Lectures on DOMAIN DECOMPOSITION METHODS

by
Ulrich Langer
Department for Numerical Mathematics and Optimization
Institute of Mathematics
Johannes Kepler University Linz
(Austria)

Mexico, February 1997

Contents

Chapter 1

Introduction

- **Domain Decomposition (DD) ?**
  = the basic tool for constructing parallel PDE solvers !

- **What should you know ?**

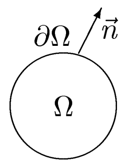
  1. PDEs
     - Sobolev spaces
     - BVP
     - Variational formulation
  2. FEM
  3. Direct and iterative methods for solving $K_h \underline{u}_h = \underline{f}_h$

$$\boxed{\begin{array}{r} -\Delta u = f \text{ in } \Omega \\ u = g \text{ on } \partial\Omega \end{array}}$$

$$\Big\downarrow$$

$$\boxed{K_h \underline{u}_h = \underline{f}_h}$$

system of algebr. eq.



However, Chapter 2 and Chapter 3 give a brief introduction to these topics.

- **<u>Literature to 1. – 3.:</u>**

[2] **Axelsson O., Barker V.A.:** Finite Element Solution of BVP. Academic Press, N.Y. 1984.

[11] **Grossmann Ch., Roos H.-G.:** Numerik partieller Differentialgleichungen. Teubner–Verlag, Stuttgart 1992.

[34] **Reusken A.:** Iterative Methods for Elliptic Boundary Value Problems. Technical University of Einhoven, Einhoven 1996.

- **<u>Literature on DD:</u>**

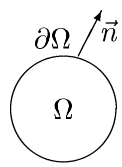
[*] **Proceedings of the DD conferences since 1986.**

[37] **Smith B., Bjorstad P., Gropp W.:** Domain Decomposition: Parallel Multilevel Methods for Elliptic PDEs. Cambridge University Press, Cambridge 1996.

■ **Contents:**

1. Introduction (1st Lecture)
2. Variational Formulation (1st Lecture)
3. Galerkin - FEM (1st Lecture)
4. History (2nd Lecture)
5. Parallelization via Non–overlapping DD (3rd Lecture)
6. Preconditioning via DD (4th Lecture)
7. DD–Preconditioners in Applications (5th Lecture)

■ **Some Notations:** (usual mathematical notation)



$\Omega \subset I\!R^m (m = 1, \underline{\underline{2}}, 3)$ – bounded domain, $\vec{n} = \vec{n}(x) = (n_1, \dots, n_m)^T$

$\Gamma = \partial\,\Omega \in C^{0,1}$ – Lipschitz–continuous boundary,

$\nabla u = grad(u) := \left(\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_m}\right)^T$ – gradient,

$\Delta u = \sum\limits_{i=1}^{m} \frac{\partial^2 u}{\partial x_i^2}$ – Laplace operator, $div(v) = \sum\limits_{i=1}^{m} \frac{\partial v_i}{\partial x_i}$, $v = (v_1, \dots, v_m)^T$;

$C^k(\Omega),\ C^k(\bar{\Omega}),\ \overset{\bullet}{C}{}^k(\Omega), \dots;$

$L_2(\Omega) := \{v : \Omega \to I\!R^1 : ||v||_0^2 := \int\limits_\Omega v^2\,dx < \infty\} = H^0(\Omega)$,

$H^k(\Omega) = W_2^k(\Omega)$ – Sobolev Hilbert–spaces:

$(\cdot,\cdot)_k = (\cdot,\cdot)_{k,\Omega}$ – inner product,

$||\cdot||_k = (\cdot,\cdot)_k^{0.5}$ – norm, $|\cdot|_k$ – semi-norm,

$k = 1 : (u, v)_1 := \int\limits_\Omega (uv + \nabla^T u \nabla v)\,dx$,

$$||u||_1^2 = ||u||_0^2 + |u|_1^2,$$

$$|u|_1^2 := \int\limits_\Omega |\nabla u|^2\,dx.$$

$\overset{\circ}{H}{}^1(\Omega) = \{v \in H^1(\Omega) :\ u = 0 \text{ on } \Gamma\} \subset H^1(\Omega)$.

Chapter 2

Variational Formulation of Elliptic Boundary Value Problems

- **Literature:** [7], [2], [4], [11], [34]

2.1 An Abstract Theory

- **Introduce**

  V – real Hilbert space (H–space) equipped with the scalar product $(\cdot,\cdot)$ and with the corresponding norm $\|\cdot\| := (\cdot,\cdot)^{0.5}$;

  $V_0 \subset V$ – some non–trivial, closed subspace of the H–space V;

  $g \in V$ – some given element defining the hyperplane (linear manifold) $V_g := g + V_0 \equiv \{v \in V : \exists\, w \in V_0,\ v = g + w\}$;

  V_0^* – dual space of all continuous (= bounded), linear functional $F : V_0 \to \mathbb{R}^1$:

  $\|F\|_* \equiv \|F\|_{V_0^*} := \sup\limits_{\substack{v \in V_0 \\ v \neq 0}} \frac{<F,v>}{\|v\|}$ – norm,

  $< \cdot,\cdot >: V_0^* \times V_0 \mapsto \mathbb{R}^1$ – duality product,

  $< F, v >:= F(v)$.

- Consider now the **abstract variational problem**

$$
\begin{array}{lll}
(1) & \text{Find } u \in V_g: & a(u,v) = < F, v > \quad \forall\, v \in V_0 \\
 & \text{Homogenization} & \Big\downarrow \quad u = w + g \\
(1)_0 & \text{Find } w \in V_0: & a(w,u) = < \widehat{F}, v > := < F, v > - a(g,v) \quad \forall\, v \in V_0,
\end{array}
$$

$$(1)_0 \qquad \boxed{u \in V_0 : \quad a(u,v) = < F, v > \quad \forall\, v \in V_0}\ ,$$

where $F \in V_0^*$ is a given linear continuous linear functional and $a(\cdot,\cdot) : V \times V \to \mathbb{R}^1$ is a given continuous (= bounded) bilinear form.

■ **Theorem 2.1:** (Lax–Milgram–Lemma)

> Assume that
>
> 0. $V_0 \subset V$ – subspace of the H–space V, $\|\cdot\|$, $(\cdot,\cdot)$,
> 1. $F \in V_0^*$,
> 2. $a(\cdot,\cdot) : V_0 \times V_0 \to I\!R^1$ – bilinear form:
>
> 2.a) V_0–elliptic, i.e. $\exists\, \mu_1 = \text{const.} > 0$: $\mu_1 \|v\|^2 \leq a(v,v) \quad \forall\, v \in V_0$,
>
> 2.b) V_0–bounded, i.e. $\exists\, \mu_2 = \text{const.} > 0$: $|a(u,v)| \leq \mu_2 \|u\| \, \|v\| \quad \forall\, u,v \in V_0$.
>
> Then there exists a unique solution ($\exists\,!$) $u \in V_0$ of the variational problem $(1)_0$.

Proof follows from Banach's fixed point theorem. Indeed,

$$
\begin{aligned}
(1)_0 \quad &\Longleftrightarrow \quad Au = F \text{ in } V_0, \text{ with } < Au, v > = a(u,v). \\
&\Longleftrightarrow \quad u = B_\tau u := u - \tau(JAu - JF) \text{ in } V_0, \text{ where } J \in L(V_0^*, V_0) \\
&\qquad\quad \text{denotes the Riesz isomorphism defined by the relation} \\
&\qquad\quad (JF, v) = < F, v > \quad \forall\, F \in V_0^* \quad \forall\, v \in V_0.
\end{aligned}
$$

It is easy to verify that, for fixed $\tau \in (0, 2\mu_1/\mu_2^2)$, the operator B_τ is a contraction. Therefore, Banach's fixed point theorem gives existence and uniqueness. Moreover, the fixed point iteration process

$$
(2) \qquad u^{n+1} = u^n - \tau(JAu^n - JF) \underset{n\to\infty}{\longrightarrow} u
$$

converges to the unique solution u of $(1)_0$ for an arbitrary initial guess $u^0 \in V_0$ such that

$$
\|u - u^n\| \leq q_{\text{opt}}^n \, \|u - u^0\|,
$$

with $q_{\text{opt}} = 1 - (\mu_2/\mu_1)^2$ for $\tau = \tau_{\text{opt}} = \mu_1/\mu_2^2$. **q.e.d.** ■

■ **Theorem 2.2:** (energy minimization problem)

> Assume that
>
> 1. $F \in V_0^*$,
> 2. the continuous bilinear form is
>   - positive: $a(u,v) > 0 \quad \forall\, v \in V_0 : v \neq \mathbf{O}$,
>   - symmetric: $a(u,v) = a(v,u) \quad \forall\, u,v \in V_0$.
>
> Then the variational problem (VP) = $(1)_0$ and the energy minimization problem (MP) are equivalent, i.e. (VP) $\Longleftrightarrow$ (MP):

(VP) Find $u \in V_0$: $a(u,v) = < F, v > \quad \forall\, v \in V_0$ $\Longleftrightarrow$ Find $u \in V_0$: $J(u) = \inf\limits_{v \in V_0} J(v)$ (MP)

with the Ritz energy functional

$$J(v) := \tfrac{1}{2}\, a(v,v) - < F, v >.$$

Proof:

- For all $u, v \in V_0$ and arbitrary $t \in \mathbb{R}^1$, we have

$$\begin{aligned} J(u + \underbrace{tv}_{=\delta u}) &= \tfrac{1}{2}\, a(u+tv, u+tv) - < F, u+tv > = \\ &= \tfrac{1}{2}\, a(u,u) + ta(u,v) + \tfrac{t^2}{2}\, a(v,v) - < F,u > - t < F,v > = \\ &= J(u) + t[a(u,v) - < F,v >] + \tfrac{t^2}{2}\, a(v,v). \end{aligned}$$

- (VP) $\Longrightarrow$ (MP):

  Let $u \in V_0$ be a solution of (VP), i.e. $a(u,v) = < F,v > \quad \forall\, v \in V_0$. For $t = 1$, we immediately obtain the relation

$$J(u+v) = J(u) + \underbrace{[a(u,v) - < F,v >]}_{=0} + \tfrac{1}{2}\, a(v,v) > J(u) \quad \forall\, v \in V_0 : v \neq \mathbf{O}$$

  form the above representation. Thus, u is obviously the unique minimizer of the energy minimization problem.

- (MP) $\Longrightarrow$ (VP):

  Let $u \in V_0$ be a minimizer of the energy functional $J(\cdot)$, i.e. $J(u) \leq J(v) \quad \forall\, v \in V_0$. Then the necessary minimization condition

$$\begin{array}{l} \frac{d}{dt}\, J(u+tv)\Big|_{t=0} = 0 \quad \forall\ (\text{fixed})\ v \in V_0 : v \neq \mathbf{O} \\ \qquad \| \\ [a(u,v) - < F,v >] + ta(v,v)|_{t=0} = a(u,v) - < F,v >, \end{array}$$

  i.e. $u \in V_0$ solves the variational problem (VP). **q.e.d.** ■

2.2 Elliptic Boundary Value Problems

2.2.1 Scalar Elliptic Second-Order PDEs

■ Let us consider the **classical formulation** of a mixed boundary value problem for a second-order linear elliptic PDE in the divergent form:

(3)

$$\boxed{\begin{array}{l}
\text{Find } u \in X := C^2(\Omega) \cap C^1(\Omega \cup \Gamma_2 \cup \Gamma_3) \cap C(\Omega \cup \Gamma_1): \\
-\sum\limits_{i,j=1}^{m} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u}{\partial x_j}\right) + \sum\limits_{i=1}^{m} a_i(x)\frac{\partial u}{\partial x_i} + a(x)u(x) = f(x),\ x \in \Omega, \\
+\text{ BC: } \bullet\ \ u(x) = g_1(x),\ x \in \Gamma_1, \\
\qquad\quad \bullet\ \ \frac{\partial u}{\partial N} := \sum\limits_{i,j=1}^{m} a_{ij}(x)\frac{\partial u(x)}{\partial x_j} n_i(x) = g_2(x),\ x \in \Gamma_2, \\
\qquad\quad \bullet\ \ \frac{\partial u}{\partial N} + \alpha(x)u(x) = \underbrace{g_3(x)}_{\alpha(x)u_A(x)},\ x \in \Gamma_3. \\
\qquad\qquad\qquad -\frac{\partial u}{\partial N} = \alpha(x)(u(x) - u_A(x))
\end{array}}$$

Remark 2.3.:

1. $u \in X$: (3) is called classical solution of the BVP (3) !

2. Note that the data $\{a_{ij}, a_i, a, \alpha, f, g, \Omega\}$ $\boxed{\text{Ex 2.1}}$ should satisfy classical smoothness assumptions, such as $a_{ij} \in C^1(\Omega) \cap C(\Omega \cup \Gamma_2 \cup \Gamma_3), \ldots$ etc. !

3. Uniform ellipticity in Ω:
   - $a_{ij}(x) = a_{ji}(x) \quad \forall x \in \bar{\Omega} \quad \forall i,j = \overline{1,m}$,
   - $\Lambda(x,\xi) := \sum\limits_{i,j=1}^{m} a_{ij}(x)\xi_i\xi_j \geq \bar{\mu}_1|\xi|^2 \quad \forall \xi \in I\!R^m \quad \forall x \in \bar{\Omega}$.

4. Example: Stationary convection-diffusion (heat-conduction – heat-transport) problem:

Find the temperature field $T \equiv T(x) \in X$:			
$-\operatorname{div}(\lambda\nabla T) + c\rho\vec{v}^T\nabla T = f$ in Ω,			
+ BC: e.g. prescribed temperature g_1 on $\Gamma_1 := \Gamma \equiv \partial\Omega$, i.e. $T(x) = g_1(x),\ x \in \Gamma$,			
where	$\lambda \equiv \lambda(x)$	–	heat conduction coefficient,
	$c \equiv c(x)$	–	heat capacity coefficient,
	$\rho \equiv \rho(x)$	–	density coefficient,
	$\vec{v}^T \equiv \vec{v}^T(x)$	$\equiv$	$(v_1(x), \ldots, v_m(x))$ – velocity field,
	$f \equiv f(x)$	–	heat intensity function.

■ Variational formulation = weak formulation = generalized formulation:

- Formal procedure for the derivation of the variational formulation (1):

1. Choose the space of test functions:

$$V_0 = \{v \in V = W_2^1(\Omega) : v = 0 \text{ on } \Gamma_1\}.$$

Basic space for scalar 2^{nd}-order PDEs

2. Multiply the PDE (3) by testfunctions $v \in V_0$ and integrate over Ω:

$$\int_\Omega \Big(-\sum_{i,j=1}^m \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + \sum_{i=1}^m a_i \frac{\partial u}{\partial x_i} + au\Big) v\,dx = \int_\Omega f v\,dx \quad \forall v \in V_0.$$

3. Use partial integration in the principal term (= term $(a_{,j}u_{,j})_{,i}$)

$$\boxed{\int_\Omega \frac{\partial w}{\partial x_i} v\,dx = -\int_\Gamma w\frac{\partial u}{\partial x_i}\,dx + \int_{\partial\Omega} wn \cdot n_i\,ds}$$

$$\int_\Omega \Big(\sum_{i,j=1}^m a_{ij}\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i} + \sum_{i=1}^m a_i \frac{\partial u}{\partial x_i} v + auv\Big)\,dx - \int_{\partial\Omega} \underbrace{\sum_{i,j=1}^m a_{ij}\frac{\partial u}{\partial x_j} n_i}_{=:\frac{\partial u}{\partial N} = \text{Co-normal derivative}} \cdot v\,ds =$$

$$= \int_\Omega f v\,dx \quad \forall v \in V_0.$$

4. Incorporate the natural BC on Γ_2 and Γ_3:

$$\int_\Gamma \frac{\partial u}{\partial N} v\,ds = \underbrace{\int_{\Gamma_1} \frac{\partial u}{\partial N} v\,ds}_{=0} + \int_{\Gamma_2} g_2 v\,ds + \int_{\Gamma_3} (g_3 - \alpha u) v\,ds$$

5. Define the linear manifold, where the solution u is searched for:

$$V_g = \{v \in V = W_2^1(\Omega) : v = g_1 \text{ on } \Gamma_1\}.$$

- Result: Variational formulation (VF)

(4)

$$\boxed{\begin{array}{lll} \text{Find } u \in V_g : a(u,v) = < F, v > \quad \forall v \in V_0, & & \\ \text{with} & & \\ a(u,v) & := & \int\limits_{\Omega} \left(\sum\limits_{i,j=1}^{m} a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + \sum\limits_{i=1}^{m} a_i \frac{\partial u}{\partial x_i} v + auv \right) dx + \int\limits_{\Gamma_3} \alpha uv \, ds, \\ < F, v > & := & \int\limits_{\Omega} fv \, dx + \int\limits_{\Gamma_2} g_2 v \, ds + \int\limits_{\Gamma_3} g_3 v \, ds, \\ V_g := \{ v \in V = W_2^1(\Omega) : v = g_1 \text{ on } \Gamma_1 \}, & & \\ V_0 := \{ v \in V : v = 0 \text{ on } \Gamma_1 \}. & & \end{array}}$$

- **Remark 2.4:**
  1. The solution $u \in V_g$ of (4) is called weak or generalized solution.
  2. For the variational formulation (4), the assumptions imposed on the data can be weakened (! the integrals involved in (4) should exist), e.g.:

(5)

$$\left\{ \begin{array}{ll} 1) & a_{ij}, a_i, a \in L_\infty(\Omega), \ \alpha \in L_\infty(\Gamma_3), \\ 2) & f \in L_2(\Omega), \ g_i \in L_2(\Gamma_i), \ i = 2,3, \\ 3) & g_1 \in H^{\frac{1}{2}}(\Gamma_1), \ \text{d.h. } \exists \, \tilde{g}_1 \in H^1(\Omega) : \tilde{g}_1|_{\Gamma_1} = g_1, \\ 4) & \Omega \subset \mathbb{R}^m \dagger : \Gamma = \partial\Omega \in C^{0,1} \ \text{(Lipschitz-continuous boundary)}, \\ 5) & \text{uniform ellipticity:} \\ & \left. \begin{array}{l} \sum\limits_{i,j=1}^{m} a_{ij}(x) \xi_i \xi_j \geq \bar{\mu}_1 |\xi|^2 \quad \forall \xi \in \mathbb{R}^m \\ a_{ij}(x) = a_{ji}(x) \quad \forall i,j = \overline{1,m} \end{array} \right\} \forall \text{ a.e. } x \in \Omega. \end{array} \right.$$

  3. Relation between classical and generalized solutions:

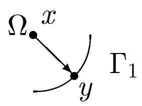
$$u \in X \cap V_g \cap W_2^2(\Omega)$$

$$(3) \quad \boxed{\begin{array}{l} u \in X : (3) \\ \text{classical solution} \end{array}} \quad \begin{array}{c} 1 \; - \; 5 \\ \Longrightarrow \\ \text{(exercise)} \\ \Longleftarrow \\ \uparrow \end{array} \quad \boxed{\begin{array}{l} u \in V_g : (4) \\ \text{generalized solution} \end{array}} \quad (4)$$

Ass.:

- $u \in V_g \cap X \cap W_2^2(\Omega)$
- classical smoothness assumptions imposed on the data in $\bar{\Omega}$

- Existence of the integrals and the correctness of the partial integration:
  ! Attention: $u \in X = C^2(\Omega) \cap C^1(\Omega \cup \Gamma_2 \cup \Gamma_3) \cap C(\Omega \cup \Gamma_1) \not\Rightarrow u \in W_2^1(\Omega)$! in general
  e.g.:
  $\frac{\partial u}{\partial x_i} = 0(|\underset{\Omega}{\underset{\cap}{x}} - \underset{\Gamma_1}{\underset{\cap}{y}}|^{-p})$ for $\overset{\Omega}{\overset{\cup}{x}} \to \overset{\Gamma_1}{\overset{\cup}{y}}$ and $p > 0$ sufficiently large !



Exercises:

Ex 2.1 Formulate the classical assumptions for the data of (3) !
Give sufficient conditions such that a generalized solution $u \in V_g \cap X \cap W_2^2(\Omega)$ of (4) is also a solution of (3) in the classical sense !
To train this, consider, at first, the Dirichlet problem for the Poisson equation:

$$(3) \quad \begin{cases} \text{Find } u \in X = C^2(\Omega) \ \cap \ C(\bar{\Omega}) : \\ -\Delta u(x) = f(x), \ x \in \Omega \subset \mathbb{R}^m \,\dagger, \\ u(x) = g(x), \ x \in \Gamma = \partial\,\Omega. \end{cases}$$

? $\Downarrow\Uparrow$?

$$(4) \quad \begin{cases} \text{Find } u \in V_g = \{v \in V = H^1(\Omega) : v = g \text{ on } \Gamma\} : \\ \underbrace{\int\limits_{\Omega} \nabla^T u \nabla v \, dx}_{=a(u,v)} = \underbrace{\int\limits_{\Omega} f v \, dx}_{=<F,v>} \quad \forall v \in V_0 = \overset{\circ}{H}{}^1(\Omega). \end{cases}$$

Ex 2.2 Show that assumptions of the Lax-Milgram-Theorem are satisfied in the following cases a) – c), and find the constants μ_1 and μ_2 !

a) In addition to assumptions (5) we assume that
$a_i = 0, \quad a(x) \geq 0 \quad \forall$ a.e. $x \in \Omega, \quad \alpha(x) \geq 0 \quad \forall$ a.e. $x \in \Gamma_3$;
$\text{meas}_{m-1}(\Gamma_1) > 0$.

b) In addition to assumptions (5) we assume that
$a_i = 0, \ a = 0, \quad \alpha(x) \geq \underline{\alpha} = \text{const.} > 0 \quad \forall$ a.e. $x \in \Gamma_3, \quad \Gamma_1 = \emptyset$.

c) In addition to assumptions (5) we assume that
$a_i = 0, \quad a(x) \geq \underline{a}$ const. $> 0 \quad \forall$ a.e. $x \in \Omega; \quad \Gamma = \Gamma_2$, d.h. $\Gamma_1 = \Gamma_3 = \emptyset$.

Ex 2.3 In addition to assumptions (5) we assume that $a(x) \geq \underline{a} = \text{const.} > 0 \quad \forall$ a.e. $x \in \Omega$, $\Gamma_1 = \Gamma_3 = \emptyset$, and $a_i \not\equiv 0$.
Formulate conditions which have to be imposed on the coefficients a_i such that assumptions of the Lax-Milgram-Theorem are fulfilled !

Hints: Use the so-called ϵ-inequality

$$|ab| \leq \frac{1}{2\epsilon}\, a^2 + \epsilon\, b^2, \quad \forall a, b \in I\!R^1 \quad \forall \epsilon > 0$$

for estimating the convection term $\sum\limits_{i=1}^{m} \int\limits_{\Omega} a_i \frac{\partial u}{\partial x_i} v\, dx$!

Ex 2.4 Find the correct variational formulation of the pure Neumann problem for the Poisson equation

$$(*) \qquad \begin{cases} -\Delta u(x) = f(x), \; x \in \Omega \\ \frac{\partial u}{\partial \vec{n}} = 0, \; x \in \Gamma = \partial\,\Omega \end{cases}$$

and clarify the problem of existence and uniqness of the generalized solution !

Hints: Obviously, $u(x) + c$ is a solution of $(*)$ for an arbitrary fixed constant $c \in I\!R^1$, provided that u solves $(*)$. The following two approaches are possible to clarify the solvability properties of Neumann problem:

1. Consider the variational formulation in $V = H^1(\Omega)$ and apply the Fredholm theory (see, e.g. [26], Chapter 2) !
2. Consider the variational formulation in the factor space $V = H^1(\Omega)|_{\text{ker}}$ with $\text{ker} = \{c : c \in I\!R^1\} = I\!R^1$ and apply the Lax-Milgram–theorem !

Ex 2.5 Consider a plane magnetic field problem (e.g. electric motor) modelled by the following Dirichlet BVP for determining the z-component $u(x_1, x_2) := A_z(x, y)$ of the magnetic vector potential:

$$\boxed{\begin{array}{l} -\operatorname{div}\left(\frac{1}{\mu(x)}\nabla u(x)\right) = S_z(x) - \frac{\partial}{\partial x_1}\left(\frac{1}{\mu(x)} B_{x_2}(x)\right) + \frac{\partial}{\partial x_2}\left(\frac{1}{\mu(x)} B_{x_1}(x)\right), \\ \qquad\qquad x \in \Omega \subset I\!R^2 \; \dagger, \\ + \text{ BC: } u(x) = 0, \; x \in \Gamma = \partial\,\Omega. \end{array}}$$

with given functions μ (permeability), S_z (current density) and given remanence induction $\vec{B} = (B_{x_1}, B_{x_2})^T$.

Give the variational formulation (1)

Find $u \in V_g : a(u,v) = < F, v > \quad \forall v \in V_0$

for this plane magnetic field problem and show that there exists a unique generalized solution $u \in V_g$ of the variational problem (1) provided that the assumptions

1. $\mu \in L_\infty : 0 < \underline{\mu} \le \mu(x) \le \bar{\mu} \quad$ a.e. $x \in \Omega$,
   with positive constants $\underline{\mu}$ and $\bar{\mu}$,
2. $S_z \in L_2(\Omega)$,
3. $B_{x_1}, B_{x_2} \in L_2(\Omega)$,
4. $\Omega \subset \mathbb{R}^2$ †, $\Gamma = \partial\,\Omega \in C^{0,1}$

are fulfilled !

2.2.2 The Linear Elasticity Problem

■ **Classical formulation of the linear 3D elasticity problem [4], [7], [27]:**

(6)

Find the displacement field $u(x) = (u_1(x), u_2(x), u_3(x))^T \in X$:

equilibrium of forces:

equilibrium of momentum

$$-\sum_{j=1}^{3} \frac{\partial}{\partial x_j} \underset{}{\overset{\sigma_{ji}}{\overset{\|}{\sigma_{ij}}}}(u(x)) = f_i(x), \quad \forall x = (x_1, x_2, x_3) \in \Omega \subset \mathbb{R}^3 \text{ †}, \quad \forall i = \overline{1,3},$$

$+$ material equations = Hook's law:

$$\sigma_{ij} = \sum_{k,l=1}^{3} D_{ijkl}\epsilon_{kl} = \lambda \left(\sum_{k=1}^{3} \epsilon_{kk} \right) \delta_{ij} + 2\mu\epsilon_{ij}, \quad \forall i,j = \overline{1,3},$$

i.g. 21 independent constants from 81

isotropic material:

$$\boxed{D_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})}$$

$\lambda, \mu = $ const. > 0 (homogeneous)

$+$ Geometrical strain–displacement relations:

$$\epsilon_{ij} = \epsilon_{ij}(u) := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_j} \right) = \epsilon_{ji} \quad \forall i,j = \overline{1,3},$$

+ boundary conditions:

$$u(x) = 0 \quad (\text{bzw.} = \bar{u}(x)) \quad \forall x \in \Gamma_1,$$

$$\sum_{j=1}^{3} \sigma_{ij}(u(x))n_j(x) = g_i(x) \quad \forall x \in \Gamma_2,$$

where

$$X = \{v = (v_1, v_2, v_3)^T : v_i \in C^2(\Omega) \cap C^1(\Omega \cup \Gamma_2) \cap C(\Omega \cup \Gamma_1),\ i = \overline{1,3}\}$$

and classical assumptions on the data $\{D_{ijkl}, f, g, \bar{u}, \Gamma, \Gamma_1, \Gamma_2\}$ are supposed.

Ex 2.6 Show that (6) is equivalent to Lamé's PDE system:

$$-\mu\Delta u(x) - (\lambda + \mu)\nabla \operatorname{div} u(x) = f(x),\ x \in \Omega$$

+ boundary conditions,

with $f = (f_1, f_2, f_3)^T$ and the differential operators

$$\Delta = \begin{pmatrix} \Delta & 0 & 0 \\ 0 & \Delta & 0 \\ 0 & 0 & \Delta \end{pmatrix} = \text{vector Laplace},$$

∇ = gradient, div = divergence.

Hints: Insert the geometrical strain–displacement relations into Hook's law, and afterwards the resulting relations for the stresses into the equilibrium equations !

- **Derivation of the variational formulation in analogy to Subsection 2.2.1:**

1 $V_0 = \{v = (v_1, v_2, v_3)^T \in V = [W_2^1(\Omega)]^3 : v = 0 \text{ on } \Gamma_1\}$.

2 $$\int_\Omega \left(-\sum_{i=1}^{3}\sum_{j=1}^{3} \frac{\partial}{\partial x_j}\sigma_{ij} v_i \right) dx = \int_\Omega \sum_{i=1}^{3} f_i v_i \, dx \quad \forall v \in V_0.$$

3 $$\int_\Omega \underbrace{\sum_{i=1}^{3}\sum_{j=1}^{3} \sigma_{ij} v_{i,j}}_{=\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}(u)\epsilon_{ij}(v)} \, dx - \int_\Gamma \sum_{i=1}^{3} \underbrace{\sum_{j=1}^{3} \sigma_{ij} n_j}_{=:\sigma_{\vec{n}\,i} = i^{\text{th}} \text{ component of the normal stress vector}} \cdot v_i \, ds = \int_\Omega f^T v \, dx \quad \forall v \in V_0.$$

④ $\int\limits_{\Gamma} \sum\limits_{i=1}^{3} \left[\sum\limits_{j=1}^{3} \sigma_{ij} n_j \right] v_i \, ds = \int\limits_{\Gamma_1} \underbrace{\sum \sum \sigma_{ij} n_j v_i}_{=0} \, ds + \int\limits_{\Gamma_2} \sum\limits_{i=1}^{3} g_i v_i \, ds.$

⑤ $V_g = \{v \in V : v = \bar{u} \text{ on } \Gamma_1\} \underset{\substack{\uparrow \\ \text{without loss of generality } \bar{u}=0}}{=} V_0.$

■ **Result:** Variational Formulation (VF):

(7)

Find $u \in V_g \equiv V_0 := \{v \in V = [W_2^1(\Omega)]^3 : v = 0 \text{ on } \Gamma_1\}$: $$a(u,v) = < F, v > \quad \forall v \in V_0,$$ where $$\begin{aligned} a(u,v) &= \int\limits_{\Omega} \sum_{i,j=1}^{3} \sigma_{ij}(u(x)) \epsilon_{ij}(v(x)) \, dx \equiv \int\limits_{\Omega} \sigma^T(u) \epsilon(v) \, dx \\ &= \int\limits_{\Omega} \sum_{ij}^{3} \sum_{k,l=1}^{3} D_{ijkl} \epsilon_{kl}(u) \epsilon_{ij}(v) \, dx \equiv \int\limits_{\Omega} \epsilon^T(u) \underset{\substack{\uparrow \\ \text{tensor of elastic constants}}}{D} \epsilon(v) \, dx = \\ &\overset{\substack{\text{isotropic} \\ \downarrow}}{=} \int\limits_{\Omega} \{\lambda \underbrace{\sum_{k=1}^{3} \epsilon_{kk}(u)}_{\text{div }(u)} \underbrace{\sum_{i=1}^{3} \epsilon_{ii}(v)}_{\text{div }(v)} + 2\mu \sum_{i,j=1}^{3} \epsilon_{ij}(u) \epsilon_{ij}(v)\} \, dx, \\ < F, v > &= \int\limits_{\Gamma} \sum_{i=1}^{3} f_i v_i \, dx + \int\limits_{\Gamma_2} \sum_{i=1}^{3} g_i v_i \, ds \equiv \int\limits_{\Omega} f^T v \, dx + \int\limits_{\Gamma_2} g^T v \, ds, \end{aligned}$$
$f = (f_1, f_2, f_3)^T \in [L_2(\Omega)]^3$, $g = (g_1, g_2, q_3)^T \in [L_2(\Gamma_2)]^3$ given.

■ **Minimum Problem:** (MP)

$(7)_{\mathrm{MP}}$

Find $u \in V_0 : J(u) = \inf\limits_{v \in V_0} J(v)$, with $$J(v) = \underbrace{\frac{1}{2} \int\limits_{\Omega} \sum_{i,j=1}^{3} \sigma_{ij}(v) \epsilon_{ij}(v) \, dx}_{\substack{= \text{deformation energy} \\ \text{(inner energy)}}} - \underbrace{\left(\int\limits_{\Omega} f^T v \, dx + \int\limits_{\Omega} g^T v \, ds \right)}_{\substack{= \text{potential energy} \\ \text{of exterior forces} \\ \text{(exterior energy)}}}.$$

■ **Exercises:**

Ex 2.7 Show that, for the 1. BVP ($\Gamma_1 = \Gamma$), the assumptions

a) $a(\cdot,\cdot)$ is symmetric, so that $a(u,v) = a(v,u)$ $\forall u, v \in V_0 = [\overset{\circ}{W}{}_2^1(\Omega)]^3$,

b) $a(\cdot,\cdot)$ is positive, so that $a(v,v) > 0 \quad \forall v \in V_0 : v \neq 0$.

of Theorem 2.2 are fulfilled $\curvearrowright$ $(7)_{\text{VF}} \Leftrightarrow (7)_{\text{MP}}$!

Ex 2.8 Show that, for the 1. BVP ($\Gamma_1 = \Gamma$) and in the case isotropic and homogeneous materials, the assumptions of the Lax-Milgram–Theorem 2.1 are fulfilled, i.e.

1) $F \in V_0^*$,

2a) $\exists\, \mu_1 = \text{const.} > 0 : a(v,v) \geq \mu_1 \|v\|_1^2 \quad \forall v \in V_0$,

2b) $\exists\, \mu_2 = \text{const.} > 0 : |a(u,v)| \leq \mu_2 \|u\|_1 \|v\|_1 \quad \forall u, v \in V_0$,

where

$$\|v\|_1^2 := \sum_{i=1}^{3} \|v_i\|_1^2 = \sum_{i=1}^{3} \int_\Omega \left(v_i^2 + \sum_{j=1}^{3} \left(\frac{\partial v_i}{\partial x_j} \right)^2 \right) dx.$$

Find the constants μ_1 and μ_2 !

<u>Hints:</u> as to the proof of the V_0–ellipticity:

- $$\begin{aligned} a(v,v) &= \int_\Omega \{\lambda(\text{div}\,(v))^2 + 2\mu \sum_{i,j=1}^{3} (\epsilon_{ij}(v))^2\}\, dx \geq \\ &\geq 2\mu \int_\Omega \sum_{i,j=2}^{3} (\epsilon_{ij}(v))^2\, dx \qquad \text{o.k.} \end{aligned}$$

- Korn's inequality (for the 1. BVP, i.e. in $V_0 = \left[\overset{\circ}{H}{}^1(\Omega)\right]^3$)

  $$\sum_{i,j=1}^{3} \int_\Omega (\epsilon_{ij}(v))^2\, dx \geq c_K \sum_{i,j=1}^{3} \int \left(\frac{\partial v_i}{\partial x_j} \right)^2 dx = c_K |v|_1^2 \quad \forall v \in V_0$$

  $c_K = ?$

- Friedrichs' inequality (see [4], [7] and [26] Chapter 3) !

■ **Remark 2.5:**

1. See literature for the discussion of the existence and uniqueness of the solution of the mixed BVP (meas $\Gamma_1 > 0$, meas $\Gamma_2 > 0$) and the 2. BVP ($\Gamma_2 = \Gamma$), e.g. [7] pp. 23 – 28 and lecture "Computational Solid Mechanics" [28].
2. The basis for this investigation is again the Lax-Milgram–Theorem 2.1. For the proof of the V_0–ellipticity,
   - Korn's inequality

     $$\|v\|_{1,\Omega} \leq C \left(\sum_{i,j=1}^{3} \|\epsilon_{ij}(v)\|_{0,\Omega}^2 + \sum_{i=1}^{3} \|v_i\|_{0,\Omega}^2 \right)^{0.5}$$

     $\forall v \in V = [W_2^1(\Omega)]^3$,

- and generalized Friedrichs' inequality (mixed BVP),
- or, for the 2. BVP,

$$\begin{aligned}\ker A_{\text{Lamé}} &= \{a \times x + b : a, b \in I\!R^3\} \\ &= \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -x^2 \\ x_1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -x_3 \\ x_2 \end{pmatrix}, \begin{pmatrix} -x_3 \\ 0 \\ x_1 \end{pmatrix} \right\} \\ &= \text{subspace of rigid body displacements}\end{aligned}$$

are needed.

2.2.3 The First Biharmonic BVP

- **Classical Formulation:**

Find $u \in X = C^4(\Omega) \cap C^1(\bar{\Omega})$:

$\Delta^2 u(x) \equiv \frac{\partial^4 u}{\partial x_1^4} + 2\frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 u}{\partial x_2^4} = f(x), \quad \forall x \in \Omega \subset I\!R^2$ †

$+$ BC: $u(x) = 0 \quad \forall x \in \Gamma,$

$\frac{\partial u(x)}{\partial n} = 0 \quad \forall x \in \Gamma,$

with $f \in C(\Omega)$ and $\Omega \subset I\!R^2$ †, $\Gamma = \partial\,\Omega$ – sufficiently smooth.

Modelled e.g. bending $u(\cdot)$ of a clamped, homogeneous and isotropic plate under vertical load ! BVP (8) can be derived from (6) by means of Kirchhoff's hypothesises.
See lecture "Numerical Solid Mechanics" [28] and literature, e.g. [4] (Braess D.: Finite Elements, p. 264).

- **Derivation of the Variational formulation:** 1 - 5

1 $V_0 = \{v \in V = W_2^2(\Omega) : u = \frac{\partial u}{\partial n} = 0 \text{ on } \Gamma\} = \overset{\circ}{W}{}_2^2(\Omega) = \overset{\circ}{H}{}^2(\Omega).$

2 $\int\limits_\Omega \Delta^2 u \cdot v\, dx = \int\limits_\Omega f v\, dx \quad \forall v \in V_0.$

2× partial integration

see [26], Subsection 3.5

3 $\int\limits_\Omega \Delta\Delta u \cdot v\, dx = \int\limits_\Omega \Delta u \Delta v\, dx + \underbrace{\int\limits_\Gamma \partial_n \Delta u \cdot v\, ds}_{=0} - \underbrace{\int\limits_\Gamma \Delta u \cdot \partial_n v\, ds}_{=0} \quad \forall v \in V_0.$

4 No natural boundary conditions ! (natural BC: $\Delta u = g_2,\ \partial_n \Delta u = g_3$).

5 $V_g = \{v \in V : v = g_0 = 0,\ \partial_n v := \frac{\partial v}{\partial n} = g_1 = 0 \text{ auf } \Gamma\} = V_0.$

- **Result:** variational formulation

Find $u \in V_0 = \overset{\circ}{W}{}_2^2(\Omega) : a(u,v) = < F, v > \quad \forall v \in V_0$ with $a(u,v) = \int\limits_\Omega \Delta u \Delta v \, dx, \;\; < F, v > = \int\limits_\Omega f v \, dx,$ $f \in L_2(\Omega)$ given, $\Gamma = \partial\,\Omega \in C^{0,1}$.

- **Ex 2.9** Show that the following statements are valid:

  1) $f \in V_0^* = W_2^{-2}(\Omega)$,
  2) $a(\cdot,\cdot) : V_0 \times V_0 \mapsto I\!R^1$ – bilinear form:

  2a) $\exists\, \mu_1 = \text{const.} > 0 : a(v,v) \geq \mu_1 \|v\|_{2,\Omega}^2 \quad \forall v \in V_0 = \overset{\circ}{W}{}_2^2(\Omega)$,

  2b) $\exists\, \mu_2 = \text{const.} > 0 : |a(u,v)| \leq \mu_2 \|u\|_{2,\Omega} \|v\|_{2,\Omega} \quad \forall u,v \in V_0$,

  2c) $a(u,v) = a(v,u) \quad \forall u,v \in V_0$.

These properties immediately yield the following results:

- $\exists\,!$ (Lax-Milgram–Theorem 2.1) (Ass. 1) – 2b) are needed, but not 2c)),
- Equivalence to the Minimum Problem (+ 2c)).

- **Remark 2.6:**

1. It follows from step 3 of the derivation of variational formulation that the following boundary conditions are feassible:

   - essential: $u = g_0$ on Γ
   - essential: $\partial_n u = g_1$ on Γ
   - natural: $\Delta u = g_2$ on Γ
   - natural: $\partial_n \Delta u = g_3$ on Γ

   possible combinations

   For Kirchhoff's plate bending problem, the corresponding bilinear form $a(\cdot,\cdot)$ coincides with biharmonic bilinear form given above only in the case of the 1. BVP: ↲ Other natural boundary condition for the plate: $l_2 u = g_2$ und $l_3 u = g_3$, where l_i denotes some i^{th} order differential operator.

2. Possible BVP for the biharmonic PDE $\Delta^2 u = f$:

   1. BVP: $u = g_0$ and $\partial_n u = g_1$ on Γ (pure Dirichlet problem),

   2. BVP: $u = g_0$ and $\Delta u = g_2$ on Γ,

   3. BVP: $\partial_n u = g_1$ and $\partial_n \Delta u = g_3$ on Γ,

   4. BVP: $\Delta u = g_2$ and $\partial_n \Delta u = g_3$ on Γ (pure Neumann problem),

   mixed BVP: e.g. $u = \partial_n u = 0$ on Γ_1 and $\Delta u = \partial_n \Delta u = 0$ on Γ_2.

■ **Ex 2.10** Give the variational formulations of the BVPs described in Remark 2.4.2, and investigate the existence and uniqueness of genaralized solutions (Lax-Milgram). Without loss of generality, assume that the essential boundary conditions are homogeneous (via homogenization) !

Chapter 3

Galerkin Finite-Element-Discretization

■ **Literature:** [2], [4], [7], [11], [34]

3.1 Galerkin–Type Methods for Abstract Variational Problems

3.1.1 The Galerkin Method

■ **Starting point:** Variational formulation:

$$(1)_g \qquad \text{Find } u \in V_g : \quad a(u,v) \; = \; < F, v > \quad \forall v \in V_0 .$$

■ **Idea:** Look for some approximation $u_h \in V_{gh} = g_h + V_{0h} \subset V_g$ to the exact solution $u \in V_g$ of $(1)_g$ such that u_h satisfies the variational equation $(1)_g$ for all test functions $v_h \in V_{0h}$.

- **Galerkin-Scheme:**

$$V_h = \operatorname{span}\{\varphi^{(i)} : i \in \overline{\omega}_h\} = \{v_h = \sum_{i \in \overline{\omega}_h} v^{(i)}\varphi^{(i)}\} = \operatorname{span}\bar{\Phi} \subset V,$$

ansatz functions
↓
linearly independent

index set
for "numbering" the ansatz functions

$$\dim V_h = |\overline{\omega}_h| = \overline{N}_h = N_h + \partial N_h < \infty,$$
$$\bar{\Phi} = [\varphi^{(i)} : i \in \overline{\omega}_h] = [\varphi_1, \ldots, \varphi_{\bar{N}_h}],$$

$$V_{gh} = \underbrace{V_h \cap V_g}_{\substack{\neq \emptyset \\ \text{Ass. (!)}}} = g_h + V_{0h} = \left\{ v_h = \underbrace{\left[\sum_{i \in \gamma_h := \bar{\omega}_h \setminus \omega_h} u_*^{(i)}\varphi^{(i)} \right]}_{\substack{=: g_h \in V_g \cap V_h \\ \text{given (fixed)}}} + \underbrace{\sum_{i \in \omega_h} v^{(i)}\varphi^{(i)}}_{\in V_{0h} \subset V_0} \right\} \subset V_g,$$

$$V_{0h} = V_h \cap V_0 = \{v_h = \sum_{i \in \omega_h} v^{(i)}\varphi^{(i)}\} = \operatorname{span}\Phi \subset V_0;\ \dim V_{0h} = N_h.$$

$(1)_{gh}$

Find $u_h \in V_{gh} : a(u_h, v_h) = < F, v_h > \quad \forall\, v_h \in V_{0h}$

$\uparrow$ test space

GS

Galerkin-isomorphism

Choose the basis: $\Phi = \{\varphi^{(i)} : i \in \bar{\omega}_h\} = [\varphi_1, \dots, \varphi_N]$

$u_h = \sum\limits_{i\in\omega_h} u^{(i)}\varphi^{(i)} + \sum\limits_{i\in\gamma_h} u_*^{(i)}\varphi^{(i)}$

$\uparrow$ look for ! $\qquad \uparrow$ given !

$v_h = \varphi^{(k)} \in V_{0h} = \operatorname{span}\{\varphi^{(j)} : j \in \omega_h\} = \operatorname{span}\Phi, \quad \forall k \in \omega_h$

$a(\cdot,\cdot)$ – bilinear, $< F, \cdot >$ linear

Find $\underline{u}_h = [u^{(i)}]_{i\in\omega_h}$: $\quad \sum\limits_{i\in\omega_h} u^{(i)} a(\varphi^{(i)}, \varphi^{(k)}) = < F, \varphi^{(k)} > - \sum\limits_{i\in\gamma_h} u_*^{(i)} a(\varphi^{(i)}, \varphi^{(k)})$

$\forall k \in \omega_h$

$\underline{(1)}_{gh}$ SS Galerkin-System

Find $\underline{u}_h \in \mathbb{R}^{N_h}$: $K_h \underline{u}_h = \underline{f}_h$ (linear system)

with $K_h = [a(\varphi^{(i)}, \varphi^{(k)})]_{k,i\in\omega_h}$ – $(N_h \times N_h)$-matrix

stiffness matrix

$\underline{f}_h = [< F, \varphi^{(k)} > - \sum\limits_{i\in\gamma_h} u_*^{(i)} a(\varphi^{(i)}, \varphi^{(k)})]_{k\in\omega_h} \in \mathbb{R}^{N_h}$

load vector

Here we use the abbreviations: GS = Galerkin Solution,
SS = Skeleton Solution (Galerkin coefficient vector)

■ **Ex 3.1** How does the Galerkin system $K_h \underline{u}_h = \underline{f}_h$ change, if one uses a new basis $[p^{(1)}, \dots, p^{(N_h)}] = [\varphi^{(1)}, \dots, \varphi^{(N_h)}]\, B$? Here the matrix $B = [b_{ki}]_{k,i=1,\dots,N_h}$ is a regular $(N_h \times N_h)$ basis tranformation matrix, where the above notation means

$$p^{(i)} = \sum_{k=1}^{N_h} b_{ki}\varphi^{(k)}.$$

3.1.2 The Galerkin-Petrow Method

- **Starting point:** Variational formulation $(1)_g$.

- **Ansatz functions:**
$$\begin{aligned} V_h &= \{v_h = \sum_{i\in\overline{\omega}_h} v^{(i)}\varphi^{(i)}\} \subset V, \\ V_{gh} &= \{v_h = \sum_{i\in\omega_h} v^{(i)}\varphi^{(i)} + \sum_{i\in\gamma_h} u_*^{(i)}\varphi^{(i)}\} \subset V_g, \\ V_{0h} &= \{v_h = \sum_{i\in\omega_h} v^{(i)}\varphi^{(i)}\} \subset V_0. \end{aligned}$$

- **Test functions:**
$$\begin{array}{lcl} U_{0h} & = & \text{span } \{q^{(i)} : i \in \omega_h\} = \{v_h = \sum\limits_{i\in\omega_h} v^{(i)}q^{(i)}\} \subset V_0, \\ \text{in general} & & U_{0h} \neq V_{0h}, \text{ but } \dim U_{0h} = \dim V_{0h} = |\omega_h| = N_h. \end{array}$$

- **Galerkin-Petrow-Scheme:**

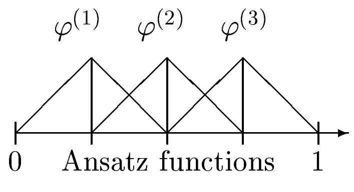
$$\begin{array}{lcl} \text{Find } u_h \in V_{gh} & : & a(u_h, v_h) = <F, v_h> \quad \forall v_h \in U_{0h}; \\ \updownarrow & & \\ \text{Find } \underline{u}_h = [u^{(i)}]_{i\in\omega_h} \in I\!R^{N_h} & : & \sum\limits_{i\in\omega_h} u^{(i)} a(\varphi^{(i)}, q^{(k)}) = <F, q^{(k)}> - \sum\limits_{i\in\gamma_h} u_*^{(i)} a(\varphi^{(i)}, q^{(k)}) \\ & & \qquad k \in \omega_h \\ \hline & & K_h \underline{u}_h = \underline{f}_h \end{array}$$

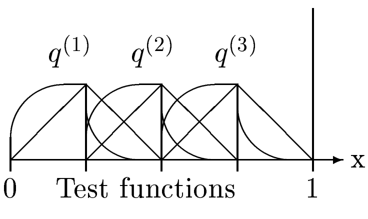
- **Advanced Application of the Galerkin-Petrow–Method:** Upwind-FEM,

e.g.: diffusion-convection problem (cf. [11], [34]):

$$\text{Find } u \in V_0 = \overset{\circ}{H}{}^1(0,1) : \int_0^1 (u'v' + \nu u'v)\, dx = \int_0^1 f v\, dx, \quad \forall v \in V_0$$

$\rightarrow \nu > 0$





<u>Goal:</u> Conservation of the M-matrix-property of K_h !

3.1.3 The Ritz Method as Special Case of the Galerkin Method

- **Starting Point:** = minimization problem (see Section 2.1, Theorem 2.2):

$$(2)_g \qquad \boxed{\text{Find } u \in V_g : J(u) = \inf_{v \in V_g} J(v)}$$

with $V_g = g + V_0 \subset V$ ($\uparrow$) and the Ritz energy functional

$$(3) \qquad J(v) = \tfrac{1}{2}\, a(v,v) - \,< F, v >,$$

and under the assumptions:

$$(4) \qquad \left\{ \begin{array}{ll} 1) & F \in V^* \ ! \ \leftarrow \text{(can be weakened !)} \\ & \\ 2) & a(\cdot,\cdot) : V \times V \mapsto I\!R^1 - \text{continuous and, on } V_0, \\ & \qquad\qquad\qquad\qquad \text{symmetric and positiv !} \end{array} \right.$$

- **Remark:** about Assumptions (4) and Theorem 2.2 $\rightarrow (2)_{g=0}$:

homogenization ansatz: $v = g + w \in V_g,\ w \in V_0$:

$$\begin{aligned} \Rightarrow \quad J(u) \quad &= \quad J(g+w) = \tfrac{1}{2}\left(a(g,g) + a(g,w) + a(w,g) + a(w,w)\right) - \\ & \qquad\qquad - < F, g > - < F, w > \\ &= \quad \tfrac{1}{2}\, a(w,w) - \underbrace{\left[< F, w > - \frac{1}{2}\left(a(g,w) + a(w,g)\right)\right]}_{=: <\hat{F}, w>} + J(g) \end{aligned}$$

$$\Rightarrow \boxed{\inf_{v \in V_g} J(v) \Leftrightarrow \inf_{w \in V_0} \left[\tfrac{1}{2}\, a(w,w) - < \hat{F}, w >\right]}$$

- **Ansatz:** see Galerkin-Method (Subsection 3.1.1):
  $\rightarrow$ Ansatz functions $\{\varphi^{(i)} : i \in \bar{\omega}_h\} \subset V \Rightarrow V_h, V_{0h}, V_{gh}$!

- **Ritz-idea:**

RS

$(2)_{gh}$ Find $u_h \in V_{gh} : J(u_h) = \min\limits_{v_h \in V_{gh}} J(v_h) \geq J(u)$

Ritz-isom. $\longleftarrow \frac{\partial J(u_h)}{\partial u^{(k)}} = 0 \quad \forall k \in \omega_h; \quad u_h = \sum\limits_{i \in \omega_h} u^{(i)} \varphi^{(i)} + \sum\limits_{i \in \gamma_h} u_*^{(i)} \varphi^{(i)}$

$(\underline{2})_{gh}$ Find $\underline{u}_h = [u^{(i)}]_{i \in \omega_h}$: $\sum\limits_{i \in \omega_h} u^{(i)} a(\varphi^{(i)}, \varphi^{(k)}) =< F, \varphi^{(k)} > - \sum\limits_{i \in \gamma_h} u_*^{(i)} a(\varphi^{(i)}, \varphi^{(k)})$ $\forall k \in \omega_h$

SS

Find $\underline{u}_h \in I\!R^{N_h}$: $K_h \underline{u}_h = \underline{f}_h$

with $K_h = [a(\varphi^{(i)}, \varphi^{(k)})]_{k,i \in \omega_h}$!

Ritz system $\qquad \underline{f}_h = [< F, \varphi^{(k)} > - \sum\limits_{i \in \gamma_h} u_*^{(i)} a(\varphi^{(i)}, \varphi^{(k)})]_{k \in \omega_h}$

Here, RS = Ritz Solution.

- **Remark:** Under the assumptions (4), we have: Ritz-System = Galerkin-System !
  $\Rightarrow$ Galerkin-Ritz-System !

- <u>**Exercises:**</u>

Ex 3.2 Prove that K_h is symmetric and positive definite !

Ex 3.3 Show that, under assumptions $(4)_{2)}$,

$$[\cdot, \cdot] := a(\cdot, \cdot)$$

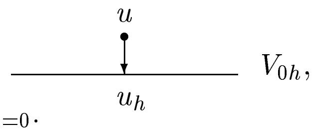
defines some scalar product on V_0 ! In contrast to the scalar product induced by the original scalar product $(\cdot, \cdot)$ of V in the subspace V_0, the scalar product $[\cdot, \cdot]$ is called **energy inner (scalar) product**, and the corresponding norm $|\cdot|$,

$$|\cdot|^2 := [\cdot, \cdot] \equiv a(\cdot, \cdot),$$

energy norm.

Ex 3.4 Prove that, under assumptions (4) with $g = 0$, the relation

$$\min_{v_h \in V_{0h}} J(v_h) \Leftrightarrow \min_{v_h \in V_{0h}} |u - v_h|,$$



is valid, where $u \in V_0$ is the unique solution of $(2)_{g=0}$.

<u>Hints:</u> 1. Theorem 2.2: $a(u, v) = < F, v > \quad \forall v \in V_0$.

2. $J(v) = \frac{1}{2}\, a(v, v) - < F, v > + \frac{1}{2}\, a(u, u) - \frac{1}{2}\, a(u, u) =$
$= \frac{1}{2}\, |u - v|^2 - \frac{1}{2}\, a(u, u)$.

Ex 3.5 Let us suppose the assumptions (4) with $g = 0$, e.g. $V_g = V_0, V_{gh} = V_{0h}$. Show that the relation

$$P_R\, u \in V_{0h} : a(P_R\, u, v_h) = a(u, v_h) \quad \forall v_h \in V_{0h} \quad \forall u \in V_0$$

uniquely defines some linear, continuous operator $P_R \in L(V_0, V_{0h})$ (cf. also Theorems 2.1 and 2.2 under the standard assumptions), which is an orthoprojection with respect to the energy inner product $[\cdot, \cdot]$, e.g. $P_R^2 = P_R$ and $[P_R\, u, v] = [u, P_R\, v] \quad \forall u, v \in V_0$. This orthoprojection is called **Ritz-Projection**. Show further that the following minimization property holds:

$$|u - P_R\, u| = \inf_{v_h \in V_{0h}} |u - v_h| \quad \forall \text{ (fix) } u \in V_0.$$

3.2 Analysis of the Galerkin Method

3.2.1 Existence and Uniqueness of the Galerkin-Ritz Solution

= Conclusion from Theorem 2.1 (Lax & Milgram) for the Ritz-Galerkin-System:
$V_0 \mapsto V_{0h}$!

■ Without loss of gererality (possibly, after homogenization, cf. Section 2.1), we consider again the **abstract variational problem**

$$(1)_0 \qquad \boxed{\text{Find } u \in V_0 : a(u,v) \;= \,< F, v > \quad \forall v \in V_0}$$

under the **standard assumptions** (see Section 2.1; $V_0 \subset V$ - closed, non-trivial subspace of the H-space $V, \|\cdot\|, (\cdot,\cdot)$):

$$(5) \quad \begin{cases} 1) & F \in V_0^*. \\ 2) & a(\cdot,\cdot) : V_0 \times V_0 \to I\!R^1 \text{ - bilinear form on } V_0: \\ 2a) & V_0\text{–elliptic:} \quad \mu_1 \|v\|^2 \le a(v,v) \quad \forall v \in V_0, \\ 2b) & V_0\text{–bounded:} \quad |a(u,v)| \le \mu_2 \|u\| \|v\| \quad \forall u, v \in V_0. \end{cases}$$

■ Let $V_{0h} = \text{span}\,\{\varphi^{(i)} : i \in \omega_h\} \subset V_0$ be some finite dimensional subspace of V_0:

→ <u>better:</u> Family of subspaces with $h \in \Theta$:
$\dim V_{0h} = N_h \to \infty$ as $h \to 0$ ($|h| \to 0$)
(see Sections 3.3 and 3.4 for Examples).

→ $\Phi = \{\varphi^{(i)} : i \in \omega_h\}$ – basis in V_0, i.e. linearly independent !

■ **Overview:** Ritz-Galerkin-Method

GALERKIN	$V_{0h} \subset V_0$
Starting point: VP $(1)_0$ Find $u \in V_0 : a(u,v) \; = \; < F, v > \;\; \forall v \in V_0$ Ass.: 1), 2), 2a), 2b)	
$u_h \in V_{0h} : a(u_h, v_h) \; = \; < F, v_h > \; \forall v_h \in V_{0h}$	

$\Longleftrightarrow$ $a(\cdot,\cdot)$ symmetric positive

RITZ	$V_{0h} \subset V_0$
Starting point: MP $(2)_0$ Find $u \in V_0 : J(u) = \inf\limits_{v \in V_0} J(v)$ Ass.: 1), 2), 2a), 2b), 2c)	
$u_h \in V_{0h} : J(u_h) = \inf\limits_{v_h \in V_{0h}} J(v_h)$	

$$V_{0h} \ni u_h \leftrightarrow \underline{u}_h = [u^{(i)}]_{i \in \omega_h} \in \mathbb{R}^{N_h}$$

$$u_h = \sum_{i \in \omega_h} u^{(i)} p^{(i)}$$

$$\updownarrow$$

$$\underline{u}_h = [u^{(i)}]_{i \in \omega_h} \in \mathbb{R}^{N_h}$$

"$\leftrightarrow$" Galerkin-Ritz–isomorphism

$V_{0h} = \underbrace{\text{span}\,\{\varphi^{(i)} : i \in \omega_h\}}_{\text{basis}}$

$N_h = |\omega_h|, \;\; h \in \Theta$

$v_h = \varphi^{(k)}$
$k \in \omega_h$

$\frac{\partial J(u_h)}{\partial u^{(k)}} = 0$
$k \in \omega_h$

GALERKIN-RITZ–system = finite dimensional system $(N_h \times N_h)$
Find $\underline{u}_h = [u^{(i)}]_{i \in \omega_h} \in \mathbb{R}^{N_h}$: $\sum\limits_{i \in \omega_h} u^{(i)} a(\varphi^{(i)}, \varphi^{(k)}) \; = \; < F, \varphi^{(k)} > \quad \forall k \in \omega_h$ $K_h \underline{u}_h \; = \; \underline{f}_h$

■ **<u>Theorem 3.1:</u>** (Lax-Milgram for Galerkin-Ritz–systems)

<u>Ass.:</u> 1. Standard assumptions (5): 1), 2), 2a), 2b),

2. $V_{0h} \subset V_0$ – finite dimensional subspace of V_0.

<u>St.:</u> 1. $\exists!\, u_h \in V_{0h} : a(u_h, v_h) \;= < F, v_h > \quad \forall v_h \in V_{0h} \qquad (1)_{0h}$

2. The sequence $\{u_h^n\} \subset V_{0h}$ uniquely defined by the relation

(6) $u_h^{n+1} \in V_{0h} : (u_h^{n+1}, v_h) = (u_h^n, v_h) - \varrho(a(u_h^n, v_h) - < F, v_h >)$
$\forall v_h \in V_{0h}$,

converges to solution $u_h \in V_{0h}$ of $(1)_{0h}$ in V_0 $(\|\cdot\|, (\cdot,\cdot))$ for an arbitrary initial guess $u_h^0 \in V_{0h}$ and for some fixed $\varrho \in (0, 2\mu_1/\mu_2^2)$.

3. The following iteration error estimates are valid:

a) $\|u_h - u_h^n\| \leq q^n \|u_h - u_h^0\|$ (a-priori)

b) $\|u_h - u_h^n\| \leq \frac{q^n}{1-q} \|u_h^1 - u_h^0\|$ (a-priori-bound)

c) $\|u_h - u_h^n\| \leq \frac{q}{1-q} \|u_h^n - u_h^{n-1}\|$ (a-posteriori-bound)

with $0 \leq q_{\text{opt}} = q(\varrho_{\text{opt}}) \leq q(\varrho) := (1 - 2\mu_1\varrho + \mu_2^2\varrho^2)^{\frac{1}{2}} < 1$
for $\varrho \in (0, 2\mu_1/\mu_2^2)$; $\varrho_{\text{opt}} = \mu_1/\mu_2^2$, $q_{\text{opt}} = \sqrt{1-\xi^2}$, $\xi = \mu_1/\mu_2$.

<u>Proof:</u> Analogous to the proof of Theorem 2.1 by Lax & Milgram: $V_{0h} \subset V_0 \subset V$ – closed, non-trivial subspace !

<u>Note:</u> Riesz-Isom.:

$$\begin{array}{llccl} J & : & V_0^* & \mapsto & V_0 \\ & & \cup & \longleftarrow & \cup \text{ — Hahn-Banach !} \\ J_h = J|_{V_{0h}} & : & V_{0h}^* & \mapsto & V_{0h} \quad \text{(ex.)} \end{array}$$

(6) $\Leftrightarrow u_h^{n+1} = u_h^n - \varrho(J_h A_h u_h^n - J_h F_h)$ in V_{0h}.

q.e.d. ■

3.2.2 Iterative Solution of the Galerkin-Ritz System

■ **Corollary 3.2:**

Because of the Galerkin-Ritz-Isomorphism $u_h \leftrightarrow \underline{u}_h$, the iteration (6) in V_{0h} is equivalent to the following iteration method in $I\!R^{N_h}$:

$$(\underline{6}) \qquad B_h \frac{\underline{u}_h^{n+1} - \underline{u}_h^n}{\varrho} + K_h \underline{u}_h^n = \underline{f}_h, \; n = 0, 1, \dots; \; \underline{u}_h^0 \in I\!R^{N_h},$$

with the symmetric and positive definite matrix (Gramer-matrix with respect to the scalar product $(\cdot, \cdot)$ in $V_0(V)$)

$$B_h = [(\varphi^{(i)}, \varphi^{(k)})]_{k, i \in \omega_h}.$$

Proof:

$$(6) \qquad u_h^{n+1} \in V_{0h} : (u_h^{n+1}, v_h) = (u_h^n, v_h) - \varrho(a(u_h^n, v_h) - < F, v_h > \qquad \forall v_h \in V_{0h}$$

$$u_h^j = \sum_{i \in \omega_h} u_j^{(i)} \varphi^{(i)}, \; j = n, n+1, \text{ where } j \text{ is here the iteration index}$$

$$v_h = \varphi^{(k)}, \; k \in \omega_h$$

$$(\underline{6}) \qquad \underline{u}_h^{n+1} \in I\!R^{N_h} : \sum_{i \in \omega_h} u_{n+1}^{(i)} (\varphi^{(i)}, \varphi^{(k)}) = \sum_{i \in \omega_h} u_n^{(i)} (\varphi^{(i)}, \varphi^{(k)}) - \varrho \sum_{i \in \omega_h} u_n^{(i)} a(\varphi^{(i)}, \varphi^{(k)}) + $$
$$+ \varrho < F, \varphi^{(k)} > \qquad \forall k \in \omega_h$$
$$B_h \cdot \underline{u}_h^{n+1} = B_h \underline{u}_h^n - \varrho K_h \underline{u}_h^n + \varrho \underline{f}_h$$

q.e.d. ■

■ **Remark 3.3:**
The iteration error estimates 3a) – 3c) of Theorem 4.1 hold, where
$\|v_h\| = \|\underline{v}_h\|_{B_h} \quad \forall v_h \leftrightarrow \underline{v}_h \in I\!R^{N_h}$ mit $\|\underline{v}_h\|_{B_h} := (B_h \underline{v}_h, \underline{v}_h)_{I\!R^{N_h}}^{0.5}$,
$(\cdot, \cdot)_{I\!R^{N_h}}$ – Euklidisches Skalarprodukt.

■ **Observation:**
The iteration ($\underline{3}$) is effective if and only if the system

$$B_h \underline{w}_h^n = \underline{d}_h^n := \underline{f}_h - K_h \underline{u}_h^n,$$
$$\underline{u}_h^{n+1} = \underline{u}_h^n + \varrho \, \underline{w}_h^n,$$

which has to be handled in each iteration step, can be solved fast (i.e. $Q(B_h^{-1} * \underline{d}_h^n) :=$ number of arithmetical operations for solving the system $B_h \underline{w}_h^n = \underline{d}_h^n = O(K_h * \underline{u}_h^n)$!!). Unfortunately, in practical applications, this is not true in general ($\exists$ a few exceptions !) !
Note that the convergence rate $q = q_{\text{opt}} = \sqrt{1 - \xi^2} < 1$ is, independently of h, less than 1, because $\xi = \mu_1 / \mu_2$ does not depend on h !

■ **Idea:** = Preconditioning C_h !

Let $C_h = C_h^T$ symmetric and positive definite (spd):

$$(7) \begin{cases} 1. & Q(C_h^{-1} * \underline{d}_h^n) = O(K_h * \underline{u}_h^n) \\ 2. & C_h \text{ is spectrally equivalent to } B_h, \\ & \text{i.e. } \exists\, \gamma_1, \gamma_2 = \text{const.} > 0 : \gamma_1 C_h \le B_h \le \gamma_2 C_h, \text{ i.e.} \\ & \gamma_1 (C_h \underline{v}_h, \underline{v}_h)_{I\!R^{N_h}} \le (B_h \underline{v}_h, \underline{v}_h)_{I\!R^{N_h}} \le \gamma_2 (C_h \underline{v}_h, \underline{v}_h)_{I\!R^{N_h}} \quad \forall \underline{v}_h \in I\!R^{N_h}. \end{cases}$$

Then we can use the iteration method ($\underline{8}$)

$$(\underline{8}) \qquad C_h \frac{\underline{u}_h^{n+1} - \underline{u}_h^n}{\tau} + K_h \underline{u}_h^n = \underline{f}_h, \; n = 0, 1, \ldots; \; \underline{u}_h^0 \in I\!R^{N_h}$$

instead of ($\underline{6}$). Moreover, the following Theorem 3.4 shows that the iteration method ($\underline{5}$) is <u>efficient</u>.

■ **Theorem 3.4:**

<u>Ass.:</u>
1. Standard assumptions (5): 1), 2), 2a), 2b),
2. $V_{0h} \subset V_0$ – finite dimensional, non-trivial subspace of V_0,
3. Assumptions (7).

<u>St.:</u>
1. The iteration ($\underline{8}$) is converging to the solution

$$\begin{array}{ll} \underline{u}_h \in I\!R^{N_h} : K_h \underline{u}_h = f_h & (\underline{1})_{0h} \\ \updownarrow & \\ u_h \in V_{0h} : \quad a(u_h, v_h) = \langle F, v_h \rangle \quad \forall v_h \in V_{0h} & (1)_{0h} \end{array}$$

   for every, fixed $\tau \in (0, 2\nu_1/\nu_2^2)$, with $\nu_1 = \mu_1\gamma_1$ and $\nu_2 = \mu_2\gamma_2$.

2. The following iteration error estimates are valid:

   (a) $\|\underline{u}_h - \underline{u}_h^n\|_{C_h} \le q^n \|\underline{u}_h - \underline{u}_h^0\|_{C_h}$

   (b) $\|\underline{u}_h - \underline{u}_h^n\|_{C_n} \le \frac{q^n}{1-q} \|\underline{u}_h^1 - \underline{u}_h^0\|_{C_h}$

   (c) $\|\underline{u}_h - \underline{u}_h^n\|_{C_n} \le \frac{q}{1-q} \|\underline{u}_h^n - \underline{u}_h^{n-1}\|_{C_h}$

   with $\|\underline{v}_h\|_{C_h} = (\underline{v}_h, \underline{v}_h)_{C_h}^{0.5} := (C_h \underline{v}_h, \underline{v}_h)_{\cancel{I\!R^{N_h}}}^{0.5}$

   and $0 \le q_{\text{opt}} = q(\tau_{\text{opt}}) \le q(\tau) := (1 - 2\nu_1\tau + \nu_2^2\tau^2)^{0.5} < 1$,
   for $\tau \in (0, 2\nu_1/\nu_2^2)$, $\tau_{\text{opt}} = \nu_1/\nu_2^2$, $q_{\text{opt}} = \sqrt{1-\xi^2}$, $\xi = \nu_1/\nu_2$.

3. If $\gamma_1, \gamma_2 \ne c(h)$, then $I(\epsilon) = O(\ln \epsilon^{-1})$ iterations and $Q(\epsilon) = O((K_h * \underline{u}_h^n) \cdot \ln \epsilon^{-1})$ arithmetical operations are needed in order to reduce the initial error by some factor $\epsilon \in (0,1)$, i.e. $\|\underline{u}_h - \underline{u}_h^{I(\epsilon)}\|_{C_h} \le \epsilon \|\underline{u}_h - \underline{u}_h^0\|_{C_h}$.

■ **Proof:**

- In analogy to Theorem 4.1: $\|v_h\| = \|\underline{v}_h\|_{B_h} \mapsto \|\underline{v}_h\|_{C_h}$.
  Then the following estimates are valid:
  a) $a(v_h, v_h) = (K_h\underline{v}_h, \underline{v}_h) \geq \mu_1 \|\underline{v}_h\|^2_{B_h} \geq \mu_1\gamma_1 \|\underline{v}_h\|^2_{C_h} \quad \forall \underline{v}_h \in I\!R^{N_h}$,
  b) $|a(u_h, v_h)| = |(K_h\underline{u}_h, \underline{v}_h)| \leq \mu_2 \|\underline{u}_h\|_{B_h} \|\underline{v}_h\|_{B_h} \leq \mu_2\gamma_2 \|\underline{u}_h\|_{C_h} \|\underline{v}_h\|_{C_h} \quad \forall \underline{u}_h, \underline{v}_h \in I\!R^{N_h}$.

#

- Direct proof (exercise).
  Hints:
  - error scheme: $\underline{z}_h^{n+1} := \underline{u}_h - \underline{u}_h^{n+1} = (I_h - \tau C_h^{-1} K_h) \underline{z}_h^n$
  - Estimate in the $\|\cdot\|_{C_h}$–norm:
  $$\begin{aligned} \|z_h^{n+1}\|^2_{C_h} &= \|(I_h - \tau C_h^{-1} K_h)\underline{z}_h^n\|^2_{c_h} = (\cdot,\cdot)_{c_h} = \\ &= \text{etc., as in the proof of Theorem 2.1 ! } \# \end{aligned}$$

- Statement 3. follows from: $q^n \leq \epsilon$, iff $n \geq \frac{\ln \epsilon^{-1}}{\ln q^{-1}}$.

q.e.d. ■

■ **Remark 3.5:** → see also literature (for PDEs) [21], [27], **[34]**

1. Candidates for C_h:
   - SSOR: [21],
   - ILU, MILU, IC, MIC: [21],
   - Multigrid–preconditioners: [21], [27], [25], [34],
   - Multilevel–preconditioners (e.g. BPX): [21], [27], [25].
2. It is <u>not</u> assumed in Theorem 3.4 that $K_h = K_h^T$. However, from the V_0-ellipticity of $a(\cdot,\cdot)$, we immediately conclude that the stiffness matrix K_h is positive definite.
   If $\boxed{K_h = K_h^T \text{ p.d.}}$, then
   (a) the convergence rate estimate (5) can be improved:
   $\nu_1 C_h \leq K_h \leq \nu_2 C_h \;\; ↳ \;\; \tau_{\text{opt}} = 2/(\nu_1 + \nu_2)$, $q_{\text{opt}} = \frac{1-\xi}{1+\xi}$, $\xi = \frac{\nu_1}{\nu_2}$; and
   (b) the conjugate gradient acceleration is feassible, and, of course, used in practice (see also Chapter 5) !

3.2.3 Discretization Error Estimates

3.2.3.1 Cea's Lemma

■ **Theorem 3.6:** (Cea's Lemma, 1964)

Ass.: 1. Standard assumptions (5): 1), 2), 2a), 2b).

2. $V_{gh} = g_h + V_{0h} \subset V_g$ - finite dimensional hyperplane with the finite dimensional subspace $V_{0h} \subset V_0$ and $g_h \in V_g \cap V_h$ given ($\uparrow$).

3.
$$\begin{array}{lll} u \in V_g: & a(u,v) = < F, v > & \forall v \in V_0 \quad (1)_g \\ u_h \in V_{gh}: & a(u_h, v_h) = < F, v_h > & \forall v_h \in V_{0h} \quad (1)_{gh} \end{array}$$

St.:

$$\underbrace{\|u - u_h\|}_{\text{discretization error}} \leq \frac{\mu_2}{\mu_1} \underbrace{\inf_{w_h \in V_{gh}} \|u - w_h\|}_{\text{approximation error}} \qquad (9)$$

Proof:

$$\begin{array}{lll} (1)_g & a(u, v_h) = < F, v_h > & \forall v_h \in V_{0h} \subset V_0 \; ! \\ -\,(1)_{gh} & a(u_h, v_h) = < F, v_h > & \forall v_h \in V_{0h} \\ \hline (10) & \underline{\underline{a(u - u_h, v_h) = 0 \quad \forall v_h \in V_{0h}}} & \end{array}$$

Choose $v_h = u - u_h - (u - w_h) = w_h - u_h \in V_{0h} \quad \forall w_h \in V_{gh}$ in (6):

$$\mu_1 \|u - u_h\|^2 \overset{2a)}{\leq} \boxed{a(u - u_h, u - u_h) = a(u - u_h, u - w_h)} \overset{2b)}{\leq} \mu_2 \|u - u_h\| \, \|u - w_h\|$$

$\Rightarrow$ (5) **q.e.d.** ■

■ **Ex 3.6** Show the improved discretization error estimate

$$\|u - u_h\| \leq \sqrt{\frac{\mu_2}{\mu_1}} \inf_{v_h \in V_{0h}} \|u - v_h\| \qquad (9)_s$$

for the homogenized problem (i.e. for $g = 0$) provided that, in addition to the standard assumption (5), the bilinear form $a(\cdot, \cdot)$ is symmetric, i.e.

2c) $\quad a(u, v) = a(v, u) \quad \forall u, v \in V_0,$

Hint: Use the equivalence which was to prove in Ex 3.4 !

■ **Remark 3.7:**

1. The "**discretization - error - problem**" is now reduced to a pure "**approximation - error - problem**" (= stability) !
2. Error estimates of the form (5) are called **quasioptimal**:

$$\inf_{v_h \in V_{gh}} \|u - v_h\| \;\leq\; \|u - u_h\| \;\leq\; \frac{\mu_2}{\mu_1} \inf_{v_h \in V_{gh}} \|u - v_h\|.$$

3. Convergence $u_h \xrightarrow[h \to 0]{} u$ in $V \Leftrightarrow \inf\limits_{v_h \in V_{gh}} \|u - v_h\| \to 0$ for $h \to 0$.
4. The family $\{V_{0h}\}_{h \in \Theta}$ of finite dimensional subspaces of the space V_0 is called **limitingly complete** if

$$\lim_{\substack{h \to 0 \\ h \in \Theta}} \inf_{v_h \in V_{0h}} \|u - v_h\| = 0 \quad \forall u \in V_0.$$

If $g = 0$, or $g \in V_{gh}$ $\forall h \in \Theta$, then the limiting completeness ensures the convergence of the Galerkin method.

Hint: $\left.\begin{array}{l} u = g + \mathbf{u},\ \mathbf{u} \in V_0 \\ v_h = g + \mathbf{v_h},\ \mathbf{v_h} \in V_{0h},\ \text{d.h. } g_h = g \in V_{gh} \end{array}\right\} \Rightarrow u - v_h = \mathbf{u} - \mathbf{v_h}$

#

3.2.3.2 Nitsche's Trick

■ Consider, without loss of generality, the homogenized **variational problem** $(1)_0$ and the corresponding Galerkin-scheme $(1)_{0h}$:

$(1)_0$ Find $u \in V_0 : a(u,v) \;= \;< F, v > \quad \forall v \in V_0,$

$(1)_{0h}$ Find $u_h \in V_{0h} : a(u_h, v_h) \;= \;< F, v_h > \quad \forall v_h \in V_{0h}.$

- Let H be some Hilbert-space equipped with some weaker norm $|\cdot|$ and the corresponding scalar product $[\cdot,\cdot]$, i.e.

$$V_0,\ \|\cdot\|, (\cdot,\cdot) \quad \text{(e.g. } \overset{\circ}{H}{}^1(\Omega) \text{ for second-order PDEs, Dirichlet's BC)}$$

$$H,\ |\cdot|, [\cdot,\cdot] \quad \text{(e.g. } L_2(\Omega),\ H^{-\lambda}(\Omega),\ 0 \le \lambda < 1),$$

and let us suppose that

$$(11)\quad \begin{cases} V_0 \subset H \subset V_0^*; \\ \overline{V_0}^{|\cdot|} = H, \text{ i.e. } V_0 \text{ is dense in } H; \\ |v| \le c\|v\| \quad \forall v \in V_0,\ c = \text{const.} > 0. \end{cases}$$

- From (11), we immediately have
$$|u-u_h| \le c\,\|u-u_h\|,$$
i.e. the Galerkin method converges in the $|\cdot|$–norm at least as fast as in the $\|\cdot\|$–norm !

- **Goal:** If the $|\cdot|$–norm is actually weaker than the $\|\cdot\|$–norm,

  i.e. $V_0 \hookrightarrow H$, or $\inf\limits_{\substack{v\in V_0\\ v\ne 0}} \frac{|v|}{\|v\|} = 0$,

  then a faster convergence should be expected, i.e. $|u-u_h| = o(\|u-u_h\|)$.

- **<u>Theorem 3.8:</u>** (Duality argument by Aubin & Nitsche, 1968)

  <u>Ass.:</u>
  1. Standard assumptions (5): 1), 2), 2a), 2b);
  2. $V_{0h} \subset V_0$ – finite dimensional, closed, non-trivial subspace of V_0;
  3. $u \in V_0 : (1)_0$ and $u_h \in V_{0h} : (1)_{0h}$;
  4. the spaces H and V_0 satisfy properties (11).

  <u>St.:</u> Then the discretization error estimate

$$(12)\qquad |u-u_h| \ \le\ \mu_2\,\|u-u_h\| \sup_{\substack{g\in H\\ g\ne 0}} \left[\frac{1}{|g|} \inf_{v_{gh}\in V_0} \|w_g - v_{gh}\|\right],$$

  holds, where $w_g \in V_0$ denotes the uniquely existing solution of the adjoint to $(1)_0$ variational problem

$$(1)_0^*\qquad a(v, w_g) = [g, v] \quad \forall v \in V_0$$

  for every fixed $g \in H,\ g \ne 0$.

Proof: Obviously, we have the representation

$$|u-u_h| = \sup_{\substack{g\in H\\ g\neq 0}} \frac{[u-u_h, g]}{|g|} \,.$$

Using the Galerkin orthogonality (10), we obtain the estimate

$$\begin{aligned} [u-u_h, g] &= [g, u-u_h] \stackrel{(1)_0^*}{=} a(u-u_h, w_g) \stackrel{\substack{\text{"Galerkin-orthogonality" !}\\(10)}}{=} a(u-u_h, w_g - v_{gh}) \leq \\ &\leq \mu_2 \|u-u_h\|\,\|w_g - v_{gh}\| \quad \forall v_{gh} \in V_{0h} \end{aligned}$$

that concludes the proof of (12). **q.e.d.** ■

■ **Remark 3.9:**

Usually, in the practice (PDEs), the following situation occurs:

$$\begin{array}{lccccccc} & \|\cdot\|_W & & \|\cdot\|_{V_0} = \|\cdot\| & & \|\cdot\|_H = |\cdot| & & \|\cdot\|_{V_0^*} = \|\cdot\|_* \\ & W & \hookrightarrow & V_0 & \hookrightarrow & H & \hookrightarrow & V_0^* \\ [\text{ e.g.} & H^{1+\lambda} & \hookrightarrow & \overset{\circ}{H}{}^1 & \hookrightarrow & H^{-1+\lambda} & \hookrightarrow & H^{-1}, \lambda \in (0,1] \text{ e.g. } \lambda = 1] \end{array}$$

with some <u>regularity result</u> (W-coerciveness) for $(1)_0^*$, i.e.

(13) $\qquad w_g \in V_0 \cap W$ and $\|w_g\|_W \leq c_R \|g\|_H$,

and with some <u>approximation result</u>

(14) $\qquad \inf\limits_{v_{gh}\in V_{0h}} \|w_g - v_{gh}\|_{V_0} \leq c_A h^\alpha \|w_g\|_W .$

Then, from (12), (13) and (14), we obtain the estimate

$(12)'$ $\qquad \|u-u_h\|_H = |u-u_h| \leq \mu_2\, c_R\, c_A\, h^\alpha \|u-u_h\|_{V_0}; \quad [\alpha = \lambda].$

3.3 Finite Element Method

- **Text books:** [7] Ciarlet (standard monography on FEM), [4], [11], [20], [23], [36].

- **FEM** = Galerkin-Ritz–method with special ansatz-function.

- **Starting point:** 1. Variational formulation (VP) = $(1)_g$ resp. minimum problem (MP).
  2. Galerkin-Ritz–method.

$$V_h \;=\; \{v_h = \sum_{i\in\bar{\omega}_h} v^{(i)}p^{(i)}(x)\} \subset \underset{\uparrow\;\text{basic space}}{V} = W_2^1(\Omega) \quad \text{(for scalar second-order PDE)},$$

$$\text{(1)}\quad V_{gh} \;=\; \underset{\substack{\pitchfork \\ V_g\cap V_h}}{g_h} + V_{0h} = \{v_h = \underbrace{\sum_{i\in\gamma_h} u_*^{(i)}p^{(i)}(x)}_{\sum_{i\in\gamma_h}|_{\Gamma_1}=g_1} + \underset{p^{(i)}(x)|_{\Gamma_1}=0,\, i\in\omega_h}{\sum_{i\in\omega_h} v^{(i)}p^{(i)}(x)}\} \subset \underset{\nwarrow\;\text{hyperspace}}{V_g} \subset V,$$

$$\text{(1)}\quad V_{0h} \;=\; V_h\cap V_0 = \{v_h = \sum_{i\in\omega_h} v^{(i)}p^{(i)}(x)\} \subset \underset{\nwarrow\;\text{infinite dimensional subspace}}{V_0} \subset V.$$

$$\text{(4)}\qquad = \operatorname{span}\{p^{(i)} : i\in\omega_h\}$$

Variational formulation		Minimum problem	
$(1)_{\mathrm{VF}}$ $u \in V_g : a(u,v) = < F, v > \;\; \forall v \in V_0$	$\Longleftrightarrow$	Find $u \in V_g : J(u) = \inf\limits_{v \in V_g} J(v)$	$(1)_{\mathrm{MP}}$
$V_{gh} \subset V_g$ $V_{0h} \subset V_0$	$a(\cdot,\cdot)$ spd	$V_{gh} \subset V_g$	
Galerkin		Ritz	
$(1)_h$ $u_h \in V_{gh} : a(u_h, v_h) = < F, v_h > \;\; \forall v_h \in V_{0h}$	$\Longleftrightarrow$	Find $u_h \in V_{gh} : J(u_h) = \min\limits_{v_h \in V_{gh}} J(v_h)$	
$v_h = p^{(k)}$ $k \in \omega_h$	$u_h = \sum\limits_{i \in \omega_h} \mathbf{u}^{(\mathbf{i})} p^{(i)} + \sum\limits_{i \in \gamma_n} u_*^{(i)} p^{(i)}$	$\frac{\partial J(\underline{u}_h)}{\partial u^{(k)}} = 0, k \in \omega_h$	

Galerkin-Ritz-System = finite dimensional system

2 Find $\underline{u}_h = [\mathbf{u}^{(\mathbf{i})}]_{i \in \omega_h} \in I\!R^{N_h}$: $\sum\limits_{i \in \omega_h} \mathbf{u}^{(\mathbf{i})} a(p^{(i)}, p^{(k)}) = < F, p^{(k)} > = \sum\limits_{i \in \gamma_h} u_*^{(i)} a(p^{(i)}, p^{(k)}), k \in \omega_h$

3 $(\underline{1})_h$ $\qquad K_h \underline{u}_h = \underline{f}_h$

- **Principle difficulties of the classical** ($\rightarrow$ **ansatz functions with global supports**) **Galerkin-Method:**

  1. Construction of $V_{0h} \subset V_0$ and $V_{gh} \subset V_g$,
  2. Assembling of the Galerkin-Ritz–systems $(\underline{1})_h$,
  3. Storing and solving $(\underline{1})_h$,
  4. Limiting completeness of the family $\{V_{0h}\}_{h \in \Theta}$ in V_0.

  Example with the Hilbert matrix: Find $u \in V = L_2(0,1) : \int\limits_0^1 uv\,dx = \int\limits_0^1 fv\,dx \;\; \forall v \in V$

  with $V_h = \text{span}\,\{x^i : i = 0, 1, \ldots, n-1\}$:

  $\Longrightarrow K_h = \left[\frac{1}{i+j+1}\right]_{i,j=\overline{0,n-1}}$

  = Hilbert matrix extremely bad conditioned !

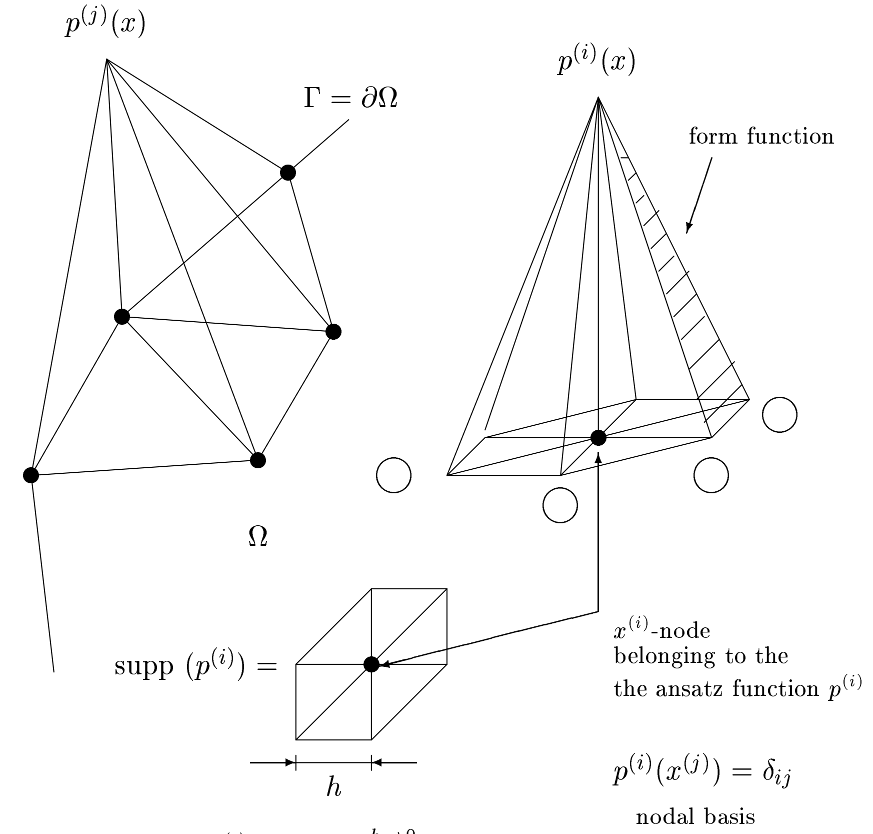
■ **Idea:** The basic idea for overcoming the principle difficulties of the classical Galerkin-method (i.e. the use of ansatz and test functions with global supports, e.g. polynomials) goes back to Richard COURANT (1943) [8], at least, in a strong mathematical sense:

"Variational Methods for the Solution of Problems
of Equilibrium and Vibrations"
Bull. Amer. Math. Soc., 49 (1943), 1-23.

> "If the variational problems contain derivatives not higher than the first order the method of finite difference can be subordinated to the Rayleigh-Ritz method by considering in the competition only functions ϕ which are linear in the meshes of a sub-division of our net into triangles formed by diagonals of the squares of the net. For such polyhedral functions the integrals become sums expressed by the finite number of values of ϕ in the net-points and the minimum conditions become our difference equations. **Such an interpretation suggests a wide generalization which provides great flexibility and seems to have considerable practical value. Instead of starting with a quadratic or rectangular net we may consider from the outset any polyhedral surfaces with edges over an arbitrarily chosen (preferably triangular) net.** Our integrals again become finite sums, and the minimum condition will be equations for the values of ϕ in the net-points. While these equations seem less simple than the original difference equations, we gain the enormous advantage of better adaptability to the data of the problem and thus we can often obtain good results with very little numerical calculation."

and was recovered by the engineers (c.f., e.g. [1] J.H. Argyris, 1955 ff; [41] M. Turner, R.W. Clough, H.C. Martin, L.J. Topp 1956 ff; u.a. ...;) in the mid-fifties (see also historical review by I. Babuška [3], [24]):

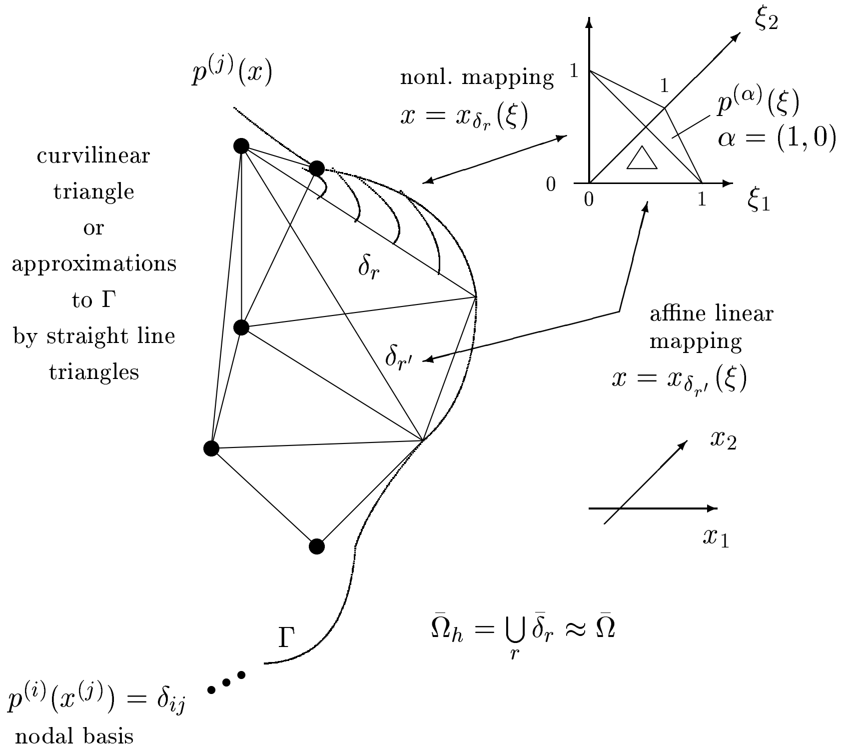
$\Rightarrow$ Use of ansatz and test functions $p^{(i)} = p^{(i)}(x)$ with local supports, where the $p^{(i)}$ can be defined elementwise by form functions:



$$\text{diam supp } (p^{(i)}) = O(h) \stackrel{h\to 0}{\longrightarrow} 0$$

$$\text{meas supp } (p^{(i)}) = O(h^2) \longrightarrow 0$$

$$\Longrightarrow \quad \boxed{u_h(x) \;=\; \sum_{i\in\bar\omega_h} \underset{u_h(x^{(i)})}{\underset{\|}{u^{(i)}}}\, p^{(i)}(x) \;\in\; W_2^1(\Omega)\cap C^0(\Omega) \quad (C^0\text{–elements !})}$$



- **This idea yields:**

1. great flexibility in satisfying the essential BC: ⤵ 1
2. sparseness of the system matrix (= stiffness matrix) because of the locality of support of the ansatz and test functions, i.e.

$$a(p^{(i)}, p^{(k)}) = 0, \text{ if } \operatorname{supp}(p^{(i)}) \cap \operatorname{supp}(p^{(k)}) = \emptyset.$$

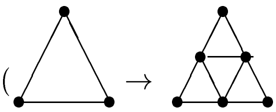
   ⤵ 2 3
3. an <u>elementwise</u> numerical integration technique for calculating the coefficients of the stiffness matrix K_h and the load vector $\underline{f}_h$: ⤵ 2

$$\bar{\Omega} \overset{=}{\approx} \bar{\Omega}_h = \bigcup_r \bar{\delta}_r \quad ⤵$$

$$\int_{\Omega_h} (\ldots)\, dx = \sum_r \int_{\delta_r} (\ldots)\, dx = \sum_r \int_{\Delta} (\ldots)|J_r|\, d\xi \approx \ldots\ .$$

$$\nearrow \qquad\qquad \nwarrow \qquad\qquad \nwarrow$$

element of f.e. mesh $\underset{x = x_{\delta_r}(\xi)}{\overset{\xi=\xi_{\delta_r}(x)}{\rightleftarrows}}$ basis element   quadrature formula (QF)

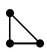
4. "limiting completeness" of the family of f.e. spaces V_{0h} ( !): ⤵ 4

- **Notations:**

$\mathcal{F}_\Delta \quad \equiv \quad \mathcal{F}(\Delta) := \{ \sum\limits_{\alpha \in A} v^{(\alpha)} p^{(\alpha)}(\xi) : \xi \in \bar{\Delta} \}$ - space spanned by the form functions;

$\mathcal{P}_k \quad := \quad \{ \sum\limits_{|\alpha| \le k} c_\alpha \xi^\alpha \}$ - space of polynomials of the degree k;

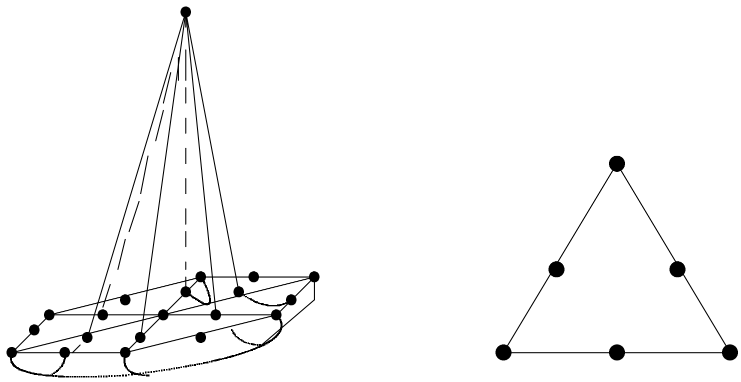
$Q_k \quad := \quad \{ \sum\limits_{|\alpha_i| \le k} c_\alpha \xi^\alpha \}$ - space of polynomials of the degree k in every variable;

$\Delta =$  (Courant's element): $\mathcal{F}_\Delta = \mathcal{P}_1 \subset Q_1$.

■ **Remark 3.10:** $\rightarrow$ Generalizations to (see also Courant !):

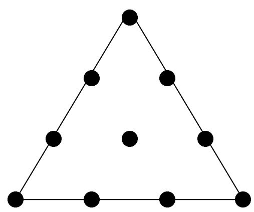
1. higher-order ansatz-functions on triangular elements:
   (a) quadratic element:

   $\mathcal{F}_\Delta = \mathcal{P}_2 = \{a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 xy + a_5 y^2\}$



   (b) cubic element:

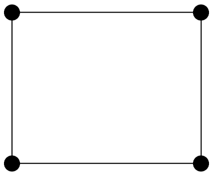
   $\mathcal{F}_\Delta = \mathcal{P}_3 = \{a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 xy + a_5 y^2 + a_6 x^3 + a_7 x^2 y + a_8 x^2 y + a_9 y^3\}$:



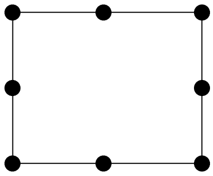
   (c) in general: Lagrangian p^{th} order triangular elements:

   $\mathcal{F}_\Delta = \mathcal{P}_p \quad \Rightarrow \quad \boxed{C^0\text{–elements}}$

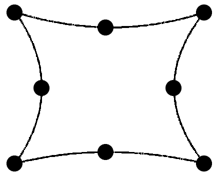
2. other elements, e.g. quadrilateral elements (C^0–elements):



bilinear element
$\mathcal{F}_\Delta = Q_1$



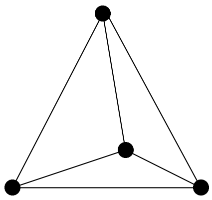
Serendipity–element
(2^{nd} order)
$\mathcal{F}_\Delta \subset Q_2$



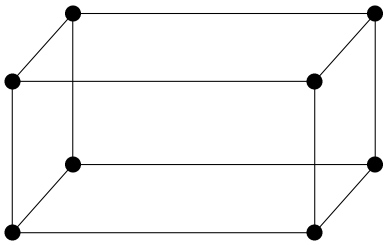
isoparametric Serendipity–element of 2^{nd} order
$x = x_{\delta_r}(\xi) = \sum_{\alpha \in A} x^{(i_\alpha)} p^{(\alpha)}(\xi)$

i.g.: $\mathcal{F}_\Delta = Q_p$ $\quad$ $\mathcal{F}_\Delta \subset Q_p$

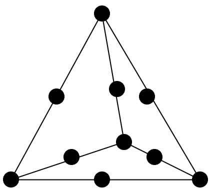
3. 3d–elements (C^0–elements):



linear tetrahedrial element
$\mathcal{F}_\Delta = \mathcal{P}_1$

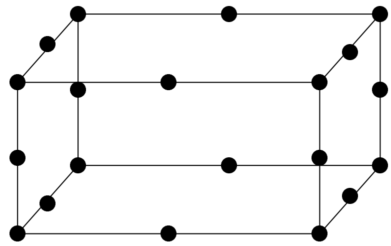


trilinear hexaedrial element
HK 24 Hexaedron
$\mathcal{F}_\Delta = Q_1$



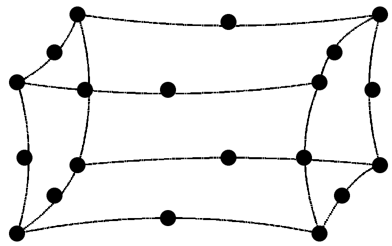
quadratic
Tetrahedron

$\mathcal{F}_\Delta = \mathcal{P}_2$



quadratic Serendipity–element
HK 60 : $\mathcal{F}_\Delta \subset Q_2$

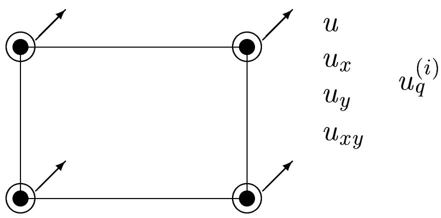
$\downarrow\uparrow \quad x = x_{\delta_r}(\xi) = \sum\limits_{\alpha \in A} x^{(i_\alpha)} p^{(\alpha)}(\xi)$



second-order isoparametric
Serendipity–element

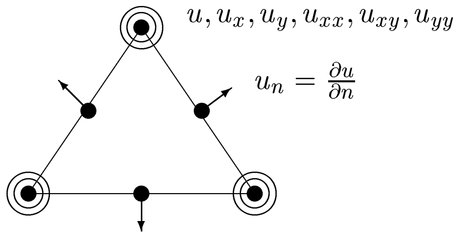
4. Elements of higher smoothness, e.g. C^1–elements for 4^{th} order PDEs:

   (a) Hermite–element



$\mathcal{F}_\Delta = Q_3$

   (b) Argyris-Ženyžek–Element



$\mathcal{F}_\Delta = \mathcal{P}_5$

3.4 Example: BVP for a Second-Order Elliptic PDE

■ Consider the following **BVP in a formal classical formulation**:

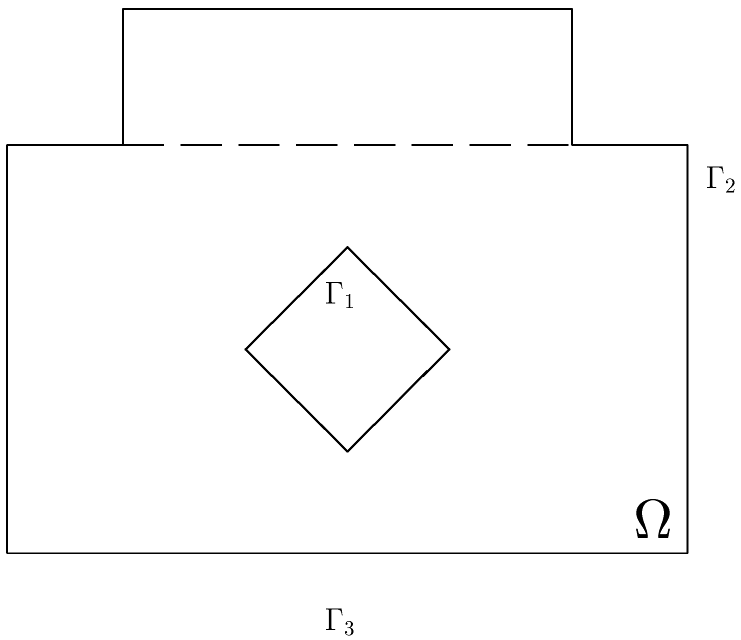
$(1)_{\mathrm{KF}}$ Find some function $u(x_1, x_2)$ satisfying the PDE

$$-\frac{\partial}{\partial x_1}\left(\lambda(x)\frac{\partial u}{\partial x_1}\right) - \frac{\partial}{\partial x_2}\left(\lambda(x)\frac{\partial u}{\partial x_2}\right) = f(x),\ x = (x_1, x_2) \in \Omega,$$

and the boundary conditions (BC)

$$\begin{aligned} u(x) &= g_1(x),\ x \in \Gamma_1\ (1^{\text{st}}\text{ kind} = \text{Dirichlet} = \text{essential BC}) \\ \frac{\partial u}{\partial N} &:= \sum_{i=1}^{2} \lambda_i(x)\frac{\partial u(x)}{\partial x_i} n_i(x) = g_2(x),\ x \in \Gamma_2\ (2^{\text{nd}}\text{ kind} = \text{Neumann}), \\ \frac{\partial u}{\partial N} &= \alpha(x)(g_3(x) - u(x)),\ x \in \Gamma_3\ (3^{\text{rd}}\text{ kind} = \text{Robin}), \end{aligned}$$

on the boundary $\Gamma = \partial\Omega = \Gamma_1 \cap \bar{\Gamma}_2 \cup \bar{\Gamma}_3$ of some given, bounded domain $\Omega \subset I\!R^2$:



$\vec{n} = (n_1, n_2)^T,\ n_i = \cos \sphericalangle (\vec{n}, x_i),\ |\vec{n}| = 1.$

- The BVP (1) models, e.g., plane **heat conduction problems** (see also Chapter 2): λ – heat conduction coefficient, α – heat transfer coefficient, f – heat intensity fuction, g_1 – given temperature on Γ_1, g_2 – given heat flux on Γ_2, g_3 – given exterior boundary temperature to Γ_3.

- **Steps for the formal derivation of the variational formulation:**

1. Choose the space of test functions: $V_0 = \{v \in V = H^1(\Omega) = W_2^1(\Omega) : v|_{\Gamma_1} = 0\}$.
2. Multiply the PDE $(1)_{\mathrm{KF}}$ by test fuctions $v \in V_0$, and integrate over Ω:
$$-\int\limits_\Omega \frac{\partial}{\partial x_1}\left(\lambda(x)\frac{\partial u}{\partial x_1}\right)\cdot v\,dx - \int\limits_\Omega \frac{\partial}{\partial x_2}\left(\lambda(x)\frac{\partial u}{\partial x_2}\right)\cdot v\,dx = \int\limits_\Omega fv\,dx, \quad \forall v \in V_0.$$
2. 2.
3. Partial integration in the principal term:
$$\int\limits_\Omega \lambda\frac{\partial u}{\partial x_1}\frac{\partial v}{\partial x_1}\,dx - \underbrace{\int\limits_\Gamma \lambda\frac{\partial u}{\partial x_1}n_1 v\,ds}_{} + \int\limits_\Omega \lambda\frac{\partial u}{\partial x_2}\frac{\partial v}{\partial x_2}\,dx - \underbrace{\int\limits_\Gamma \lambda\frac{\partial u}{\partial x_2}n_2 v\,ds}_{} = \int\limits_\Omega fv\,dx, \quad \forall v \in V_0,$$
d.h.
$$\int\limits_\Omega \lambda\left(\frac{\partial u}{\partial x_1}\frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2}\frac{\partial v}{\partial x_2}\right)dx - \int\limits_{\Gamma=\Gamma_1\cup\Gamma_2\cup\Gamma_3} \frac{\partial u}{\partial N}v\,ds = \int\limits_\Omega fv\,dx \quad \forall v \in V_0$$
4. Incorporate the natural boundary conditions on Γ_2 and Γ_3:
$$\int\limits_\Gamma \frac{\partial u}{\partial N}v\,ds = \underbrace{\int\limits_{\Gamma_1} \frac{\partial u}{\partial N}\underset{=0}{v}\,ds}_{=0} + \int\limits_{\Gamma_2} \underbrace{\frac{\partial u}{\partial N}}_{=g_2}v\,ds + \int\limits_{\Gamma_3} \underbrace{\frac{\partial u}{\partial N}}_{=\alpha(g_3-u)}v\,ds =$$
$$= \int\limits_{\Gamma_2} g_2 v\,dx + \int\limits_{\Gamma_3} \alpha g_3 v\,ds - \int\limits_{\Gamma_3} \alpha uv\,ds$$
5. Define the linear manifold (hyperplane) V_g of admissible functions = the set, where the solution u is searched for (= incorporation of the essential boundary condition):
$$V_g = \{v \in V : v|_{\Gamma_1} = 0\}$$

- **Result:** = Variational Formulation:

$(1)_{\mathrm{VF}}$ Find $u \in V_g : a(u,v) \ = < F, v > \quad \forall v \in V_0$

$$\underbrace{\int\limits_\Omega \lambda\left(\frac{\partial u}{\partial x_1}\frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2}\frac{\partial v}{\partial x_2}\right)dx + \int\limits_{\Gamma_3} \alpha uv\,ds}_{\text{bilinear form}} = \underbrace{\int\limits_\Omega fv\,dx + \int\limits_{\Gamma_2} g_2 v\,ds + \int\limits_{\Gamma_3} \alpha g_3 v\,ds}_{\text{linear form}}$$

■ **Assumptions:**

1. $\lambda \in L_\infty(\Omega) : \exists\ \underline{\lambda}, \bar{\lambda} = \text{const.} > 0 :\ \underline{\lambda} \leq \lambda(x) \leq \bar{\lambda} \quad \forall$ a.e. $x \in \Omega$;
2. $\alpha \in L_\infty(\Gamma_3) : \exists\ \underline{\alpha}, \bar{\alpha} = \text{const.} > 0 :\ \underline{\alpha} \leq \alpha(x) \leq \bar{\alpha} \quad \forall$ a.e. $x \in \Gamma_3$;
3. $f \in L_2(\Omega)$;
4. $g_2 \in L_2(\Gamma_2),\ g_3 \in L_2(\Gamma_3)$;
5. $g_1 \in H^{\frac{1}{2}}(\Gamma_1)$, d.h. $\exists\ \tilde{g}_1 \in H^1(\Omega) : \tilde{g}_1|_{\Gamma_1} = g_1$;
6. $\Omega \subset \mathbb{R}^2\ \dagger : \Gamma = \partial\Omega \in C^{0,1}$ (Lipschitz-continuous boundary);
7. $\text{meas}_{\mathbb{R}^{n-1}}(\Gamma_1) > 0$.

■ **Existence and uniqueness:** Theorem 2.2 (Lax & Milgram):

- We first homogenize the variational problem $(1)_{\text{VF}}$:
  Ansatz: $u = \tilde{g}_1 + w$ with unknown $w \in V_0$ and given $\tilde{g}_1 \in H^1(\Omega)$ (Ass. 5)):
  $(1)_{\text{VF}} \to (2)_0 \qquad$ Find $w \in V_0 : a(w, v) \ = \ < \hat{F}, v > \quad \forall v \in V_0$
  with $< \hat{F}, v > = < F, v > - a(\tilde{g}_1, v)$.
- We now check the assumptions of Lax-Milgram's Theorem 2.1:
  1. $\hat{F} \in V_0^*$:
     * Obviously, $< \hat{F}, \cdot >$ is a linear form on V_0.
     * $\hat{F}$ is bounded (continuous), because

     $$\begin{aligned}(a)\ | < F, v > | &\leq \|f\|_{0,\Omega}\|v\|_{0,\Omega} + \|g_2\|_{0,\Gamma_2}\|v\|_{0,\Gamma_2} + \|\alpha\|_{\infty,\Gamma_3}\|g_3\|_{0,\Gamma_3}\|v\|_{0,\Gamma_3} \\ &\leq (\|f\|_{0,\Omega} + c_1\|g_2\|_{0,\Gamma_2} + c_1\|\alpha\|_{\infty,\Gamma_3}\|g_3\|_{0,\Gamma_3})\ \|v\|_{1,\Omega} \\ &\uparrow\end{aligned}$$

     $$\|v\|_{0,\Omega} \leq \|v\|_{1,\Omega}$$
     $$\|v\|_{0,\Gamma_i} \leq \|v\|_{0,\Gamma} \leq c_1\|v\|_{1,\Omega} \ \text{(embedding: } H^1(\Omega) \hookrightarrow L_2(\Gamma)\ !)$$

     i.e. $\|F\|_* := \|F\|_{V_0^*} \leq \|f\|_{0,\Omega} + c_1(\|g_2\|_{\sigma,\Gamma_2} + \bar{\alpha}\|g_3\|_{0,\Gamma_3})$,

     $$\begin{aligned}(b)\ |a(\tilde{g}_1, v)| &\leq \bar{\lambda}|\tilde{g}_1|_{1,\Omega}|v|_{1,\Omega} + \bar{\alpha}\|\tilde{g}_1\|_{0,\Gamma_3}\|v\|_{0,\Gamma_3} \\ &\leq (\bar{\lambda}|\tilde{g}_1|_{1,\Omega} + c_1\bar{\alpha}\|\tilde{g}_1\|_{0,\Gamma_3})\|v\|_{1,\Omega},\end{aligned}$$

     (a) + (b) $\Rightarrow \|\tilde{F}\|_* \leq \|f\|_{0,\Omega} + c_1(\|g_2\|_{0,\Gamma_2} + \bar{\alpha}\|g_3\|_{a\Gamma_3}) + \bar{\lambda}|\tilde{g}_1|_{1,\Omega} + c_1\bar{\alpha}\|\tilde{g}_1\|_{0,\Gamma_3}$.

  2. $a(\cdot, \cdot) : V_0 \times V_0 \to \mathbb{R}^1$ – bilinear form (trivial):

2a) V_0–elliptic, i.e.. $\exists\, \mu_1 = \text{const.} > 0 : \mu_1 \|v\|^2_{1,\Omega} \leq a(v,v) \quad \forall v \in V_0$.
Indeed,
$$a(v,u) \geq \underline{\lambda} |v|^2_{1,\Omega} \underset{\uparrow}{\geq} \underline{\lambda}(\bar{c}^2+1)^{-1}\|v\|^2_{1,\Omega},$$
Friedrichs' inequality $\|v\|_{0,\Omega} \leq \bar{c}|v|_{1,\Omega}$, thus, $\|v\|^2_{1,\Omega} \leq (\bar{c}^2+1)|v|^2_{1,\Omega}$.
i.e. $\mu_1 = \underline{\lambda}(\bar{c}^2+1)^{-1}$.

2b) V_0–bounded, i.e. $\exists\, \mu_2 = \text{const.} > 0 : |a(u,v)| \leq \mu_2 \|u\|_{1,\Omega}\, \|v\|_{1,\Omega} \;\; \forall u,v \in V_0$.
Indeed,
$$\begin{aligned} |a(u,v)| &\leq \bar{\lambda}|u|_{1,\Omega}|v|_{1,\Omega} + \bar{\alpha}|u|_{0,\Gamma_3}\,|v|_{0,\Gamma_3} \leq \\ &\leq \bar{\lambda}\|u\|_{1,\Omega}\|v\|_{1,\Omega} + \bar{\alpha}c_1\|u\|_{1,\Omega}\|v\|_{1,\Omega} = (\bar{\lambda}+\bar{\alpha}c_1)\|u\|_{1,\Omega}\|v\|_{1,\Omega}, \end{aligned}$$
i.e. $\mu_2 = \bar{\lambda} + \alpha c_1$.

- $\exists\,!\; w \in V_0 : a(w,v) = < \hat{F}, v > \quad \forall v \in V_0$
  $\Rightarrow \exists\,!\; u \in V_g : a(u,v) = < F, v > \quad \forall v \in V_0$ #

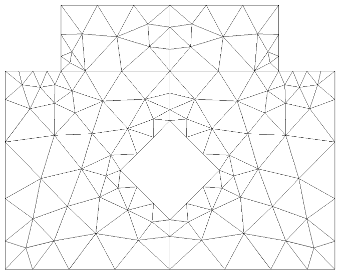
Remark:

1. The proof of Theorem 2.1 is based on Banach's fixed point iteration (2.2) converging linearly to the solution of the variational problem (1). Rewrite this iteration process for the example considered in this subsection !
2. Open analytic questions: (see [27])
   $\rightarrow$ regularity and structur of the solution, e.g. $u \in V_g \cap H^{1+s}(\Omega)$ with some $s > 0$?

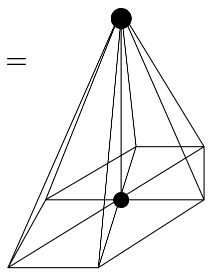
Galerkin-FEM-Discretization:

$$V_h = \text{span}\left\{p^{(i)}(x) : i \in \bar{\omega}_h\right\} \subset V$$
$$V_{0h} = \text{span}\left\{p^{(i)}(x) : i \in \omega_h\right\} \subset V_0$$
$$V_{gh} = \{v_h \underbrace{\sum_{i\in\gamma_h} u_*^{(i)} p^{(i)}(x)}_{g_{1h}|_{\Gamma_1} = g_1} + \sum_{i\in\omega_h} v^{(i)} p^{(i)}(x)\} \subset V_g$$

E.g.: piecewise linear ansatz functions $p^{(i)}$ on the model domain: $p^{(i)}(x^{(j)}) = \delta_{ij}$:



$\omega_h = \{\bullet\} = \{1, 2, \dots, 112\},\ N_h = 112,$
$\gamma_h = \{*\} = \{113, \cdots, 124\}$
$u_*^{(i)} = g_1(x^{(i)}) = 500$

$p^{(i_*)}(x) =$ 

$$(1)_h \quad \text{Find } u_h \in V_{gh} : a(u_h, v_h) = < F, v_h > \quad \forall v_h \in V_{0h}$$
$$\updownarrow$$
$$(\underline{1})_h \quad \text{Find } \underline{u}_h = \{u^{(i)}\}_{i\in\omega_h} \in I\!R^{N_h}: \ \sum_{i\in\omega_h} u^{(i)} a(p^{(i)}, p^{(k)}) = \\ = < F, p^{(k)} > - \sum_{i\in\gamma_h} u_*^{(i)} a(p^{(i)}, p^{(k)}) \ \ \forall k \in \omega_h$$

$$K_h \underline{u}_h = \underline{f}_h$$

- **Result:**
  1. Theorem 3.1: $\Rightarrow \exists!$
  2. Corollary 3.2 and Theorem 3.4: $\Rightarrow$ Iteration for solving $K_h \underline{u}_h = \underline{f}_h$.

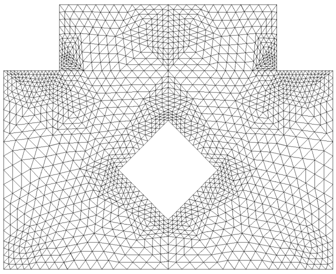
- **Questions:**
  1. Efficient techniques for generating K_h and $\underline{f}_h$ (see e.g. [2], [27]) !
  2. Properties of K_h, e.g. condition number $\kappa(K_h) := \frac{\lambda_{\max}(K_h)}{\lambda_{\min}(K_h)} = O(h^{-2})$ etc. (see, e.g., [27]) !
  3. Preconditioning C_h (see Chapters 4 – 6, and literature [2], [27]) !

4. A-priori discretization error estimates [7], [11], [27]:
   (a) $\|u - u_h\|_{1,\Omega} \leq \frac{\mu_2}{\mu_1} \inf\limits_{v_h \in V_{gh}} \|u - v_h\|_{1,\Omega} \leq \ldots$? (Bramble-Hilbert)
   (b) $\|u - u_n\|_{0,\Omega} \leq$? (Nitsche-Trick)
   (c) $\|u - u_n\|_{\infty,\Omega} \leq$?
5. A-posteriori discretization error estimates: [42].
6. Other discretization techniques such like FDM, BEM, ...: [11], [27].

$\vdots$

- **Results of finite element calculations:**

1. Finer mesh



2. Isolines of the temperature for the following data:

   λ (material I) $= 0.01 \left[\frac{W}{\mathrm{cm}\,K}\right]$ (silicon)

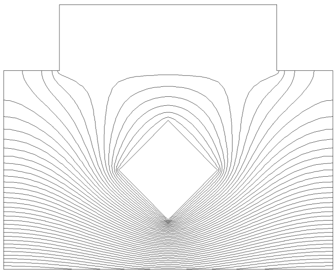
   λ (material II) $= 3.95 \left[\frac{W}{\mathrm{cm}\,K}\right]$ (copper)

   $f = 0$ (no interior heat sources)

   $\Gamma_1 : g_1 = \text{const.} = 500\,K$ (boundary temperature)

   $\Gamma_2 : g_2 = 0$ (isolation)

   Γ_3: $\frac{\partial u}{\partial N} := \lambda$ (material I) $\frac{\partial u}{\partial \vec{n}} \equiv -\lambda(\cdot)\frac{\partial u}{\partial y} = \alpha(g_3 - u)$
   $\alpha = 0.2 \left[\frac{W}{\mathrm{cm}^2 K}\right]$
   $g_3 = 300\,K$ (exterior boundary temperature)



Chapter 4

Domain Decomposition Methods: An Historical Review

4.1
Overlapping DD
(1869)

4.2
Non–overlapping DD
(since the 60ties)

Parallel DD
(since the 80ties)

4.1 The Overlapping Domain Decomposition Method of Herman Schwarz

- **In 1869, H. Schwarz investigated the existence ($\exists$) of harmonic functions in domains with non–smooth boundaries [35]:**

$$(1) \quad \begin{cases} -\Delta u(x) = 0 & , x \in \Omega : \\ u(x) = g(x) & , x \in \partial\Omega \\ \text{(Laplace – Equation)} & \end{cases} \qquad \bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2$$

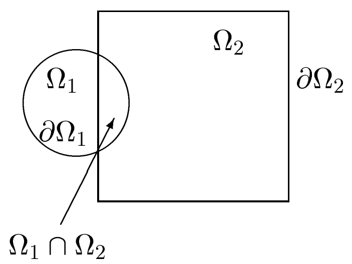
complicated one $\exists$ unknown!

simple ones $\exists$ known! e.g. Fourier–Method

$\exists$

recurrently !

Example from the paper of Schwarz:



■ **Let us make the "Ansatz":**

$$\begin{array}{ccccc} & & \text{known} & & \text{unknown} \\ & & C^2(\Omega)\cap C(\bar{\Omega}) & & C^2(\Omega)\cap C(\bar{\Omega}) \\ u(x) & = & g(x) & + & w(x): \end{array}$$

$$\begin{array}{l} (1) \longrightarrow (1)_0 \\ \text{Poisson–Eq.} \end{array} \left\{ \begin{array}{l} \text{Find } w \in C^2(\Omega)\cap C(\bar{\Omega}) : \\ -\Delta w(x) = f(x) := \Delta g(x),\ x \in \Omega \\ \qquad w(x) = 0 \qquad\qquad\qquad ,\ x \in \Gamma = \partial\Omega \end{array} \right.$$

■ **Thus without loss of generality, we consider the Poisson equation under homogeneous Dirichlet boundary conditions:**

$$(1)_0 \qquad \left\{ \begin{array}{l} \text{Find } u \in C^2(\Omega)\cap C(\bar{\Omega}) : \\ -\Delta u(x) = f(x),\ x \in \Omega, \\ u(x) = 0,\ x \in \Gamma, \\ \text{with given } f \in C(\Omega). \end{array} \right.$$

- **In 1869, H. Schwarz proposed the following <u>Alternating Schwarz Algorithm</u> called now <u>Multiplicative Schwarz Method (MSM) [35]:</u>**

$(2)_{CF}$

0. $u^0 \in C^2(\Omega) \cap C(\bar{\Omega}) : u^0 \,|_{\Gamma} = 0$ given :
   <u>Iteration</u> : $n = 0, 1, \ldots$
1. $\tilde{u}^{n+1/2} \in C^2(\Omega_1) \cap C(\bar{\Omega}_1) : -\Delta \tilde{u}^{n+1/2}(x) = f(x), x \in \Omega_1$
   $\tilde{u}^{n+1/2}(x) = u^n(x), x \in \partial\Omega_1$
   $$u^{n+1/2}(x) = \begin{cases} \tilde{u}^{n+1/2}(x), & x \in \bar{\Omega}_1 \\ u^n(x) & , x \in \bar{\Omega}_2 \setminus \Omega_1 \end{cases}$$
2. $\tilde{u}^{n+1} \in C^2(\Omega_2) \cap C(\bar{\Omega}_2) : -\Delta \tilde{u}^{n+1}(x) = f(x), x \in \Omega_2$
   $\tilde{u}^{n+1}(x) = u^{n+1/2}(x), x \in \partial\Omega_2$
   $$u^{n+1}(x) = \begin{cases} \tilde{u}^{n+1}(x) & , x \in \bar{\Omega}_2, \\ u^{n+1/2}(x), & x \in \bar{\Omega}_1 \setminus \Omega_2. \end{cases}$$

<u>Convergence analysis:</u> via maximum principle!

- **In 1936, S.L. Sobolev gave the <u>variational formulation</u> ($\Leftrightarrow$ minimum problem) of $(\mathbf{2})_{\mathrm{CF}}$ [39] :**

- $u^{n+1/2} = u^n + w^{n+1/2} \quad : \quad w^{n+1/2} = 0$ on $\bar{\Omega} \setminus \Omega_1$
- $\to (2)_{\mathrm{CF}}$:

$$-\Delta w^{n+1/2} = f + \Delta u^n \leftarrow v \in \boldsymbol{V}_1 = \overset{\circ}{H}{}^1(\Omega_1) \subset \overset{\circ}{H}{}^1(\Omega) =: \boldsymbol{V}$$

$$-\int_{\Omega_1} \Delta w^{n+1/2} \cdot v \, dx = \int_{\Omega_1} fv \, dx + \int_{\Omega_1} \Delta u^n \cdot v \, dx \quad \forall v \in \boldsymbol{V}_1$$

$$\int_{\Omega_1} \nabla^T w^{n+1/2} \nabla v \, dx - \int_{\partial\Omega_1} \frac{\partial w^{n+1/2}}{\partial \vec{n}} \cdot \underbrace{v}_{0} \, ds = \int_{\Omega_1} fv \, dx -$$

$$- \int_{\Omega_1} \nabla^T u^n \cdot \nabla v \, dx + \int_{\partial\Omega_1} \frac{\partial u^n}{\partial \vec{n}} \underbrace{v}_{0} \, ds \quad \forall v \in \boldsymbol{V}_1$$

$$\int_{\Omega} \nabla^T w^{n+1/2} \nabla v \, dx = \int_{\Omega} fv \, dx - \int_{\Omega} \nabla^T u^n \nabla v \, dx \quad \forall v \in \boldsymbol{V}_1$$

$$\boxed{\overset{!!}{a(w^{n+1/2}, v)} = \overset{!!}{< F, v >} - \overset{!!}{a(u^n, v)} \quad \forall v \in \boldsymbol{V}_1,}$$

with $\quad a(u, v) := \int_{\Omega} \nabla^T u \nabla v \, dx, \quad < F, v > := \int_{\Omega} fv \, dx.$

$a(u-u^n,v)$

- $w^{n+1/2} \in \boldsymbol{V}_1 : \quad a(w^{n+1/2},v) = \boxed{\begin{array}{c} <F,v> \\ \| \\ a(u,v) \end{array}} - a(u^n,v) \quad \forall v \in \boldsymbol{V}_1$

  variational formul. of $(1)_0$

  $w^{n+1/2} = P_1 \underbrace{(u-u^n)}_{z^n}$ – orthoprojection of the error z^n on $\boldsymbol{V}_1 \subset \boldsymbol{V}$

  $J(u^n + w^{n+1/2}) = \inf\limits_{w \in \boldsymbol{V}_1} J(u^n + w),$

  with $J(v) := \frac{1}{2} a(v,v) - <F,v>$ – energy functional.

- Applying the same technique to the 2nd equation of $(2)_{\mathrm{CF}}$, we arrive at the variational formulation $(2)_{\mathrm{VF}}$ of $(2)_{\mathrm{CF}}$:

$(2)_{VF}$

0. $u^0 \in \boldsymbol{V} = \overset{\circ}{H}{}^1(\Omega) = W_2^1(\Omega)$ given.
   Iteration: $n = 0, 1, \ldots$
1. $u^{n+1/2} = u^n + \boldsymbol{w}^{n+1/2}$ with $\boldsymbol{w}^{n+1/2} \in \boldsymbol{V}_1 = \overset{\circ}{H}{}^1(\Omega_1) \subset \overset{\circ}{H}{}^1(\Omega)$:
   $$\begin{array}{ccccc} a(\boldsymbol{w}^{n+1/2},v) & = & <F,v> & - & a(u^n,v) \quad \forall v \in \boldsymbol{V}_1 \\ \| & & \| & & \| \\ \int\limits_{\Omega_1} \nabla^T w^{n+1/2} \nabla v \, dx & = & \int\limits_{\Omega_1} f v \, dx & - & \int\limits_{\Omega_1} \nabla^T u^n \nabla v \, dx \end{array}$$
   $\boldsymbol{w}^{n+1/2} = \boldsymbol{P}_1(\boldsymbol{u} - \boldsymbol{u}^n)$
2. $u^{n+1} = u^{n+1/2} + \boldsymbol{w}^{n+1}$ with $\boldsymbol{w}^{n+1} \in \boldsymbol{V}_2 = \overset{\circ}{H}{}^1(\Omega_2) \subset \overset{\circ}{H}{}^1(\Omega)$:
   $$\begin{array}{ccccc} a(\boldsymbol{w}^{n+1},v) & = & <F,v> & - & a(u^{n+1/2},v) \quad \forall v \in \boldsymbol{V}_2 \\ \| & & \| & & \| \\ \int\limits_{\Omega_2} \nabla^T w^{n+1} \nabla v \, dx & = & \int\limits_{\Omega_2} f v \, dx & - & \int\limits_{\Omega_2} \nabla^T u^{n+1/2} \nabla v \, dx \end{array}$$
   $\boldsymbol{w}^{n+1} = \boldsymbol{P}_2(\boldsymbol{u} - \boldsymbol{u}^{n+1/2})$

where $P_i : \boldsymbol{V} \to \boldsymbol{V}_i$ – orthoprojection w.r.t. the energy inner product $a(\cdot,\cdot)$:
$a(P_i u, v) = a(u,v) \quad \forall v \in \boldsymbol{V}_i \quad \forall u \in \boldsymbol{V}$.

- Convergence analysis: 1) $J(u^n) := \frac{1}{2} a(u^n,u^n) - <F,u^n>$ decreasing and bounded from below.
  2) $J(u^{n+1} - u^n) \underset{n \to \infty}{\longrightarrow} 0$

■ **Remarks:** to the MSM in variational formulation:

1. Consider the error $z^{n+1} = u - u^{n+1}$:

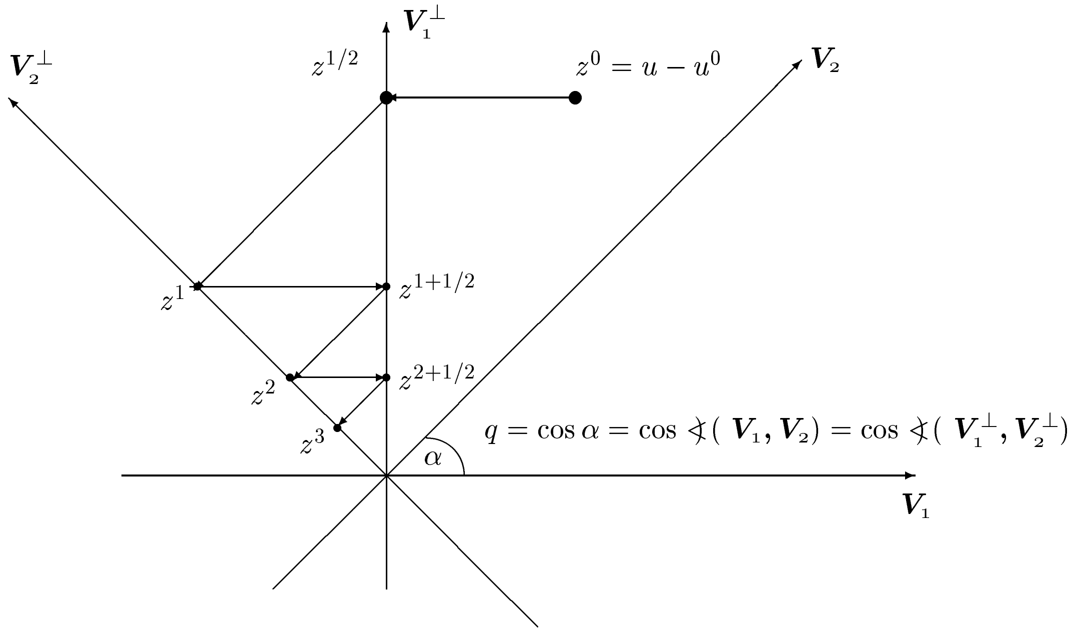
$$\begin{aligned} z^{n+1} &= u - u^{n+1} = u - u^{n+1/2} - P_2 z^{n+1/2} = (I - P_2) z^{n+1/2} = \\ &= (I - P_2)(u - u^{n+1/2}) = (I - P_2)(\underbrace{u - u^n}_{= z^n} - P_1 z^n) \\ &= \underbrace{(I - P_2)}_{= Q_2} \underbrace{(I - P_1)}_{= Q_1} z^n = Q_2 Q_1 z^n \end{aligned}$$

$||z^{n+1}|| \ \leq \ ||Q_2\, Q_1||\, ||z^n||$, with $||\cdot||^2 = a(\cdot,\cdot)$.

Interpretation: $\boldsymbol{V} = \boldsymbol{V}_1 + \boldsymbol{V}_2$: $\underline{\boldsymbol{V}_1 \cap \boldsymbol{V}_2 = \{\boldsymbol{O}\}\ !}$

! This property is not valid for the overlapping DD of H. Schwarz!

$P_i : \boldsymbol{V} \longmapsto \boldsymbol{V}_i$– orthoprojection w.r.t. $a(\cdot,\cdot)$
$Q_i = I - P_i : \boldsymbol{V} \mapsto \boldsymbol{V}_i^{\perp}$



$$\begin{array}{ll} & ||z^1|| = q\,||z^{1/2}|| \le q\,||z^0||, \\ \Longrightarrow & ||z^{1+1/2}|| = q\,||z^1|| \le q^2\,||z^0||, \\ & ||z^2|| = q\,||z^{1+1/2}|| \le q^3\,||z^0||, \ldots, \end{array}$$

$$\Longrightarrow \boxed{||z^n|| \le q^{2n-1}\,||z^0||}$$

2. 2 subdomain $\longmapsto$ p subdomains

   $\bar\Omega = \bar\Omega_1 \cup \bar\Omega_2$ $\boxed{\text{no problem}}$ $\bar\Omega = \bar\Omega_1 \cup \bar\Omega_2 \cup \ldots \cup \bar\Omega_p$

   $\boldsymbol{V} = \boldsymbol{V}_1 + \boldsymbol{V}_2$ $\qquad$ $\boldsymbol{V} = \boldsymbol{V}_1 + \boldsymbol{V}_2 + \ldots + \boldsymbol{V}_p$

3. Variational Formulation $\longmapsto$ Finite Element Model

   $\boldsymbol{V} = \boldsymbol{V}_1 + \ldots + \boldsymbol{V}_p$ $\boxed{\text{no problem}}$ $\boldsymbol{V}_h = \boldsymbol{V}_{1h} + \ldots + \boldsymbol{V}_{ph}$

   $\supset$

4. Parallelization of the MSM ??

5. <u>However:</u> [29] A.M. Matsokin & S.V. Nepomnyaschikh (1985),
   [9] M. Dryja & O. Widlund (1987):
   $\longrightarrow$ Additive Schwarz Method (ASM)

M S M $\longrightarrow$ **A S M**

MSM:

$$u - u^{n+1} = (I - P_2)(I - P_1)(u - u^n) = (I - P_1 - P_2 + P_2P_1)(u - u^n)$$

$u^{n+1} = (I - \wp)u^n + \wp\, u$,
with $\wp = P_1 + P_2 - P_2P_1$

ASM:

$$u - u^{n+1} = (I - \tau P)(u - u^n)$$

$u^{n+1} = (I - \tau P)u^n + \tau P u$,
with $P = P_1 + P_2$ and some relaxation parameter $\tau > 0$.

$\Big\downarrow$ generalization to $\boldsymbol{V} = \boldsymbol{V}_1 + \ldots + \boldsymbol{V}_p$ $\Big\downarrow$

MSM:

$u^{n+1} = (I - \wp)u^n + Pu$,
with $\wp = I - (I - P_p)(I - P_{p-1}) : . : (I - P_1)$

$\longrightarrow$ difficult to parallelize!

ASM:

$u^{n+1} = (I - \tau P)u^n + \tau P u$,
with $P = P_1 + P_2 + \ldots + P_p$

$\longrightarrow$ easy to parallelize!

6. <u>Recent and current research:</u>
   – Convergence analysis
   – Applications
   – Realization and Implementation of MSM and ASM on parallel computers.

■ **Exercise 4.1:**

Let

$$(*)\left\{\begin{array}{ll} \boldsymbol{V} & = I\!R^N = \boldsymbol{V}_1 + \ldots + \boldsymbol{V}_N \text{ , with} \\ \boldsymbol{V}_i & = span\left\{ e_i = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} i^{th} \right\}, N_i = dim\ \boldsymbol{V}_i = 1. \end{array}\right.$$

Show that

ASM corresponding to the splitting $(*)$ = Jacobi
MSM corresponding to the splitting $(*)$ = Gauss–Seidel

4.2 The Finite Element Substructuring Technique and the Non-overlapping Domain Decomposition

■ **The FE–Substructuring Technique has been used by the engineers in the finite element analysis of structures since the 60ties:**

$\longrightarrow$ [33] Prezemieniecki J.S.: Matrix Structural Analysis of Substructures. Am.Inst.Aeor.Astro. J., v.1, 1963.

■ **Let us consider the variational formulation of some symmetric, elliptic bvp for a 2nd order PDE:**

(3) Find $u \in \boldsymbol{V} : a(u,v) =< F, v > \quad \forall v \in \boldsymbol{V} \subset H^1(\Omega)$,

e.g. the Dirichlet problem $(1)_0$ for the Poisson equation:

$(1)_{0,VF}$ Find $u \in \overset{\circ}{H}{}^1(\Omega) : \int\limits_\Omega \nabla^T u \nabla v \, dx = \int\limits_\Omega f v \, dx \quad \forall v \in \overset{\circ}{H}{}^1(\Omega)$.

■ **Idea of the FE–Substructuring Technique:**

$$\bar{\Omega} = \bigcup_{i=1}^{p} \bar{\Omega}_i \quad \xleftarrow[\text{triangulation}]{\text{conform}} \quad \bar{\Omega}_i = \bigcup_{r \in I\!R_{h,i}} \bar{\delta}_{i,r} \quad , i = \overline{1,p}$$

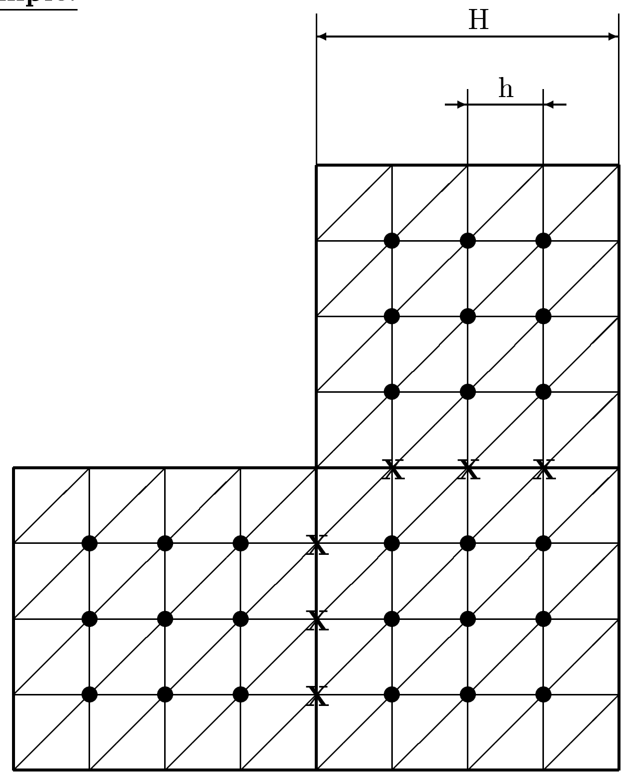
$$\Omega_i \cap \Omega_j = \emptyset, \; i \neq j$$

$$\Gamma_C = \bigcup_{i=1}^{p} \partial\Omega_i \setminus \Gamma_D \supset \quad \{x\} \quad \widehat{=} \text{ "}C\text{"} \quad : \text{coupling boundary ("interface")}$$

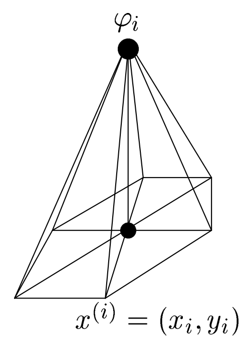
$$\Omega_I = \bigcup_{i=1}^{p} \Omega_i \qquad \supset \quad \{\bullet\} \quad \widehat{=} \text{ "}I\text{"}$$

$$\Gamma_D = \text{Dirichlet boundary } (u|_{\Gamma_D} = \text{given} \overset{\text{e.g.}}{=} 0, \Gamma_D \subset \Gamma)$$

■ **Example:**



$\Gamma_D = \Gamma = \partial\Omega$



Model Problem "L–shaped domain"

- **Consider the standard FE–basis = <u>nodal basis</u>:**

$$\Phi = \Big[\underbrace{\varphi_1,\dots,\varphi_{N_C}}_{\substack{\Gamma_C\\ N_C}};\underbrace{\varphi_{N_C+1},\dots,\varphi_{N_C+N_{I_1}}}_{\substack{\Omega_1\\ N_{I_1}}};\underbrace{\varphi_{N_C+N_{I_1}+1},\dots}_{\substack{\Omega_2\\ N_{I_2}}}\ \dots\ \underbrace{\dots,\varphi_{N=N_C+N_I}}_{\substack{\Omega_p\\ N_{I_p}}}\Big]$$

$$\Omega_I = \Omega_1 \cup \Omega_2 \cup \dots \cup \Omega_p$$
$$N_I = N_{I_1} + N_{I_2} + \dots + N_{I_p}$$

$x \,\widehat{=}\, "C" \qquad \bullet \,\widehat{=}\, "I"$

- <u>**FE–space:**</u> $\boldsymbol{V} = \boldsymbol{V}_h = span\,\Phi = \boldsymbol{V}_C + \boldsymbol{V}_I, \quad \boldsymbol{V}_I = \boldsymbol{V}_{I_1} \oplus \dots \oplus \boldsymbol{V}_{I_2}$

- <u>**FE–Scheme $\overset{\Phi}{\longleftrightarrow}$ FE–Equation:**</u>

$(3)_h$ Find $u_h = \Phi\underline{u} \in \boldsymbol{V} :\ a(\Phi\underline{u}, \Phi\underline{v}) =< F, \Phi\underline{v} > \quad \forall v = \Phi\underline{v} \equiv \sum_{i=1}^{N} v_i\varphi_i \in \boldsymbol{V}$

$\updownarrow \Phi$

$(\underline{3})_h \qquad \underline{u} = [u_i]_{i=\overline{1,N}} \in I\!R^N : \qquad K\underline{u} = \underline{f} \qquad \underline{\underline{(K = K^T p.d.)}}$

$$\begin{bmatrix} K_C & K_{CI} \\ K_{IC} & K_{\overline{I}} \end{bmatrix} \begin{bmatrix} \underline{u}_C \\ \underline{u}_I \end{bmatrix} = \begin{bmatrix} \underline{f}_C \\ \underline{f}_I \end{bmatrix},$$

where $\qquad K_I = diag\,[K_{I_i}]_{i=\overline{1,p}} = \begin{bmatrix} K_{I_1} & & \boldsymbol{O} \\ & K_{I_2} & \\ \boldsymbol{O} & & \ddots \\ & & & K_{I_p} \end{bmatrix}$

- <u>**FE–Substructuring Technique**</u> **= Block Gaussian elimination of $\underline{u}_I$ from the equations, i.e.**

$$\underline{u}_I = K_I^{-1}\,\underline{f}_I - K_I^{-1}\,K_{IC}\,\underline{u}_C$$
$$K_C\underline{u}_C + K_{CI}\,(K_I^{-1}\,\underline{f}_I - K_I^{-1}\,K_{IC}\,\underline{u}_C) = \underline{f}_C$$

$\Longrightarrow$ <u>FE–Substructuring Algorithm:</u>

$$
(4) \qquad \boxed{
\begin{array}{l}
\overset{*}{\underline{u}}_I = K_I^{-1} \underline{f}_I, \text{ i.e. solve } K_I \overset{*}{\underline{u}}_I = \underline{f}_I, \\
\underline{g}_C = \underline{f}_C - K_{CI} \overset{*}{\underline{u}}_I \\
S_C = K_C - K_{CI} K_I^{-1} K_{IC} \quad - \quad \text{FE–Schur–Complement} \\
\underline{u}_C = S_C^{-1} \underline{g}_C \\
\underline{u}_I = \overset{*}{\underline{u}}_I - K_I^{-1} K_{IC} \underline{u}_C
\end{array}
}
$$

<u>operation count</u>

$\Longrightarrow$ (4) = too expensive $\Rightarrow$ 2D: $O(h^{-4})$ instead of $O(N) = O(h^{-2})$
3D: $O(h^{-7})$ instead of $O(N) = O(h^{-3})$

$\Longrightarrow$ <u>However:</u> you get an idea of: 1) parallelization: data partitioning!
2) preconditioning for CG!

Chapter 5

The Parallelization of the Conjugate Gradient (CG) Method via Non–overlapping Domain Decomposition

■ **Let us define the <u>subdomain connectivity matrix</u>**

$$
(1)\qquad A_i \equiv
\begin{array}{cc}
 & \begin{array}{cc} \quad N_C \quad & \quad N_I \quad \end{array} \\
\begin{array}{c} N_{C_i} \\ \\ N_{I_i} \end{array} &
\left[\begin{array}{c|c}
\begin{matrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \end{matrix} & \boldsymbol{O} \\
\hline
\boldsymbol{O} & \begin{matrix} 1 & & & & & & & & & & \\ & 1 & & & & & & & & & \\ & & & & 1 & & & & & & \\ & & & 1 & & & & & & & \\ & & & & & 1 & & & & & \\ & & & & & & 1 & & & & \\ & & & & & & & 1 & & & \end{matrix}
\end{array}\right]_{N_i \times N}
\end{array}
\equiv
\left[\begin{array}{c|c} A_{C_i} & \boldsymbol{O} \\ \hline \boldsymbol{O} & A_{I_i} \end{array}\right]
: I\!R^N \longrightarrow I\!R^{N_i}
$$

$$N_i = N_{C_i} + N_{I_i},\; N = N_C + N_I$$

that maps an overall vector $\underline{u} \in I\!R^N$ into the vector $\underline{u}_i = A_i \underline{u} \in I\!R^{N_i}$ the components of which corresponding to all nodal values in the subdomain Ω_i (see also examples below). The matrices A_i are Boolean matrices.
If

$$
(2)\qquad K_i = \begin{bmatrix} K_{C_i} & K_{C_i I_i} \\ K_{I_i C_i} & K_{I_i} \end{bmatrix}_{N_i \times N_i}
$$

denotes the subdomain stiffness matrix, then the stiffness matrix K can be represented in the form

$$
(3)\qquad K = \sum_{i=1}^{p} A_i^T \, K_i \, A_i
$$

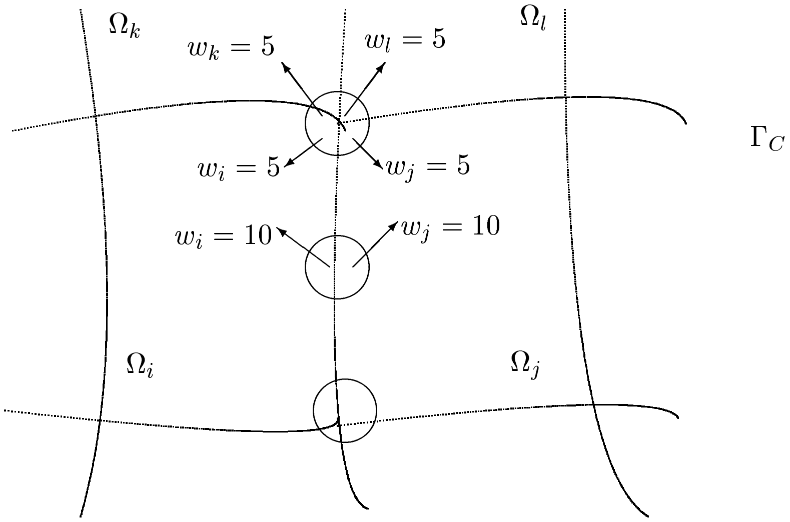
reflecting the subdomain assembling process.

■ **Distribution of the data to the processors $P_i,\ i = \overline{1,p}$:**

We use two types of distributing the data (vectors, matrices) needed in the CG to the processors:

Type I: = overlapping:

$\underline{u}$, $\underline{w}$, $\underline{s}$ are stored in the processor P_i ($\widehat{=}\ \bar{\Omega}_i$)
as $\underline{u}_i = A_i\underline{u}$, $\underline{w}_i = A_i\underline{w}$, $\underline{s}_i = A_i\underline{s} \in \mathbb{R}^{N_i}$
($\Rightarrow$ overlap in the C–components !):

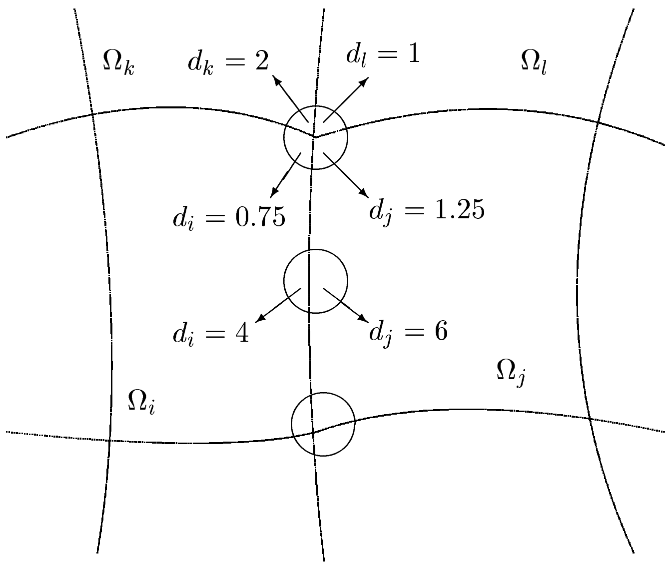


Type II: = adding = assembling:

$\underline{\boldsymbol{d}}$, $\underline{\boldsymbol{x}}$, $\underline{\boldsymbol{f}}$ are stored in P_i ($\widehat{=} \bar{\Omega}_i$)

as $\underline{\boldsymbol{d}}_i$, $\underline{\boldsymbol{x}}_i$, $\underline{\boldsymbol{f}}_i$ such that

$\underline{d} = \sum_{i=1}^{p} A_i^T \, \underline{d}_i, \;\; \underline{x} = \sum_{i=1}^{p} A_i^T \, \underline{x}_i, \;\; \underline{f} = \sum_{i=1}^{p} A_i^T \, \underline{f}_i$



Stiffness matrix: $\Rightarrow$ C–components $\boldsymbol{K}_{C_i}$: $K_C = \sum_{i=1}^{p} A_{C_i}^T \, K_{C_i} \, A_{C_i}$

(P) CG		DD Par (P) CG
serial		FOR ALL $i = 1, p$ DO IN PARALLEL
$\underline{u} = \mathbf{O}$ $\underline{x} = K\,\underline{u}$ $\underline{d} = \underline{f} - \underline{x}$	0. Initial step	Define $\underline{u}_i$, e.g. $\underline{u}_i = \mathbf{O}$ or otherwise $\underline{\boldsymbol{x}}_i = \boldsymbol{K}_i\ \underline{u}_i$ $\underline{\boldsymbol{d}}_i = \underline{\boldsymbol{f}}_i - \underline{\boldsymbol{x}}_i$
$\underline{w} = C^{-1}\,\underline{d}$		$\underline{w}_C = \sum A_{C_i}^T\,\underline{\boldsymbol{d}}_{C_i}$ $\underline{w}_{I_i} = \underline{d}_{I_i}$ $= C^{-1}\begin{bmatrix} \underline{d}_{C_i} \\ \underline{d}_{I_i} \end{bmatrix}$
$\underline{s} = \underline{w}$ $\beta = \underline{w}^T\,\underline{d}$ $\beta\emptyset = \beta$		$\underline{s}_i = \underline{w}_i$ $\beta_i = \underline{w}_i^T\,\underline{\boldsymbol{d}}_i$ $\beta 1 = \sum \beta_i;\ \beta\emptyset = \beta 1$
$\underline{x} = K\,\underline{s}$ $\alpha = \dfrac{\underline{w}^T\,\underline{d}}{\underline{x}^T\,\underline{s}}$	Iteration 1.	$\underline{\boldsymbol{x}}_i = \boldsymbol{K}_i\ \underline{s}_i$ $\alpha_i = \underline{\boldsymbol{x}}_i^T\ \underline{s}_i$ $\alpha 1 = \sum \alpha_i$ $\alpha = \beta 1/\alpha 1$
$\widehat{\underline{u}} = \underline{u} + \alpha\underline{s}$ $\widehat{\underline{d}} = \underline{d} - \alpha\underline{x}$	2.	$\widehat{\underline{u}}_i = \underline{u}_i + \alpha\underline{s}_i$ $\widehat{\underline{\boldsymbol{d}}}_i = \underline{\boldsymbol{d}}_i - \alpha\,\underline{\boldsymbol{x}}_i$
$\widehat{\underline{w}} = C^{-1}\,\widehat{\underline{d}}$	3.	$\widehat{\underline{w}}_C = \sum A_{C_i}^T\,\widehat{\underline{\boldsymbol{d}}}_{C_i}$ $\widehat{\underline{w}}_{I_i} = \widehat{\underline{d}}_{I_i}$ $= C^{-1}\begin{bmatrix} \underline{d}_{C_i} \\ \underline{d}_{I_i} \end{bmatrix}$
$\beta = \dfrac{\widehat{\underline{w}}^T\,\widehat{\underline{d}}}{\underline{w}^T\,\underline{d}}$ $\widehat{\underline{s}} = \widehat{\underline{w}} + \beta\underline{s}$ $\left(\widehat{\underline{w}}, \widehat{\underline{d}}\right) \le \varepsilon^2 * \beta\emptyset$? no yes STOP	4.	$\beta_i = \widehat{\underline{w}}_i^T\,\widehat{\underline{\boldsymbol{d}}}_i$ $\beta 2 = \sum \beta_i$ $\beta = \beta 2/\beta 1;\ \beta 1 = \beta 2$ $\widehat{\underline{s}}_i = \widehat{\underline{w}}_i + \beta\underline{s}_i$ $\beta 2 \le \varepsilon^2 * \beta\emptyset$? no yes STOP

where $\hat{\underline{u}}, \hat{\underline{d}}, \ldots$ are the new iterates, $\sum = \sum\limits_{i=1}^{p}$ means communication.

■ **Convergence:** $||\underline{u} - \underline{u}^n||_K \;\leq\; q_n \, ||\underline{u} - \underline{u}^0||_K$,

$\boxed{\mathrm{C = I}}$ with $|| \cdot ||_K^2 := (\cdot, \cdot)_K = (K\cdot, \cdot)$,

$$q_n = \frac{2\rho^n}{1+\rho^{2n}}, \quad \rho = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \underline{\gamma}/\overline{\gamma},$$

$K\underline{v} = \lambda \underline{v}$

$\underline{\gamma} \leq \lambda_{min} \leq \lambda_{max} \leq \overline{\gamma}$

$$\boxed{\begin{array}{l} \underline{\gamma} I \;\leq\; K \;\leq\; \overline{\gamma} I \\ \underline{\gamma}\,(I\underline{v}, \underline{v}) \;\leq\; (K\underline{v}, \underline{v}) \;\leq\; \overline{\gamma}\,(I\underline{v}, \underline{v}) \quad \forall \underline{v} \in I\!R^N \end{array}}$$

$$\Rightarrow I(\varepsilon) = \text{Number of iterations} = O(\sqrt{\kappa(K)}\;\ln \varepsilon^{-1}) = \underbrace{O(h^{-1} \ln \varepsilon^{-1})}_{\text{2nd order PDE}}$$

relative accuracy: $0 < \varepsilon < 1 : q_{I(\varepsilon)} \;\leq\; \varepsilon;\; \underbrace{\kappa(K)}_{\text{condition number}} = \dfrac{\lambda_{max}}{\lambda_{min}} \;\leq\; \overline{\gamma}/\underline{\gamma}$

■ **Communication:** C = I

1. $\sum$ = "inner product communication"
2. Type I = $\sum A_{C_i}^T$ **Type II** = "**N**earst **N**eighbour communication"

■ **Preconditioner:** $C = C^T\, p.d.$ (spd):

1. $\kappa(C^{-1}K) << \kappa(K)$, possibly $\kappa(C^{-1}K) = O(1)$ for $h \to 0$.
   Spectral equivalence inequalities:
   $\underline{\gamma}\, C \;\leq\; K \;\leq\; \overline{\gamma}\, C \;\Rightarrow\; \kappa(C^{-1}K) \;\leq\; \overline{\gamma}/\underline{\gamma}$,
   $K\underline{v} = \lambda C \underline{v} \;:\; \underline{\gamma} \;\leq\; \lambda_{min} \;\leq\; \lambda_{max} \;\leq\; \overline{\gamma}$
2. $\underline{w} = C^{-1} * \underline{d} \;\Rightarrow$ fast, i.e. $O\,(h^{-m}(\ln h^{-1})^r)$ arithm. operations,
   possibly $r = 0 \Rightarrow$ cost $\approx N$.
3. Communication in 2. $\underline{w} = C^{-1} * \underline{d}$:
   a) Type I = $\sum A_{c_i}^T$ **Type II** – one "**N**earst **N**eighbour communication"
   b) + one "**C**ross–**P**oint communication" (global information transport !)

Chapter 6

Preconditioning via DD

6.1 The Discrete Harmonic Basis and the Simple DD Preconditioner

■ **Introduce the discrete harmonic basis**

$$\overset{*}{\Phi} \equiv \left[\overset{*}{\varphi}_1, \ldots, \overset{*}{\varphi}_{N_C},\ \overset{*}{\varphi}_{N_C+1}, \ldots, \overset{*}{\varphi}_N\right] = \Phi\ \overset{*}{V},$$

with the basis transformation matrix

$$\overset{*}{V} = \left[\overset{*}{V}_C\ \overset{*}{V}_I\right] = \left[\begin{array}{c|c} I_C & \boldsymbol{O} \\ -K_I^{-1} K_{IC} & I_I \end{array}\right],$$

and define the subspaces

$$\overset{*}{\boldsymbol{V}}_C = span\,\Phi\ \overset{*}{V}_C, \quad \overset{*}{\boldsymbol{V}}_I = span\,\Phi\ \overset{*}{V}_I = \boldsymbol{V}_I \subset \boldsymbol{V} \equiv \boldsymbol{V}_h\ .$$

Then we have the orthogonal splitting

$$\boldsymbol{V} = \overset{*}{\boldsymbol{V}}_C \underset{a(\cdot,\cdot)}{\oplus} \overset{*}{\boldsymbol{V}}_I$$

- **FE–Scheme $\overset{\overset{*}{\Phi}}{\longleftrightarrow}$ FE–Equation:**

$(1)_h$ Find $u_h = \underbrace{\Phi\, \overset{*}{V}}\, \overset{*}{\underline{u}} \in \boldsymbol{V} : a(\Phi\, \overset{*}{V}\overset{*}{\underline{u}}, \Phi\, \overset{*}{V}\overset{*}{\underline{v}}) = < F, \Phi\, \overset{*}{V}\overset{*}{\underline{v}} > \quad \forall \Phi\, \overset{*}{V}\overset{*}{\underline{v}} \in \boldsymbol{V}$

$\updownarrow \qquad \overset{*}{\Phi}$

$(\underline{1})_{DH} \qquad \overset{*}{\underline{u}} \in I\!R^N : \boxed{\overset{*}{K}\, \overset{*}{\underline{u}} = \overset{*}{\underline{f}}}$ with $\overset{*}{K} = \overset{*}{V}{}^T K\, \overset{*}{V},\ \overset{*}{\underline{f}} = \overset{*}{V}{}^T \underline{f},\ \underline{u} = \overset{*}{V}\overset{*}{\underline{u}}$

NB DHB

$$\downarrow$$

$$\begin{bmatrix} S_C & \boldsymbol{O} \\ \boldsymbol{O} & K_I \end{bmatrix} \begin{bmatrix} \overset{*}{\underline{u}}_C \\ \overset{*}{\underline{u}}_I \end{bmatrix} = \begin{bmatrix} \underline{f}_C - K_{CI} K_I^{-1} \underline{f}_I \\ \underline{f}_I \end{bmatrix},$$

with the Schur–Complement $S_C = K_C - K_{CI} K_I^{-1} K_{IC}$.

- **From $(\underline{1})_{DH}$, we observe two important things:**

1. The solution of $(\underline{1})_{DH}$ is equivalent to the classical FE–Substructuring Algorithm (4.4) from Subsect. 4.2. !
2. From $(\underline{1})_{DH}$, we obtain the following factorization of K:

(2) $\qquad K = \overset{*}{V}{}^{-T} \overset{*}{K} \overset{*}{V}{}^{-1} =$

$$= \begin{bmatrix} I_C & K_{CI} K_I^{-1} \\ \boldsymbol{O} & I_I \end{bmatrix} \begin{bmatrix} S_C & \boldsymbol{O} \\ \boldsymbol{O} & K_I \end{bmatrix} \begin{bmatrix} I_C & \boldsymbol{O} \\ K_I^{-1} K_{IC} & I_I \end{bmatrix}$$

- **Replacing in (2)**

K_I by some blockpreconditioner C_I (spd) and
S_C by some Schur–Complement–Preconditioner C_C (spd),

then we arrived at the simple DD–preconditioner

(3) $$C = \begin{bmatrix} I_C & K_{CI} C_I^{-1} \\ \boldsymbol{O} & I_I \end{bmatrix} \begin{bmatrix} C_C & \boldsymbol{O} \\ \boldsymbol{O} & C_I \end{bmatrix} \begin{bmatrix} I_C & \boldsymbol{O} \\ C_I^{-1} K_{IC} & I_I \end{bmatrix},$$

leading to the preconditioning equation

CP–Communication NN–Communication

$\underline{\widehat{w}} = C^{-1}\underline{\hat{d}}$	$\underline{\widehat{w}}_C = C_C^{-1} \sum\limits_{i=1}^{p} A_{C_i}^T (\underline{\hat{d}}_{C_i} - K_{C_i I_i} C_{I_i}^{-1} \underline{\hat{d}}_{I_i})$ $\underline{\widehat{w}}_{I_i} = C_{I_i}^{-1} (\underline{\hat{d}}_{I_i} - K_{I_i C_i} \underline{\widehat{w}}_{C_i}),\ \underline{\widehat{w}}_{C_i} = A_{C_i} \underline{w}_C,\ i = \overline{1,p}$

in the DD Par (P) CG of Chapter 5.

In [17] Haase G., Langer U., Meyer A.:
Domain Decomposition Preconditioners with Inexact Subdomain Solvers.
J. Num. Lin. Alg. Appl., 1991, v.1, N.1, 27 – 41,

we could show the spectral equivalence inequalities

$$\underline{\gamma}\, C \;\le\; K \;\le\; \overline{\gamma}\, C, \quad \text{with}$$

$$\underline{\gamma} = min\{\underline{\gamma}_C, \underline{\gamma}_I\}\left(1 + \frac{1}{2}\left(\mu - \sqrt{\mu^2 + 4\mu}\right)\right),$$

$$\overline{\gamma} = max\{\overline{\gamma}_C, \overline{\gamma}_I\}\left(1 + \frac{1}{2}\left(\mu + \sqrt{\mu^2 - 4\mu}\right)\right),$$

where

$$\underline{\gamma}_C\, C_C \;\le\; S_C \;\le\; \overline{\gamma}_C\, C_C,$$

$$\underline{\gamma}_I\, C_I \;\le\; K_I \;\le\; \overline{\gamma}_I\, C_I,$$

$$\mu = \rho\left(S_C^{-1}\, T_C\right) = \sup_{\underline{v}_C \in I\!R^{N_C} \setminus \{\boldsymbol{O}\}} \frac{(T_C\, \underline{v}_C, \underline{v}_C)}{(S_C\, \underline{v}_C, \underline{v}_C)},$$

$$T_C = K_{CI}\left(C_I^{-1} - K_I^{-1}\right) K_I \left(C_I^{-1} - K_I^{-1}\right) K_{IC}.$$

Thus, we get the estimate

$$\kappa(C^{-1}K) \le \frac{max\{\overline{\gamma}_C, \overline{\gamma}_I\}}{min\{\underline{\gamma}_C, \underline{\gamma}_I\}}\left[1 + \frac{1}{2}\left(\mu + \sqrt{\mu^2 + 4\mu}\right)\right]^2.$$

for the relative condition number.

6.2 The Approximate Discrete Harmonic Basis and the ASM – DD – Preconditioner

■ **Introduce the approximate discrete harmonic basis**

$$\widetilde{\Phi} \equiv [\widetilde{\varphi}_1, \ldots, \widetilde{\varphi}_{N_C}, \widetilde{\varphi}_{N_C+1}, \ldots, \widetilde{\varphi}_N] = \Phi\, \widetilde{V},$$

with the basis transformation matrix

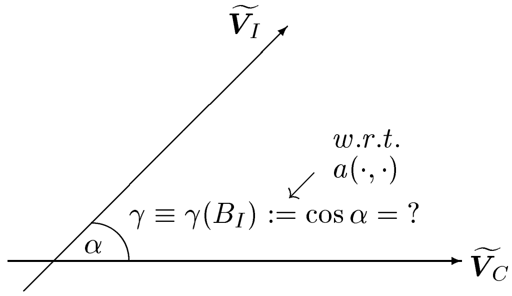
$$\widetilde{V} = \left[\widetilde{V}_C\, \widetilde{V}_I\right] = \left[\begin{array}{c|c} I_C & \boldsymbol{O} \\ -B_I^{-1} K_{IC} & I_I \end{array}\right], \quad B_I^{-1} = diag\,[B_{I_i}]_{i=\overline{1,p}}$$

and define the subspaces

$$\widetilde{\boldsymbol{V}_C} = span\,\Phi\, \widetilde{V}_C, \;\; \widetilde{\boldsymbol{V}_I} = span\,\Phi\, \widetilde{V}_I = \boldsymbol{V}_I \subset \boldsymbol{V} \equiv \boldsymbol{V}_h\ .$$

Then we have the representation

$$\boldsymbol{V} = \widetilde{\boldsymbol{V}}_C + \widetilde{\boldsymbol{V}}_I$$



■ <u>**FE–Scheme $\overset{\widetilde{\boldsymbol{\Phi}}}{\longleftrightarrow}$ FE–Equation:**</u>

$(1)_h$ Find $u_h = \Phi\widetilde{V}\underline{\tilde{u}} \in \boldsymbol{V} : a(\Phi\widetilde{V}\underline{\tilde{u}}, \Phi\widetilde{V}\underline{\tilde{v}}) =< F, \Phi\widetilde{V}\underline{\tilde{v}} > \quad \forall v_h = \Phi\widetilde{V}\underline{\tilde{v}} \in \boldsymbol{V}$

$\updownarrow$

$(\underline{1})_{ADH}$ $\underline{\tilde{u}} \in I\!R^N : \boxed{\tilde{K}\,\underline{\tilde{u}} = \underline{\tilde{f}}}$, with $\tilde{K} = \widetilde{V}^T K \widetilde{V}$, $\underline{\tilde{f}} = \widetilde{V}^T \underline{f}$, $\underbrace{\underline{u}}_{\text{NB}} = \widetilde{V}\underbrace{\underline{\tilde{u}}}_{\text{ADHB}}$

$\downarrow$

$$\boxed{\begin{bmatrix} S_C + T_C & K_{CI} - K_{CI}\,B_I^{-T}\,K_I \\ K_{IC} - K_I\,B_I^{-1}\,K_{IC} & K_I \end{bmatrix} \begin{bmatrix} \underline{\tilde{u}}_C \\ \underline{\tilde{u}}_I \end{bmatrix} = \begin{bmatrix} \underline{f}_C - K_{CI}\,B_I^{-T}\underline{f}_I \\ \underline{f}_I \end{bmatrix}}$$

with $S_C + T_C = K_C - K_{CI}\,K_I^{-1}\,K_{IC} + K_{CI}\,(K_I^{-1} - B_I^{-T})\,K_I\,(K_I^{-1} - B_I^{-1})\,K_{IC}$

■ <u>**Lemma 6.1:**</u>

$$(4) \qquad \boxed{\gamma = \cos \sphericalangle (\widetilde{\boldsymbol{V}}_C, \widetilde{\boldsymbol{V}}_I) = \sqrt{\frac{\mu}{1+\mu}} < 1, \text{ with } \mu = \underbrace{\rho}_{\text{spectral radius}}(S_C^{-1}\,T_C)}$$

<u>Proof</u> see [15] Haase G., Langer U., Meyer A.:
The Approximate Dirichlet Domain Decomposition
Method I. Computing, 1991, v. 47, 137 – 151. ■

From (4), we obtain the strengthened Cauchy inequality

$$|a(u,v)| \leq \gamma\,||u||\,||v|| \quad \forall u \in \widetilde{\boldsymbol{V}}_C \quad \forall v \in \widetilde{\boldsymbol{V}}_I$$

with $||\cdot||^2 = a(\cdot,\cdot)$.

■ **The ASM – DD – Preconditioner:**

Theorem 6.2:

(5)
$$(1-\gamma)\,\widetilde{D} \;\le\; \widetilde{K} \;\le\; (1+\gamma)\,\widetilde{D}, \;\; \text{with } \widetilde{D} = \begin{bmatrix} S_C + T_C & \boldsymbol{O} \\ \boldsymbol{O} & K_I \end{bmatrix}$$

Proof follows immediately from Lemma 6.1. ■

Corollary 6.3:

$$(1-\gamma)\,D \;\le\; K \;\le\; (1+\gamma)\,D, \;\; \text{with } D = \widetilde{V}^{-T}\,\widetilde{D}\,\widetilde{V}^{-1}$$

Theorem 6.4:

If $C_C = C_C^T\, p.d. : \underline{\gamma}_C\, C_C \;\le\; S_C + T_C \;\le\; \overline{\gamma}_C\, C_C$ and
$C_I = C_I^T\, p.d. : \underline{\gamma}_I\, C_I \;\le\; K_I \;\le\; \overline{\gamma}_I\, C_I,$

then

(6) $\underline{\gamma}\, C \;\le\; K \;\le\; \overline{\gamma}\, C,$

with

$$C = \begin{bmatrix} I_C & K_{CI}\, B_I^{-T} \\ \boldsymbol{O} & I_I \end{bmatrix} \begin{bmatrix} C_C & \boldsymbol{O} \\ \boldsymbol{O} & C_I \end{bmatrix} \begin{bmatrix} I_C & \boldsymbol{O} \\ B_I^{-1} K_{IC} & I_I \end{bmatrix},$$

$$\underline{\gamma} = min\{\underline{\gamma}_C, \underline{\gamma}_I\} \left(1 - \sqrt{\mu/(1+\mu)}\right),$$

$$\overline{\gamma} = max\{\overline{\gamma}_C, \overline{\gamma}_I\} \left(1 + \sqrt{\mu/(1+\mu)}\right).$$

■ **The preconditioning equation $\underline{\widehat{w}} = C^{-1}\underline{\widehat{d}}$ takes now the form:**

CP–Communication (in general)

NN–Communication

$$\underline{\widehat{w}}_C = C_C^{-1} \sum_{i=1}^{p} A_{C_i}^T \left(\underline{\hat{d}}_{C_i} - K_{C_i I_i} B_{I_i}^{-T} \underline{\hat{d}}_{I_i} \right),$$

$$\underline{\widehat{w}}_{C_i} = A_{C_i} \underline{\widehat{w}}_C,$$

$$\underline{\widehat{w}}_{I_i} = C_{I_i}^{-1} \underline{\hat{d}}_{I_i} - B_{I_i}^{-1} K_{I_i C_i} \underline{\widehat{w}}_{C_i}$$

$$i = \overline{1,p}$$

■ **The <u>ASM</u> ($\boldsymbol{\tau} = \boldsymbol{\tau}_{opt} = 1$) corresponding to the splitting**

$$\boldsymbol{V} = \widetilde{\boldsymbol{V}}_C + \widetilde{\boldsymbol{V}}_I$$

is nothing else but the Richardson – Iteration with the preconditioners $\widetilde{D}$ ($w.r.t.\, \widetilde{\Phi}$) or D ($w.r.t.\, \Phi$):

$$u^{n+1} = u^n + P_C\, z^n + P_I\, z^n \quad \text{in } \boldsymbol{V} \quad \text{with } z^n = u - u^n,$$

$$\updownarrow$$

$$\widetilde{\Phi}: \qquad \underline{\tilde{u}}^{n+1} = \underline{\tilde{u}}^n - \widetilde{D}^{-1} \tilde{K}\, \underline{\tilde{u}}^n + \widetilde{D}^{-1} \underline{\tilde{f}} \text{ in } \mathbb{R}^N,$$

$$\updownarrow$$

$$\Phi: \qquad \underline{u}^{n+1} = \underline{u}^n - D^{-1} K\, \underline{u}^n + D^{-1} \underline{f} \text{ in } \mathbb{R}^N,$$

where $P_C : \boldsymbol{V} \to \boldsymbol{V}_C$, $P_I : \boldsymbol{V} \to \boldsymbol{V}_I$ – orthoprojections $w.r.t.\, a(\cdot,\cdot)$.
Obviousely, we have the estimate

$$||u - u^{n+1}|| \;\le\; \gamma\, ||u - u^n||, \quad \text{with } ||\cdot||^2 = a(\cdot,\cdot).$$

- **Ex 6.1** Show that the symmetric MSM

$$u - u^{n+1} = (I - P_I)\,(I - P_C)\,(I - P_I)\,(u - u^n)$$

is nothing else but the Richardson – Iteration with some preconditioner $\widetilde{G}\,(w.r.t.\,\widetilde{\Phi})$ or $G\,(w.r.t.\,\Phi)$.
Derive these preconditioners in an explicite form.
These preconditioners and approximations ($\widetilde{P}_C \approx P_C,\ \widetilde{P}_I \approx P_I$) of which are called MSM – DD – Preconditioners.

- **Lemma 6.5:**

Assume that C_C is a spd preconditioner for the Schur-complement $S_C = K_C - K_{CI}K_I^{-1}K_{IC}$, i.e. $\exists\, \underline{\delta}_C, \bar{\delta}_C = \text{const.} > 0$:

(7) $\underline{\delta}_C C_C \le S_C \le \bar{\delta}_C S_C.$

Then the spectral equivalence inequalities

(8) $\underline{\gamma}_C C_C \le S_C + T_C \le \bar{\gamma}_C C_C$

hold with $\underline{\gamma}_C = \underline{\delta}_C$ and $\bar{\gamma}_C = (1+\mu)\,\bar{\delta}_C$.

Proof: is trivial. Indeed, combining inequalities (7) and $S_C \le S_C + T_C \le (1+\mu)\,S_C$ yields (8). ■ **q.e.d.**

Remark:
Lemma 6.5 means that a good Schur-complement–preconditioner S_C is also a good preconditioner for $S_C + T_C$ provided that μ is small !

- **Lemma 6.6:** (estimation of μ)

Assume that there exists some positive constant $c_E = \text{const.} > 0$ (≥ 1 because of Exercise 6.2) such that

(9) $\left\| \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} \right\|_K \le c_E\, \|\underline{v}_C\|_{S_C} \quad \forall\, \underline{v}_C \in I\!R^{N_C},$

where $\|\cdot\|_K^2 := (K\cdot,\cdot)$, $\|\cdot\|_{S_C}^2 := (S_C\cdot,\cdot)$, and E_{IC} replaces $-B_I^{-1}K_{IC}$ in C.
Then the estimate

(10) $\mu := \rho\,(S_C^{-1}T_C) \le c_E^2 - 1$

holds.

Proof: Using the identity

$$\left\| \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} - \begin{bmatrix} \underline{v}_C \\ -K_I^{-1}K_{IC}\underline{v}_C \end{bmatrix} \right\|_K^2 =$$

$$= \left(\begin{bmatrix} K_C & K_{CI} \\ K_{IC} & K_I \end{bmatrix} \left\{ \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} - \begin{bmatrix} \underline{v}_C \\ -K_I^{-1}K_{IC}\underline{v}_C \end{bmatrix} \right\}, \right.$$

$$\left. \left\{ \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} - \begin{bmatrix} \underline{v}_C \\ -K_I^{-1}K_{IC}\underline{v}_C \end{bmatrix} \right\} \right) =$$

$$= \|E_{IC}\underline{v}_C + K_I^{-1}K_{IC}\underline{v}_C\|_{K_I}^2 \quad \forall\, \underline{v}_C \in \mathbb{R}^{N_C},$$

we arrive at the following representation of μ:

$$\text{(11)} \qquad \mu = \max_{\underline{v}_C \in \mathbb{R}^{N_C}\setminus\{\mathbf{O}\}} \frac{(T_C\underline{v}_C, \underline{v}_C)}{(S_C\underline{v}_C, \underline{v}_C)} =$$

$$= \max_{\underline{v}_C} \frac{(K_{CI}\,(K_I^{-1} - B_I^{-T})\,K_I\,(K_I^{-1} - B_I^{-1})\,K_{IC}\underline{v}_C, \underline{v}_C)}{(S_C\underline{v}_C, \underline{v}_C)} =$$

$$= \max_{\underline{v}_C} \frac{\|K_I^{-1}K_{IC}\underline{v}_C - B_I^{-1}K_{IC}\underline{v}_C\|_{K_I}^2}{\|\underline{v}_C\|_{S_C}^2} =$$

$$= \max_{\underline{v}_C} \frac{\|E_{IC}\underline{v}_C + K_I^{-1}K_{IC}\underline{v}_C\|_{K_I}^2}{\|\underline{v}_C\|_{S_C}^2} =$$

$$= \max_{\underline{v}_C} \frac{\left\| \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} - \begin{bmatrix} \underline{v}_C \\ -K_I^{-1}K_{IC}\underline{v}_C \end{bmatrix} \right\|_K^2}{\|\underline{v}_C\|_{S_C}^2}$$

The identities

$$\left(\begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix}, \begin{bmatrix} \underline{v}_C \\ -K_I^{-1}K_{IC}\underline{v}_C \end{bmatrix} \right)_K =$$

$$= \left(\underbrace{\begin{bmatrix} K_C & K_{CI} \\ K_{IC} & K_I \end{bmatrix} \begin{bmatrix} \underline{v}_C \\ -K_I^{-1}K_{IC}\underline{v}_C \end{bmatrix}}, \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} \right) =$$

$$= \begin{bmatrix} (K_C - K_{Ci}K_I^{-1}K_{IC})\,\underline{v}_C \\ (K_{IC} - K_IK_I^{-1}K_{IC})\,\underline{v}_C \end{bmatrix} = \begin{bmatrix} S_C\underline{v}_C \\ \mathbf{O} \end{bmatrix}$$

$$= (S_C\underline{v}_C, \underline{v}_C) = \|\underline{v}_C\|_{S_C}^2$$

and

$$\left\| \begin{bmatrix} \underline{v}_C \\ -K_I^{-1} K_{IC} \underline{v}_C \end{bmatrix} \right\|_K^2 = (S_C \underline{v}_C, \underline{v}_C) = \|\underline{v}_C\|_{S_C}^2$$

together with assumption (29) yield the estimate

$$\begin{aligned} &\left\| \begin{bmatrix} \underline{v}_C \\ E_{IC} \underline{v}_C \end{bmatrix} - \begin{bmatrix} \underline{v}_C \\ -K_I^{-1} K_{IC} \underline{v}_C \end{bmatrix} \right\|_K^2 = \\ &= \left\| \begin{bmatrix} \underline{v}_C \\ E_{IC} \underline{v}_C \end{bmatrix} \right\|_K^2 - 2 \left(\begin{bmatrix} \underline{v}_C \\ E_{IC} \underline{v}_C \end{bmatrix} \begin{bmatrix} \underline{v}_C \\ -K_I^{-1} K_{IC} \underline{v}_C \end{bmatrix} \right)_K + \left\| \begin{bmatrix} \underline{v}_C \\ -K_I^{-1} K_{IC} \underline{v}_C \end{bmatrix} \right\|_K^2 = \\ &= \left\| \begin{bmatrix} \underline{v}_C \\ E_{IC} \underline{v}_C \end{bmatrix} \right\|_K^2 - 2 \|\underline{v}_C\|_{S_C}^2 + \|\underline{v}_C\|_{S_C}^2 \le \\ &\le c_E^2 \|\underline{v}_C\|_{S_C}^2 - \|\underline{v}_C\|_{S_C}^2 = (c_E^2 - 1) \|\underline{v}_C\|_{S_C}^2 \end{aligned}$$

that completes the proof. **q.e.d.**

■ $\boxed{\textbf{Ex 6.2}}$ Prove the inequality

$$\|\underline{v}_C\|_{S_C} \le \left\| \begin{bmatrix} \underline{v}_C \\ E_{IC} \underline{v}_C \end{bmatrix} \right\|_K \qquad \forall\, \underline{v}_C \in \mathbb{R}^{N_C}.$$

■ **Corollary 6.7:**

<u>Ass.:</u>
1. Let $E_{IC} : \mathbb{R}^{N_C} \to \mathbb{R}^{N_I}$ be a linear mapping ($N_C \times N_I$-matrix) such that inequality (9) holds, i.e. E_{IC} is an energy-preserving extension operator.
2. The block preconditioners C_C and C_I satisfy the spectral equivalence inequalities assumed in Theorem 6.4.

<u>St.:</u> Then the DD-preconditioner

(12) $$C = \begin{bmatrix} I_C & -E_{CI} \\ \mathbf{O} & I_I \end{bmatrix} \begin{bmatrix} C_C & \mathbf{O} \\ \mathbf{O} & C_I \end{bmatrix} \begin{bmatrix} I_C & \mathbf{O} \\ -E_{IC} & I_I \end{bmatrix}$$

is spectrally equivalent to K and the spectral equivalence inequalities (6) of Theorem 6.4 are valid with μ replaced by $c_E^2 - 1$ and $E_{CI} = E_{IC}^T$.

Chapter 7

DD Preconditioners in Practical Applications

- The **aim of this chapter** consists in choosing the components C_I, C_C and B_I of the ASM-DD–Preconditioner C for real-life applications such that high efficiency can be expected in practical computations on massively parallel computers.

7.1 The Subdomain Preconditioner C_I

- $C_I = C_I^T$ p.d.: $\underline{\gamma}_I C_I \le K_I \le \bar{\gamma}_I C_I$:

  $C_I = \text{diag}\,[C_{I_i}]_{i=\overline{1,p}}$

  $\simeq \qquad \simeq \qquad$ i.e. $\underline{\gamma}_{I_i} C_{I_i} \le K_{I_i} \le \bar{\gamma}_{I_i} C_{I_i}, \;\; i = \overline{1,p}$

  $K_I = \text{diag}\,[K_{I_i}]_{i=\overline{1,p}}$

  $\Rightarrow \underline{\gamma}_I = \min\{\underline{\gamma}_{I_i}\}, \quad \bar{\gamma}_I = \max\{\bar{\gamma}_{I_i}\}$

  $\Rightarrow$ Every good preconditioner C_{I_i} for the FE-Dirichlet-Subdomain-Stiffness–Matrix K_{I_i} ($\approx$ PDE in Ω_i with homogeneous Dirichlet boundary conditions on $\partial\,\Omega_i$) can be used (e.g. MG [19], BPX [6], MFC [31], ...)

- $C_I \equiv C_{I_i} = K_I (I_I - M_I^k)^{-1}$ (Index i will be omitted !):

  - We implicitely define C_I by some stationary, regular, two-level iteration method (e.g. **M**ulti-**G**rid **M**ethod [19, 34]) applied to

    (1) $\qquad K_I\, \underline{w}_I \;=\; \underline{d}_I$

    with the initial guess $\underline{w}_I^0 = \mathbf{O}$:

(2)	$\underline{w}_I := C_I^{-1}\underline{d}_I$	$\underline{w}_I^0 = \mathbf{O}$ $n = 0, 1, \ldots, k-1$ $\underline{w}_I^{n+1} = M_I \underline{w}_I^n + (I_I - M_I) K_I^{-1} \underline{d}_I$ $\underline{w}_I := \underline{w}_I^k$

Here M_I denotes the iteration operator (matrix) of the iteration method used (see, e.g. [19, 34] for the representation of the multigrid iteration operator).

- $\boxed{\textbf{Ex 7.1}}$ Show that C_I^{-1} has the representation

  (3) $$C_I^{-1} = (I_I - M_I^k)\, K_I^{-1}.$$

- **Lemma 7.1:**

Let us assume that the iteration operator satisfies the following two conditions:

1. $M_I = (M_I)^{*K_I}$ is supposed to be self-adjoint w.r.t. the K_I-energy inner product, i.e.

   (4) $$(M_I \underline{u}_I, \underline{v}_I)_{K_I} = (\underline{u}_I, M_I \underline{v}_I)_{K_I} \quad \forall\, \underline{u}_I, \underline{v}_I \in \mathbb{R}^{N_I}.$$

2. $$\begin{aligned} \rho\,(M_I) &= \text{spectral radius of } M_I \stackrel{1.}{=} \\ &= \|M_I\|_{K_I} \le \eta_I < 1. \end{aligned}$$

Then the subdomain preconditioner

(5) $$C_I = K_I\,(I_I - M_I^k)^{-1}$$

is spd, and satisfies the spectral equivalence inequalities

(6) $$\underline{\gamma}_I C_I \le K_I \le \bar{\gamma}_I C_I,$$

with $\underline{\gamma}_I = 1 - \eta_I^k$ and $\bar{\gamma}_I = 1 + \eta_I^k$. Moreover, if

(7) $$(M_I \underline{u}_I, \underline{u}_I)_{K_I} \ge 0 \quad \forall\, \underline{u}_I \in \mathbb{R}^{N_I},$$

then $\bar{\gamma}_I = 1$.

Proof: see [22], [25]. **q.e.d.** ■

7.2 The Basis Transformations B_I or B_{IC}

■ $B_I = \text{diag}\,[B_{I_i}]_{i=\overline{1,p}}$ " $\simeq$ " K_I:

1. regular, but non-necessarily spd !
2. $\mu = \rho(S_C^{-1}T_C(B_I)) \searrow 0$ "small" !
3. the operations
   $B_I^{-T} * \underline{d}_I$ resp. $-K_{CI}B_I^{-T} * \underline{d}_I$ resp. $B_{CI} * \underline{d}_I$ and
   $B_I^{-1} * \underline{v}_I$ resp. $-B_I^{-1}K_{IC}\underline{w}_C$ resp. $B_{IC} * \underline{w}_C$
   should be fast executable, i.e. via $O(N_I)$ or $O(N_I \ln N_I)$ arithmetical operations.

■ Implicite definition by some iteration method (e.g. MGM) with the initial guess $\mathbf{O}$:

- Define B_I by

$$(8) \qquad B_I = K_I(I_I - \bar{M}_I^s)^{-1}, \;\; B_I^T = K_I\left(I_I - \overset{*}{\bar{M}}{}_I^s\right)^{-1},$$

where $\overset{*}{\bar{M}}_I = (\bar{M}_I)^{*K_I}$ is the K_I-energy adjoint iteration operator, i.e.

$$(9) \qquad (\bar{M}_I\underline{u}_I, \underline{v}_I)_{K_I} = \left(\underline{u}_I, \overset{*}{\bar{M}}_I\,\underline{v}_I\right)_{K_I} \qquad \forall\, \underline{u}_I, \underline{v}_I \in \mathbb{R}^{N_I}.$$

- **<u>Lemma 7.2:</u>**

> Assume that
>
> $$(10) \qquad \|\bar{M}_I\|_{K_I} \le \bar{\eta}_I < 1.$$
>
> Then the estimate
>
> $$(11) \qquad \mu := \rho(S_C^{-1}T_C) \le \bar{\eta}^{2s}(\lambda_{\max}(S_C^{-1}K_C) - 1)$$
>
> holds, where $\lambda_{\max}(S_C^{-1}K_C)$ denotes the maximal eigenvalue of the generalized eigenvalue problem $K_C\underline{v}_C = \lambda S_C\underline{v}_C$.

<u>Proof:</u> Taking into account the representation (6.11), we arrive at the the estimates

$$\begin{aligned}
\mu &= \max_{\underline{v}_C \in \mathbb{R}^{N_C}\setminus\{\mathbf{O}\}} \frac{\|K_I^{-1}K_{IC}\underline{v}_C - B_I^{-1}K_{IC}\underline{v}_C\|_{K_I}^2}{\|\underline{v}_C\|_{S_C}^2} = \\
&\overset{(8)}{=} \max_{\underline{v}_C} \frac{\|K_I^{-1}K_{IC}\underline{v}_C - (I_I - \bar{M}_I^s)\,K_I^{-1}K_{IC}\underline{v}_C\|_{K_I}^2}{\|\underline{v}_C\|_{S_C}^2} \le \\
&\le \|\bar{M}_I\|_{K_I}^{2s} \max_{\underline{v}_C} \frac{\|K_I^{-1}K_{IC}\underline{v}_C\|_{K_I}^2}{\|\underline{v}_C\|_{S_C}^2} =
\end{aligned}$$

$$
\begin{aligned}
&= \|\bar{M}_I\|_{K_I}^{2s} \max_{\underline{v}_C} \frac{(K_{CI}K_I^{-1}K_{IC}\underline{v}_C, \underline{v}_C)}{(S_C\underline{v}_C, \underline{v}_C)} = \\
&= \|\bar{M}_I\|_{K_I}^{2} \max_{\underline{v}_C} \frac{(K_C\underline{v}_C, \underline{v}_C) - (S_C\underline{v}_C, \underline{v}_C)}{(S_C\underline{v}_C, \underline{v}_C)} = \\
&= \|\bar{M}_I\|_{K_I}^{2} \left(\max_{\underline{v}_C} \frac{(K_C\underline{v}_C, \underline{v}_C)}{(S_C\underline{v}_C, \underline{v}_C)} - 1 \right) = \\
&= \|\bar{M}_I\|_{K_I}^{2} (\lambda_{\max}(S_C^{-1}K_C) - 1) \leq \bar{\eta}_I^{2s} (\lambda_{\max}(S_C^{-1}K_C) - 1).
\end{aligned}
$$

q.e.d. ■

- **Remark 7.3:**

1. For second-order PDEs, i.e. $V_0 \subset W_2^1(\Omega)$ (scalar case) or $V_0 \subset [W_2^1(\Omega)]^l$ (systems of PDEs), $\lambda_{\max}(S_C^{-1}K_C)$ behaves like h^{-1}, i.e.

   (12) $$\lambda_{\max}(S_C^{-1}K_C) = \frac{1}{\lambda_{\min}(K_C^{-1}S_C)} = O(h^{-1}) \leq c\,h^{-1}.$$

   This behaviour of the maximal eigenvalue can be concluded from the behaviour of the eigenvalue of S_C and K_C [5]:

   (13) $$\lambda_{\min}(S_C) = O(h^{d-1}),\ \ \lambda_{\max}(S_C) = O(h^{d-2}),\ \ \kappa(S_C) = O(h^{-1}),$$

   (14) $$\lambda_{\min}(K_C) = O(h^{d-2}),\ \ \lambda_{\max}(K_C) = O(h^{d-2}),\ \ \kappa(K_C) = O(1).$$

2. The estimates (11) and (12) immediately yield

   (15) $$\mu \leq \bar{\eta}_I^{2s}\,(\tfrac{c}{h} - 1),$$

   i.e. $\mu \leq$ const. $\neq c(h)$ requires

   (16) $$\bar{\eta}_I^{s} \leq \sqrt{h}\ .$$

   Condition (16) means that

   (17) $$s = O(\ln h^{-1}).$$

3. To imagine the consequences of (15) in practice, we consider the model case $s = 1$ (e.g. 1 MG-cycle) and

$$\boxed{\bar{\eta}_I \leq \sqrt{h}}$$

$$\begin{array}{lcc}
 & \boxed{\text{V11}} \quad \text{MG-V-Cycle} & \boxed{\text{V22}} \\
 & \downarrow & \downarrow \\
 & \bar{\eta}_I = \frac{1}{10} \leq \sqrt{\frac{1}{100}} & \bar{\eta}_I = \frac{1}{16} \leq \sqrt{\frac{1}{256}} \\
 & \downarrow & \downarrow \\
 & h = 10^{-2} & h = \frac{1}{256} \\
 & \Downarrow & \Downarrow \\
\text{2D:} & N_i \approx h^{-2} = 10^4 & N_i \approx h^{-2} = 65536 \\
\text{3D:} & N_i \approx h^{-3} = 10^6 & N_i \approx h^{-3} = 16.8 \text{ Mill.}
\end{array}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{unknowns per subdomain } \bar{\Omega}_i \text{ !!}}$$

Here the notation Vkl means 1 V-cycle with k pre-smoothing sweeps and l post-smoothing sweeps (e.g. Gauss-Seidel sweeps) [25].
Obviously, the condition (15) is rather restrictive in 2D, but surprisingly not in 3D !

■ Implicite definition by some iteration method (e.g. MG) with a non-zero initial guess:

- Define the operator

(18) $\quad B_{IC} := \text{"} - B_I^{-1} K_{IC} \text{"}$

implicitely by some stationary, regular, two-level iteration method (e.g. MGM) applied to

(19) $\quad K_I \underline{w}_I = -K_{IC} \underline{\hat{w}}_C$

with the non-zero initial

(20) $\quad \underline{w}_I^0 = E_{IC} \underline{\hat{w}}_C,$

where $E_{IC} : \mathbb{R}^{N_C} \to \mathbb{R}^{N_I}$ is some energy preserving extension operator (see (6.9)):

(21)

$\underline{w}_I := B_{IC} \underline{\hat{w}}_C$	$\underline{w}_I^0 = E_{IC} \underline{\hat{w}}_C$ $n = 0, 1, \ldots, s-1$ $\underline{w}_I^{n+1} = \bar{M}_I \underline{w}_I^n + (I_I - \bar{M}_I)\, K_I^{-1}(-K_{IC} \underline{\hat{w}}_C)$ $\underline{w}_I := \underline{w}_I^n$

Here $\bar{M}_I$ denotes the iteration operator of the iteration method applied.

- Obviously, B_{IC} has the representation

(22) $$B_{IC} = \bar{M}_I^s E_{IC} + (I_I - \bar{M}_I^s)\, K_I^{-1}\,(-K_{IC}).$$

Indeed, (21) yields

$$\begin{aligned}\underline{w}_I^s &= \bar{M}_I \underline{w}_I^{s-1} + (I_I - \bar{M}_I)\, K_I^{-1}\,(-K_{IC}\underline{\hat{w}}_C) = \\ &= \bar{M}_I \left[\bar{M}_I \underline{w}_I^{s-2} + (I_I - \bar{M}_I)\, K_I^{-1}\underline{\hat{v}}_I\right] + (I_I - \bar{M}_I)\, K_I^{-1}\underline{\hat{v}}_I = \\ &= \ldots = \bar{M}_I^s \underline{w}_I^0 + (I_I - \bar{M}_I^s)\, K_I^{-1}\underline{\hat{v}}_I,\end{aligned}$$

with the notation $\underline{\hat{v}}_I = -K_{IC}\underline{\hat{w}}_C$. ■

- **Ex 7.2** Show that the matrix $B_{CI} := B_{IC}^T$ that is transposed to the matrix B_{IC} defined by (22) has the form $B_{CI} = E_{CI}(\bar{M}_I^T)^s + (-K_{CI})\left(I_I - \overset{*}{\bar{M}}{}_I^s\right) K_I^{-1}$, where $E_{CI} = E_{IC}^T$ and $\overset{*}{\bar{M}}_I = (\bar{M}_I)^{*K_I}$. Rewrite the operation $\underline{\hat{w}}_C = B_{CI}\underline{\hat{d}}_I$ as algorithm in analogy to (21) !

- Replacing formally the matrix "$B_I^{-1}K_{IC}$" by "$-B_{IC}$" in our ASM-DD-preconditioner C defined in Theorem 6.4, we arrive at the ASM-DD-preconditioner in the more general form

(23) $$C = \begin{bmatrix} I_C & -B_{CI} \\ \mathbf{O} & I_I \end{bmatrix} \begin{bmatrix} C_C & \mathbf{O} \\ \mathbf{O} & C_I \end{bmatrix} \begin{bmatrix} I_C & \mathbf{O} \\ -B_{IC} & I_I \end{bmatrix}.$$

- **Lemma 7.4:**

Assume that

(24) $$\begin{cases} 1. \quad \|\bar{M}_I\|_{K_I} \le \bar{\eta}_I < 1 \ \text{ and} \\ 2. \quad \left\| \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} \right\|_k \le c_E\, \|\underline{v}_C\|_{S_C} \quad \forall\, \underline{v}_C \in I\!R^{N_C}. \end{cases}$$

Then the estimate

(25) $$\mu = \rho\,(S_C^{-1}T_C) \le \bar{\eta}_I^{2s}(c_E^2 - 1)$$

holds.

Proof: In analogy to the proof of Lemma 6.6, we conclude:

$$\begin{aligned}\mu &= \max_{\underline{v}_C} \frac{\|K_I^{-1}K_{IC}\underline{v}_C + B_{IC}\underline{v}_C\|_{K_I}^2}{\|\underline{v}_C\|_{S_C}^2} = \\ &= \max_{\underline{v}_C} \frac{\|K_I^{-1}K_{IC}\underline{v}_C + \bar{M}_I^s E_{IC}\underline{v}_C - (I_I - \bar{M}_I^s)\, K_I^{-1}K_{IC}\underline{v}_C\|_{K_I}^2}{\|\underline{v}_C\|_{S_C}^2} \le\end{aligned}$$

$$\leq \; \bar{\eta}_I^{2s} \max_{\underline{v}_C} \frac{\|E_{IC}\underline{v}_C + K_I^{-1}K_{IC}\underline{v}_C\|_{K_I}^2}{\|\underline{v}_C\|_{S_C}^2} =$$

$$= \; \bar{\eta}_I^{2s} \max_{\underline{v}_C} \frac{\left\| \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} - \begin{bmatrix} \underline{v}_C \\ -K_I^{-1}K_{IC}\underline{v}_C \end{bmatrix} \right\|_k^2}{\|\underline{v}_C\|_{S_C}^2} \;\leq\; \bar{\eta}_I^{2s}(c_E^2 - 1).$$

q.e.d. ■

■ **Energy preserving extension operators:**

$$E_{IC} : I\!R^{N_C} \to I\!R^{N_I} : \left\| \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} \right\|_k \leq c_E \|\underline{v}_C\|_{S_C} \quad \forall\, \underline{v}_C \in I\!R^{N_C}. \tag{26}$$

- Extension by averaging

  → [30] Nepomnyaschikh (1990).

  → 2D and 3D; it is complicated to realize, however, it doesn't require any mesh hierarchy !

  → $O(N)$ arithmetical operations in parallel.

  → $c_E = O(1)$

  → Motivation:

  $$\left\| \Phi \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} \right\|_{H^1(\Omega)} \leq \tilde{c}_E \left\| \Phi \begin{bmatrix} \underline{v}_C \\ \mathbf{O} \end{bmatrix} \right\|_{H^{1/2}(\Gamma_C)}$$

  ↳ (26) with $c_E = c_E(\tilde{c}_E)$! (Exercise !)

- Hierarchical extension in the 2D case:

  → [18] Haase-Langer-Meyer-Nepomnyaschikh (1994).

  → 2D; it is easy and cheap to realize, but it requires an hierarchical mesh structure !

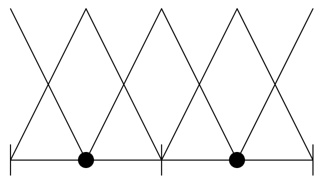
  → $O(N)$ arithmetical operations in parallel.

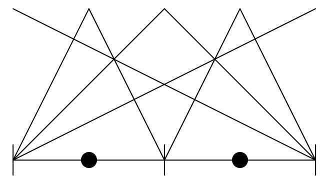
  → $c_E = O(l)$, with $l = O(\ln(H/h))$ = number of levels.

→ Motivation:

(27) $\hat{\Phi} = \Phi H$, with $H = \begin{bmatrix} H_C & \mathbf{O} \\ H_{IC} & H_I \end{bmatrix}$, $H^{-1} = \begin{bmatrix} H_C^{-1} & \mathbf{O} \\ -H_I^{-1} H_{IC} H_C^{-1} & H_I^{-1} \end{bmatrix}$.

nodal basis Φ $\xrightarrow{H}$ hierarchical basis $\hat{\Phi}$





(28) $E_{IC} = H_{IC} H_C^{-1}$.

Indeed,

$$v_h(x) = \Phi \underbrace{\begin{bmatrix} H_C \\ H_{IC} \end{bmatrix} \underline{\hat{v}}_C}_{\underline{v}} = \Phi_C H_C \underline{\hat{v}}_C + \Phi_I H_{IC} \underline{\hat{v}}_C$$

$$\Rightarrow \quad \underline{v}_C = [v_h(x^{(i)})]_{i \in \omega_C} = H_C \underline{\hat{v}}_C, \text{ i.e. } \underline{\hat{v}}_C = H_C^{-1} \underline{v}_C,$$

$$\underline{v}_I = [v_h(x^{(i)})]_{i \in \omega_I} = H_{IC} \underline{\hat{v}}_C = \underbrace{H_{IC} H_C^{-1}}_{=: E_{IC}} \underline{v}_C.$$

- <u>BPX-extension:</u>
  - → [32] Nepomnyaschikh (1995).
  - → 2D and 3D, it is easy to realize, but it requires an hierarchical mesh structure !
  - → $O(N)$ arithmetical operations in parallel.
  - → $c_E = O(1)$.
- <u>Further references:</u> [12]

7.3 The $(S_C + T_C)$–Preconditioner C_C

■ $C_C = C_C^T$ p.d.: $\underline{\gamma}_C C_C \le S_C + T_C \le \bar{\gamma}_C C_C$:

- $$\begin{aligned} S_C + T_C &= K_C - K_{CI}B_I^{-1}K_{IC} - K_{CI}B_I^{-T}K_{IC} + K_{CI}B_I^{-T}K_I B_I^{-1}K_{IC} = \\ &= K_C + K_{CI}B_{IC} + B_{CI}K_{IC} + B_{CI}K_I B_{IC} \end{aligned}$$

  $\Rightarrow$ The operation $(S_C + T_C) * \underline{v}_C$ can be carried out fast provided that the operations $B_I^{-1} * K_{IC}\underline{v}_C$ and $K_{CI}B_I^{-T} * \underline{v}_I$, or $B_{IC} * \underline{v}_C$ and $B_{CI}\underline{v}_I$ are fast !

- Question: How can we exploit this observation ?

■ **Schur-Complement–Preconditioners:**

- Simple Result:
  Every good Schur-Complement–Preconditioner C_C is also a good preconditioner for $S_C + T_C$ provided that μ is small (cf. Lemma 6.5) !
  Indeed, the spectral equivalence inequalities

  (29) $\quad \underline{\delta}_C C_C \le S_C \le \bar{\delta}_C C_C$

  immediately yield the inequalities

  (30) $\quad \underline{\delta}_C C_C \le S_C + T_C \le \bar{\delta}_C\,(1+\mu)\,C_C,$

  i.e. $\underline{\gamma}_C = \underline{\delta}_C$ and $\bar{\gamma}_C = \bar{\delta}_C\,(1+\mu)$.

- BPS-Preconditioner:
  → [5] Bramble–Pasciak–Schatz (1986).
  → $C_C = \delta\,(J_C^T * \mathrm{BPS} * J_C)$,
  where $J_C = (\mathrm{diag}\, K_C)^{-1/2}$, δ-scaling parameter.
  → Result: (2D-case, i.e. $d = 2$; second-order PDE)

(31)

$\underline{\gamma}_C = \delta\underline{c}\,(1 + \ln^2(H/h))^{-1}, \;\; \bar{\gamma}_C = \delta\bar{c}\,(1+\mu);$
$Q\,(C_C^{-1} * \underline{v}_C) = O(h_{\mathrm{loc}}^{-1}\ln h_{\mathrm{loc}}^{-1} + H^{-4}),$
with $h_{\mathrm{loc}} = h/H, \;\; h = h_{\mathrm{glo}}, \;\; H = O(p^{-(1/d)}),$
provided that the cross-point systems are solved directly.

- Hierarchical Schur-Complement–Preconditioner:

  → [14] Haase–Langer–Meyer (1990), see also [16],
  [38] Smith–Widlund (1990).

  → $C_C = \delta H_C^{-T} D_C H_C^{-1}$,
  where δ – scaling parameter,
  H_C – hierarchical basis transformation matrix of dimension $N_C \times N_C$, cf. (27),
  D_C – scaling matrix of dimension $N_C \times N_C$ with cross-point stiffness matrix block K_H.

  $$D_C = \left[\begin{array}{c|c} \backslash & \\ \hline & K_H \end{array}\right]$$

  → Result: (2D case, i.e. $d = 2$; second-order PDE)

(32)
$$\underline{\gamma}_C = \delta \underline{c}(1 + \underbrace{\ln^2(H/h)}_{=(\text{levels})^2})^{-1}), \quad \bar{\gamma}_C = \delta \bar{c}\,(1+\mu),$$
$$Q\,(C_C^{-1} * \underline{v}_C) = O(h_{\text{loc}}^{-1} + H^{-4}),$$
with $h_{\text{loc}} = h/H, \;\; h = h_{\text{glo}}, \;\; H = h_0 = O(p^{-(1/d)})$,
provided that the cross-point systems $K_H \underline{w}_H = \underline{d}_H$ are solved by some direct method.

Let us verify this result for the simplified case $D_C = I_C$ and $\delta = 1$. We show that the spectral inequalities

$$\lambda_{\min}(\hat{K})\, H_C^{-T} H_C^{-1} \le S_C \le \lambda_{\max}(\hat{K})\, H_C^{-T} H_C^{-1}$$

follows directly from Yserentant's spectral estimates [43]

$$\underline{c}\,(1 + \ln^2 h^{-1})^{-1} \le \lambda_{\min}(\hat{K}) \le \bar{c},$$

where $\hat{K} = H^T K H$ denotes the stiffness matrix w.r.t the hierarchical basis $\hat{\Phi} = \Phi H$, cf. also (27).
Indeed, let us denote the Schur-Complement of the hierarchical basis stiffness matrix $\hat{K}$ by

$$\hat{S}_C = \hat{K}_C - \hat{K}_{CI} \hat{K}_I^{-1} \hat{K}_{IC}.$$

It is easy to show that

$$\hat{S}_C = H_C^T S_C H_C,$$

where $S_C = K_C - K_{CI} K_I^{-1} K_{CI}$ denotes again the Schur-Complement of the nodal basis stiffness matrix K.

Estimate now $(\hat{S}_C \underline{u}_C, \underline{u}_C)$ from below (a) $\geq$) and from above (b) $\leq$):

$$
\begin{aligned}
\text{a)}\quad (\hat{S}_C \underline{u}_C, \underline{u}_C) &= \inf_{\underline{u}_I \in \mathbb{R}^{N_I}} (\hat{K}\underline{u}, \underline{u}) \geq \\
&\geq \lambda_{\min}(\hat{K}) \inf_{\underline{u}_I} (\underline{u}, \underline{u}) = \lambda_{\min}(\hat{K})\,(\underline{u}_C, \underline{u}_C), \\
\text{b)}\quad (\hat{S}_C \underline{u}_C, \underline{u}_C) &= (\hat{K}_C \underline{u}_C, \underline{u}_C) - (\hat{K}_{CI}\hat{K}_I\hat{K}_{IC}\underline{u}_C, \underline{u}_C) \leq \\
&\leq (\hat{K}_C \underline{u}_C, \underline{u}_C) \leq \lambda_{\max}(\hat{K}_C)\,(\underline{u}_C, \underline{u}_C) \leq \\
&\leq \lambda_{\max}(\hat{K})\,(\underline{u}_C, \underline{u}_C)
\end{aligned}
$$

for all $\underline{u}_C \in \mathbb{R}^{N_C}$. Therefore, we have proved the spectral inequalities:

$$\lambda_{\min}(\hat{K})\,(\underline{u}_C, \underline{u}_C) \leq (\underbrace{H_C^T S_C H_C}_{=\hat{S}_C}\underbrace{\underline{u}_C}_{\underline{v}_C = H_C \underline{u}_C}, \underline{u}_C) \leq \lambda_{\max}(\hat{K})\,(\underline{u}_C, \underline{u}_C)$$

$$\lambda_{\min}(\hat{K})\,(H_C^{-T}H_C^{-1}\underline{v}_C, \underline{v}_C) \leq (S_C \underline{v}_C, \underline{v}_C) \leq \lambda_{\max}(\hat{K})\,(H_C^{-T}H_C^{-1}\underline{v}_C, \underline{v}_C),$$

i.e. $\underline{\delta}_C = \lambda_{\min}(\hat{K})$ and $\bar{\delta}_C = \lambda_{\max}(\hat{K})$. Together with (29) and (30), these estimates yield (32) for the case $D_C = I_C$ and $\delta = 1$.

- BPX-Schur-Complement–Preconditioner:
  - $\rightarrow$ [40] Tong–Chan–Kao (1991).
  - $\rightarrow$ $C_C^{-1} = \delta^{-1} M_C D_C^{-1} M_C^T$,

    with δ – scaling parameter,

    M_C – multilevel "basis" transformation matrix of dimension $N_C \times \bar{N}_C$ with $\bar{N}_C > N_C$ (see [10]),

    $D_C = \left[\begin{array}{c|c} \backslash & \\ \hline & K_H \end{array}\right]$ – scaling matrix of dimension $\bar{N}_C \times \bar{N}_C$ with cross-point stiffness matrix block K_H

→ Result: (2D and 3D, second-order PDE)

(33)
$$\underline{\gamma}_C = O(1), \quad \bar{\gamma}_C = O(1),$$
$$Q\,(C_C^{-1} * \underline{v}_C) = \begin{cases} O(h_{\text{loc}}^{-1} + H^{-4}), & \text{2D}, \\ O(h_{\text{loc}}^{-2} + H^{-7}), & \text{3D}, \end{cases}$$

with $h_{\text{loc}} = h/H, \quad h = h_{\text{glo}}, \quad H = h_0 = O(p^{-(1/d)})$,
provided that the cross-point systems $K_H \underline{w}_H = \underline{d}_H$ are solved by some direct method.

7.4 The MSM-DD–Preconditioner

■ **Consider the DD–Preconditioner (23)**

$$C = \begin{bmatrix} I_C & -B_{CI} \\ \mathbf{O} & I_I \end{bmatrix} \begin{bmatrix} C_C & \mathbf{O} \\ \mathbf{O} & C_I \end{bmatrix} \begin{bmatrix} I_C & \mathbf{O} \\ -B_{IC} & I_I \end{bmatrix}$$

with the components

$$\begin{cases} C_I &=& K_I \left(I_I - \overset{*}{\tilde{M}}{}_I^k \, \tilde{M}_I^k \right)^{-1}, \quad \overset{*}{\tilde{M}}_I = \left(\tilde{M}_I\right)^{*K_I}; \\ B_{IC} &=& \tilde{M}_I^k \bar{M}_I^s E_{IC} + \left(I_I - \tilde{M}_I^k \bar{M}_I^s \right) K_I^{-1}(-K_{IC}); \\ E_{IC} &:& I\!R^{N_C} \to I\!R^{N_I} - \text{ norm-preserving extension operator: (24) } (\uparrow); \\ B_{CI} &=& B_{IC}^T = ?? \qquad \text{(exercise)}; \\ C_C &=& C_C^T \text{ p.d.: } \underline{\gamma}_C C_C \le S_C + T_C \le \bar{\gamma}_C C_C \quad (\uparrow). \end{cases}$$

The derivation of this MSM-DD–preconditioner and its interpretation as ASM-DD–preconditioner is given in [13], cf. also Exercise 6.1.

■ **Ex 7.3** **Solve the following problems:**

1. Rewrite the preconditioning algorithm

$\hat{\underline{w}} = C^{-1}\hat{\underline{d}}$	(exercise)

in detail, and show that no more than $k\,\tilde{M}_I$ –, $k\,\overset{*}{\tilde{M}}_I$ –, $s\,\bar{M}_I$ – and $s\,\overset{*}{\bar{M}}_I$ – iterations are necessary !

2. Prove that, under the assumptions
   - $\|\tilde{M}_I\|_{K_I} \le \tilde{\eta}_I < 1$,
   - $\|\bar{M}_I\|_{K_I} \le \tilde{\eta}_I < 1$,
   - $\left\| \begin{bmatrix} \underline{v}_C \\ E_{IC}\underline{v}_C \end{bmatrix} \right\|_K \le c_E \, \|\underline{v}_C\|_{S_C} \quad \forall\, \underline{v}_C \in I\!R^{N_C}$ (24)

   the following estimates are valid:
   - $\underline{\gamma}_I = 1 - \bar{\eta}_I^{2k}, \ \ \bar{\gamma}_I = 1$,
   - $\mu = \rho(S_C^{-1}T_C) \le \bar{\eta}_I^{2s}\, \tilde{\eta}_I^{2k}\, (c_E^2 - 1)$.

■ In **practice**, the following components result in an asymptotically optimal preconditioner:

- $k = 1, \ \tilde{M}_I = V20, \ \overset{*}{\tilde{M}_I} = V02$;
- $s = 0$;
- E_{IC} = BPX–extension;
- C_C = BPX–Schur–complement–Preconditioner.

However, in 2D codes, the hierarchical extension should be used instead of the BPX-extension. Although the hierarchical extension leads to a log-grow of c_E, it is cheaper than the BPX-extension in its realization. For μ, we obtain the following estimate:

$$\mu \le \bar{\eta}_I^{2s}\, \tilde{\eta}_I^{2k}(c^2 l^2 - 1), \quad l = \ln(H/h) + 1.$$

In the case of $s = 0$ and $k = 1$, we have

$$\mu \le \tilde{\eta}_I^2(c^2 l^2 - 1) < c^2, \text{ if } \tilde{\eta}_I \le 1/l.$$

Suppose that $\tilde{\eta}_I = 1/4$, then $l = 4$ levels are possible !

Bibliography

[1] J.H. Argyris. Energy theorems and structural analysis. *Aircraft Engng.*, 27:125–154, 1955.

[2] O. Axelsson and V.A. Barker. *Finite Element Solution of Boundary Value Problems.* Academic Press, New York, 1984.

[3] I. Babuška. Courant element: Before and after. In M. Krizek, P. Neittaanmaki, and R. Stenberg, editors, *Finite Element Methods: Fifty years of the Courant Element*, pages 37–51. Marcel Dekker, Inc., New York, Basel, Hong Kong, 1994.

[4] D. Braess. *Finite Elemente.* Springer-Verlag, Berlin - Heidelberg, 1992.

[5] J. H. Bramble, J. E. Pasciak, and A. H. Schatz. The construction of preconditioners for elliptic problems by substructuring. I-IV. *Math. Comput.*, 47, 49, 51, 53(175, 179):103–234, 1–16, 415–430, 1–24, 1986, 1987, 1988, 1989.

[6] J. H. Bramble, J. E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comput.*, 55:1–22, 1990.

[7] Ph. Ciarlet. *The finite element method for elliptic problems.* North-Holland Publishing Company, Amsterdam-New York-Oxford, 1978.

[8] R. Courant. Variational methods for the solution of problems of equilibrium and vibrations. *Bull. Amer. Math.*, 49:1–23, 1943.

[9] M. Dryja and O. B. Widlund. An additive variant of Schwarz alternating method for the case of many subregions. Technical Report 339, Department of Computer Science, Courant Institute, New York, 1987. also Ultracomputer Note 131.

[10] M. Griebel. *Multilevelmethoden als Iterationsverfahren "uber Erzeugendensystemen.* B.G. Teubner, Stuttgart, 1994.

[11] Ch. Grossmann and H.-G. Roos. *Numerik partieller Differentialgleichungen.* B.G. Teubner, Stuttgart, 1992.

[12] G. Haase. Hierarchical extension operators plus smoothing in domain decomposition preconditioners. *Appl. Num. Math.*, 1:111–111, 1996. (to appear).

[13] G. Haase and U. Langer. The non-overlapping domain decomposition multiplicative Schwarz method. *International Journal of Computer Mathemathics*, 44:223–242, 1992.

[14] G. Haase, U. Langer, and A. Meyer. A new approach to the Dirichlet decomposition method. In Sabine Hengst, editor, *Proceedings of the "5-th Multigrid Seminar" held at Eberswalde, GDR, May 14-18, 1990*, pages 1–59, Berlin, 1990. Academy of Science. Report-Nr. R-MATH-09/90.

[15] G. Haase, U. Langer, and A. Meyer. The approximate dirichlet decomposition method. part I : An algebraic approach. *Computing*, 47:137–151, 1991.

[16] G. Haase, U. Langer, and A. Meyer. The approximate dirichlet decomposition method. part II : Application to 2nd-order elliptic B.V.P.s. *Computing*, 47:153–167, 1991.

[17] G. Haase, U. Langer, and A. Meyer. Domain decomposition preconditioners with inexact subdomain solvers. *J. of Num. Lin. Alg. with Appl.*, 1:27–42, 1992.

[18] G. Haase, U. Langer, A. Meyer, and S.V. Nepomnyaschikh. Hierarchical extension operators and local multigrid methods in domain decomposition preconditioners. *East-West Journal of Numerical Mathematics*, 2:173–193, 1994.

[19] W. Hackbusch. *Multi-Grid Methods and Applications.* Springer, Berlin, 1980.

[20] W. Hackbusch. *Theorie und Numerik elliptischer Differentialgleichungen.* B.G. Teubner, Stuttgart, 1986.

[21] W. Hackbusch. *Iterative Lösung grosser schwachbesetzter Gleichungssysteme.* B. G. Teubner, Stuttgart, 1991.

[22] M. Jung, U. Langer, A. Meyer, W. Queck, and M. Schneider. Multigrid preconditioners and their applications. In G. Telschow, editor, *Third Multigrid Seminar, Biesenthal 1988*, pages 11–52, Berlin, 1989. Karl–Weierstrass–Institut. Report R–MATH–03/89.

[23] N. Kikuchi. *Finite Element Method in Mechanics.* Cambridge University Press, Cambridge, 1986.

[24] M. Krizek, P. Neittaanmaki, and R. Stenberg, editors. *Finite Element Methods: Fifty years of the Courant Element*, New York, Basel, Hong Kong, 1994. Marcel Dekker, Inc. Jyvaskyla, September 1993.

[25] U. Langer. *Skriptum zur Vorlesung MULTIGRID – METHODEN.* Johannes Kepler Universität, Institut für Mathematik, Linz, 1996.

[26] U. Langer. *Skriptum zur Vorlesung NUMERIK I (Operatorgleichungen).* Johannes Kepler Universität, Institut für Mathematik, Linz, 1996.

[27] U. Langer. *Skriptum zur Vorlesung NUMERIK II (RWA).* Johannes Kepler Universität, Institut für Mathematik, Linz, 1996.

[28] U. Langer. *Skriptum zur Vorlesung "Numerische Festkörpermechanik".* Johannes Kepler Universität, Institut für Mathematik, Linz, 1997.

[29] A. M. Matsokin and S. V. Nepomnyaschikh. A Schwarz alternating method in a subspace. *Soviet Mathemathics*, 29(10):78–84, 1985.

[30] S. Nepomnyaschikh. Mesh theorems on traces, normalization of function traces and their inversion. *Sov. J. Numer. Anal. Math. Modelling*, 6(3):223–242, 1991.

[31] S. Nepomnyaschikh. Method of splitting into subspaces for solving elliptic boundary value problems in complex-from domains. *Sov. J. Numer. Anal. Math. Modelling*, 6(2):111–111, 1991.

[32] S. Nepomnyaschikh. Optimal multilevel extension operators. Technical Report 3, TU Chemnitz-Zwickau, Faculty for Mathematics, Cemnitz, 1995.

[33] J. S. Przemieniecki. Matrix structural analysis of substructures. *AIAA J.*, 1:138–147, 1963.

[34] A. Reusken. *Iterative methods for elliptic boundary value problems.* Technical University of Einhoven, Einhoven, 1996.

[35] H. A. Schwarz. *Über einige Abbildungsaufgaben. Gesammelte Mathemathische Abhandlungen*, volume 2. Springer, Bonn, 1890. First published in "Vierteljahresschrift der Naturforschenden Gesellschaft in Zürich" 1870, v.15, pp 272-286.

[36] H.R. Schwarz. *Methode der finiten Elemente.* B.G. Teubner, Stuttgart, 1991.

[37] B. Smith, P. Bjorstad, and W. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic PDEs.* Cambridge University Press, Cambridge, 1996.

[38] B. Smith and O. Widlund. A domain decompostion algorithm using a hierarchical basis. *SIAM J. Sci. Stat. Comput.*, 11:1212–1220, 1990.

[39] S. Sobolev. Schwarz's algorithm in the theory of elasticity. *Doklady Akademii Hauk SSSR*, 4(6):235–238, 1936. (in Russian).

[40] C.H. Tong, T.F. Chan, and C.C.J. Kuo. A decomposition preconditioner based on a change to a multilevel nodal basis. *J. Sci. Stat. Comput.*, 12(6):1486–1495, Nov. 1991.

[41] M. Turner, R.W. Clough, H.C. Martin, and L.J. Topp. Stiffness and deflection analysis of complex structures. *J. Aeronaut. Sci.*, 23:805–823, 1956.

[42] R. Verfuerth. *A Review of A-Posteriori Error Estimation and Adaptive Mesh-Refinement.* Wiley & Teubner, Chichester, Stuttgart, 1996.

[43] H. Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math.*, 49(4):379–412, 1986.