# Scientific Computing WS 2019/2020 

## Lecture 18

Jürgen Fuhrmann<br>juergen.fuhrmann@wias-berlin.de

## The Galerkin method II

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

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

- Let $V_{h} \subset V$ be a finite dimensional subspace of $V$
- "Discrete" problem $\equiv$ Galerkin approximation:

Search $u_{h} \in V_{h}$ such that

$$
a\left(u_{h}, v_{h}\right)=f\left(v_{h}\right) \forall v_{h} \in V_{h}
$$

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

## Céa's lemma

- What is the connection between $u$ and $u_{h}$ ?
- Let $v_{h} \in V_{h}$ be arbitrary. Then

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

- As a result

$$
\left\|u-u_{h}\right\| \leq \frac{\gamma}{\alpha} \inf _{v_{h} \in V_{h}}\left\|u-v_{h}\right\|
$$

- Up to a constant, the error of the Galerkin approximation is the error of the best approximation of the solution in the subspace $V_{h}$.


## From the Galerkin method to the matrix eqation

- Let $\phi_{1} \ldots \phi_{n}$ be a set of basis functions of $V_{h}$.
- Then, we have the representation $u_{h}=\sum_{j=1}^{n} u_{j} \phi_{j}$
- In order to search $u_{h} \in V_{h}$ such that

$$
a\left(u_{h}, v_{h}\right)=f\left(v_{h}\right) \forall v_{h} \in V_{h}
$$

it is actually sufficient to require

$$
\begin{aligned}
a\left(u_{h}, \phi_{i}\right) & =f\left(\phi_{i}\right)(i=1 \ldots n) \\
a\left(\sum_{j=1}^{n} u_{j} \phi_{j}, \phi_{i}\right) & =f\left(\phi_{i}\right)(i=1 \ldots n) \\
\sum_{j=1}^{n} a\left(\phi_{j}, \phi_{i}\right) u_{j} & =f\left(\phi_{i}\right)(i=1 \ldots n) \\
A U & =F
\end{aligned}
$$

with $A=\left(a_{i j}\right), a_{i j}=a\left(\phi_{i}, \phi_{j}\right), F=\left(f_{i}\right), f_{i}=F\left(\phi_{i}\right), U=\left(u_{i}\right)$.

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


## Obtaining a finite dimensional subspace

- Let $\Omega=(a, b) \subset \mathbb{R}^{1}$
- Let $a(u, v)=\int_{a}^{b} \delta \vec{\nabla} u \vec{\nabla} v d \vec{x}+\alpha u(a) v(a)+\alpha u(b) v(b)$
- Calculus 101 provides a finite dimensional subspace: the space of $\sin / \cos$ functions up to a certain frequency $\Rightarrow$ spectral method
- Ansatz functions have global support $\Rightarrow$ full $n \times n$ matrix
- OTOH: rather fast convergence for smooth data
- Generalization to higher dimensions possible
- Big problem in irregular domains: we need the eigenfunction basis of some operator...
- Spectral methods are successful in cases where one has regular geometry structures and smooth/constant coefficients - e.g. "Spectral Einstein Code"


## The finite element idea I

- Choose basis functions with local support. $\Rightarrow$ only integrals of basis function pairs with overlapping support contribute to matrix.
- Linear finite elements in $\Omega=(a, b) \subset \mathbb{R}^{1}$ :
- Partition $a=x_{1} \leq x_{2} \leq \cdots \leq x_{n}=b$
- Basis functions (for $i=1 \ldots n$ )

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



## FE matrix elements for 1D heat equation I

- Any function $u_{h} \in V_{h}=\operatorname{span}\left\{\phi_{1} \ldots \phi_{n}\right\}$ is piecewise linear, and the coefficients in the representation $u_{h}=\sum_{i=1}^{n} u_{i} \phi_{i}$ are the values $u_{h}\left(x_{i}\right)$.
- Fortunately, we are working with a weak formulation, and weak derivatives are well defined (and coincide with the classical derivatives where the basis functions are smooth)
- Let $\phi_{i}, \phi_{j}$ be two basis functions, regard

$$
s_{i j}=\int_{a}^{b} \vec{\nabla} \phi_{i} \cdot \vec{\nabla} \phi_{j} d x
$$

- We have supp $\phi_{i} \cap \operatorname{supp} \phi_{j}=\emptyset$ unless $i=j, i+1=j$ or $i-1=j$.
- Therefore $s_{i j}=0$ unless $i=j, i+1=j$ or $i-1=j$.


## FE matrix elements for 1D heat equation II

- Let $j=i+1$. Then supp $\phi_{i} \cap \operatorname{supp} \phi_{j}=\left(x_{i}, x_{i+1}\right), \phi_{i}^{\prime}=-\frac{1}{h}$, $\phi_{j}^{\prime}=f r a c 1 h$ where $h=x_{i+1}-x_{i}$

$$
\int_{a}^{b} \vec{\nabla} \phi_{i} \vec{\nabla} \phi_{j} d \vec{x}=\int_{x_{i}}^{x_{i+1}} \phi_{i}^{\prime} \phi_{j}^{\prime} d \vec{x}=-\int_{x_{i}}^{x_{i+1}} \frac{1}{h^{2}} d \vec{x}=-\frac{1}{h}
$$

- Similarly, for $j=i-1$ : $\int_{a}^{b} \vec{\nabla} \phi_{i} \vec{\nabla} \phi_{j} d \vec{x}=-\frac{1}{h}$
- For $1<i<N$ :

$$
\int_{a}^{b} \vec{\nabla} \phi_{i} \vec{\nabla} \phi_{i} d \vec{x}=\int_{x_{i-1}}^{x_{i+1}}\left(\phi_{i}^{\prime}\right)^{2} d \vec{x}=\int_{x_{i-1}}^{x_{i+1}} \frac{1}{h^{2}} d \vec{x}=\frac{2}{h}
$$

- For $i=1$ or $i=N: \int_{a}^{b} \vec{\nabla} \phi_{i} \vec{\nabla} \phi_{i} d \vec{x}=\frac{1}{h}$
- For the right hand side, calculate vector elements $f_{i}=\int_{a}^{b} f(x) \phi_{i} d x$ using a quadrature rule.


## FE matrix elements for 1D heat equation III

Adding the boundary integrals yields

$$
A=\left(\begin{array}{cccccc}
\alpha+\frac{1}{h} & -\frac{1}{h} & & & & \\
-\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & & \\
& -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\
& \ddots & \ddots & \ddots & \ddots & \\
& & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\
& & & -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} \\
& & & & -\frac{1}{h} & \frac{1}{h}+\alpha
\end{array}\right)
$$

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

## Simplices

- Let $\left\{\vec{a}_{1} \ldots \vec{a}_{d+1}\right\} \subset \mathbb{R}^{d}$ such that the $d$ vectors $\vec{a}_{2}-\vec{a}_{1} \ldots \vec{a}_{d+1}-\vec{a}_{1}$ are linearly independent. Then the convex hull $K$ of $\vec{a}_{1} \ldots \vec{a}_{d+1}$ is called simplex, and $\vec{a}_{1} \ldots \vec{a}_{d+1}$ are called vertices of the simplex.
- Unit simplex: $\vec{a}_{1}=(0 \ldots 0), \vec{a}_{1}=(0,1 \ldots 0) \ldots \vec{a}_{d+1}=(0 \ldots 0,1)$.

$$
K=\left\{\vec{x} \in \mathbb{R}^{d}: x_{i} \geq 0(i=1 \ldots d) \text { and } \sum_{i=1}^{d} x_{i} \leq 1\right\}
$$

- A general simplex can be defined as an image of the unit simplex under some affine transformation
- $F_{i}$ : face of $K$ opposite to $\vec{a}_{i}$
- $\vec{n}_{i}$ : outward normal to $F_{i}$


## Simplex characteristics

- Diameter of $K: h_{K}=\max _{\vec{x}_{1}, \vec{x}_{2} \in K}\left\|\vec{x}_{1}-\vec{x}_{2}\right\|$ $\equiv$ length of longest edge if $K$
- $\rho_{K}$ diameter of largest ball that can be inscribed into $K$
- $\sigma_{K}=\frac{h_{K}}{\rho_{K}}$ : local shape regularity measure
- $\sigma_{K}=2 \sqrt{3}$ for equilateral triangle
- $\sigma_{K} \rightarrow \infty$ if largest angle approaches $\pi$.


## Barycentric coordinates

Definition: Let $K \subset \mathbb{R}^{d}$ be a $d$-simplex given by the points $\vec{a}_{1} \ldots \vec{a}_{d+1}$. Let $\Lambda(x)=\left(\lambda_{1}(\vec{x}) \ldots \lambda_{d+1}(\vec{x})\right)$ be a vector such that for all $\vec{x} \in \mathbb{R}^{d}$

$$
\sum_{j=1}^{d+1} \vec{a}_{j} \lambda_{j}(\vec{x})=\vec{x}, \quad \sum_{j=1}^{d+1} \lambda_{j}(\vec{x})=1
$$

This vector is called the vector of barycentric coordinates of $\vec{x}$ with respect to $K$.

## Barycentric coordinates II

Lemma The barycentric coordinates of a given point is well defined and unique. Moreover, for the simplex edges $\vec{a}_{i}$, one has

$$
\lambda_{j}\left(\vec{a}_{i}\right)=\delta_{i j}
$$

Proof: The definition of $\Lambda$ given by a $d+1 \times d+1$ system of equations with the matrix

$$
M=\left(\begin{array}{cccc}
a_{1,1} & a_{2,1} & \ldots & a_{d+1,1} \\
a_{1,2} & a_{2,2} & \ldots & a_{d+1,2} \\
\vdots & \vdots & & \vdots \\
a_{1, d} & a_{2, d} & \ldots & a_{d+1, d} \\
1 & 1 & \ldots & 1
\end{array}\right)
$$

Subtracting the first column from the others gives

$$
M^{\prime}=\left(\begin{array}{cccc}
a_{1,1} & a_{2,1}-a_{1,1} & \ldots & a_{d+1,1}-a_{1,1} \\
a_{1,2} & a_{2,2}-a_{1,2} & \ldots & a_{d+1,2}-a_{1,2} \\
\vdots & \vdots & & \vdots \\
a_{1, d} & a_{2, d}-a_{1, d} & \ldots & a_{d+1, d}-a_{1, d} \\
1 & 0 & \ldots & 0
\end{array}\right)
$$

## Barycentric coordinates III

$\operatorname{det} M=\operatorname{det} M^{\prime}$ is the determinant of the matrix whose columns are the edge vectors of $K$ which are linearly independent.

For the simplex edges one has

$$
\sum_{j=1}^{d+1} \vec{a}_{j} \lambda_{j}\left(\vec{a}_{i}\right)=\vec{a}_{i}
$$

which is fulfilled if $\lambda_{j}\left(\vec{a}_{i}\right)=1$ for $i=j$ and $\lambda_{j}\left(\vec{a}_{i}\right)=0$ for $i \neq j$. And we have uniqueness.

At the same time, the measure (area) is calculated as $|K|=\frac{1}{d!}\left|\operatorname{det} M^{\prime}\right|$.

## Barycentric coordinates IV

- Let $K_{j}(\vec{x})$ be the subsimplex of $K$ made of $\vec{x}$ and $\vec{a}_{1} \ldots \vec{a}_{d+1}$ with $\vec{a}_{j}$ omitted.
- Its measure $\left|K_{j}(\vec{x})\right|$ is established from its determinant and a linear function of the coordiates for $\vec{x}$.
- One has $\frac{\left|K_{j}\left(\vec{a}_{i}\right)\right|}{|K|}=\delta_{i j}$ and therefore,

$$
\lambda_{j}(\vec{x})=\frac{\left|K_{j}(\vec{x})\right|}{|K|}
$$


is the ratio of the measures of $K_{j}(\vec{x})$ and $K$.

## Conformal triangulations II

- $d=1$ : Each intersection $F=K_{m} \cap K_{n}$ is either empty or a common vertex
- $d=2$ : Each intersection $F=K_{m} \cap K_{n}$ is either empty or a common vertex or a common edge

- $d=3$ : Each intersection $F=K_{m} \cap K_{n}$ is either empty or a common vertex or a common edge or a common face
- Delaunay triangulations are conformal


## Shape regularity

- Now we discuss a family of meshes $\mathcal{T}_{h}$ for $h \rightarrow 0$.
- For given $\mathcal{T}_{h}$, assume that $h=\max _{K \in \mathcal{T}_{h}} h_{K}$
- A family of meshes is called shape regular if

$$
\forall h, \forall K \in \mathcal{T}_{h}, \sigma_{K}=\frac{h_{K}}{\rho_{K}} \leq \sigma_{0}
$$

- $\ln 1 \mathrm{D}, \sigma_{K}=1$
- In 2D, $\sigma_{K} \leq \frac{2}{\sin \theta_{K}}$ where $\theta_{K}$ is the smallest angle


## Polynomial space $\mathbb{P}_{k}$

- Space of polynomials in $x_{1} \ldots x_{d}$ of total degree $\leq k$ with real coefficients $\alpha_{i_{1} \ldots i_{d}}$ :

$$
\mathbb{P}_{k}=\left\{p(x)=\sum_{\substack{0 \leq i_{1} \ldots i_{d} \leq k \\ i_{1}+\cdots+i_{d} \leq k}} \alpha_{i_{1} \ldots i_{d}} x_{1}^{i_{1}} \ldots x_{d}^{i_{d}}\right\}
$$

- Dimension:

$$
\begin{aligned}
\operatorname{dim} \mathbb{P}_{k}=\binom{d+k}{k} & = \begin{cases}k+1, & d=1 \\
\frac{1}{2}(k+1)(k+2), & d=2 \\
\frac{1}{6}(k+1)(k+2)(k+3), & d=3\end{cases} \\
\operatorname{dim} \mathbb{P}_{1} & =d+1 \\
\operatorname{dim} \mathbb{P}_{2} & = \begin{cases}3, & d=1 \\
6, & d=2 \\
10, & d=3\end{cases}
\end{aligned}
$$

## $\mathbb{P}_{k}$ simplex finite elements

- $K$ : simplex spanned by $\vec{a}_{1} \ldots \vec{a}_{d+1}$ in $\mathbb{R}^{d}$
- For $0 \leq i_{1} \ldots i_{d+1} \leq k, i_{1}+\cdots+i_{d+1}=k$, let the set of nodes $\Sigma=\left\{\vec{\sigma}_{1} \ldots \vec{\sigma}_{s}\right\}$ be defined by the points $\vec{a}_{i_{1} \ldots i_{d} ; k}$ with barycentric coordinates $\left(\frac{i_{1}}{k} \ldots \frac{i_{d+1}}{k}\right)$.

- $s=\operatorname{card} \Sigma=\operatorname{dim} \mathbb{P}_{K} \Rightarrow$ there exists a basis $\theta_{1} \ldots \theta_{s}$ of $\mathbb{P}_{k}$ such that $\theta_{i}\left(\vec{\sigma}_{j}\right)=\delta_{i j}$


## $\mathbb{P}_{1}$ simplex finite elements

- $K$ : simplex spanned by $a_{1} \ldots a_{d+1}$ in $\mathbb{R}^{d}$
- $s=d+1$
- Nodes $\equiv$ vertices
- Basis functions $\theta_{1} \ldots \theta_{d+1} \equiv$ barycentric coordinates $\lambda_{1} \ldots \lambda_{d+1}$



## Global degrees of freedom

- Given a triangulation $\mathcal{T}_{h}$
- Let $\left\{\vec{a}_{1} \ldots \vec{a}_{N}\right\}=\bigcup_{K \in \mathcal{T}_{h}}\left\{\vec{\sigma}_{K, 1} \ldots \vec{\sigma}_{K, s}\right\}$ be the set of global degrees of freedom.
- Degree of freedom map

$$
j: \mathcal{T}_{h} \times\{1 \ldots s\} \rightarrow\{1 \ldots N\}
$$

$(K, m) \mapsto j(K, m)$ the global degree of freedom number

## Lagrange finite element space

- Given a triangulation $\mathcal{T}_{h}$ of $\Omega$, define the spaces

$$
\begin{aligned}
P_{h}^{k} & =\left\{v_{h} \in \mathbb{C}^{0}(\Omega):\left.v_{h}\right|_{K} \in \mathbb{P}_{K} \forall K \in \mathcal{T}_{j}\right\} \subset H^{1}(\Omega) \\
P_{0, h}^{k} & =\left\{v_{h} \in P_{h}^{k}:\left.v_{h}\right|_{\partial \Omega}=0\right\} \subset H_{0}^{1}(\Omega)
\end{aligned}
$$

- Global shape functions $\theta_{1}, \ldots, \theta_{N} \in P_{h}^{k}$ defined by

$$
\phi_{i} \left\lvert\, K\left(\vec{a}_{K, m}\right)= \begin{cases}\delta_{m n} & \text { if } \exists n \in\{1 \ldots s\}: j(K, n)=i \\ 0 & \text { otherwise }\end{cases}\right.
$$

- $\left\{\phi_{1}, \ldots, \phi_{N}\right\}$ is a basis of $P_{h}$, and $\gamma_{1} \ldots \gamma_{N}$ is a basis of $\mathcal{L}\left(P_{h}, \mathbb{R}\right)$ :
- $\left\{\phi_{1}, \ldots, \phi_{N}\right\}$ are linearly independent: if $\sum_{j=1}^{N} \alpha_{j} \phi_{j}=0$ then evaluation at $\vec{a}_{1} \ldots \vec{a}_{N}$ yields that $\alpha_{1} \ldots \alpha_{N}=0$.
- Let $v_{h} \in P_{h}$. Let $w_{h}=\sum_{j=1}^{N} v_{h}\left(\vec{a}_{j}\right) \phi_{j}$. Then for all $K \in \mathcal{T}_{h},\left.v_{h}\right|_{K}$ and $\left.w_{h}\right|_{K}$ coincide in the local nodes $\vec{a}_{K, 1} \ldots \vec{a}_{K, 2},\left.\Rightarrow v_{h}\right|_{K}=\left.w_{h}\right|_{K}$.


## Finite element approximation space

We have

| d | k | $N=\operatorname{dim} P_{h}^{k}$ |
| :--- | :--- | :--- |
| 1 | 1 | $N_{v}$ |
| 1 | 2 | $N_{v}+N_{e l}$ |
| 1 | 3 | $N_{v}+2 N_{e l}$ |
| 2 | 1 | $N_{v}$ |
| 2 | 2 | $N_{v}+N_{e d}$ |
| 2 | 3 | $N_{v}+2 N_{e d}+N_{e l}$ |
| 3 | 1 | $N_{v}$ |
| 3 | 2 | $N_{v}+N_{e d}$ |
| 3 | 3 | $N_{v}+2 N_{e d}+N_{f}$ |

## Local Lagrange interpolation operator

- Let $\{K, P, \Sigma\}$ be a finite element with shape function bases $\left\{\theta_{1} \ldots \theta_{s}\right\}$. Let $V(K)=\mathbb{C}^{0}(K)$ and $P \subset V(K)$
- local interpolation operator

$$
\begin{aligned}
\mathcal{I}_{K}: V(K) & \rightarrow P \\
v & \mapsto \sum_{i=1}^{s} v\left(\vec{\sigma}_{i}\right) \theta_{i}
\end{aligned}
$$

- $P$ is invariant under the action of $\mathcal{I}_{K}$, i.e. $\forall p \in P, \mathcal{I}_{K}(p)=p$ : - Let $p=\sum_{j=1}^{s} \alpha_{j} \theta_{j}$ Then,

$$
\begin{aligned}
\mathcal{I}_{K}(p) & =\sum_{i=1}^{s} p\left(\vec{\sigma}_{i}\right) \theta_{i}=\sum_{i=1}^{s} \sum_{j=1}^{s} \alpha_{j} \theta_{j}\left(\vec{\sigma}_{i}\right) \theta_{i} \\
& =\sum_{i=1}^{s} \sum_{j=1}^{s} \alpha_{j} \delta_{i j} \theta_{i}=\sum_{j=1}^{s} \alpha_{j} \theta_{j}
\end{aligned}
$$

## Global Lagrange interpolation operator

Let $V_{h}=P_{h}^{k}$

$$
\begin{array}{r}
\mathcal{I}_{h}: \mathcal{C}^{0}\left(\bar{\Omega}_{h}\right) \rightarrow V_{h} \\
v(\vec{x}) \mapsto v_{h}(\vec{x})=\sum_{i=1}^{N} v\left(\vec{a}_{i}\right) \phi_{i}(\vec{x})
\end{array}
$$

## Local interpolation error estimate I

Theorem: Let $\{\widehat{K}, \widehat{P}, \widehat{\Sigma}\}$ be a finite element with associated normed vector space $V(\widehat{K})$. Assume that

$$
\mathbb{P}_{k} \subset \widehat{P} \subset H^{2}(\widehat{K}) \subset V(\widehat{K})
$$

Then there exists $c>0$ such that for all $m=0 \ldots 2, K \in \mathcal{T}_{h}, v \in H^{2}(K)$ :

$$
\left|v-\mathcal{I}_{K}^{1} v\right|_{m, K} \leq c h_{K}^{2-m} \sigma_{K}^{m}|v|_{2, K} .
$$

l.e. the the local interpolation error can be estimated through $h_{K}, \sigma_{K}$ and the norm of a higher derivative.

## Local interpolation: special cases

- $m=0:\left|v-\mathcal{I}_{K}^{1} v\right|_{0, K} \leq c h_{K}^{2}|v|_{2, K}$
- $m=1:\left|v-\mathcal{I}_{K}^{1} v\right|_{1, K} \leq c h_{K} \sigma_{K}|v|_{2, K}$


## Global interpolation error estimate for Lagrangian finite elements, $k=1$

- Assume $v \in H^{2}(\Omega)$, e.g. if problem coefficients are smooth and the domain is convex

$$
\begin{aligned}
\left\|v-\mathcal{I}_{h}^{1} v\right\|_{0, \Omega}+h\left|v-\mathcal{I}_{h}^{1} v\right|_{1, \Omega} & \leq c h^{2}|v|_{2, \Omega} \\
\left|v-\mathcal{I}_{h}^{1} v\right|_{1, \Omega} & \leq c h|v|_{2, \Omega} \\
\lim _{h \rightarrow 0}\left(\inf _{v_{h} \in V_{h}^{1}}\left|v-v_{h}\right|_{1, \Omega}\right) & =0
\end{aligned}
$$

- If $v \in H^{2}(\Omega)$ cannot be guaranteed, estimates become worse.

Example: L-shaped domain.

- These results immediately can be applied in Cea's lemma.


## Error estimates for homogeneous Dirichlet problem

- Search $u \in H_{0}^{1}(\Omega)$ such that

$$
\int_{\Omega} \delta \vec{\nabla} u \vec{\nabla} v d \vec{x}=\int_{\Omega} f v d \vec{x} \forall v \in H_{0}^{1}(\Omega)+
$$

Then, $\lim _{h \rightarrow 0}\left\|u-u_{h}\right\|_{1, \Omega}=0$. If $u \in H^{2}(\Omega)$ (e.g. on convex domains) then

$$
\begin{aligned}
& \left\|u-u_{h}\right\|_{1, \Omega} \leq c h|u|_{2, \Omega} \\
& \left\|u-u_{h}\right\|_{0, \Omega} \leq c h^{2}|u|_{2, \Omega}
\end{aligned}
$$

Under certain conditions (convex domain, smooth coefficients) one also has

$$
\left\|u-u_{h}\right\|_{0, \Omega} \leq c h|u|_{1, \Omega}
$$

("Aubin-Nitsche-Lemma")

## $H^{2}$-Regularity

- $u \in H^{2}(\Omega)$ may be not fulfilled e.g.
- if $\Omega$ has re-entrant corners
- if on a smooth part of the domain, the boundary condition type changes
- if problem coefficients ( $\delta$ ) are discontinuos
- Situations differ as well between two and three space dimensions
- Delicate theory, ongoing research in functional analysis
- Consequence for simuations
- Deterioration of convergence rate
- Remedy: local refinement of the discretization mesh
- using a priori information
- using a posteriori error estimators + automatic refinement of discretizatiom mesh


## Weak formulation of homogeneous Dirichlet problem

- Search $u \in V=H_{0}^{1}(\Omega)$ such that

$$
\int_{\Omega} \delta \vec{\nabla} u \vec{\nabla} v d \vec{x}=\int_{\Omega} f v d \vec{x} \forall v \in H_{0}^{1}(\Omega)
$$

- Then,

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

is a self-adjoint bilinear form defined on the Hilbert space $H_{0}^{1}(\Omega)$.

## Galerkin ansatz

- Let $V_{h} \subset V$ be a finite dimensional subspace of $V$
- "Discrete" problem $\equiv$ Galerkin approximation: Search $u_{h} \in V_{h}$ such that

$$
a\left(u_{h}, v_{h}\right)=f\left(v_{h}\right) \forall v_{h} \in V_{h}
$$

- E.g. $V_{h}$ is the space of P1 Lagrange finite element approximations


## Stiffness matrix for Laplace operator for P1 FEM

- Element-wise calculation:

$$
a_{i j}=a\left(\phi_{i}, \phi_{j}\right)=\int_{\Omega} \vec{\nabla} \phi_{i} \vec{\nabla} \phi_{j} d \vec{x}=\left.\left.\int_{\Omega} \sum_{K \in \mathcal{T}_{h}} \vec{\nabla} \phi_{i}\right|_{K} \vec{\nabla} \phi_{j}\right|_{K} d \vec{x}
$$

- Standard assembly loop:
for $i, j=1 \ldots N$ do

$$
\text { set } a_{i j}=0
$$

end
for $K \in \mathcal{T}_{h}$ do

$$
\text { for } m, n=0 \ldots d \text { do }
$$

$$
s_{m n}=\int_{K} \vec{\nabla} \lambda_{m} \vec{\nabla} \lambda_{n} d \vec{x}
$$

$$
a_{j_{d o f}(K, m), j_{d o f}(K, n)}=a_{j_{d o f}(K, m), j_{d o f}(K, n)}+s_{m n}
$$

end
end

- Local stiffness matrix:

$$
S_{K}=\left(s_{K ; m, n}\right)=\int_{K} \vec{\nabla} \lambda_{m} \vec{\nabla} \lambda_{n} d \vec{x}
$$

## Local stiffness matrix calculation for P1 FEM

- $a_{0} \ldots a_{d}$ : vertices of the simplex $K, a \in K$.
- Barycentric coordinates: $\lambda_{j}(a)=\frac{\left|K_{j}(a)\right|}{|K|}$
- For indexing modulo $\mathrm{d}+1$ we can write

$$
\begin{array}{r}
|K|=\frac{1}{d!} \operatorname{det}\left(a_{j+1}-a_{j}, \ldots a_{j+d}-a_{j}\right) \\
\left|K_{j}(a)\right|=\frac{1}{d!} \operatorname{det}\left(a_{j+1}-a, \ldots a_{j+d}-a\right)
\end{array}
$$

- From this information, we can calculate explicitely $\vec{\nabla} \lambda_{j}(x)$ (which are constant vectors due to linearity) and the corresponding entries of the local stiffness

$$
s_{i j}=\int_{K} \vec{\nabla} \lambda_{i} \vec{\nabla} \lambda_{j} d \vec{x}
$$

## Local stiffness matrix calculation for P1 FEM in 2D

- $a_{0}=\left(x_{0}, y_{0}\right) \ldots a_{d}=\left(x_{2}, y_{2}\right)$ : vertices of the simplex $K$, $a=(x, y) \in K$.
- Barycentric coordinates: $\lambda_{j}(x, y)=\frac{\left|K_{j}(x, y)\right|}{|K|}$
- For indexing modulo $d+1$ we can write

$$
\begin{array}{r}
|K|=\frac{1}{2} \operatorname{det}\left(\begin{array}{ll}
x_{j+1}-x_{j} & x_{j+2}-x_{j} \\
y_{j+1}-y_{j} & y_{j+2}-y_{j}
\end{array}\right) \\
\left|K_{j}(x, y)\right|
\end{array}=\frac{1}{2} \operatorname{det}\left(\begin{array}{ll}
x_{j+1}-x & x_{j+2}-x \\
y_{j+1}-y & y_{j+2}-y
\end{array}\right) .
$$

- Therefore, we have

$$
\begin{aligned}
\left|K_{j}(x, y)\right| & =\frac{1}{2}\left(\left(x_{j+1}-x\right)\left(y_{j+2}-y\right)-\left(x_{j+2}-x\right)\left(y_{j+1}-y\right)\right) \\
\partial_{x}\left|K_{j}(x, y)\right| & =\frac{1}{2}\left(\left(y_{j+1}-y\right)-\left(y_{j+2}-y\right)\right)=\frac{1}{2}\left(y_{j+1}-y_{j+2}\right) \\
\partial_{y}\left|K_{j}(x, y)\right| & =\frac{1}{2}\left(\left(x_{j+2}-x\right)-\left(x_{j+1}-x\right)\right)=\frac{1}{2}\left(x_{j+2}-x_{j+1}\right)
\end{aligned}
$$

## Local stiffness matrix calculation for P1 FEM in 2D II

$$
s_{i j}=\int_{K} \vec{\nabla} \lambda_{i} \vec{\nabla} \lambda_{j} d \vec{x}=\frac{|K|}{4|K|^{2}}\left(y_{i+1}-y_{i+2}, x_{i+2}-x_{i+1}\right)\binom{y_{j+1}-y_{j+2}}{x_{j+2}-x_{j+1}}
$$

- So, let $V=\left(\begin{array}{ll}x_{1}-x_{0} & x_{2}-x_{0} \\ y_{1}-y_{0} & y_{2}-y_{0}\end{array}\right)$
- Then

$$
\begin{aligned}
& x_{1}-x_{2}=V_{00}-V_{01} \\
& y_{1}-y_{2}=V_{10}-V_{11}
\end{aligned}
$$

and

$$
\begin{aligned}
2|K| \vec{\nabla} \lambda_{0} & =\binom{y_{1}-y_{2}}{x_{2}-x_{1}}
\end{aligned}=\binom{V_{10}-V_{11}}{V_{01}-V_{00}} .
$$

## Degree of freedom map representation for P1 finite elements

- List of global nodes $a_{0} \ldots a_{N}$ : two dimensional array of coordinate values with $N$ rows and $d$ columns
- Local-global degree of freedom map: two-dimensional array $C$ of index values with $N_{e l}$ rows and $d+1$ columns such that $C(i, m)=j_{d o f}\left(K_{i}, m\right)$.
- The mesh generator triangle generates this information directly


## Practical realization of boundary conditions

- Robin boundary value problem

$$
\begin{array}{rlrl}
-\nabla \cdot \delta \vec{\nabla} u=f & \text { in } \Omega \\
\delta \vec{\nabla} u+\alpha(u-g)=0 & & \text { on } \partial \Omega
\end{array}
$$

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

$$
\int_{\Omega} \delta \vec{\nabla} u \vec{\nabla} v d \vec{x}+\int_{\partial \Omega} \alpha u v d s=\int_{\Omega} f v d \vec{x}+\int_{\partial \Omega} \alpha g v d s \forall v \in H^{1}(\Omega)
$$

- In 2D, for $P^{1}$ FEM, boundary integrals can be calculated by trapezoidal rule without sacrificing approximation order
- Use Dirichlet penalty method to handle Dirichlet boundary conditions


## More complicated integrals

- Assume non-constant right hand side $f$, space dependent heat conduction coefficient $\delta$.
- Right hand side integrals

$$
f_{i}=\int_{K} f(x) \lambda_{i}(x) d \vec{x}
$$

- $P^{1}$ stiffness matrix elements

$$
a_{i j}=\int_{K} \delta(x) \vec{\nabla} \lambda_{i} \vec{\nabla} \lambda_{j} d \vec{x}
$$

- $P^{k}$ stiffness matrix elements created from higher order ansatz functions


## Quadrature rules

- Quadrature rule:

$$
\int_{K} g(x) d \vec{x} \approx|K| \sum_{l=1}^{I_{q}} \omega_{l} g\left(\xi_{l}\right)
$$

- $\xi_{l}$ : nodes, Gauss points
- $\omega_{l}$ : weights
- The largest number $k$ such that the quadrature is exact for polynomials of order $k$ is called order $k_{q}$ of the quadrature rule, i.e.

$$
\forall k \leq k_{q}, \forall p \in \mathbb{P}^{k} \int_{K} p(x) d \vec{x}=|K| \sum_{l=1}^{l_{q}} \omega_{l} p\left(\xi_{l}\right)
$$

- Error estimate:

$$
\begin{aligned}
\forall \phi \in \mathcal{C}^{k_{q}+1}(K), \left\lvert\, \frac{1}{|K|} \int_{K} \phi(x) d \vec{x}\right. & -\sum_{l=1}^{I_{q}} \omega_{l} g\left(\xi_{l}\right) \mid \\
& \leq c h_{K}^{k_{q}+1} \sup _{x \in K,|\alpha|=k_{q}+1}\left|\partial^{\alpha} \phi(x)\right|
\end{aligned}
$$

## Some common quadrature rules

Nodes are characterized by the barycentric coordinates

| $d$ | $k_{q}$ | $I_{q}$ | Nodes | Weights |
| :--- | :--- | :--- | :--- | :--- |
| 1 | 1 | 1 | $\left(\frac{1}{2}, \frac{1}{2}\right)$ | 1 |
|  | 1 | 2 | $(1,0),(0,1)$ | $\frac{1}{2}, \frac{1}{2}$ |
|  | 3 | 2 | $\left(\frac{1}{2}+\frac{\sqrt{3}}{6}, \frac{1}{2}-\frac{\sqrt{3}}{6}\right),\left(\frac{1}{2}-\frac{\sqrt{3}}{6}, \frac{1}{2}+\frac{\sqrt{3}}{6}\right)$ | $\frac{1}{2}, \frac{1}{2}$ |
|  | 5 | 3 | $\left(\frac{1}{2},\right),\left(\frac{1}{2}+\sqrt{\frac{3}{20}}, \frac{1}{2}-\sqrt{\frac{3}{20}}\right),\left(\frac{1}{2}-\sqrt{\frac{3}{20}}, \frac{1}{2}+\sqrt{\frac{3}{20}}\right)$ | $\frac{8}{18}, \frac{5}{18}, \frac{5}{18}$ |
| 2 | 1 | 1 | $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ | 1 |
|  | 1 | 3 | $(1,0,0),(0,1,0),(0,0,1)$ | $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$ |
|  | 2 | 3 | $\left(\frac{1}{2}, \frac{1}{2}, 0\right),\left(\frac{1}{2}, 0, \frac{1}{2}\right),\left(0, \frac{1}{2}, \frac{1}{2}\right)$ | $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$ |
|  | 3 | 4 | $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right),\left(\frac{1}{5}, \frac{1}{5}, \frac{3}{5}\right),\left(\frac{1}{5}, \frac{3}{5}, \frac{1}{5}\right),\left(\frac{3}{5}, \frac{1}{5}, \frac{1}{5}\right)$, | $\frac{9}{16}, \frac{25}{48}, \frac{25}{48}, \frac{25}{48}$ |
| 3 | 1 | 1 | $\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ |  |
|  | 1 | 4 | $(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)$ | $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$ |
|  | 2 | 4 | $\left(\frac{5-\sqrt{5}}{20}, \frac{5-\sqrt{5}}{20}, \frac{5-\sqrt{5}}{20}, \frac{5+3 \sqrt{5}}{20}\right) \ldots$ | $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$ |

## Matching of approximation order and quadrature order

- "Variational crime": instead of

$$
a\left(u_{h}, v_{h}\right)=f\left(v_{h}\right) \forall v_{h} \in V_{h}
$$

we solve

$$
a_{h}\left(u_{h}, v_{h}\right)=f_{h}\left(v_{h}\right) \forall v_{h} \in V_{h}
$$

where $a_{h}, f_{h}$ are derived from their exact counterparts by quadrature

- For $P^{1}$ finite elements, zero order quadrature for volume integrals and first order quadrature for surface intergals is sufficient to keep the convergence order estimates stated before
- The rule of thumb for the volume quadrature is that the highest order terms must be evaluated exactly if the coefficients of the PDE are constant.


## Practical realization of integrals

- Integral over barycentric coordinate function

$$
\int_{K} \lambda_{i}(x) d \vec{x}=\frac{1}{3}|K|
$$

- Right hand side integrals. Assume $f(x)$ is given as a piecewise linear function with given values in the nodes of the triangulation

$$
f_{i}=\int_{K} f(x) \lambda_{i}(x) d \vec{x} \approx \frac{1}{3}|K|\left(f\left(a_{0}\right)+f\left(a_{1}\right)+f\left(a_{2}\right)\right)
$$

- Integral over space dependent heat conduction coefficient: Assume $\delta(x)$ is given as a piecewise linear function with given values in the nodes of the triangulation

$$
a_{i j}=\int_{K} \delta(x) \vec{\nabla} \lambda_{i} \vec{\nabla} \lambda_{j} d \vec{x}=\frac{1}{3}\left(\delta\left(a_{0}\right)+\delta\left(a_{1}\right)+\delta\left(a_{2}\right)\right) \int_{K} \vec{\nabla} \lambda_{i} \vec{\nabla} \lambda_{j} d \vec{x}
$$

## P1 FEM stiffness matrix condition number

- Homogeneous dirichlet boundary value problem

$$
\begin{aligned}
& -\nabla \cdot \delta \vec{\nabla} u=f \text { in } \Omega \\
& \left.u\right|_{\partial \Omega}=0
\end{aligned}
$$

- Lagrange degrees of freedom $a_{1} \ldots a_{N}$ corresponding to global basis functions $\phi_{1} \ldots \phi_{N}$ such that $\phi_{i} \mid \partial \Omega=0$ aka $\phi_{i} \in V_{h} \subset H_{0}^{1}(\Omega)$
- Stiffness matrix $A=\left(a_{i j}\right)$ :

$$
a_{i j}=a\left(\phi_{i}, \phi_{j}\right)=\int_{\Omega} \delta \vec{\nabla} \phi_{i} \vec{\nabla} \phi_{j} d \vec{x}
$$

- bilinear form $a(\cdot, \cdot)$ is self-adjoint, therefore $A$ is symmetric, positive definite
- Condition number estimate for $P^{1}$ finite elements on quasi-uniform triangulation:

$$
\kappa(A) \leq c h^{-2}
$$

## The problem with Dirichlet boundary conditions

- Homogeneous Dirichlet $B C \Rightarrow$ include boundary condition into set of basis functions
- Inhomogeneous Dirichlet, may be only at a part of the boundary
- Use exact approach from as in continous formulation (with lifting $u_{g}$ etc) $\Rightarrow$ highly technical
- Eliminate Dirichlet BC algebraically after building of the matrix, i.e. fix "known unknowns" at the Dirichlet boundary $\Rightarrow$ highly technical
- Modifiy matrix such that equations at boundary exactly result in Dirichlet values $\Rightarrow$ loss of symmetry of the matrix
- Penalty method


## Dirichlet BC: Algebraic manipulation

- Assume 1D situation with $\mathrm{BC} u_{1}=g$
- From integration in $H^{1}$ regardless of boundary values:

$$
A U=\left(\begin{array}{ccccc}
\frac{1}{h} & -\frac{1}{h} & & & \\
-\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\
& -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\
& & \ddots & \ddots & \ddots
\end{array}\right)\left(\begin{array}{c}
u_{1} \\
u_{2} \\
u_{3} \\
\vdots
\end{array}\right)=\left(\begin{array}{c}
f_{1} \\
f_{2} \\
f_{3} \\
\vdots
\end{array}\right)
$$

- Fix $u_{1}$ and eliminate:

$$
A^{\prime} U=\left(\begin{array}{cccc}
\frac{2}{h} & -\frac{1}{h} & & \\
-\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\
& \ddots & \ddots & \ddots
\end{array}\right)\left(\begin{array}{c}
u_{2} \\
u_{3} \\
\vdots
\end{array}\right)=\left(\begin{array}{c}
f_{2}+\frac{1}{h} g \\
f_{3} \\
\vdots
\end{array}\right)
$$

- $A^{\prime}$ becomes idd and stays symmetric
- operation is quite technical


## Dirichlet BC: Modify boundary equations

- From integration in $H^{1}$ regardless of boundary values:

$$
A U=\left(\begin{array}{ccccc}
\frac{1}{h} & -\frac{1}{h} & & & \\
-\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\
& -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\
& & \ddots & \ddots & \ddots
\end{array}\right)\left(\begin{array}{c}
u_{1} \\
u_{2} \\
u_{3} \\
\vdots
\end{array}\right)=\left(\begin{array}{c}
f_{1} \\
f_{2} \\
f_{3} \\
\vdots
\end{array}\right)
$$

- Modify equation at boundary to exactly represent Dirichlet values

$$
A^{\prime} U=\left(\begin{array}{ccccc}
\frac{1}{h} & 0 & & & \\
-\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\
& -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\
& & \ddots & \ddots & \ddots
\end{array}\right)\left(\begin{array}{c}
u_{1} \\
u_{2} \\
u_{3} \\
\vdots
\end{array}\right)=\left(\begin{array}{c}
\frac{1}{h} g \\
f_{2} \\
f_{3} \\
\vdots
\end{array}\right)
$$

- $A^{\prime}$ becomes idd
- loses symmetry $\Rightarrow$ problem e.g. with CG method


## Dirichlet BC: Discrete penalty trick

- From integration in $H^{1}$ regardless of boundary values:

$$
A U=\left(\begin{array}{ccccc}
\frac{1}{h} & -\frac{1}{h} & & & \\
-\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\
& -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\
& & \ddots & \ddots & \ddots
\end{array}\right)\left(\begin{array}{c}
u_{1} \\
u_{2} \\
u_{3} \\
\vdots
\end{array}\right)=\left(\begin{array}{c}
f_{1} \\
f_{2} \\
f_{3} \\
\vdots
\end{array}\right)
$$

- Add penalty terms

$$
A^{\prime} U=\left(\begin{array}{ccccc}
\frac{1}{\varepsilon}+\frac{1}{h} & -\frac{1}{h} & & & \\
-\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & & \\
& -\frac{1}{h} & \frac{2}{h} & -\frac{1}{h} & \\
& & \ddots & \ddots & \ddots
\end{array}\right)\left(\begin{array}{c}
u_{1} \\
u_{2} \\
u_{3} \\
\vdots
\end{array}\right)=\left(\begin{array}{c}
f_{1}+\frac{1}{\varepsilon} g \\
f_{2} \\
f_{3} \\
\vdots
\end{array}\right)
$$

- $A^{\prime}$ becomes idd, keeps symmetry, and the realization is technically easy.
- If $\varepsilon$ is small enough, $u_{1}=g$ will be satisfied exactly within floating point accuracy.
- Iterative methods should be initialized with Dirichlet values.
- Works for nonlinear problems, finite volume methods


## Dirichlet penalty trick, general formulation

- Dirichlet boundary value problem

$$
\begin{aligned}
& -\nabla \cdot \delta \vec{\nabla} u=f \quad \text { in } \Omega \\
& \left.u\right|_{\partial \Omega}=g
\end{aligned}
$$

- We discussed approximation of Dirichlet problem by Robin problem
- Practical realization uses discrete approach for Lagrange degrees of freedom $a_{1} \ldots a_{N}$ corresponding to global basis functions $\phi_{1} \ldots \phi_{N}$ :
- Search $u_{h}=\sum_{i=1}^{N} u_{i} \phi_{i} \in V_{h}=\operatorname{span}\left\{\phi_{1} \ldots \phi_{N}\right\}$ such that

$$
A U+\Pi U=F+\Pi G
$$

where

- $U=\left(u_{1} \ldots u_{N}\right)$
- $A=\left(a_{i j}\right):$ stiffness matrix with $a_{i j}=\int_{\Omega} \delta \vec{\nabla} \phi_{i} \vec{\nabla} \phi_{j} d \vec{x}$
- $F=\int_{\Omega} f \vec{\nabla} \phi_{i} d \vec{x}$
- $G=\left(g_{i}\right)$ with $g_{i}= \begin{cases}g\left(a_{i}\right), & a_{i} \in \partial \Omega \\ 0, & \text { else }\end{cases}$
- $\Pi=\left(\pi_{i j}\right)$ is a diagonal matrix with $\pi_{i j}= \begin{cases}\frac{1}{\varepsilon}, & i=j, a_{i} \in \partial \Omega \\ 0, & \text { else }\end{cases}$


## P1 FEM Stiffness matrix row sums

Row sums:

$$
\begin{aligned}
\sum_{j=1}^{N} a_{i j} & =\sum_{j=1}^{N} \int_{\Omega} \vec{\nabla} \phi_{i} \vec{\nabla} \phi_{j} d x=\int_{\Omega} \vec{\nabla} \phi_{i} \vec{\nabla}\left(\sum_{j=1}^{N} \phi_{j}\right) d x \\
& =\int_{\Omega} \vec{\nabla} \phi_{i} \vec{\nabla}(1) d x \\
& =0
\end{aligned}
$$

## P1 FEM stiffness matrix entry signs

Local stiffness matrix $S_{K}$

$$
s_{i j}=\int_{K} \vec{\nabla} \lambda_{i} \vec{\nabla} \lambda_{j} d x=\frac{|K|}{2|K|^{2}}\left(y_{i+1}-y_{i+2}, x_{i+2}-x_{i+1}\right)\binom{y_{j+1}-y_{j+2}}{x_{j+2}-x_{j+1}}
$$

- Main diagonal entries positive
- Local contributions from element stiffness matrices: Scalar products of vectors orthogonal to edges. These are nonpositive if the angle between the edges are $\leq 90^{\circ}$
- weakly acute triangulation: all triangle angles are less than $\leq 90^{\circ}$
- In fact, for constant coefficients, in 2D, Delaunay is sufficient, as contributions from opposite angles compensate each other
- All row sums are zero $\Rightarrow A$ is singular
- Matrix becomes irreducibly diagonally dominant if we add at least one positive value to the main diagonal, e.g. from Dirichlet BC


## Stationary linear reaction-diffusion

- Assume additional process in each REV which produces or destroys species depending on the amount of species present with given rate $r$.
Search function $u: \Omega \rightarrow \mathbb{R}$ such that

$$
\begin{aligned}
-\nabla \cdot \delta \vec{\nabla} u+r u & =f & & \text { in } \Omega \\
\delta \vec{\nabla} u \cdot \vec{n}+\alpha(u-g) & =0 & & \text { on } \partial \Omega
\end{aligned}
$$

- Coercivity guaranteed e.g. for $\alpha \geq 0, r>0$ which means species destruction FEM formulation: search $u_{h} \in V_{h}=\operatorname{span}\left\{\phi_{1} \ldots \phi_{N}\right\}$ such that

$$
\begin{aligned}
& \underbrace{\int_{\Omega} \delta \vec{\nabla} u_{h} \vec{\nabla} v_{h} d \vec{x}}_{\text {"stiffness matrix" }}+\underbrace{\int_{\Omega} r u_{h} v_{h} d \vec{x}}_{\text {"mass matrix" }}+\underbrace{\int_{\partial \Omega} \alpha u_{h} v_{h} d s}_{\text {"boundary mass matrix" }} \\
&=\int_{\Omega} f v_{h} d \vec{x}+\int_{\partial \Omega} \alpha g v_{h} d s \forall v_{h} \in V_{h}
\end{aligned}
$$

- Coercivity + symmetry $\Rightarrow$ positive definiteness


## Mass matrix properties

- Mass matrix (for $r=1$ ): $M=\left(m_{i j}\right)$ :

$$
m_{i j}=\int_{\Omega} \phi_{i} \phi_{j} d x
$$

- Self-adjoint, coercive bilinear form $\Rightarrow M$ is symmetric, positiv definite
- For a family of quasi-uniform, shape-regular triangulations, for every eigenvalue $\mu$ one has the estimate

$$
c_{1} h^{d} \leq \mu \leq c_{2} h^{d}
$$

$\mathrm{T} \Rightarrow$ condition number $\kappa(M)$ bounded by constant independent of $h$ :

$$
\kappa(M) \leq c
$$

- How to see this ? Let $u_{h}=\sum_{i=1}^{N} U_{i} \phi_{i}$, and $\mu$ an eigenvalue (positive,real!) Then

$$
\left\|u_{h}\right\|_{0}^{2}=(U, M U)_{\mathbb{R}^{N}}=\mu(U, U)_{\mathbb{R}^{N}}=\mu\|U\|_{\mathbb{R}^{N}}^{2}
$$

From quasi-uniformity we obtain

$$
c_{1} h^{d}\|U\|_{\mathbb{R}^{N}}^{2} \leq\left\|u_{h}\right\|_{0}^{2} \leq c_{2} h^{d}\|U\|_{\mathbb{R}^{N}}^{2}
$$

## Mass matrix M-Property (P1 FEM) ?

- For $P^{1}$-finite elements, all integrals $m_{i j}=\int_{\Omega} \phi_{i} \phi_{j} d x$ are zero or positive, so we get positive off diagonal elements.
- No M-Property!


## Mass matrix lumping (P1 FEM)

- Local mass matrix for P1 FEM on element K (calculated by 2nd order exact edge midpoint quadrature rule):

$$
M_{K}=|K|\left(\begin{array}{lll}
\frac{1}{6} & \frac{1}{12} & \frac{1}{12} \\
\frac{1}{12} & \frac{1}{6} & \frac{1}{12} \\
\frac{1}{12} & \frac{1}{12} & \frac{1}{6}
\end{array}\right)
$$

- Lumping: sum up off diagonal elements to main diagonal, set off diagonal entries to zero

$$
\tilde{M}_{K}=|K|\left(\begin{array}{ccc}
\frac{1}{3} & 0 & 0 \\
0 & \frac{1}{3} & 0 \\
0 & 0 & \frac{1}{3}
\end{array}\right)
$$

- Interpretation as change of quadrature rule to first order exact vertex based quadrature rule
- Loss of accuracy, gain of stability


## Effect of mass matrix lumping

- For P1 FEM on weakly acute or Delaunay triangulations, mass matrix lumping can guarantee M-Property of system matrix
- Adding a mass matrix which is not lumped yields a positive definite matrix (due to coercivity) and thus nonsingularity, but destroys $M$-property unless the absolute values of its off diagonal entries are less than those of $A$, i.e. for small $r$.
- Same situation witb Robin boundary conditions and boundary mass matrix
- Introducing the Dirichlet Penalty trick at continuous level without mass lumping would be disastrous.

The values of a piecewise linear function $u=u(\boldsymbol{x})$ can be defined in dependence of the values $u_{i}=u\left(\boldsymbol{p}_{i}\right)$ at the vertices of the simplex by the expression

$$
u(\boldsymbol{x})=\sum_{k=0}^{n} u_{k} \lambda_{k}(\boldsymbol{x})
$$

The gradient of this function is given by

$$
\text { (7) } \quad \nabla u(\boldsymbol{x})=\sum_{k=0}^{n} u_{k} \nabla \lambda_{k}(\boldsymbol{x})
$$

where the $\nabla \lambda_{k}$ can be determinated from the equations

$$
\delta_{i j}=\frac{\partial x_{i}}{\partial x_{j}}=\sum_{k=0}^{n} p_{k i} \frac{\partial \lambda_{k}}{\partial x_{j}}, \quad 0=\sum_{k=0}^{n} \frac{\partial \lambda_{k}}{\partial x_{j}} \quad(i, j=1, \ldots, n)
$$

(cf. (3)), i.e. we have to solve the linear systems

$$
\left(\begin{array}{cccc}
1 & 1 & \ldots & 1 \\
\boldsymbol{p}_{0} & \boldsymbol{p}_{1} & \ldots & \boldsymbol{p}_{n}
\end{array}\right)\left(\begin{array}{cccc}
\lambda_{0}(\mathbf{0}) & \frac{\partial \lambda_{0}}{\partial x_{1}} & \ldots & \frac{\partial \lambda_{0}}{\partial x_{n}} \\
\lambda_{1}(\mathbf{0}) & \frac{\partial \lambda_{1}}{\partial x_{1}} & \ldots & \frac{\partial \lambda_{1}}{\partial x_{n}} \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_{n}(\mathbf{0}) & \frac{\partial \lambda_{n}}{\partial x_{1}} & \ldots & \frac{\partial \lambda_{n}}{\partial x_{n}}
\end{array}\right)=\left(\begin{array}{cccc}
1 & 0 & \ldots & 0 \\
0 & 1 & & \vdots \\
\vdots & & \ddots & 0 \\
0 & \ldots & 0 & 1
\end{array}\right) .
$$

So we get

$$
\begin{aligned}
& \left(\begin{array}{cccc}
1 & 1 & \ldots & 1 \\
p_{0} & p_{1} & \ldots & p_{n}
\end{array}\right)\left(\begin{array}{c}
\partial_{i} \lambda_{0} \\
\partial_{i} \lambda_{1} \\
\vdots \\
\partial_{i} \lambda_{n}
\end{array}\right)=e_{i}=\left(\begin{array}{c}
0 \\
\vdots \\
1 \\
\vdots \\
0
\end{array}\right) \\
& \partial_{i} \lambda_{0}+\partial_{i} \lambda_{1}+\cdots+\partial_{i} \lambda_{n}=e_{i, 0} \\
& p_{0} \partial_{i} \lambda_{0}+p_{1} \partial_{i} \lambda_{1}+\cdots+p_{n} \partial_{i} \lambda_{n}=e_{i,(1 \ldots n)}
\end{aligned}
$$

Comparing this with the determination of the normals of the faces (cf.ecture 18 slide 56
(2)) we get the identities

$$
\nabla \lambda_{i}=\boldsymbol{g}_{i}, \quad \lambda_{i}(\mathbf{0})=d_{i} \quad(i=0, \ldots, n) .
$$

and therefore

$$
\nabla u(\boldsymbol{x})=\sum_{i=0}^{n} \boldsymbol{g}_{i} u_{i}=\left(\begin{array}{llll}
\boldsymbol{g}_{0} & \boldsymbol{g}_{1} & \ldots & \boldsymbol{g}_{n}
\end{array}\right)\left(\begin{array}{c}
u_{0} \\
u_{1} \\
\vdots \\
u_{n}
\end{array}\right)
$$

holds. That means that

$$
\sum_{i=0}^{n} \boldsymbol{g}_{i} u\left(\boldsymbol{p}_{i}\right)=\sum_{i=0}^{n} \frac{1}{h_{i}} \boldsymbol{a}_{i} u\left(\boldsymbol{p}_{i}\right)
$$

can be considered as a discrete gradient of a function $u$.

