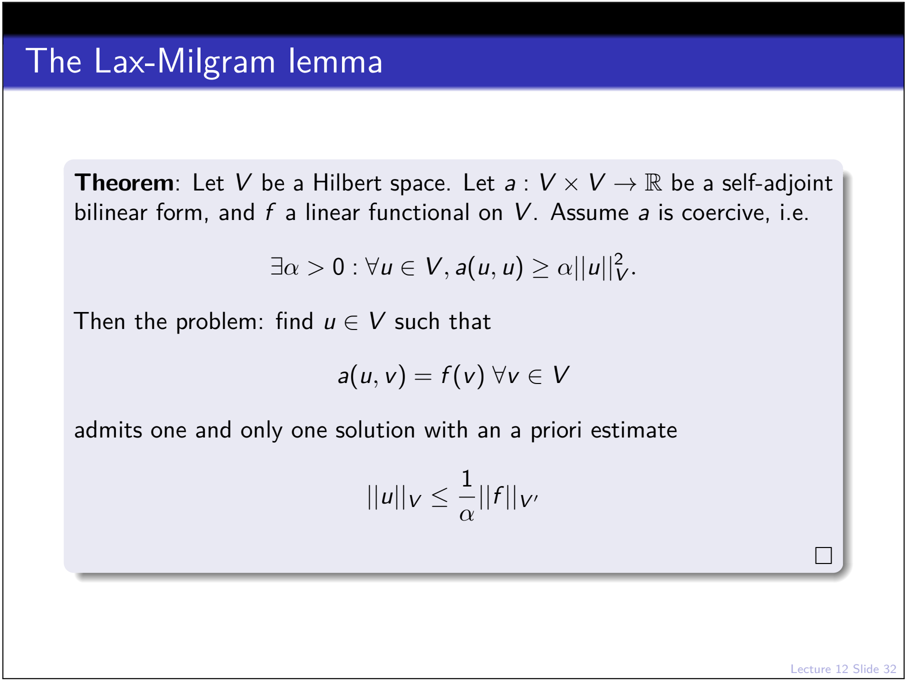
  is a self-adjoint bilinear form defined on the Hilbert space $H^1_0(\Omega)$.

- It is bounded due to Cauchy-Schwarz:

  $$|a(u,v)| = \lambda \cdot \left| \int_\Omega \vec{\nabla} u \cdot \vec{\nabla} v \, d\vec{x} \right| \leq \lambda ||u||_{H^1_0(\Omega)} \cdot ||v||_{H^1_0(\Omega)}$$

- $f(v) = \int_\Omega fv \, d\vec{x}$ is a linear functional on $H^1_0(\Omega)$. For Hilbert spaces V the dual space V' (the space of linear functionals) can be identified with the space itself.

Lecture 12 Slide 31



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.

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

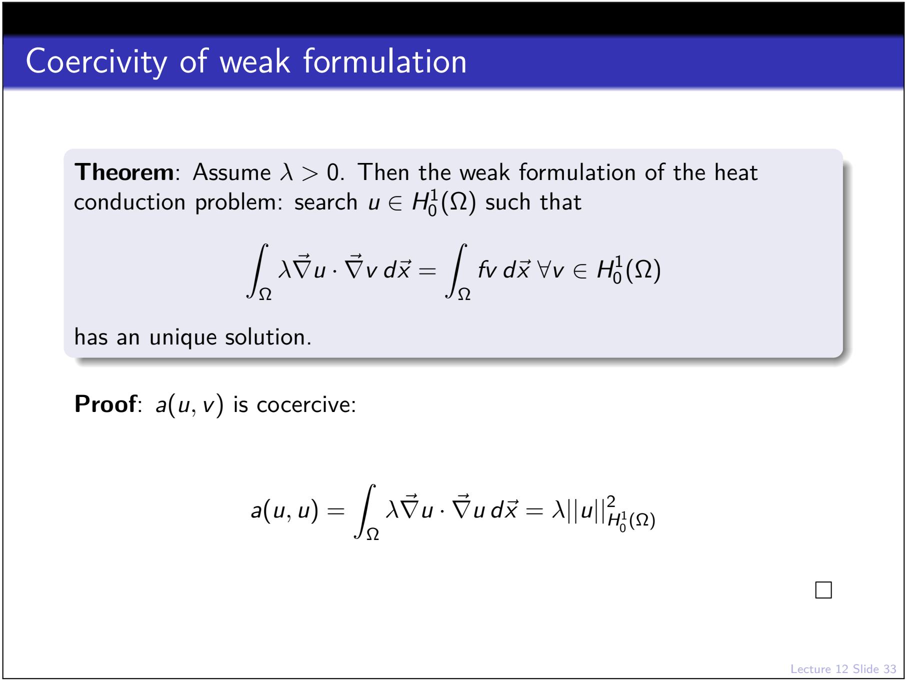
Then the problem: find $u \in V$ such that

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

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

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

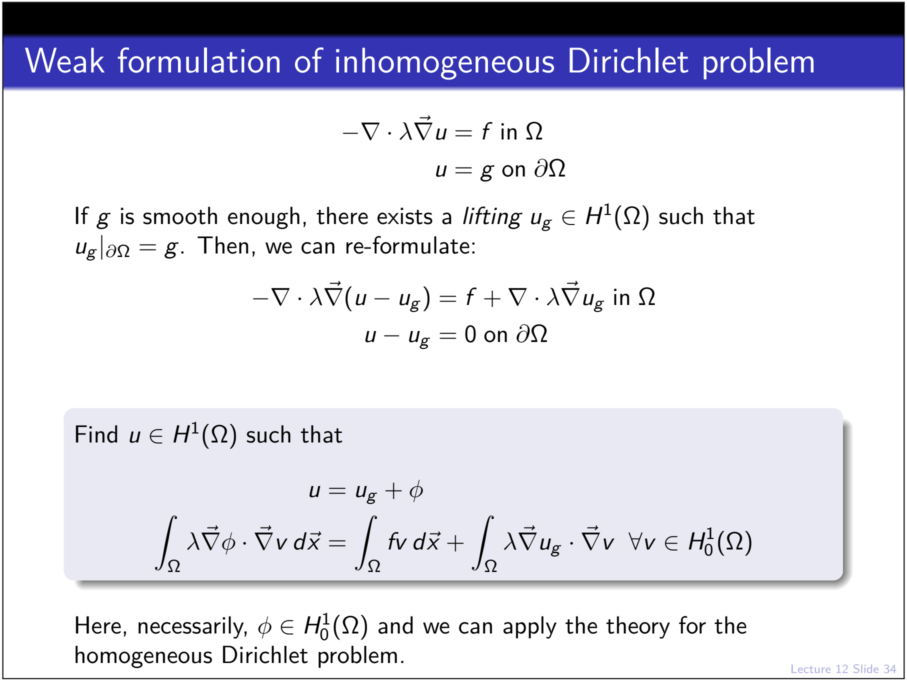
has an unique solution.

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

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

□

Lecture 12 Slide 33



Weak formulation of inhomogeneous Dirichlet problem

$$\begin{aligned} -\nabla \cdot \lambda \vec{\nabla} u &= f \text{ in } \Omega \\ u &= g \text{ on } \partial\Omega \end{aligned}$$

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:

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

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

$$u = u_g + \phi$$

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

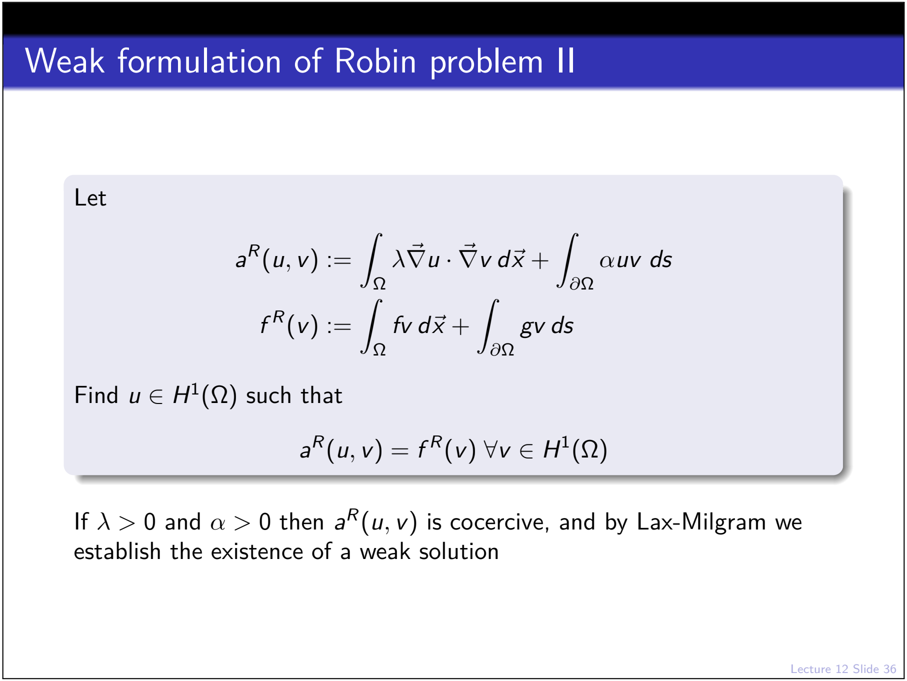
Here, necessarily, $\phi \in H_0^1(\Omega)$ and we can apply the theory for the homogeneous Dirichlet problem.

Weak formulation of Robin problem

$$
\begin{aligned}
-\nabla \cdot \lambda \vec{\nabla} u &= f \text{ in } \Omega \\
\lambda \vec{\nabla} u \cdot \vec{n} + \alpha u &= g \text{ on } \partial\Omega
\end{aligned}
$$

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

$$
\begin{aligned}
-\int_\Omega (\nabla \cdot \lambda \vec{\nabla} u) v \, d\vec{x} &= \int_\Omega fv \, d\vec{x} \\
\int_\Omega \lambda \vec{\nabla} u \cdot \vec{\nabla} v \, d\vec{x} + \int_{\partial\Omega} \lambda \vec{\nabla} u \cdot \vec{n} v ds &= \int_\Omega fv \, d\vec{x} \\
\int_\Omega \lambda \vec{\nabla} u \cdot \vec{\nabla} v \, d\vec{x} + \int_{\partial\Omega} \alpha uv \; ds &= \int_\Omega fv \, d\vec{x} + \int_{\partial\Omega} gv \; ds
\end{aligned}
$$



Weak formulation of Robin problem II

Let

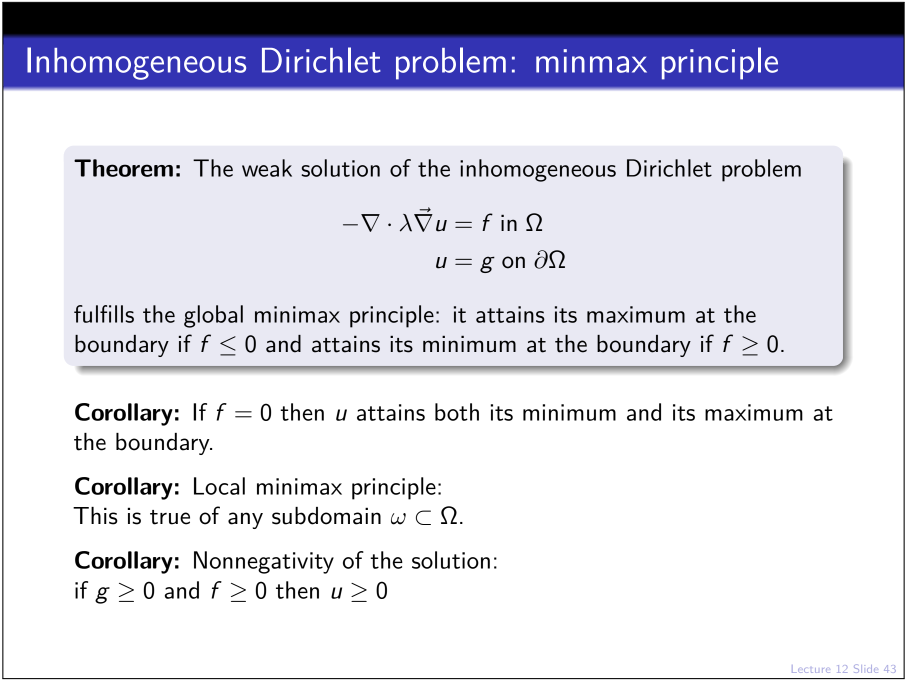
$$a^R(u,v) := \int_\Omega \lambda \vec{\nabla} u \cdot \vec{\nabla} v \, d\vec{x} + \int_{\partial\Omega} \alpha uv \; ds$$

$$f^R(v) := \int_\Omega fv \, d\vec{x} + \int_{\partial\Omega} gv \, ds$$

Find $u \in H^1(\Omega)$ 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, and by Lax-Milgram we establish the existence of a weak solution



Inhomogeneous Dirichlet problem: minmax principle

Theorem: The weak solution of the inhomogeneous Dirichlet problem

$$
\begin{aligned}
-\nabla \cdot \lambda \vec{\nabla} u &= f \text{ in } \Omega \\
u &= g \text{ on } \partial\Omega
\end{aligned}
$$

fulfills the global minimax principle: it attains its maximum at the boundary if $f \leq 0$ and attains its minimum at the boundary if $f \geq 0$.

Corollary: If $f = 0$ then u attains both its minimum and its maximum at the boundary.

Corollary: Local minimax principle:
This is true of any subdomain $\omega \subset \Omega$.

Corollary: Nonnegativity of the solution:
if $g \geq 0$ and $f \geq 0$ then $u \geq 0$

Interpretation of minimax principle

- Positive right hand side ⇒ "production" of heat, matter . . .
- No local minimum in the interior of domain if matter is produced.
- Also, positivity/nonnegativity of solutions if boundary conditions are positive/nonnegative
- Negative right hand side ⇒ "consumption" of heat, matter . . .
- No local maximum in the interior of domain if matter is consumed.
- Basic physical principle !

The Finite volume method

Finite volumes: motivation

Regard stationary second order PDE with Robin boundary conditions as a system of two first order equations in a Lipschitz domain Ω:

$$
\begin{aligned}
\nabla\cdot\vec{j} &= f && \text{continuity equation in } \Omega\\
\vec{j} &= -\delta\vec{\nabla}u && \text{flux law in } \Omega\\
\vec{j}\cdot\vec{n} &= \alpha u - g && \text{on } \Gamma
\end{aligned}
$$

- Derivation of the continuity equation was based on the consideration of species balances of an representative elementary volume (REV)
- Why not just subdivide the computational domain into a finite number of REV's ?
  - Assign a value of u to each REV
  - Approximate $\vec{\nabla}u$ by finite differece of u values in neigboring REVs
  - ... call REVs “control volumes” or “finite volumes”

Constructing control volumes I

Assume $\Omega \subset \mathbb{R}^d$ is a polygonal domain such that $\partial\Omega = \bigcup_{m \in \mathcal{G}} \Gamma_m$, where Γ_m are planar such that $\vec{n}|_{\Gamma_m} = \vec{n}_m$.

Subdivide Ω into into a finite number of **control volumes** $\bar{\Omega} = \bigcup_{k \in \mathcal{N}} \bar{\omega}_k$ such that

- ω_k are open convex domains such that $\omega_k \cap \omega_l = \emptyset$ if $\omega_k \neq \omega_l$
- $\sigma_{kl} = \bar{\omega}_k \cap \bar{\omega}_l$ are either empty, points or straight lines.
  If $|\sigma_{kl}| > 0$ we say that ω_k, ω_l are neighbours.
- $\vec{\nu}_{kl} \perp \sigma_{kl}$: normal of $\partial\omega$ at σ_{kl}
- $\mathcal{N}_k = \{l \in \mathcal{N} : |\sigma_{kl}| > 0\}$: set of neighbours of ω_k
- $\gamma_{km} = \partial\omega_k \cap \Gamma_m$: domain boundary part of $\partial\omega_k$
- $\mathcal{G}_k = \{m \in \mathcal{G} : |\gamma_{km}| > 0\}$: set of non-empty boundary parts of ω_k.

$\Rightarrow \partial\omega_k = (\cup_{l \in \mathcal{N}_k} \sigma_{kl}) \bigcup (\cup_{m \in \mathcal{G}_k} \gamma_{km})$

Constructing control volumes II

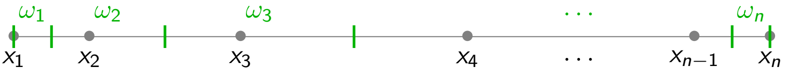
To each control volume ω_k assign a **collocation point**: $\vec{x}_k \in \bar{\omega}_k$ such that

- **Admissibility condition**:
  if $l \in \mathcal{N}_k$ then the line $\vec{x}_k\vec{x}_l$ is orthogonal to σ_{kl}
    - For a given function $u : \Omega \to \mathbb{R}$ this will allow to associate its value $u_k = u(\vec{x}_k)$ as the value of an unknown at $\vec{x}_k$.
    - For two neigboring control volumes ω_k, ω_l , this will allow to approximate $\vec{\nabla} u \cdot \vec{\nu}_{kl} \approx \frac{u_l - u_k}{h}$
- **Placement of boundary unknowns at the boundary**:
  if ω_k is situated at the boundary, i.e. for $|\partial\omega_k \cap \partial\Omega| > 0$, then $\vec{x}_k \in \partial\Omega$
    - This will allow to apply boundary conditions in a direct manner

Constructing Control Volumes in 1D

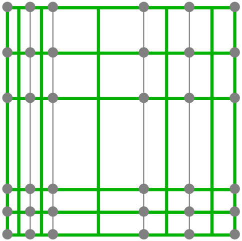
Let $\Omega = (a, b)$ be subdivided into intervals by $x_1 = a < x_2 < x_3 < \cdots < x_{n-1} < x_n = b$. Then we set

$$\omega_k = \begin{cases} \left(x_1, \frac{x_1+x_2}{2}\right), & k = 1 \\ \left(\frac{x_{k-1}+x_k}{2}, \frac{x_k+x_{k+1}}{2}\right), & 1 < k < n \\ \left(\frac{x_{n-1}+x_n}{2}, x_n\right), & k = n \end{cases}$$



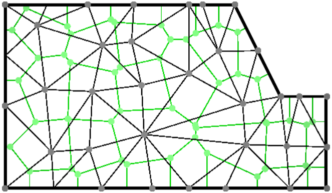
Control Volumes for a 2D tensor product mesh

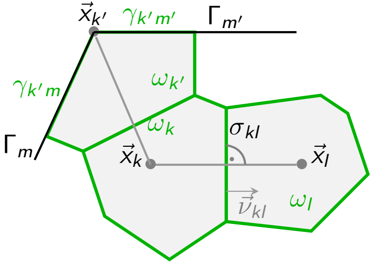
- Let $\Omega = (a, b) \times (c, d) \subset \mathbb{R}^2$.
- Assume subdivisions $x_1 = a < x_2 < x_3 < \dots < x_{n-1} < x_n = b$ and $y_1 = c < y_2 < y_3 < \dots < y_{n-1} < y_n = d$
- $\Rightarrow$ 1D control volumes ω_k^x and ω_k^y
- Set $\vec{x}_{kl} = (x_k, y_l)$ and $\omega_{kl} = \omega_k^x \times \omega_l^y$.



- Gray: original grid lines and points
- Green: boundaries of control volumes

Control volumes on boundary conforming Delaunay mesh





- Obtain a boundary conforming Delaunay triangulation with vertices $\vec{x}_k$
- Construct restricted Voronoi cells ω_k with $\vec{x}_k \in \omega_k$
  - Corners of Voronoi cells are either cell circumcenters of midpoints of boundary edges
  - Admissibility condition $\vec{x}_k\vec{x}_l \perp \sigma_{kl} = \bar{\omega}_k \cap \bar{\omega}_l$ fulfilled in a natural way
- Triangulation edges: connected neigborhood graph of Voronoi cells
- Boundary placement of collocation points of boundary control volumes

Discretization of continuity equation

- Stationary continuity equation: $\nabla \cdot \vec{j} = f$
- Integrate over control volume ω_k:

$$
\begin{aligned}
0 &= \int_{\omega_k} \nabla \cdot \vec{j}\, d\omega - \int_{\omega_k} f \; d\omega \\
&= \int_{\partial\omega_k} \vec{j} \cdot \vec{n}_\omega \; ds - \int_{\omega_k} f \; d\omega \\
&= \sum_{l \in \mathcal{N}_k} \int_{\sigma_{kl}} \vec{j} \cdot \vec{\nu}_{kl} \; ds + \sum_{m \in \mathcal{G}_k} \int_{\gamma_{km}} \vec{j} \cdot \vec{n}_m \; ds - \int_{\omega_k} f d\omega
\end{aligned}
$$

Approximation of flux between control volumes

Utilize flux law: $\vec{j} = -\delta\vec{\nabla}u$

- Admissibility condition $\Rightarrow \vec{x}_k\vec{x}_l \parallel \nu_{kl}$
- Let $u_k = u(\vec{x}_k)$, $u_l = u(\vec{x}_l)$
- $h_{kl} = |\vec{x}_k - \vec{x}_l|$: distance between neigboring collocation points
- Finite difference approximation of normal derivative:

$$\vec{\nabla}u \cdot \vec{\nu}_{kl} \approx \frac{u_l - u_k}{h_{kl}}$$

$\Rightarrow$ flux between neigboring control volumes:

$$\begin{aligned}\int_{\sigma_{kl}} \vec{j} \cdot \vec{\nu}_{kl}\ ds &\approx \frac{|\sigma_{kl}|}{h_{kl}}\delta(u_k - u_l)\\ &=: \frac{|\sigma_{kl}|}{h_{kl}}g(u_k, u_l)\end{aligned}$$

where $g(\cdot,\cdot)$ is called flux function

Approximation of Robin boundary conditions

Utilize boundary condition $\vec{j} \cdot \vec{n} = \alpha u - g$

- Assume $\alpha|_{\Gamma_m} = \alpha_m$
- Approximation of $\vec{j} \cdot \vec{n}_m$ at the boundary of ω_k:

$$\vec{j} \cdot \vec{n}_m \approx \alpha_m u_k - g$$

- Approximation of flux from ω_k through Γ_m:

$$\int_{\gamma_{km}} \vec{j} \cdot \vec{n}_m \, ds \approx |\gamma_{km}|(\alpha_m u_k - g)$$

Discrete system of equations

- Let $f_k = f(\vec{x}_k)$ (or $f_k = \frac{1}{|\omega_k|}\int_{\omega_k} f(\vec{x}_k)\ d\omega$)
- The discrete system of equations then writes for $k \in \mathcal{N}$:

$$\sum_{l\in\mathcal{N}_k} \frac{|\sigma_{kl}|}{h_{kl}}\delta(u_k - u_l) + \sum_{m\in\mathcal{G}_k} |\gamma_{km}|\alpha_m u_k = |\omega_k|f_k + \sum_{m\in\mathcal{G}_k} |\gamma_{km}|g$$

$$u_k\left(\delta\sum_{l\in\mathcal{N}_k}\frac{|\sigma_{kl}|}{h_{kl}} + \alpha_m\sum_{m\in\mathcal{G}_k}|\gamma_{km}|\right) - \delta\sum_{l\in\mathcal{N}_k}\frac{|\sigma_{kl}|}{h_{kl}}u_l = |\omega_k|f_k + \sum_{m\in\mathcal{G}_k}|\gamma_{km}|g$$

$$a_{kk}u_k + \sum_{l=1\dots|\mathcal{N}|, l\neq k} a_{kl}u_l = b_k$$

with $b_k = |\omega_k|f_k + \sum_{m\in\mathcal{G}_k}|\gamma_{km}|g$ and

$$a_{kl} = \begin{cases} \sum_{l'\in\mathcal{N}_k}\delta\frac{|\sigma_{kl'}|}{h_{kl'}} + \sum_{m\in\mathcal{G}_k}|\gamma_{km}|\alpha_m, & l = k \\ -\delta\frac{\sigma_{kl}}{h_{kl}}, & l \in \mathcal{N}_k \\ 0, & else \end{cases}$$

Discretization matrix properties

- $N = |\mathcal{N}|$ equations (one for each control volume ω_k)
- $N = |\mathcal{N}|$ unknowns (one for each collocation point $x_k \in \omega_k$)
- Matrix is sparse: nonzero entries only for neighboring control volumes
- Matrix graph is connected: nonzero entries correspond to edges in Delaunay triangulation $\Rightarrow$ irreducible
- 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 algorithm

- Due to the connection between Voronoi diagram and Delaunay triangulation, one can assemble the discrete system based on the triangulation
- Necessary information:
  - List of point coordinates $\vec{x}_K$
  - List of triangles which for each triangle describes indices of points belonging to triangle
    - This induces a mapping of local node numbers of a triangle T to the global ones: $\{1, 2, 3\} \to \{k_{T,1}, k_{T,2}, k_{T,3}\}$
  - List of (boundary) segments which for each segment describes indices of points belonging to segment
- Assembly in two loops:
  - Loop over all triangles, calculate triangle contribution to matrix entries
  - Loop over all boundary segments, calculate contribution to matrix entries

Matrix assembly – main part

- Loop over all triangles $T \in \mathcal{T}$, add up edge contributions

```
for $k, l = 1 \dots N$ do
    set $a_{kl} = 0$
end
for $T \in \mathcal{T}$ do
    for $i, j = 1 \dots 3, i \neq j$ do
        $\sigma = \sigma_{k_{T,j},k_{T,i}} \cap T$
        $s = \frac{|\sigma|}{h_{k_{T,j},k_{T,i}}}$
        $a_{k_{T,j},k_{T,j}} += \delta\sigma_h$
        $a_{k_{T,j},k_{T,i}} -= \delta\sigma_h$
        $a_{k_{T,i},k_{T,j}} -= \delta\sigma_h$
        $a_{k_{T,i},k_{T,n}} += \delta\sigma_h$
    end
end
```

Matrix assembly – boundary part

- Keep list of global node numbers per boundary element γ mapping local node element to the global node numbers: $\{1,2\} \to \{k_{\gamma,1}, k_{\gamma,2}\}$
- Keep list of boundary part numbers m_γ per boundary element
- Loop over all boundary elements $\gamma \in \mathcal{G}$ of the discretization, add up contributions

```
for $\gamma \in \mathcal{G}$ do
    for $i = 1, 2$ do
        $a_{k_{\gamma_i}, k_{\gamma_i}} += \alpha_{m_\gamma} |\gamma \cap \partial\omega_{k_{\gamma_i}}|$
    end
end
```

RHS assembly: calculate control volumes

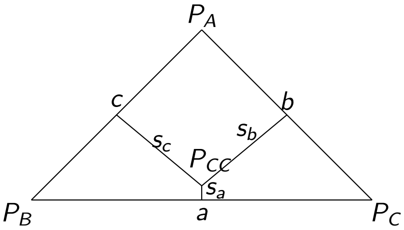
- Denote $w_k = |\omega_k|$
- Loop over triangles, add up contributions

```
for $k \dots N$ do
    set $w_k = 0$
end
for $\tau \in \mathcal{T}$ do
    for $n = \dots 3$ do
        $w_k += |\omega_{k_{\tau,j}} \cap \tau|$
    end
end
```

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

Finite volume local stiffness matrix calculation I



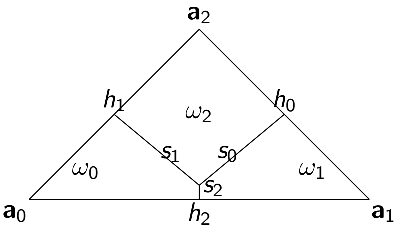
- Need to calculate s_a, s_b, s_c
- Triangle edge lengths: a, b, c
- Semiperimeter: $s = \frac{a}{2} + \frac{b}{2} + \frac{c}{2}$
- Square area (from Heron's formula):
  $16A^2 = 16s(s-a)(s-b)(s-c) =$
  $(-a+b+c)\,(a-b+c)\,(a+b-c)\,(a+b+c)$
- Square circumradius: $R^2 = \frac{a^2b^2c^2}{(-a+b+c)(a-b+c)(a+b-c)(a+b+c)} = \frac{a^2b^2c^2}{16A^2}$

Finite volume local stiffness matrix calculation II

- Square of the Voronoi surface contribution via Pythagoras:
  $s_a^2 = R^2 - \left(\frac{1}{2}a\right)^2 = -\frac{a^2\left(a^2-b^2-c^2\right)^2}{4(a-b-c)(a-b+c)(a+b-c)(a+b+c)}$
- Square of edge contribution in the finite volume method:
  $e_a^2 = \frac{s_a^2}{a^2} = -\frac{\left(a^2-b^2-c^2\right)^2}{4(a-b-c)(a-b+c)(a+b-c)(a+b+c)} = \frac{(b^2+c^2-a^2)^2}{64A^2}$
- Edge contribution. $e_a = \frac{s_a}{a} = \frac{b^2+c^2-a^2}{8A}$
- The sign chosen implies a positive value if the angle $\alpha < \frac{\pi}{2}$, and a negative value if it is obtuse. In the latter case, this corresponds to the negative length of the line between edge midpoint and circumcenter, which is exactly the value which needs to be added to the corresponding amount from the opposite triangle in order to obtain the measure of the Voronoi face.

Finite volume local stiffness matrix calculation III

$a_0 = (x_0, y_0) \dots a_d = (x_2, y_2)$: vertices of the simplex K Calculate the contribution from triangle to $\frac{\sigma_{kl}}{h_{kl}}$ in the finite volume discretization



Let $h_i = |a_{i+1} - a_{i+2}|$ (i counting modulo 2) be the lengths of the discretization edges. Let A be the area of the triangle. Then for the contribution from the triangle to the form factor one has

$$\frac{|s_i|}{h_i} = \frac{1}{8A}(h_{i+1}^2 + h_{i+2}^2 - h_i^2)$$

$$|\omega_i| = (|s_{i+1}|h_{i+1} + |s_{i+2}|h_{i+2})/4$$

Variations of the discretization ansatz

- 3D: tetrahedron based
- $\delta = \delta(x) \Rightarrow \delta(x)\nabla u \approx \delta_{kl}\frac{u_l - u_k}{h_{kl}}$
- Non-constant α_i, g
- Nonlinear dependencies …

Interpretation of results

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

The problem with Dirichlet boundary conditions

- Eliminate Dirichlet BC algebraically after building of the matrix, i.e. fix “known unknowns” at the Dirichlet boundary ⇒ highly technical
- Modifiy matrix such that equations at boundary exactly result in Dirichlet values ⇒ loss of symmetry of the matrix
- Penalty method

Dirichlet BC: Algebraic manipulation

- Assume 1D situation with BC $u_1 = g$
- From handling of control volumes without regard of boundary values:

$$AU = \begin{pmatrix} \frac{1}{h} & -\frac{1}{h} & & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\ & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \end{pmatrix}$$

- Fix u_1 and eliminate:

$$A'U = \begin{pmatrix} \frac{2}{h} & -\frac{1}{h} & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\ & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_2 \\ u_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} f_2 + \frac{1}{h}g \\ f_3 \\ \vdots \end{pmatrix}$$

- A' becomes idd and stays symmetric
- operation is quite technical

Dirichlet BC: Modify boundary equations

- From handling of control volumes without regard of boundary values:

$$AU = \begin{pmatrix} \frac{1}{h} & -\frac{1}{h} & & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\ & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \end{pmatrix}$$

- Modify equation at boundary to exactly represent Dirichlet values

$$A'U = \begin{pmatrix} \frac{1}{h} & 0 & & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\ & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} \frac{1}{h}g \\ f_2 \\ f_3 \\ \vdots \end{pmatrix}$$

- A' becomes idd
- loses symmetry ⇒ problem e.g. with CG method

Dirichlet BC: Discrete penalty trick

- From handling of control volumes without regard of boundary values:

$$AU = \begin{pmatrix} \frac{1}{h} & -\frac{1}{h} & & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\ & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \end{pmatrix}$$

- Add penalty terms

$$A'U = \begin{pmatrix} \frac{1}{\varepsilon} + \frac{1}{h} & -\frac{1}{h} & & & \\ -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\ & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} f_1 + \frac{1}{\varepsilon} g \\ f_2 \\ f_3 \\ \vdots \end{pmatrix}$$

- A' becomes idd, keeps symmetry, and the realization is technically easy.
- If ε is small enough, $u_1 = g$ will be satisfied exactly within floating point accuracy.
- Iterative methods should be initialized with Dirichlet values.
- Works for nonlinear problems, finite volume methods

Dirichlet penalty trick, general formulation

- Dirichlet boundary value problem

$$
\begin{aligned}
-\nabla \cdot \delta \nabla u &= f \quad \text{in } \Omega \\
u|_{\Gamma_m} &= g_m
\end{aligned}
$$

- Approximate Dirichlet boundary condition by

$$
\delta \nabla u \cdot \vec{n}_m + \frac{1}{\epsilon} u|_{\Gamma_m} = \frac{1}{\epsilon} g_m
$$