Discrete Maximum Principles in the Finite Element Method

Tomáš Vejchodský (vejchod@math.cas.cz)

Institute of Mathematics, Academy of Sciences
Žitná 25, 115 67 Praha 1
Czech Republic

November 4, 2009, FSv ČVUT, Praha

Outline

- Introduction/Motivation
- (Continuous) maximum principles
- Discrete maximum principles
- Lowest-order FEM
- Higher-order FEM
- Numerical test

Elliptic problems

- Classical formulation:

$$\begin{aligned} -\operatorname{div}(\mathcal{A}\nabla u) + cu &= f && \text{in } \Omega, \\ u &= g_{\mathrm{D}} && \text{on } \Gamma_{\mathrm{D}}, \\ \alpha u + (\mathcal{A}\nabla u)\cdot n &= g_{\mathrm{N}} && \text{on } \Gamma_{\mathrm{N}}, \end{aligned}$$

- Weak formulation: $u = u^0 + \widetilde{g}_{\mathrm{D}}$

$$u^0 \in V: \quad a(u^0, v) = F(v) - a(\widetilde{g}_{\mathrm{D}}, v) \quad \forall v \in V$$

  - $V = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_{\mathrm{D}}\}$
  - $a(u,v) = \int_\Omega (\mathcal{A}\nabla u)\cdot\nabla v + cuv \,\mathrm{d}x + \int_{\Gamma_{\mathrm{N}}} \alpha uv \,\mathrm{d}s$
  - $F(v) = \int_\Omega fv \,\mathrm{d}x + \int_{\Gamma_{\mathrm{N}}} g_{\mathrm{N}} v \,\mathrm{d}s$
  - $\widetilde{g}_{\mathrm{D}} \in H^1(\Omega)$, $\widetilde{g}_{\mathrm{D}} = g_{\mathrm{D}}$ on Γ_{D}

(Continuous) maximum principles – weak problem

- Maximum principle (MaxP)

$$f \le 0 \text{ and } g_{\mathrm{N}} \le 0 \quad \Rightarrow \quad \max_{\overline{\Omega}} u \le \max_{\Gamma_{\mathrm{D}}} \max\{0, u\}$$

- Minimum principle (MinP)

$$f \ge 0 \text{ and } g_{\mathrm{N}} \ge 0 \quad \Rightarrow \quad \min_{\overline{\Omega}} u \ge \min_{\Gamma_{\mathrm{D}}} \min\{0, u\}$$

- Conservation of nonnegativity (ConN)

$$f \ge 0, \; g_{\mathrm{D}} \ge 0, \text{ and } g_{\mathrm{N}} \ge 0 \quad \Rightarrow \quad u \ge 0$$

- Comparison principle (CmpP)

$$f_1 \ge f_2, \; g_{\mathrm{D},1} \ge g_{\mathrm{D},2}, \text{ and } g_{\mathrm{N},1} \ge g_{\mathrm{N},2} \quad \Rightarrow \quad u_1 \ge u_2$$

Theorem

MaxP $\Leftrightarrow$ *MinP* $\Leftrightarrow$ *ConN* $\Leftrightarrow$ *CmpP*

(Continuous) maximum principles – weak problem

- Conservation of nonnegativity (ConN)

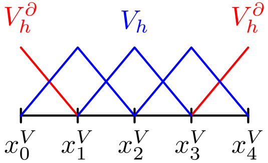
$$f \geq 0, \; g_{\mathrm{D}} \geq 0, \text{ and } g_{\mathrm{N}} \geq 0 \quad \Rightarrow \quad u \geq 0$$

Finite element method (FEM)

- *hp*-FEM: $u_h = u_h^0 + \widetilde{g}_{\mathrm{D}h}$

$$u_h^0 \in V_h : \quad a(u_h^0, v_h) = F(v_h) - a(\widetilde{g}_{\mathrm{D}h}, v_h) \quad \forall v_h \in V_h$$

  - $\mathcal{T}_h$ triangulation of Ω
  - p_K polynomial degree on $K \in \mathcal{T}_h$
  - $X_h = \{v_h \in H^1(\Omega) : v_h|_K \in P^{p_K}(K), K \in \mathcal{T}_h\}$
  - $V_h = X_h \cap V$
  - $X_h = V_h \oplus V_h^\partial$
  - $\widetilde{g}_{\mathrm{D}h} \in V_h^\partial$, $\widetilde{g}_{\mathrm{D}h} \approx g_{\mathrm{D}}$ on Γ_{D}



Discrete maximum principles – V_h fixed

- Discrete maximum principle (DMaxP)

$$f \leq 0 \text{ and } g_{\mathrm{N}} \leq 0 \quad \Rightarrow \quad \max_{\overline{\Omega}} u_h \leq \max_{\Gamma_{\mathrm{D}}} \max\{0, u_h\}$$

- Discrete minimum principle (DMinP)

$$f \geq 0 \text{ and } g_{\mathrm{N}} \geq 0 \quad \Rightarrow \quad \min_{\overline{\Omega}} u_h \geq \min_{\Gamma_{\mathrm{D}}} \min\{0, u_h\}$$

- Discrete conservation of nonnegativity (DConN)

$$f \geq 0,\ \widetilde{g}_{\mathrm{D}h} \geq 0, \text{ and } g_{\mathrm{N}} \geq 0 \quad \Rightarrow \quad u_h \geq 0$$

- Discrete comparison principle (DCmpP)

$$f_1 \geq f_2,\ \widetilde{g}_{\mathrm{D}h,1} \geq \widetilde{g}_{\mathrm{D}h,2}, \text{ and } g_{\mathrm{N},1} \geq g_{\mathrm{N},2} \quad \Rightarrow \quad u_{h,1} \geq u_{h,2}$$

Theorem

DMaxP $\Leftrightarrow$ *DMinP* $\Leftrightarrow$ *DConN* $\Leftrightarrow$ *DCmpP*

Discrete maximum principles – V_h fixed

- Discrete conservation of nonnegativity (DConN) $\equiv$ DMP

$$f \geq 0, \ \widetilde{g}_{\mathrm{D}h} \geq 0, \text{ and } g_{\mathrm{N}} \geq 0 \quad \Rightarrow \quad u_h \geq 0$$

Discrete Green's function (DGF)

Theorem (main)

$$DMP \quad \Leftrightarrow \quad \begin{array}{l} (a)\ G_h(x,y) \geq 0 \quad \forall (x,y) \in \Omega^2 \\ (b)\ G_h^\partial(s,y) \geq 0 \quad \forall s \in \Gamma_\mathrm{D},\ \forall y \in \Omega \end{array}$$

- $G_{h,y} \in V_h \quad :$
  $a(v_h, G_{h,y}) = v_h(y) \quad \forall v_h \in V_h,\ y \in \Omega$
- $G_{h,y}^\partial \in V_h^\partial \quad :$
  $\int_{\Gamma_\mathrm{D}} w_h(s) G_{h,y}^\partial(s)\,\mathrm{d}s = w_h(y) - a(w_h, G_{h,y}) \quad \forall w_h \in X_h,\ y \in \Omega$
- $G_h(x,y) = G_{h,y}(x) \qquad G_h^\partial(s,y) = G_{h,y}^\partial(s)$

$$u_h(y) = \int_\Omega f(x) G_h(x,y)\,\mathrm{d}x + \int_{\Gamma_\mathrm{N}} g_\mathrm{N}(s) G_h(s,y)\,\mathrm{d}s + \int_{\Gamma_\mathrm{D}} \widetilde{g}_{\mathrm{D}h}(s) G_h^\partial(s,y)\,\mathrm{d}s$$

Expressions for the DGF

- $\varphi_1, \varphi_2, \ldots, \varphi_N$ basis in V_h
- $\varphi_1^\partial, \varphi_2^\partial, \ldots, \varphi_{N^\partial}^\partial$ basis in V_h^∂
- $A \in \mathbb{R}^{N\times N}$, $A_{ij} = a(\varphi_j, \varphi_i)$ $i,j = 1,2,\ldots,N$
- $A^\partial \in \mathbb{R}^{N\times N^\partial}$, $A_{ik}^\partial = a(\varphi_k^\partial, \varphi_i)$ $i = 1,\ldots,N,\ k = 1,\ldots,N^\partial$
- $M^\partial \in \mathbb{R}^{N^\partial\times N^\partial}$, $M_{k\ell}^\partial = \int_{\Gamma_\mathrm{D}} \varphi_\ell^\partial \varphi_k^\partial \,\mathrm{d}s$ $k,\ell = 1,2,\ldots,N^\partial$
- $$G_h(x,y) = \sum_{i=1}^N \sum_{j=1}^N \varphi_i(y)(A^{-1})_{ij}\varphi_j(x)$$
- $$G_h^\partial(s,y) = \sum_{k=1}^{N^\partial} \sum_{\ell=1}^{N^\partial} \varphi_k^\partial(s)(M^\partial)_{k\ell}^{-1} \left[\varphi_\ell^\partial(y) - \sum_{i=1}^N \sum_{j=1}^N \varphi_i(y)(A^{-1})_{ij} A_{j\ell}^\partial\right]$$

Linear FEM, $g_D = 0$

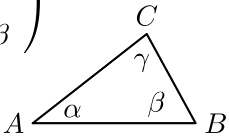
- DMP $\Leftrightarrow$ $G_h \geq 0$ $\Leftrightarrow$ $A^{-1} \geq 0$ $\Leftrightarrow$ A monotone
- A s.p.d., off-diag$(A) \leq 0$ $\Rightarrow$ A M-matrix $\Rightarrow$ A monotone
- Element matrices: off-diag$(A^K) \leq 0$ $\Rightarrow$ off-diag$(A) \leq 0$
  - $A = \sum_{K \in \mathcal{T}_h} A^K, \quad A^K_{ij} = a_K(\varphi_j, \varphi_i) = \int_K (\mathcal{A} \nabla \varphi_i) \cdot \nabla \varphi_j \, \mathrm{d}x + \cdots$

Linear FEM, $g_{\mathrm{D}} = 0$

- DMP $\Leftrightarrow$ $G_h \geq 0$ $\Leftrightarrow$ $A^{-1} \geq 0$ $\Leftrightarrow$ A monotone
- A s.p.d., off-diag$(A) \leq 0$ $\Rightarrow$ A M-matrix $\Rightarrow$ A monotone
- Element matrices: off-diag$(A^K) \leq 0$ $\Rightarrow$ off-diag$(A) \leq 0$
  - $A = \sum_{K \in \mathcal{T}_h} A^K, \quad A^K_{ij} = a_K(\varphi_j, \varphi_i) = \int_K (\mathcal{A} \nabla \varphi_i) \cdot \nabla \varphi_j \, \mathrm{d}x + \cdots$

Example (triangles):

$$A^K = \frac{1}{2} \begin{pmatrix} \cot\beta + \cot\gamma & -\cot\gamma & -\cot\beta \\ -\cot\gamma & \cot\alpha + \cot\gamma & -\cot\alpha \\ -\cot\beta & -\cot\alpha & \cot\alpha + \cot\beta \end{pmatrix}$$

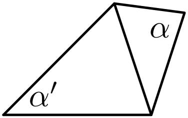


Linear FEM, $g_D = 0$

- DMP $\quad\Leftrightarrow\quad G_h \geq 0 \quad\Leftrightarrow\quad A^{-1} \geq 0 \quad\Leftrightarrow\quad A$ monotone
- A s.p.d., off-diag$(A) \leq 0 \quad\Rightarrow\quad A$ M-matrix $\quad\Rightarrow\quad A$ monotone
- Element matrices: off-diag$(A^K) \leq 0 \quad\Rightarrow\quad$ off-diag$(A) \leq 0$
  - $A = \sum_{K \in \mathcal{T}_h} A^K, \quad A^K_{ij} = a_K(\varphi_j, \varphi_i) = \int_K (\mathcal{A}\nabla\varphi_i) \cdot \nabla\varphi_j \,\mathrm{d}x + \cdots$

Example (triangles):

- $\alpha_{\max} \leq \pi/2 \quad\Rightarrow\quad$ off-diag$(A^K) \leq 0$
- $\alpha + \alpha' \leq \pi \quad\Leftrightarrow\quad$ off-diag$(A) \leq 0$



Linear FEM, $g_{\mathrm{D}} = 0$

- DMP $\Leftrightarrow$ $G_h \geq 0$ $\Leftrightarrow$ $A^{-1} \geq 0$ $\Leftrightarrow$ A monotone
- A s.p.d., off-diag$(A) \leq 0$ $\Rightarrow$ A M-matrix $\Rightarrow$ A monotone
- Element matrices: off-diag$(A^K) \leq 0$ $\Rightarrow$ off-diag$(A) \leq 0$
  - $A = \sum_{K \in \mathcal{T}_h} A^K, \quad A^K_{ij} = a_K(\varphi_j, \varphi_i) = \int_K (\mathcal{A} \nabla \varphi_i) \cdot \nabla \varphi_j \, \mathrm{d}x + \cdots$

Gap: A monotone but not M-matrix
$\Rightarrow$ numerical tests

Higher order FEM, $g_{\mathrm{D}} = 0$

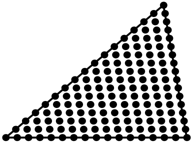
- DMP $\Leftrightarrow$ $G_h \geq 0$ $\not\Leftarrow$ $A^{-1} \geq 0$
- $G_h(x_i^{\mathrm{V}}, x_j^{\mathrm{V}}) \geq 0 \quad \Leftrightarrow \quad S^{-1} \geq 0$
  - x_i^{V}, $i = 1, \dots, N_{\mathrm{vert}}$ vertices (nodes) in $\mathcal{T}_h$
  - $S = A_{\mathrm{VV}} - A_{\mathrm{VN}} A_{\mathrm{NN}}^{-1} A_{\mathrm{NV}}$
  - $A = \begin{pmatrix} A_{\mathrm{VV}} & A_{\mathrm{VN}} \\ A_{\mathrm{NV}} & A_{\mathrm{NN}} \end{pmatrix}$
  - $A_{\mathrm{VV}} \in \mathbb{R}^{N_{\mathrm{vert}} \times N_{\mathrm{vert}}}$, $\quad A_{\mathrm{NN}} \in \mathbb{R}^{N_{\mathrm{nonv}} \times N_{\mathrm{nonv}}}$, $\quad N_{\mathrm{dof}} = N_{\mathrm{vert}} + N_{\mathrm{nonv}}$
  - *hp*-FEM basis: $\underbrace{\varphi_1^{\mathrm{V}}, \dots, \varphi_{N_{\mathrm{vert}}}^{\mathrm{V}}}_{\text{vertex fun.}}, \underbrace{\varphi_{N_{\mathrm{vert}}+1}^{\mathrm{N}}, \dots, \varphi_{N_{\mathrm{dof}}}^{\mathrm{N}}}_{\text{edge, bubble fun.}}$

Higher order FEM, $g_\mathrm{D} = 0$

- DMP $\quad \Leftrightarrow \quad G_h \geq 0 \quad \not\Leftarrow \quad A^{-1} \geq 0$
- $G_h(x_i^\mathrm{V}, x_j^\mathrm{V}) \geq 0 \quad \Leftrightarrow \quad S^{-1} \geq 0$
- $G_h|_{K\times L}(x,y) = \sum_{i\in\iota_K} \sum_{j\in\iota_L} \varphi_j|_L(y)\,(A^{-1})_{ij}\,\varphi_i|_K(x), \quad (x,y) \in K \times L$
  - $\iota_K = \{i : \mathrm{meas}(K \cap \operatorname{supp}\varphi_i) > 0\}$
    (indices of basis functions supported in K)
  - $K, L \in \mathcal{T}_h$

Higher order FEM, $g_{\mathrm{D}} = 0$

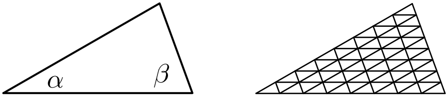
- DMP $\Leftrightarrow$ $G_h \ge 0$ $\not\Leftarrow$ $A^{-1} \ge 0$
- $G_h(x_i^{\mathrm{V}}, x_j^{\mathrm{V}}) \ge 0$ $\Leftrightarrow$ $S^{-1} \ge 0$
- $G_h|_{K\times L}(x,y) = \sum_{i\in\iota_K}\sum_{j\in\iota_L} \varphi_j|_L(y)\,(A^{-1})_{ij}\,\varphi_i|_K(x), \quad (x,y)\in K\times L$
- $G_h|_{K\times L}(x,y) \ge 0$ in $K \times L$? (cf. 17th Hilbert problem)



Numerical test

$$-\Delta u = f \quad \text{in } \Omega$$
$$u = 0 \quad \text{on } \partial\Omega$$

$\alpha = 30^\circ$
$\beta = 70^\circ$

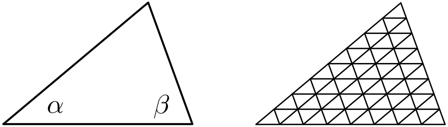


$\alpha = 1^\circ, 2^\circ, \ldots, 179^\circ \qquad \beta = 1^\circ, 2^\circ, \ldots, 179^\circ \qquad \alpha + \beta < 180^\circ$

Numerical test

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

$\alpha = 40^\circ$
$\beta = 70^\circ$

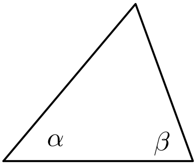


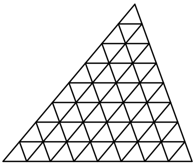
$\alpha = 1^\circ, 2^\circ, \ldots, 179^\circ \qquad \beta = 1^\circ, 2^\circ, \ldots, 179^\circ \qquad \alpha + \beta < 180^\circ$

Numerical test

$$-\Delta u = f \quad \text{in } \Omega$$
$$u = 0 \quad \text{on } \partial\Omega$$

$\alpha = 50^\circ$
$\beta = 70^\circ$



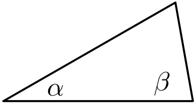


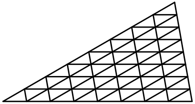
$\alpha = 1^\circ, 2^\circ, \ldots, 179^\circ$ $\quad\beta = 1^\circ, 2^\circ, \ldots, 179^\circ$ $\quad\alpha + \beta < 180^\circ$

Numerical test

$$-\Delta u = f \quad \text{in } \Omega$$
$$u = 0 \quad \text{on } \partial\Omega$$

$\alpha = 30^\circ$
$\beta = 80^\circ$



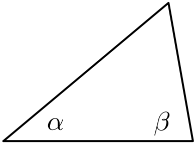


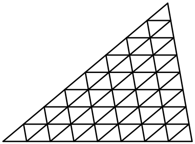
$\alpha = 1^\circ, 2^\circ, \dots, 179^\circ \qquad \beta = 1^\circ, 2^\circ, \dots, 179^\circ \qquad \alpha + \beta < 180^\circ$

Numerical test

$$-\Delta u = f \quad \text{in } \Omega$$
$$u = 0 \quad \text{on } \partial\Omega$$

$\alpha = 40^\circ$
$\beta = 80^\circ$



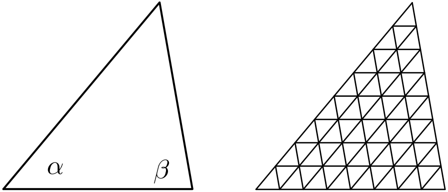


$\alpha = 1^\circ, 2^\circ, \dots, 179^\circ \qquad \beta = 1^\circ, 2^\circ, \dots, 179^\circ \qquad \alpha + \beta < 180^\circ$

Numerical test

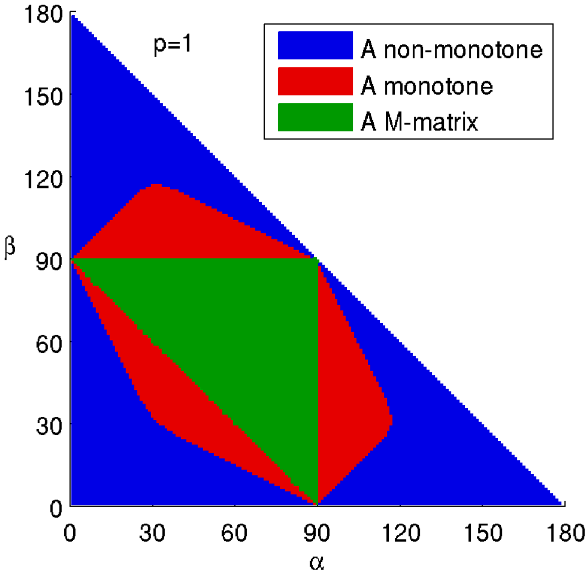
$$-\Delta u = f \quad \text{in } \Omega$$
$$u = 0 \quad \text{on } \partial\Omega$$

$\alpha = 50^\circ$
$\beta = 80^\circ$

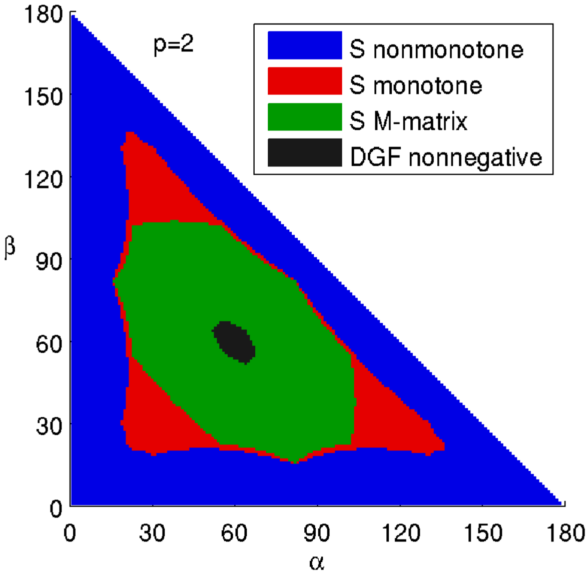


$\alpha = 1^\circ, 2^\circ, \ldots, 179^\circ \qquad \beta = 1^\circ, 2^\circ, \ldots, 179^\circ \qquad \alpha + \beta < 180^\circ$

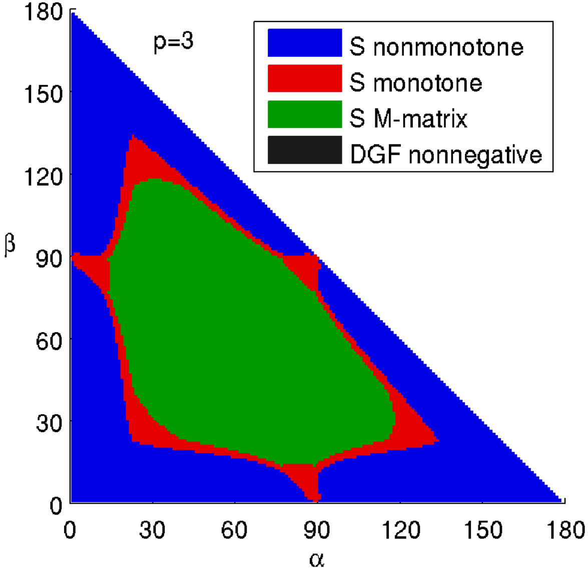
Numerical test – $p = 1$



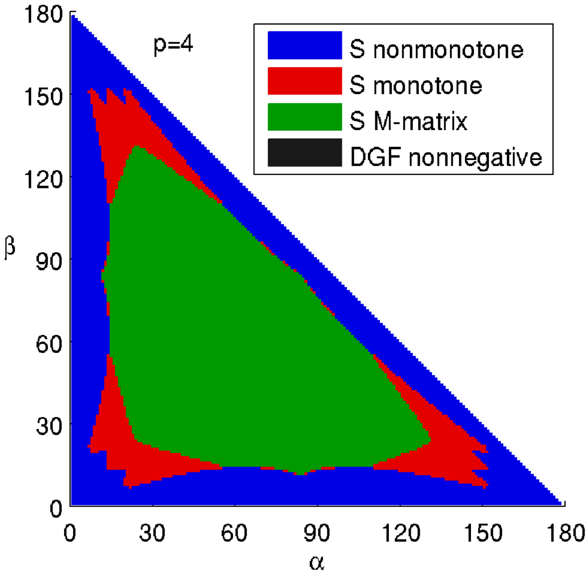
Numerical test – $p = 2$



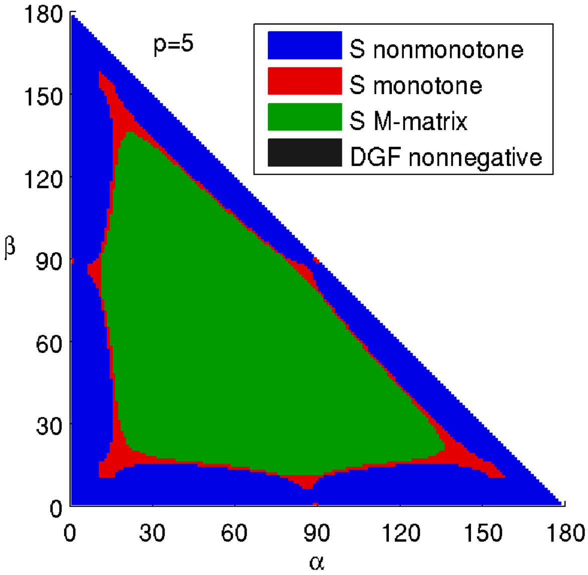
Numerical test – $p = 3$



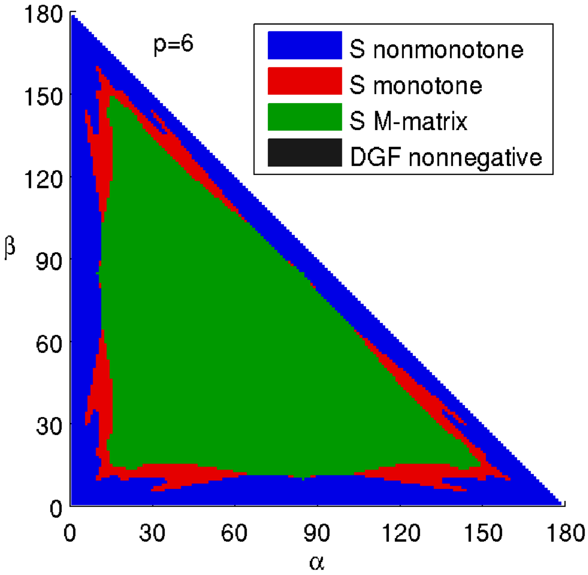
Numerical test – $p = 4$



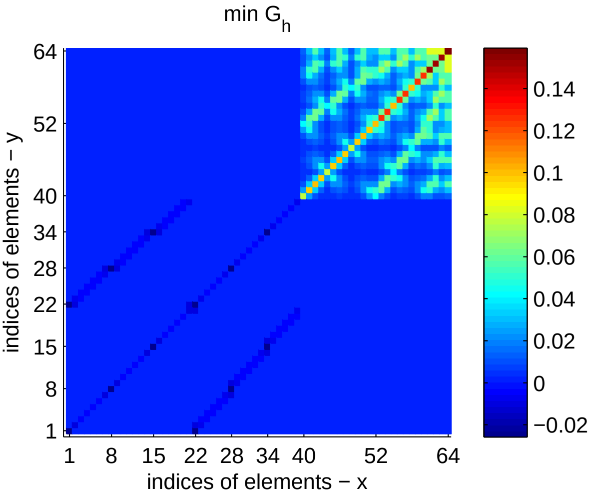
Numerical test – $p = 5$



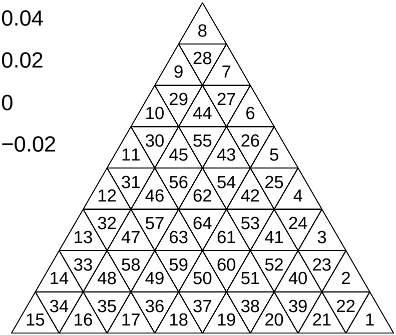
Numerical test – $p = 6$



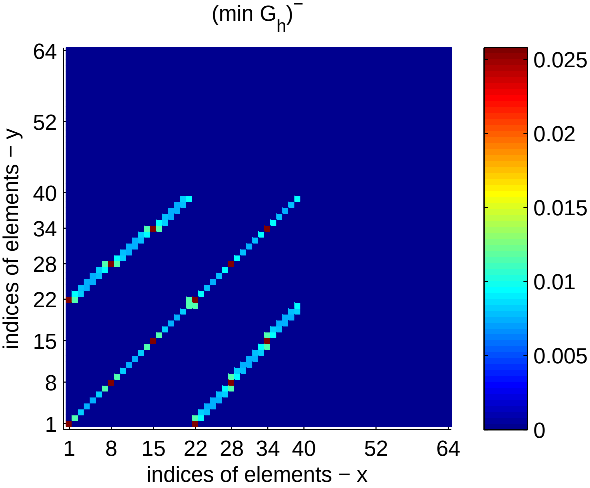
Visualization of DGF



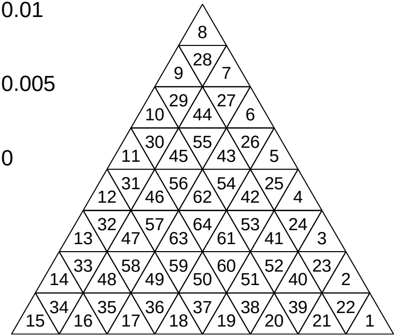
$\alpha = 60°$
$\beta = 60°$
$p = 3$



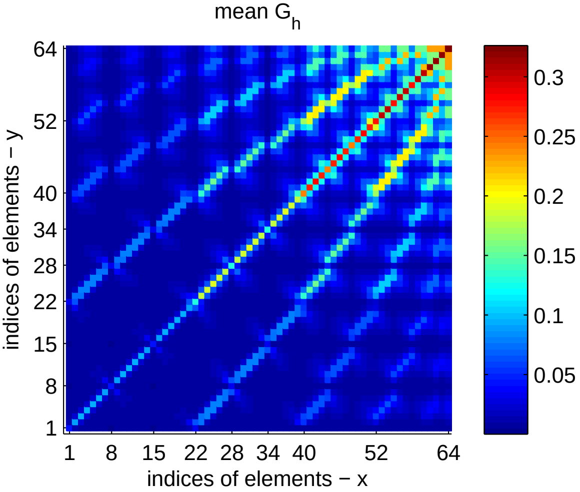
Visualization of DGF



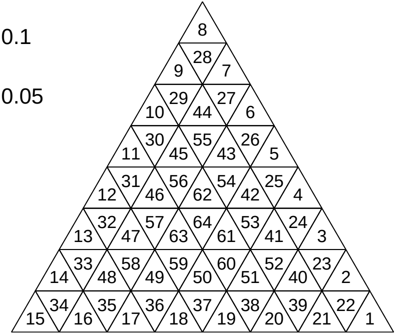
$\alpha = 60^\circ$
$\beta = 60^\circ$
$p = 3$



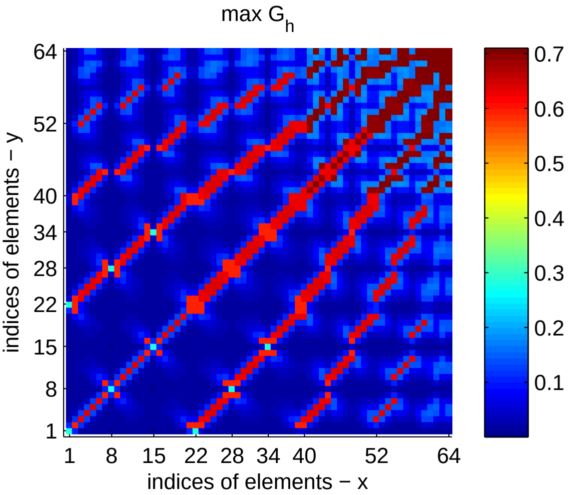
Visualization of DGF



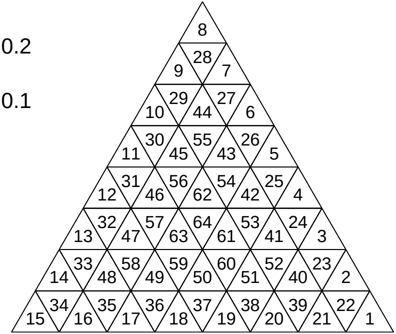
$\alpha = 60^\circ$
$\beta = 60^\circ$
$p = 3$



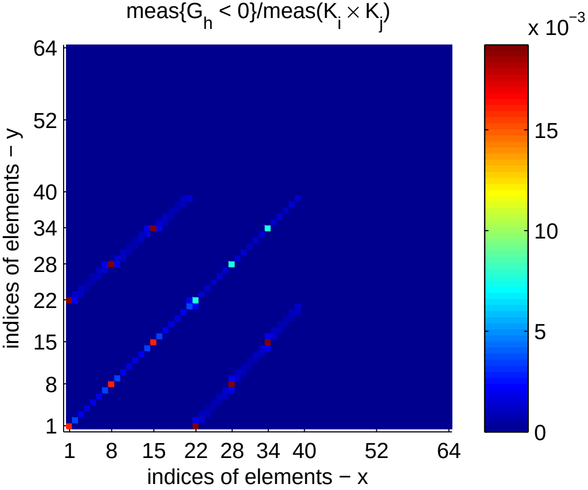
Visualization of DGF



$\alpha = 60^\circ$
$\beta = 60^\circ$
$p = 3$



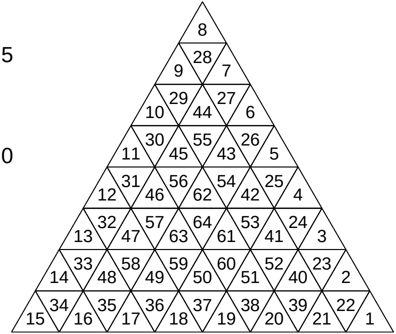
Visualization of DGF



$\alpha = 60^\circ$

$\beta = 60^\circ$

$p = 3$



Conclusions

- $G_h \not\geq 0$ on uniform triangular meshes for $p \geq 3$
- $G_h \geq 0$ for $p \geq 2$ for triangles close to equilateral
- $G_h < 0$ close to the boundary
- $\text{meas}\{(x,y) : G_h(x,y) < 0\}$ is small
- $|\min G_h| << |\max G_h|$
- $f \geq 0$ such that $u_h \not\geq 0$ is weird
- If f is well approximated on $\mathcal{T}_h$ then $u_h \geq 0$.
- Polynomials not suitable – try sin and cos

Thank you for your attention

Tomáš Vejchodský (vejchod@math.cas.cz)

Institute of Mathematics, Academy of Sciences
Žitná 25, 115 67 Praha 1
Czech Republic

November 4, 2009, FSv ČVUT, Praha