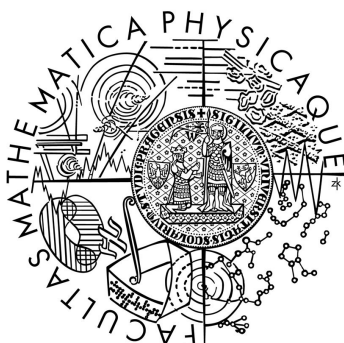


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Bc. Jiří Kohutka

Volba parametru metody SUPG pro konečné prvky vyššího řádu přesnosti

Katedra numerické matematiky

Vedoucí diplomové práce: doc. Mgr. Petr Knobloch, Dr.

Studijní program: Matematika

Studijní obor: Matematické modelování ve fyzice a technice

Praha 2013

Poděkování: Především děkuji vedoucímu této práce doc. Mgr. Petru Knoblochovi, Dr., a to jednak za poskytnutí studijních materiálů a jednak za četné konzultace. Dále za konzultace děkuji prof. Ing. Zdeňku Strakošovi, DrSc., doc. RNDr. Petru Lachoutovi, CSc., RNDr. Václavu Kučerovi, Ph.D., prof. RNDr. Josefu Málkovi, CSc., DSc. a prof. RNDr. Miloslavu Feistauerovi, DrSc., dr. h. c. Speciálně za velkorysost, díky které jsem mohl tvořit tuto diplomovou práci i při zaměstnání, děkuji Ing. Otakaru Frýbortovi.

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V Praze dne 6. 12. 2013

Podpis autora

Název práce: Volba parametru metody SUPG pro konečné prvky vyššího řádu přesnosti

Autor: Bc. Jiří Kohutka

Katedra: Katedra numerické matematiky

Vedoucí diplomové práce: doc. Mgr. Petr Knobloch, Dr., Katedra numerické matematiky

Abstrakt:

V této práci se zabýváme metodou konečných prvků Streamline Upwind/Petrov-Galerkin (SUPG) a používáme ji k řešení stacionární okrajové úlohy pro rovnici konvekce-difuze s převažující konvekcí s Dirichletovou okrajovou podmínkou na celé hranici omezené polyedrické výpočetní oblasti dimenze 1 resp. 2. Uvažujeme lagrangeovské kvadratické konečné prvky na úsečkách resp. trojúhelnících. Jádrem práce je návrh volit stabilizační parametr metody SUPG ve výtokové hraniční vstvě jako funkci afinní po elementech a ve zbytku výpočetní oblasti jako funkci konstantní po elementech. Ukážeme, že tato volba dává přesnější řešení než volba stabilizačního parametru konstantního na všech elementech.

Klíčová slova: konvekce-difuze, Streamline Upwind/Petrov-Galerkin, SUPG, MKP, nefyzikální oscilace, mezní vrstva

Title: Choice of parameter of SUPG method for finite elements of higher order of accuracy

Author: Bc. Jiří Kohutka

Department: Department of Numerical Mathematics

Supervisor: doc. Mgr. Petr Knobloch, Dr., Department of Numerical Mathematics

Abstract:

In this work, we deal with the finite element method Streamline Upwind/Petrov-Galerkin (SUPG) and use it to solve boundary value problem for the stationary convection-diffusion equation with dominant convection with Dirichlet boundary condition on the whole boundary of bounded polyhedral computational domain of dimension 1 and 2, respectively. We consider a quadratic Lagrangian finite elements on the line segments and triangles, respectively. The core of the work is a proposition of choice of stabilizing parameter of SUPG method as an elementwise affine function in outflow boundary layer and as an elementwise constant function in the rest of the computational domain. We show that this choice gives a more accurate solution than the choice of the stabilization parameter as a constant in each element.

Keywords: convection-diffusion, Streamline Upwind/Petrov-Galerkin, SUPG, FEM, nonphysical oscillations, boundary layer

Obsah

Úvod	2
1 Formulace problému	4
1.1 Značení	5
1.2 Klasická formulace	6
1.3 Slabá formulace	6
1.4 Diskretizace	7
1.5 Galerkinova formulace	16
1.6 SUPG formulace	16
1.7 Galerkinova a SUPG formulace v maticovém tvaru	18
1.8 Existence a jednoznačnost řešení SUPG formulace	19
1.9 Odhad chyby řešení SUPG formulace	21
2 Situace pro $d = 1$	32
2.1 Modelový problém – klasická a Galerkinova formulace	32
2.2 Modelový problém – SUPG formulace, volba stabilizačního parametru . . .	35
2.2.1 Volba stabilizačního parametru metodou C	36
2.2.2 Volba stabilizačního parametru metodou 1	36
2.2.3 Volba stabilizačního parametru metodou 2	37
2.2.4 Rozdíl mezi metodou 1 a 2	38
2.3 Modelový problém – numerické výsledky	39
2.4 Obecný problém – volba stabilizačního parametru	41
2.5 Další numerické výsledky	42
3 Situace pro $d = 2$	45
3.1 Modelový problém – klasická a Galerkinova formulace	45
3.2 Obecný problém – SUPG formulace, volba stabilizačního parametru	49
3.2.1 Volba stabilizačního parametru metodou C	52
3.2.2 Volba stabilizačního parametru metodou 1	52
3.2.3 Volba stabilizačního parametru metodou 2	56
3.3 Modelový problém – numerické výsledky	57
3.4 Další numerické výsledky	69
Závěr	73
Seznam použité literatury	74

Úvod

Rovnice typu konvekce-difuze popisuje například přenos nějaké veličiny proudící tekutinou. Tou veličinou může být například koncentrace nějaké látky rozpuštěné v tekutině nebo teplota tekutiny. Obecný stacionární matematický model je formulován rovnicí

$$-\varepsilon\Delta u + b \cdot \nabla u = f, \quad \square$$

kde ε , b a f jsou zadané parametry a u je neznámá funkce. Parametr ε udává míru difuze, tj. procesu, při kterém je přenos u dán gradientem u , a parametr b popisuje konvekci, tj. proces, při kterém je přenos u dán rychlostí pohybu proudící tekutiny.

Je-li $\|b\|$ mnohem větší než ε , převažuje konvekce nad difuzí a v přibližných řešeních rovnice \square získaných Galerkinovou metodou konečných prvků se pak vyskytují nefyzikální oscilace.

V této práci použijeme k řešení rovnice \square metodou konečných prvků místo metody Galerkinovy metodu Streamline Upwind/Petrovovu-Galerkinovu, označovanou zkratkou SUPG. Tato metoda je popsána například v pracích [1], [3], [4], [5], [6], [7], [8], [10] a [13]. Metodu popíšeme i v této práci. Budeme pracovat s lagrangeovskými kvadratickými konečnými prvky, a to v dimenzi 1 na úsečkách a v dimenzi 2 na trojúhelnících. Metoda SUPG obsahuje tzv. *stabilizační parametr*, jehož vhodnou volbou lze dosáhnout toho, že tato metoda dává takové přibližné řešení rovnice \square , ve kterém jsou výrazně potlačeny oscilace zmíněné výše.

Například v pracích [5], [6], [7], [8], [9], [10], [11], [12] a [13] je stabilizační parametr volen alespoň jedním z následujících způsobů:

- buď jako funkce definovaná na množině výpočetních uzlů, tedy nějaká konstantní hodnota stabilizačního parametru je přiřazena každému výpočetnímu uzlu,
- anebo jako funkce definovaná na všech elementech, která je konstantní na každém elementu, tedy nějaká konstantní hodnota stabilizačního parametru je přiřazena každému elementu.

My budeme stabilizační parametr volit jako funkci definovanou na všech elementech, a to obecně afinní na každém elementu. Při tom se budeme inspirovat prací [13], ve které je stabilizační parametr volen ve výtokové hraniční vrstvě elementů odlišným způsobem než na zbývajících elementech.

V práci jsou předvedeny výsledky několika numerických experimentů, v nichž jsou použity navrhované metody volby stabilizačního parametru. Pro získání těchto výsledků byly v jazyce C vytvořeny počítačové programy, ty jsou přiloženy na CD – nutno podotknout, že programy jsou pouze v pracovní verzi. Výsledky byly zobrazeny programem Lines 3D 1.4.3 a pro tisk byly editovány v programech MWSnap 3.0.0.74 a GIMP 2.8.4. Výsledky jsou zhodnoceny v Závěru.

Grafy všech veličin vyobrazené v této práci jsou grafy pouze lineárních interpolací hodnot těchto veličin ve vrcholech a středech hran, přičemž tyto hodnoty jsou interpolovány pouze na hranách triangulace.

Pro připomínky, dotazy či konzultace je možné spojit se s autorem této práce prostřednictvím e-mailové adresy jirikohutka@seznam.cz.

1. Formulace problému

1.1 Značení

Uvažujeme následující pojmy a jejich označení:

Označení:	Pojem:
$d \in \{1, 2\}$	dimenze
$\Omega \subset \mathbb{R}^d$	výpočetní oblast = otevřená souvislá množina
$\partial\Omega$	hranice výpočetní oblasti
$x_{[i]}, i \in \{1, 2, \dots, d\}$	i -tá složka prvku $x \in \mathbb{R}^d$
$\ x\ = \left(\sum_{i=1}^d x_{[i]}^2 \right)^{1/2}$	euklidovská norma prvku $x \in \mathbb{R}^d$
$\varepsilon \in \mathbb{R}, \varepsilon > 0$	difuzivita
$b : \bar{\Omega} \rightarrow \mathbb{R}^d$	vektor konvekce
$f : \Omega \rightarrow \mathbb{R}$	pravá strana
$u^B : \partial\Omega \rightarrow \mathbb{R}$	funkce Dirichletovy okrajové podmínky
$\nu : \partial\Omega \rightarrow \mathbb{R}^d$	jednotková vnější normála k $\partial\Omega$
$\text{int}(Y)$	vnitřek množiny $Y \subset \mathbb{R}^d$
$\text{relint}(Y)$	relativní vnitřek množiny $Y \subset \mathbb{R}^d$
$\text{meas}(Y)$	Lebesgueova míra množiny $Y \subset \mathbb{R}^d$
$\text{card}(Y)$	mohutnost množiny $Y \subset \mathbb{R}^d$
HN	Hölderova nerovnost
YN	Youngova nerovnost
FV	Fubiniova věta
VS	věta o substituci
PP	per partes
GV	Greenova věta
$\frac{DF}{Dx}$	Jacobiova matice zobrazení $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$
$D^\alpha u$	α -tá parciální derivace funkce u , kde α je multiindex
$H^k(Y)$	prostor funkcí z $L^2(Y)$, $Y \subset \mathbb{R}^d$, jejichž slabé derivace až do řádu k včetně jsou prvky $L^2(Y)$, tj. Sobolevův prostor $W^{k,2}(Y)$
$\ u\ _{k,Y} := \left(\sum_{ \alpha \leq k} \int_Y (D^\alpha u(x))^2 dx \right)^{1/2}$	norma funkce $u \in H^k(Y)$
$ u _{k,Y} := \left(\sum_{ \alpha =k} \int_Y (D^\alpha u(x))^2 dx \right)^{1/2}$	seminorma funkce $u \in H^k(Y)$
$P_k(Y)$	prostor polynomů stupně k definovaných na množině $Y \subset \mathbb{R}^d$

Bude-li z kontextu zřejmé, že $d = 1$, budeme používat značení $x = x_{[1]}$ a $|x| = \|x\|$. Dále pro měřitelnou funkci $u : Y \subset \mathbb{R}^d \rightarrow \mathbb{R}$ budeme používat značení

$$\int_Y u = \int_Y u(x) dx.$$

Uvažujeme Ω s lipschitzovskou hranicí. Předpokládáme, že Ω je interval, je-li $d = 1$, a mnohoúhelník, je-li $d = 2$. Vzhledem k tomu je potřeba dodefinovat jednotkovou vnější normálu ve vrcholech Ω , je-li $d = 2$.

Buď tedy $d = 2$ a $x \in \partial\Omega$ vrchol oblasti Ω a $E_1, E_2 \subset \partial\Omega$ dvě hrany takové, že $x \in E_1 \wedge x \in E_2$. Buďte dále $x_{E_1} \in \text{relint}(E_1), x_{E_2} \in \text{relint}(E_2)$ libovolné body. V nich je jednotková vnější normála k $\partial\Omega$ dobře definována, můžeme ji tedy pomocí nich definovat v bodě x . To provedeme následujícím způsobem.

$$\nu(x) := \frac{1}{\|\nu(x_{E_1}) + \nu(x_{E_2})\|} (\nu(x_{E_1}) + \nu(x_{E_2})). \quad (1.1)$$

Dále předpokládáme nestlačitelnost tekutiny, tj.

$$\text{div } b = 0 \text{ v } \Omega. \quad (1.2)$$

1.2 Klasická formulace

Klasická formulace stacionární úlohy pro rovnici konvekce-difuze s Dirichletovou okrajovou podmínkou na celé $\partial\Omega$ je následující:

Úloha 1: Najít $u \in C^2(\Omega) \cap C(\bar{\Omega})$ splňující

$$-\varepsilon \Delta u + b \cdot \nabla u = f \text{ v } \Omega, \quad (1.3)$$

$$u = u^B \text{ na } \partial\Omega.$$

1.3 Slabá formulace

Nyní odvodíme slabou formulaci. Rovnici (1.3) vynásobíme libovolnou testovací funkcí $v \in C_0^\infty(\Omega)$, zintegrujeme přes Ω a použijeme linearitu integrálu. Dostaneme

$$-\varepsilon \underbrace{\int_{\Omega} (\Delta u) v}_{I_1} + \int_{\Omega} (b \cdot \nabla u) v = \int_{\Omega} f v \quad \forall v \in C_0^\infty(\Omega). \quad (1.4)$$

Na integrál I_1 použijeme Greenovu větu s uvážením, že funkce v má z definice prostoru $C_0^\infty(\Omega)$ nulovou hodnotu na $\partial\Omega$. Rovnice (1.4) tak přejde na tvar

$$\varepsilon \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (b \cdot \nabla u) v = \int_{\Omega} f v \quad \forall v \in C_0^\infty(\Omega). \quad (1.5)$$

Definujeme množinu

$$U := \{u \in H^1(\Omega) : u = u^B \text{ na } \partial\Omega\},$$

ve které budeme hledat řešení, a prostor

$$V := \{v \in H^1(\Omega) : v = 0 \text{ na } \partial\Omega\},$$

což je prostor testovacích funkcí. Uvažujeme

$$b \in L^\infty(\Omega), \quad f \in L^2(\Omega).$$

Definujeme bilineární formu

$$a_G : H^1(\Omega) \times H^1(\Omega) \longrightarrow \mathbb{R}, \quad a_G(u, v) := \varepsilon \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (b \cdot \nabla u)v \quad (1.6)$$

a lineární funkcional

$$F_G : L^2(\Omega) \longrightarrow \mathbb{R}, \quad F_G(v) := \int_{\Omega} f v. \quad (1.7)$$

V rovnici (1.5) přejdeme k funkcím $u \in U$ a $v \in V$ a napíšeme ji pomocí definic (1.6) a (1.7). Dostáváme tak slabou formulaci úlohy 1.

Úloha 2: Najít $u \in U$ splňující

$$a_G(u, v) = F_G(v) \quad \forall v \in V. \quad (1.8)$$

1.4 Diskretizace

Uvažujeme *parametr triangulace* $h > 0$, *množinu vrcholů triangulace* N_T a *triangulaci* \mathcal{T}_h , jejímiž elementy jsou d -simplexy. Pro každý element $K \in \mathcal{T}_h$ označíme h_K resp. ρ_K průměr elementu K resp. průměr koule vepsané elementu K . Předpokládáme, že existuje $\sigma > 0$ takové, že pro každý element $K \in \mathcal{T}_h$ je splněno

$$\frac{h_K}{\rho_K} \leq \sigma.$$

Uvažujeme

$$h = \max_{K \in \mathcal{T}_h} h_K. \quad (1.9)$$

Vrcholy daného d -simplexu K budeme značit $a_{K,j}$, $j \in \mathcal{J}$, $\text{card}(\mathcal{J}) = d+1$. Přiřadíme-li vrcholům daného d -simplexu K indexy $j = 1, 2, \dots, d+1$, definujeme *těžiště d -simplexu K*

$$c_K := \frac{1}{d+1} \sum_{j=1}^{d+1} a_{K,j}. \quad (1.10)$$

Budeme se zabývat lagrangeovskými kvadratickými prvky, proto jako N_S označíme *množinu středů hran triangulace* \mathcal{T}_h . Nejprve pro definici zvolíme indexy vrcholů daného elementu $K \in \mathcal{T}_h$ jako $1, 2, \dots, d+1$. Pak definujeme

$$N_S := \bigcup_{K \in \mathcal{T}_h} \left\{ \frac{1}{2}(a_{K,i} + a_{K,j}), i, j \in \{1, 2, \dots, d+1\}, i \neq j \right\},$$

Dále uvažujeme *množinu výpočetních uzlů*

$$N_N := N_T \cup N_S.$$

Výpočetní uzel budeme značit standardně n_j .

Množina N_N je konečná, označíme

$$\mathcal{N}_h := \text{card}(N_N).$$

Bodům z N_N přiřadíme indexy z množiny

$$J := \{0, 1, \dots, \mathcal{N}_h - 1\}.$$

Je tedy

$$N_N = \{n_j \in \bigcup_{K \in \mathcal{T}_h} \partial K, j \in J\}.$$

Buď $Y \subset \bar{\Omega}$ libovolná. Definujeme množinu indexů výpočetních uzlů náležících do Y ,

$$J(Y) := \{j \in J : n_j \in N_N \cap Y\}.$$

Platí

$$J(\Omega) \cup J(\partial\Omega) = J.$$

Pro vektor $w \in \mathbb{R}^2$, $\|w\| \neq 0$, a element $K \in \mathcal{T}_h$ definujeme h_K^w jako *průměr elementu* K *ve směru vektoru* w , tj.

$$h_K^w := \max\{|x - y| : x, y \in \partial K \wedge \exists c \in \mathbb{R} \setminus \{0\} : y = x + cw\}.$$

Dále definujeme množinu elementů obsahujících daný výpočetní uzel,

$$\mathcal{T}_h^j := \{K \in \mathcal{T}_h : n_j \in K\}, \quad j \in J,$$

Uvažujeme prostor konečných prvků

$$X_h := \{u_h \in L^2(\Omega) : u_h \in P_2(K) \forall K \in \mathcal{T}_h \wedge \\ \wedge \forall n_j \in N_N \forall K, T \in \mathcal{T}_h^j : u_h|_K(n_j) = u_h|_T(n_j)\}.$$

Jako aproximaci množiny U uvažujeme množinu

$$U_h := \{u_h \in X_h : \forall n_j \in N_N \cap \partial\Omega : u_h(n_j) = u^B(n_j)\}$$

a jako aproximaci prostoru V uvažujeme konečně rozměrný prostor

$$V_h := \{v_h \in X_h : v_h = 0 \text{ na } \partial\Omega\}.$$

Platí inkluze $V_h \subset V, V_h \subset X_h, U_h \subset X_h$.

Nadále budeme uvažovat, že $\delta_{ij}, i, j \in \mathbb{N}$, je Kroneckerovo delta, tedy že

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad i, j \in \mathbb{N}.$$

Uvažujeme báze prostorů X_h a V_h

$$\begin{aligned} \text{base}(X_h) &:= \{\varphi_j \in X_h : \varphi_j(n_k) = \delta_{jk}, j, k \in J\}, \\ \text{base}(V_h) &:= \{\varphi_i \in \text{base}(X_h) : i \in J(\Omega)\}. \end{aligned} \quad (1.11)$$

Platí inkluze $\text{base}(V_h) \subset \text{base}(X_h)$.

Buď $d \in \mathbb{N}$ a K libovolný d -simplex s vrcholy $a_{K,j}, j \in \mathcal{J}, \text{card}(\mathcal{J}) = d + 1$. Pak *barycentrické souřadnice na d -simplexu K příslušné jeho vrcholům* jsou funkce

$$\lambda_{K,j} : K \in \mathcal{T}_h \longrightarrow [0, 1], \quad j \in \mathcal{J}, \quad (1.12)$$

splňující

$$\begin{aligned} \lambda_{K,j} &\in P_1(K) && \forall j \in \mathcal{J}, \\ \lambda_{K,i}(a_{K,j}) &= \delta_{ij} && \forall i, j \in \mathcal{J}. \end{aligned}$$

Buď dán d -simplex K s vrcholy s indexy $1, 2, \dots, d + 1$. Platí

$$\sum_{i=1}^{d+1} \lambda_{K,i} = 1. \quad (1.13)$$

Z (1.13) plyne

$$\sum_{i=1}^{d+1} \nabla \lambda_{K,i} = \underbrace{(0, 0, \dots, 0)}_{d \text{ složek}}. \quad (1.14)$$

Pomocí barycentrických souřadnic vyjádříme prvky $\text{base}(X_h)$ a jejich gradienty na libovolném elementu $K \in \mathcal{T}_h$.

	$n_j \in N_T \cap K$	$n_j \in N_S \cap K \wedge$ $\wedge \exists n_k, n_l \in N_T \cap K : n_j = (1/2)(n_k + n_l)$	
$\varphi_j _K =$	$2\lambda_{K,j}^2 - \lambda_{K,j}$	$4\lambda_{K,k}\lambda_{K,l}$	(1.15)
$\nabla \varphi_j _K =$	$(4\lambda_{K,j} - 1)\nabla \lambda_{K,j}$	$4\lambda_{K,k}\nabla \lambda_{K,l} + 4\lambda_{K,l}\nabla \lambda_{K,k}$	

Platí

$$\sum_{j \in J(K)} \varphi_j|_K = 1 \quad \forall K \in \mathcal{T}_h,$$

z čehož plyne

$$\sum_{j \in J(K)} \nabla \varphi_j|_K = \underbrace{(0, 0, \dots, 0)}_{d \text{ složek}} \quad \forall K \in \mathcal{T}_h.$$

Definujeme *referenční element dimenze d*.

$$\widehat{K} := \{\widehat{x} \in \mathbb{R}^d : 0 \leq \widehat{x}_{[i]} \leq 1, i \in \{1, 2, \dots, d\}, \sum_{i=1}^d \widehat{x}_{[i]} \leq 1\}. \quad (1.16)$$

Platí

$$\widehat{K} := \{\widehat{x} \in \mathbb{R}^d : \widehat{x}_{[1]} \in [0, 1], \widehat{x}_{[i]} \in [0, 1 - m_i(\widehat{x})], i \in \{2, 3, \dots, d\}\}, \quad (1.17)$$

kde

$$m_i(\widehat{x}) = \sum_{j=1}^{i-1} \widehat{x}_{[j]}, \quad (1.18)$$

a dále pro vrcholy referenčního elementu, jimž teď přiřadíme indexy $1, 2, \dots, d+1$, platí

$$a_{\widehat{K},j} := \underbrace{(0, 0, \dots, 0, \overbrace{1}^{j\text{-tá složka}}, 0, \dots, 0)}_{d \text{ složek}}, \quad j \in \{1, 2, \dots, d\}, \quad (1.19)$$

$$a_{\widehat{K},d+1} := \underbrace{(0, 0, \dots, 0)}_{d \text{ složek}}.$$

Pro daný element $K \in \mathcal{T}_h$ uvažujeme *afinní regulární zobrazení*

$$\mathcal{F}_K : \widehat{K} \xrightarrow{\text{na}} K. \quad (1.20)$$

Protože \mathcal{F} je afinní a regulární, platí

$$\det \frac{D\mathcal{F}_K}{D\widehat{x}} = c \neq 0 \quad \forall K \in \mathcal{T}_h, \quad (1.21)$$

takže je

$$\text{meas}(K) = \int_K 1 dx \stackrel{VS}{=} \int_{\mathcal{F}_K^{-1}(K)} \left| \det \frac{D\mathcal{F}_K}{D\widehat{x}} \right| d\widehat{x} \stackrel{(1.21)}{=} \left| \det \frac{D\mathcal{F}_K}{D\widehat{x}} \right| \int_{\widehat{K}} 1 d\widehat{x} = \left| \det \frac{D\mathcal{F}_K}{D\widehat{x}} \right| \text{meas}(\widehat{K}),$$

tedy

$$\left| \det \frac{D\mathcal{F}_K}{D\widehat{x}} \right| = \frac{\text{meas}(K)}{\text{meas}(\widehat{K})}. \quad (1.22)$$

Dokážeme následující pomocné lemma.

Lemma 3: Pro libovolná $p, q \in \mathbb{N}_0$ a $c > 0$ platí

$$\underbrace{\int_0^c \xi^p (c - \xi)^q d\xi}_{I_5} = \frac{c^{p+q+1}}{(p+q+1)!} p! q!.$$

Důkaz:

- $q = 0$

$$I_5 = \left[\frac{\xi^{p+1}}{p+1} \right]_{\xi=0}^c = \frac{c^{p+1}}{p+1} = \frac{c^{p+q+1}}{(p+q+1)!} p! q!.$$

- $q \neq 0$

$$\begin{aligned} I_5 &\stackrel{PP}{=} \left[\frac{\xi^{p+1}}{p+1} (c - \xi)^q \right]_{\xi=0}^c + \int_0^c \frac{\xi^{p+1}}{p+1} q (c - \xi)^{q-1} d\xi = \\ &\stackrel{PP}{=} \left[\frac{\xi^{p+2}}{(p+2)(p+1)} q (c - \xi)^{q-1} \right]_{\xi=0}^c + \int_0^c \frac{\xi^{p+2}}{(p+2)(p+1)} q (q-1) (c - \xi)^{q-2} d\xi = \\ &\stackrel{PP}{=} \dots = \left[\frac{\xi^{p+q}}{\prod_{i=1}^q (p+i)} \prod_{j=2}^q j (c - \xi) \right]_{\xi=0}^c + \int_0^c \frac{\xi^{p+q}}{\prod_{i=1}^q (p+i)} q! d\xi = \frac{c^{p+q+1}}{(p+q+1)!} p! q!. \end{aligned}$$

■

Dále platí následující Lemma.

Lemma 4: Buď dána libovolná dimenze $d \in \mathbb{N}$. Pak platí

$$\text{meas}(\widehat{K}) = \frac{1}{d!}. \tag{1.23}$$

Důkaz:

$$\begin{aligned}
\text{meas}(\widehat{K}) &= \int_{\widehat{K}} 1 d\widehat{x} \stackrel{\text{FV, (1.17)}}{=} \\
&= \int_0^1 \int_0^{1-m_2(\widehat{x})} \dots \\
\dots & \int_0^{1-m_{d-3}(\widehat{x})} \int_0^{1-m_{d-2}(\widehat{x})} \int_0^{1-m_{d-1}(\widehat{x})} \int_0^{1-m_d(\widehat{x})} 1 d\widehat{x}_{[d]} d\widehat{x}_{[d-1]} d\widehat{x}_{[d-2]} d\widehat{x}_{[d-3]} \dots d\widehat{x}_{[2]} d\widehat{x}_{[1]} = \\
&= \int_0^1 \int_0^{1-m_2(\widehat{x})} \dots \\
\dots & \int_0^{1-m_{d-3}(\widehat{x})} \int_0^{1-m_{d-2}(\widehat{x})} \int_0^{1-m_{d-1}(\widehat{x})} (1 - m_d(\widehat{x})) d\widehat{x}_{[d-1]} d\widehat{x}_{[d-2]} d\widehat{x}_{[d-3]} \dots d\widehat{x}_{[2]} d\widehat{x}_{[1]} \stackrel{(1.18)}{=}
\end{aligned}$$

$$\begin{aligned}
&= \int_0^1 \int_0^{1-m_2(\hat{x})} \dots \\
\dots &\int_0^{1-m_{d-3}(\hat{x})} \int_0^{1-m_{d-2}(\hat{x})} \int_0^{1-m_{d-1}(\hat{x})} \left((1-m_{d-1}(\hat{x})) - \hat{x}_{[d-1]} \right) d\hat{x}_{[d-1]} d\hat{x}_{[d-2]} d\hat{x}_{[d-3]} \dots d\hat{x}_{[2]} d\hat{x}_{[1]} =
\end{aligned}$$

lemma 3,
 $p := 0, q := 1,$
 $c := 1 - m_{d-1}(\hat{x}),$
 $\xi := \hat{x}_{[d-1]}$

$$\underbrace{\hspace{10em}}_{=} \frac{1}{2} \int_0^1 \int_0^{1-m_2(\hat{x})} \dots$$

$$\dots \int_0^{1-m_{d-3}(\hat{x})} \int_0^{1-m_{d-2}(\hat{x})} (1-m_{d-1}(\hat{x}))^2 d\hat{x}_{[d-2]} d\hat{x}_{[d-3]} \dots d\hat{x}_{[2]} d\hat{x}_{[1]} \underbrace{\hspace{2em}}_{=} \tag{1.18}$$

$$= \frac{1}{2} \int_0^1 \int_0^{1-m_2(\hat{x})} \dots \int_0^{1-m_{d-3}(\hat{x})} \int_0^{1-m_{d-2}(\hat{x})} \left((1-m_{d-2}(\hat{x})) - \hat{x}_{[d-2]} \right)^2 d\hat{x}_{[d-2]} d\hat{x}_{[d-3]} \dots d\hat{x}_{[2]} d\hat{x}_{[1]} =$$

lemma 3,
 $p := 0, q := 2,$
 $c := 1 - m_{d-2}(\hat{x}),$
 $\xi := \hat{x}_{[d-2]}$

$$\underbrace{\hspace{10em}}_{=} \frac{1}{2 \cdot 3} \int_0^1 \int_0^{1-m_2(\hat{x})} \dots \int_0^{1-m_{d-3}(\hat{x})} (1-m_{d-2}(\hat{x}))^3 d\hat{x}_{[d-3]} \dots d\hat{x}_{[2]} d\hat{x}_{[1]} =$$

$$= \dots = \frac{1}{(d-1)!} \int_0^1 (1-m_2(\hat{x}))^{d-1} d\hat{x}_{[1]} \underbrace{\hspace{2em}}_{=} \tag{1.18}$$

lemma 3,
 $p := 0, q := d-1,$
 $c := 1, \xi := \hat{x}_{[1]}$

$$= \frac{1}{(d-1)!} \int_0^1 (1-\hat{x}_{[1]})^{d-1} d\hat{x}_{[1]} \underbrace{\hspace{2em}}_{=} \frac{1}{d!}$$

■

Barycentrické souřadnice na referenčním elementu \widehat{K} příslušné jeho vrcholům, tj. bodům (1.19), lze vyjádřit vzorci

$$\begin{aligned}\lambda_{\widehat{K},i}(\widehat{x}) &= \widehat{x}_{[i]}, & i \in \{1, 2, \dots, d\}, \\ \lambda_{\widehat{K},d+1}(\widehat{x}) &= 1 - m_{d+1}(\widehat{x}).\end{aligned}\tag{1.24}$$

Dále platí následující věta.

Věta 5: Buď dána libovolná dimenze $d \in \mathbb{N}$, buď K_d libovolný d -simplex s indexy svých vrcholů $1, 2, \dots, d+1$ a buďte $\alpha_i \in \mathbb{N} \cup \{0\}, i \in \{1, 2, \dots, d+1\}$. Pak

$$\int_{K_d} \prod_{i=1}^{d+1} \lambda_{K_d,i}^{\alpha_i} = \frac{\prod_{i=1}^{d+1} (\alpha_i!) d!}{\left(\sum_{i=1}^{d+1} \alpha_i + d\right)!} \text{meas}(K_d).\tag{1.25}$$

Důkaz:

Referenční element \widehat{K} z definice (1.16) nyní označíme \widehat{K}_d , protože budeme postupovat indukcí podle dimenze.

Platí

$$\text{levá strana (1.25)} = \int_{K_d} \prod_{i=1}^{d+1} \lambda_{K_d,i}^{\alpha_i} \underbrace{\text{VS, (1.22), (1.23)}}_{=} \text{meas}(K_d) d! \int_{\widehat{K}_d} \prod_{i=1}^{d+1} \lambda_{\widehat{K}_d,i}^{\alpha_i} =: \chi_d$$

Dále budeme postupovat indukcí.

- Buď nejprve $d = 1$. Pak

$$\chi_d = \text{meas}(K_1) \int_{\widehat{K}_1} \lambda_{\widehat{K}_1,1}^{\alpha_1} \lambda_{\widehat{K}_1,2}^{\alpha_2} \underbrace{\text{(1.24)}}_{=} \text{meas}(K_1) \int_0^1 \widehat{x}_{[1]}^{\alpha_1} (1 - \widehat{x}_{[1]})^{\alpha_2} d\widehat{x}_{[1]} =$$

lemma 3,

$$\underbrace{\begin{aligned} p &:= \alpha_1, q := \alpha_2, \\ c &:= 1, \xi := \widehat{x}_{[1]} \end{aligned}}_{=} \frac{\alpha_1! \alpha_2!}{(\alpha_1 + \alpha_2 + 1)!} \text{meas}(K_1) = \begin{array}{l} \text{pravá strana} \\ \text{(1.25)} \\ \text{pro } d = 1 \end{array}.$$

- Nyní buď $d = s \in \mathbb{N} \setminus \{1\}$ a předpokládáme, že věta 5 platí pro $d = s - 1$. Pak

$$\begin{aligned}
\chi_d &= \text{meas}(K_s) s! \int_{\widehat{K}_s} \prod_{i=1}^{s+1} \lambda_{\widehat{K}_{s,i}}^{\alpha_i} \underbrace{\text{FV, (1.17), (1.24)}}_{=} \\
&= \text{meas}(K_s) s! \int_0^1 \int_0^{1-m_2(\widehat{x})} \dots \\
&\dots \int_0^{1-m_{s-1}(\widehat{x})} \int_0^{1-m_s(\widehat{x})} \prod_{i=1}^s \widehat{x}_{[i]}^{\alpha_i} (1 - m_{s+1}(\widehat{x}))^{\alpha_{s+1}} d\widehat{x}_{[s]} d\widehat{x}_{[s-1]} \dots d\widehat{x}_{[2]} d\widehat{x}_{[1]} \underbrace{(1.18)}_{=} \\
&\dots \text{meas}(K_s) s! \int_0^1 \int_0^{1-m_2(\widehat{x})} \dots \\
&\dots \underbrace{\int_0^{1-m_{s-1}(\widehat{x})} \prod_{i=1}^{s-1} \widehat{x}_{[i]}^{\alpha_i} \int_0^{1-m_s(\widehat{x})} \widehat{x}_{[s]}^{\alpha_s} ((1 - m_s(\widehat{x})) - \widehat{x}_{[s]})^{\alpha_{s+1}} d\widehat{x}_{[s]} d\widehat{x}_{[s-1]} \dots d\widehat{x}_{[2]} d\widehat{x}_{[1]}}_{I_5} =: \chi_s
\end{aligned}$$

Na integrál I_5 použijeme Lemma 3 s definicemi $p := \alpha_s, q := \alpha_{s+1}, c := 1 - m_s(\widehat{x})$ a $\xi := \widehat{x}_{[s]}$. Pak je

$$I_5 = \frac{(1 - m_s(\widehat{x}))^{\alpha_s + \alpha_{s+1} + 1}}{(\alpha_s + \alpha_{s+1} + 1)!} \alpha_s! \alpha_{s+1}!,$$

takže

$$\begin{aligned}
\chi_s &= \frac{\text{meas}(K_s) \alpha_s! \alpha_{s+1}! s!}{(\alpha_s + \alpha_{s+1} + 1)!} \int_0^1 \int_0^{1-m_2(\widehat{x})} \dots \\
&\dots \int_0^{1-m_{s-1}(\widehat{x})} \prod_{i=1}^{s-1} \widehat{x}_{[i]}^{\alpha_i} (1 - m_s(\widehat{x}))^{\alpha_s + \alpha_{s+1} + 1} d\widehat{x}_{[s-1]} \dots d\widehat{x}_{[2]} d\widehat{x}_{[1]} \underbrace{\text{FV, (1.24)}}_{=} \\
&= \frac{\text{meas}(K_s) \alpha_s! \alpha_{s+1}! s!}{(\alpha_s + \alpha_{s+1} + 1)!} \int_{\widehat{K}_{s-1}} \prod_{i=1}^{s-1} \lambda_{\widehat{K}_{s-1,i}}^{\alpha_i} \lambda_{\widehat{K}_{s-1,s}}^{\alpha_s + \alpha_{s+1} + 1} \underbrace{\text{indukční předpoklad}}_{=} \\
&= \frac{\text{meas}(K_s) \alpha_s! \alpha_{s+1}! s!}{(\alpha_s + \alpha_{s+1} + 1)!} \cdot \frac{\prod_{i=1}^{s-1} (\alpha_i!) (\alpha_s + \alpha_{s+1} + 1)! (s-1)!}{\left(\sum_{i=1}^{s+1} \alpha_i + s \right)!} \text{meas}(\widehat{K}_{s-1}) \underbrace{\text{lemma 4}}_{=} \text{pravá strana} \\
&\hspace{25em} (1.25) \quad \text{pro } d = s
\end{aligned}$$

■

1.5 Galerkinova formulace

Přechodem od funkcí $u \in U, v \in V$ k funkcím $u_h \in U_h, v_h \in V_h$ v rovnici (1.8) dostáváme následující diskrétní formulaci úlohy 1 získanou Galerkinovou metodou:

Úloha 6: Najít $u_h \in U_h$ splňující

$$a_G(u_h, v_h) = F_G(v_h) \quad \forall v_h \in V_h. \quad (1.26)$$

1.6 SUPG formulace

Nyní odvodíme diskrétní formulaci metodou SUPG.

Stabilizační parametr budeme chápat jako omezenou měřitelnou funkci

$$\tau : \bigcup_{K \in \mathcal{T}_h} \text{int}(K) \longrightarrow [0, \infty).$$

Začneme podobně jako při odvození slabé formulace. Rovnici (1.3) vynásobíme libovolnou testovací funkcí, ale tentokrát ve tvaru

$$v + \tau b \cdot \nabla v, \quad v \in C_0^\infty(\Omega),$$

zintegrujeme přes Ω a použijeme linearitu integrálu. Dostaneme

$$\begin{aligned} -\varepsilon \underbrace{\int_{\Omega} (\Delta u)v}_{I_2} + \int_{\Omega} (b \cdot \nabla u)v + \int_{\Omega} (-\varepsilon \Delta u + b \cdot \nabla u)(\tau b \cdot \nabla v) &= \\ &= \int_{\Omega} f v + \int_{\Omega} f(\tau b \cdot \nabla v) \quad \forall v \in C_0^\infty(\Omega). \end{aligned} \quad (1.27)$$

Na integrál I_2 použijeme Greenovu větu s uvážením, že funkce v má z definice prostoru $C_0^\infty(\Omega)$ nulovou hodnotu na $\partial\Omega$. Rovnice (1.27) tak přejde na tvar

$$\begin{aligned} \varepsilon \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (b \cdot \nabla u)v + \underbrace{\int_{\Omega} (-\varepsilon \Delta u + b \cdot \nabla u)(\tau b \cdot \nabla v)}_{I_3} &= \\ &= \int_{\Omega} f v + \int_{\Omega} f(\tau b \cdot \nabla v) \quad \forall v \in C_0^\infty(\Omega). \end{aligned} \quad (1.28)$$

Pro diskrétní formulaci budeme chtít v rovnici (1.28) opět přejít k funkci $u_h \in U_h$, ta však neleží v prostoru $H^2(\Omega)$, ale pouze v prostoru

$$H^2(\mathcal{T}_h) := \{u \in H^1(\Omega) : u|_K \in H^2(K) \forall K \in \mathcal{T}_h\},$$

takže na ni nelze aplikovat operátor Δ v integrálu I_3 na hranách triangulace. Proto nejprve v I_3 přejdeme k integraci přes jednotlivé elementy triangulace. Rovnice (1.28) tak přejde na tvar

$$\begin{aligned} \varepsilon \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (b \cdot \nabla u)v + \sum_{K \in \mathcal{T}_h} \int_{\text{int}(K)} (-\varepsilon \Delta u + b \cdot \nabla u)(\tau b \cdot \nabla v) &= \\ = \int_{\Omega} f v + \underbrace{\int_{\Omega} f(\tau b \cdot \nabla v)}_{I_4} \quad \forall v \in C_0^\infty(\Omega). \end{aligned} \quad (1.29)$$

Opět uvažujeme $b \in L^\infty(\Omega)$, $f \in L^2(\Omega)$.

Definujeme bilineární formu metody SUPG

$$\begin{aligned} a_{SUPG} : H^2(\mathcal{T}_h) \times H^1(\Omega) &\longrightarrow \mathbb{R}, \\ a_{SUPG}(u, v) &= a_G(u, v) + \sum_{K \in \mathcal{T}_h} \int_{\text{int}(K)} (-\varepsilon \Delta u + b \cdot \nabla u)(\tau b \cdot \nabla v) \end{aligned} \quad (1.30)$$

a lineární funkcionál metody SUPG

$$\begin{aligned} F_{SUPG} : H^1(\Omega) &\longrightarrow \mathbb{R}, \\ F_{SUPG}(v) &= F_G(v) + \int_{\Omega} f(\tau b \cdot \nabla v). \end{aligned} \quad (1.31)$$

V rovnici (1.29) přejdeme k funkcím $u_h \in U_h$ a $v_h \in V_h$ a napíšeme ji pomocí definic (1.30) a (1.31). Dostáváme tak diskrétní formulaci úlohy 1 získanou metodou SUPG:

Úloha 7: Najít $u_h \in U_h$ splňující

$$a_{SUPG}(u_h, v_h) = F_{SUPG}(v_h) \quad \forall v_h \in V_h. \quad (1.32)$$

Poznámka: Převědeme-li v rovnici (1.29) integrál I_4 na levou stranu, přidáme-li ho použitím linearitity sumace a integrálu do sumy na levé straně a použijeme-li definic (1.6) a (1.7), můžeme diskrétní úlohu s rovnicí (1.32) zformulovat následujícím způsobem:

Najít $u_h \in U_h$ splňující

$$a_G(u_h, v_h) + \sum_{K \in \mathcal{T}_h} \int_{\text{int}(K)} (-\varepsilon \Delta u_h + b \cdot \nabla u_h - f)(\tau b \cdot \nabla v_h) = F_G(v_h) \quad \forall v_h \in V_h. \quad (1.33)$$

Z toho je vidět, že diskrétní formulace metodou SUPG je vlastně diskrétní formulace Galerkinovou metodou obohacená o reziduum rovnice (1.3) vážené funkcí $\tau b \cdot \nabla v_h$ integrálně na každém elementu $K \in \mathcal{T}_h$.

Takže na metodu SUPG se lze dívat jako na Petrovovu-Galerkinovu metodu s testovacími funkcemi tvaru

$$v_h + \tau b \cdot \nabla v_h, \quad v_h \in V_h. \quad (1.34)$$

1.7 Galerkinova a SUPG formulace v maticovém tvaru

Uvažujeme řešení úlohy 6 respektive 7 ve tvaru lineární kombinace bázových funkcí z množiny $\text{base}(X_h)$, viz (1.11), tedy

$$u_h = \sum_{j \in J} u_j \varphi_j, \quad u_j \in \mathbb{R}. \quad (1.35)$$

Na pravé straně rovnice (1.35) od sebe s využitím (1.11) oddělíme funkce nulové na $\partial\Omega$ a funkce nenulové na $\partial\Omega$. Dostaneme

$$u_h = \sum_{j \in J(\Omega)} u_j \varphi_j + \sum_{j \in J(\partial\Omega)} u_j \varphi_j. \quad (1.36)$$

V druhé sumě na pravé straně (1.36) se sčítá přes $j \in J(\partial\Omega)$, takže z definice prostoru U_h pro její sčítanec platí

$$u_j \varphi_j(n_j) = u^B(n_j), \quad j \in J(\partial\Omega). \quad (1.37)$$

Z definice množiny $\text{base}(X_h)$ plyne, že $\varphi_j(n_j) = 1$, takže když toto použijeme v (1.37), dostaneme

$$u_j = u^B(n_j), \quad j \in J(\partial\Omega). \quad (1.38)$$

Dosazení z rovnice (1.38) do (1.36) dává, že řešení úlohy s rovnicí 1.26 respektive 1.32 má tvar

$$u_h = \sum_{j \in J(\Omega)} u_j \varphi_j + \sum_{j \in J(\partial\Omega)} u^B(n_j) \varphi_j. \quad (1.39)$$

Dosadíme u_h ve tvaru (1.39) do rovnice (1.26) respektive (1.32), použijeme bilinearitu formy a_G respektive a_{SUPG} a linearitu sumace a místo všech testovacích funkcí z V_h uvažujeme jen funkce z množiny $\text{base}(V_h)$. Rovnice (1.26) respektive (1.32) tak přejde na tvar

$$\sum_{j \in J(\Omega)} a_G(\varphi_j, \varphi_i) u_j + \sum_{j \in J(\partial\Omega)} u^B(n_j) a_G(\varphi_j, \varphi_i) = F_G(\varphi_i), \quad i \in J(\Omega) \quad (1.40)$$

respektive

$$\sum_{j \in J(\Omega)} a_{SUPG}(\varphi_j, \varphi_i) u_j + \sum_{j \in J(\partial\Omega)} u^B(n_j) a_{SUPG}(\varphi_j, \varphi_i) = F_{SUPG}(\varphi_i), \quad i \in J(\Omega). \quad (1.41)$$

Druhý člen levé strany rovnice (1.40) respektive (1.41) neobsahuje žádný neznámý koeficient u_j , takže tento člen převedeme na pravou stranu, a tak konečně dostáváme diskrétní formulaci úlohy 1 získanou Galerkinovou metodou respektive metodou SUPG v maticovém tvaru:

Úloha 8: Najít $u_j \in \mathbb{R}, j \in J(\Omega)$, splňující

$$\sum_{j \in J(\Omega)} a_G(\varphi_j, \varphi_i) u_j = F_G(\varphi_i) - \sum_{j \in J(\partial\Omega)} u^B(n_j) a_G(\varphi_j, \varphi_i), \quad i \in J(\Omega), \quad (1.42)$$

a pak zkonstruovat funkci u_h podle (1.35),

respektive

Úloha 9: Najít $u_j \in \mathbb{R}, j \in J(\Omega)$, splňující

$$\sum_{j \in J(\Omega)} a_{SUPG}(\varphi_j, \varphi_i) u_j = F_{SUPG}(\varphi_i) - \sum_{j \in J(\partial\Omega)} u^B(n_j) a_{SUPG}(\varphi_j, \varphi_i), \quad i \in J(\Omega), \quad (1.43)$$

a pak zkonstruovat funkci u_h podle (1.35).

1.8 Existence a jednoznačnost řešení SUPG formule

Lemma 10: Existuje takové $\mu > 0$ nezávislé na h_K a na $K \in \mathcal{T}_h$, že platí následující tzv. *inverzní nerovnost*

$$\|\Delta v_h\|_{0,K} \leq \mu h_K^{-1} |v_h|_{1,K} \quad \forall v_h \in V_h \quad \forall K \in \mathcal{T}_h. \quad (1.44)$$

Důkaz: Viz např. [3]. ■

Nadále budeme předpokládat, že stabilizační parametr splňuje nerovnost

$$\tau \leq \frac{h_K^2}{\varepsilon \mu^2} \quad \forall K \in \mathcal{T}_h, \quad (1.45)$$

kde μ je konstanta z inverzní nerovnosti (1.44) z lemmatu 10.

Na prostoru V definujeme tzv. *SD-normu* $||| \cdot |||$ (SD jako Streamline Diffusion) předpisem

$$|||v||| := (\varepsilon |v|_{1,\Omega}^2 + \|\tau^{1/2} b \cdot \nabla v\|_{0,\Omega}^2)^{1/2}. \quad (1.46)$$

což je norma na prostoru V .

Lemma 11: Buď μ konstanta z lemmatu 10 o inverzní nerovnosti a předpokládejme platnost odhadu (1.45). Pak bilineární forma a_{SUPG} je V_h -eliptická vzhledem k SD-normě s konstantou $1/2$, tj.

$$a_{SUPG}(v_h, v_h) \geq \frac{1}{2} |||v_h|||^2 \quad \forall v_h \in V_h. \quad (1.47)$$

Důkaz: Platí

$$\begin{aligned} \int_{\Omega} (b \cdot \nabla v_h) v_h &= \sum_{i=1}^d \int_{\Omega} b_{[i]} \frac{\partial v_h}{\partial x^{[i]}} v_h \stackrel{GV, (1.2)}{=} \sum_{i=1}^d \left(\underbrace{\int_{\partial\Omega} b_{[i]} v_h^2 \nu^{[i]}}_{=0, \text{ neboť } v_h \in V_h} - \int_{\Omega} b_{[i]} v_h \frac{\partial v_h}{\partial x^{[i]}} \right) = \\ &= - \sum_{i=1}^d \int_{\partial\Omega} b_{[i]} \frac{\partial}{\partial x^{[i]}} v_h v_h = - \int_{\Omega} (b \cdot \nabla v_h) v_h \implies \int_{\Omega} (b \cdot \nabla v_h) v_h = 0, \end{aligned} \quad (1.48)$$

takže je

$$\begin{aligned}
a_{SUPG}(v_h, v_h) &\stackrel{(1.30), (1.48)}{=} \varepsilon |v_h|_{1,\Omega}^2 + \|\tau^{1/2}b \cdot \nabla v_h\|_{0,\Omega}^2 - \sum_{K \in \mathcal{T}_h} \varepsilon \int_K \Delta v_h \tau b \cdot \nabla v_h \stackrel{(1.46)}{=} \\
&= |||v_h|||^2 - \underbrace{\sum_{K \in \mathcal{T}_h} \varepsilon \int_K \Delta v_h \tau b \cdot \nabla v_h}_{I_6}.
\end{aligned}$$

Dále platí

$$\begin{aligned}
I_6 &\stackrel{\text{HN}}{\leq} \sum_{K \in \mathcal{T}_h} \varepsilon \|\Delta v_h\|_{0,K} \|\tau b \cdot \nabla v_h\|_{0,K} \stackrel{(1.44)}{\leq} \sum_{K \in \mathcal{T}_h} \varepsilon \mu h_K^{-1} |v_h|_{1,K} \|\tau b \cdot \nabla v_h\|_{0,K} \stackrel{(1.45)}{\leq} \\
&\leq \sum_{K \in \mathcal{T}_h} \varepsilon^{1/2} |v_h|_{1,K} \|\tau^{1/2}b \cdot \nabla v_h\|_{0,K} \sum_{K \in \mathcal{T}_h} \stackrel{\text{YN}}{\leq} \frac{1}{2} (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2}b \cdot \nabla v_h\|_{0,K}^2) \stackrel{(1.46)}{=} \frac{1}{2} |||v_h|||^2,
\end{aligned}$$

z čehož už plyne dokazované tvrzení. ■

Věta 12: Existuje právě jedno řešení úlohy 9.

Důkaz: Označíme

$$\mathbf{A} := \{a_{SUPG}(\varphi_j, \varphi_i)\}_{i,j \in J(\Omega)}, \quad \mathbf{u} := \{u_j\}_{j \in J(\Omega)}^T, \quad \mathbf{0} := \underbrace{\{0, 0, \dots, 0\}^T}_{d \text{ složek}}. \quad (1.49)$$

Zdůrazňujeme, že vektory \mathbf{u} a $\mathbf{0}$ jsou chápány jako sloupcové. Pak levou stranu rovnice (1.43) v úloze 9 lze s využitím značení (1.49) zapsat jako $\mathbf{A}\mathbf{u}$. Matice \mathbf{A} je čtvercová, takže úloha 9 je jednoznačně řešitelná, právě když pouze nulový vektor \mathbf{u} vyhovuje rovnici

$$\mathbf{A}\mathbf{u} = \mathbf{0}. \quad (1.50)$$

To vyšetříme v následujícím. Rovnici (1.50) vynásobíme zleva transpozicí vektoru \mathbf{u} a provedeme jisté úpravy.

$$\begin{aligned}
0 &= \mathbf{u}^T \mathbf{A}\mathbf{u} = \sum_{i \in J(\Omega)} \sum_{j \in J(\Omega)} u_i u_j a_{SUPG}(\varphi_j, \varphi_i) = \\
&= a_{SUPG}\left(\sum_{j \in J(\Omega)} u_j \varphi_j, \sum_{i \in J(\Omega)} u_i \varphi_i\right) \stackrel{\text{lemma 11}}{\geq} \frac{1}{2} |||\sum_{k \in J(\Omega)} u_k \varphi_k|||^2 \implies \\
\implies |||\sum_{k \in J(\Omega)} u_k \varphi_k||| &= 0 \stackrel{\text{|||} \cdot \text{||| je norma na } V}{\implies} \sum_{k \in J(\Omega)} u_k \varphi_k \equiv 0 \stackrel{\varphi_k \text{ lineárně nezávislé}}{\implies} \mathbf{u} = \mathbf{0}
\end{aligned} \quad (1.51)$$

Lemma 11 lze použít proto, že funkce $\sum_{k \in J(\Omega)} u_k \varphi_k$ je z prostoru V_h vzhledem k definici (1.11) a vzhledem k tomu, že v sumách v (1.51) se vždy sčítá přes index náležející pouze do množiny $J(\Omega)$. Tyto skutečnosti jsou odůvodněním také pro předposlední implikaci v (1.51).

V (1.51) jsme tedy zjistili, že pouze nulový vektor \mathbf{u} vyhovuje rovnici (1.50), takže matice \mathbf{A} je regulární, a tedy existuje právě jedno řešení úlohy 9. ■

1.9 Odhad chyby řešení SUPG formulace

Lemma 13: Buď

- a) u řešení úlohy 2,
 - b) $u \in H^2(\mathcal{T}_h)$,
 - c) u_h řešení úlohy 7.
- Pak platí

$$\| \|u - u_h\| \| \leq \inf_{z_h \in U_h} \left\{ \underbrace{\| \|u - z_h\| \| + 2 \sup_{v_h \in V_h} \frac{a_{SUPG}(u - z_h, v_h)}{\| \|v_h\| \|}}_{\zeta} \right\}. \quad (1.52)$$

Důkaz: Z předpokladů a) a b) a z linearitě integrálu a jeho aditivity vzhledem k integračnímu oboru plyne

$$\begin{aligned} 0 &\stackrel{a)}{=} a_G(u, v) - F_G(v) \stackrel{(1.6)}{=} \varepsilon \int_{\Omega} \nabla u \cdot \nabla v + \int_{\Omega} (b \cdot \nabla u)v - \int_{\Omega} f v = \\ &= \sum_{K \in \mathcal{T}_h} \varepsilon \int_K \nabla u \cdot \nabla v + \int_K (b \cdot \nabla u)v - \int_K f v \stackrel{b), \text{GV}}{=} \\ &= \sum_{K \in \mathcal{T}_h} -\varepsilon \int_K (\Delta u)v + \int_K (b \cdot \nabla u)v - \int_K f v = \int_{\Omega} (-\varepsilon \Delta u + b \cdot \nabla u - f)v \quad \forall v \in V. \end{aligned}$$

Z toho s využitím předpokladu c) plyne

$$-\varepsilon \Delta u + b \cdot \nabla u = f \quad \text{skoro všude v } \Omega. \quad (1.53)$$

Potom platí

$$\begin{aligned} a_{SUPG}(u, v_h) &\stackrel{(1.30)}{=} a_G(u, v_h) + \sum_{K \in \mathcal{T}_h} \int_{\text{int}(K)} (-\varepsilon \Delta u + b \cdot \nabla u)(\tau b \cdot \nabla v_h) \stackrel{a), (1.53)}{=} \\ F_G(v_h) + \sum_{K \in \mathcal{T}_h} \int_{\text{int}(K)} f(\tau b \cdot \nabla v_h) &\stackrel{(1.31)}{=} F_{SUPG}(v_h) \stackrel{c)}{=} a_{SUPG}(u_h, v_h) \quad \forall v_h \in V_h. \end{aligned}$$

Tedy forma a_{SUPG} je tzv. *konzistentní*, tj.

$$a_{SUPG}(u, v_h) = a_{SUPG}(u_h, v_h) \quad \forall v_h \in V_h. \quad (1.54)$$

Dále označme z_h libovolnou funkci z U_h . Pak $u_h - z_h \in V_h \forall z_h \in U_h$, a tedy pro z_h různou od u_h je

$$\begin{aligned} \frac{1}{2} |||u_h - z_h|||^2 &\stackrel{(1.47)}{\leq} a_{SUPG}(u_h - z_h, u_h - z_h) = a_{SUPG}(u_h, u_h - z_h) - a_{SUPG}(z_h, u_h - z_h) = \\ &\stackrel{(1.54)}{=} a_{SUPG}(u, u_h - z_h) - a_{SUPG}(z_h, u_h - z_h) = a_{SUPG}(u - z_h, u_h - z_h) \frac{|||u_h - z_h|||}{|||u_h - z_h|||} \leq \\ &\leq \sup_{v_h \in V_h} \frac{a_{SUPG}(u - z_h, v_h)}{|||v_h|||} |||u_h - z_h||| \quad \forall z_h \in U_h. \end{aligned} \quad (1.55)$$

Po zkrácení $|||u_h - z_h|||$ v (1.55) dostáváme

$$|||u_h - z_h||| \leq 2 \sup_{v_h \in V_h} \frac{a_{SUPG}(u - z_h, v_h)}{|||v_h|||} \quad \forall z_h \in U_h. \quad (1.56)$$

Nakonec máme

$$\begin{aligned} |||u - u_h||| &\leq |||u - z_h||| + |||z_h - u_h||| \stackrel{(1.56)}{\leq} \\ &\leq |||u - z_h||| + 2 \sup_{v_h \in V_h} \frac{a_{SUPG}(u - z_h, v_h)}{|||v_h|||} \quad \forall z_h \in U_h. \end{aligned} \quad (1.57)$$

Protože nerovnost (1.57) platí pro všechna $z_h \in U_h$, lze na její pravé straně přejít k infimu přes všechna $z_h \in U_h$, z čehož plyne dokazované tvrzení. ■

Buď K libovolný d -simplex a buďte $j = 1, 2, \dots, d + 1$ indexy jeho vrcholů. Pak definujeme body

$$y_{K,j} := \frac{1}{2}a_{K,j} + \frac{1}{2}c_K, \quad j \in \{1, 2, \dots, d + 1\}, \quad (1.58)$$

kde c_K je těžiště d -simplexu K podle definice (1.10). Dále definujeme množinu

$$I(K) := \left\{ \sum_{j=1}^{d+1} \alpha_j y_{K,j}, \sum_{j=1}^{d+1} \alpha_j = 1, \alpha_j \geq 0 \forall j \in \{1, 2, \dots, d + 1\} \right\}.$$

Je $I(K) \subsetneq \text{int}(K)$.

Uvažujeme referenční element \widehat{K} , definovaný předpisem (1.16), libovolný element $K \in \mathcal{T}_h$ afinní regulární zobrazení $\mathcal{F}_K : \widehat{K} \xrightarrow{\text{na}} K$, libovolnou funkci $v \in P_{2k+2}(K)$ a definujeme funkci $\widehat{v} := v \circ \mathcal{F}_K$. Je $\widehat{v} \in P_{2k+2}(\widehat{K})$. Protože $P_{2k+2}(\mathbb{R}^d)$ je konečně dimenzionální prostor, jsou na něm všechny normy ekvivalentní, a tedy existuje konstanta $C_E > 0$ taková, že

$$\begin{aligned} \int_K |v| &\stackrel{VS, (1.22)}{=} \frac{\text{meas}(K)}{\text{meas}(\widehat{K})} \int_{\widehat{K}} |\widehat{v}| = \frac{\text{meas}(I(K))}{\text{meas}(I(\widehat{K}))} \int_{\widehat{K}} |\widehat{v}| \leq \\ &\leq C_E \frac{\text{meas}(I(K))}{\text{meas}(I(\widehat{K}))} \int_{I(\widehat{K})} |\widehat{v}| \stackrel{VS}{=} C_E \int_{I(K)} |v|. \end{aligned} \quad (1.59)$$

Protože jsme přešli k referenčnímu elementu \widehat{K} , konstanta C_E nezávisí na volbě elementu K .

Pro odhad členu ζ z nerovnosti (1.52) dokážeme následující lemma.

Lemma 14: Buď $w \in H^2(\mathcal{T}_h)$ libovolná funkce, μ konstanta z lemmatu 10 o inverzní nerovnosti a předpokládejme platnost nerovnosti (1.45). Dále předpokládejme, že

$$\tau|_K \in P_1(K) \quad \forall K \in \mathcal{T}_h, \quad (1.60)$$

$$\exists k \in \mathbb{N} \cup \{0\} : b \in (P_k(K))^d \quad \forall K \in \mathcal{T}_h. \quad (1.61)$$

Uvažujeme těžiště elementu K podle definice (1.10). Dále definujeme konstantu

$$\gamma_K := \varepsilon + (d+1)\tau(c_K) \|b\|_{0,\infty,K}^2 + \|b\|_{0,\infty,K}^2 \left(\max \left\{ \varepsilon, \|b\|_{0,\infty,K}^2 \frac{\tau(c_K)}{2C_E} \right\} \right)^{-1} h_K^2, \quad (1.62)$$

kde C_E je konstanta z nerovnosti (1.59). Pak platí

$$\|w\| + 2 \sup_{v_h \in V_h} \frac{a_{SUPG}(w, v_h)}{\|v_h\|} \leq \underbrace{\left(5 + \frac{2}{\mu} \right) \left(\sum_{K \in \mathcal{T}_h} \gamma_K (h_K^{-2} \|w\|_{0,K}^2 + |w|_{1,K}^2 + h_K^2 \|\Delta w\|_{0,K}^2) \right)^{1/2}}_{\Xi_1}. \quad (1.63)$$

Důkaz: V celém důkazu budeme uvažovat, že daný element $K \in \mathcal{T}_h$ má indexy svých vrcholů $1, 2, \dots, d+1$.

Z předpokladu (1.60) vyplývá, že pro každý element $K \in \mathcal{T}_h$ existují $p_K^i \in \mathbb{R}, i \in \{1, 2, \dots, d+1\}^1$, a $q_K \in \mathbb{R}$ taková, že

$$\tau|_K(x) = \sum_{i=1}^d p_K^i x_{[i]} + q_K, \quad K \in \mathcal{T}_h, \quad x \in K. \quad (1.64)$$

¹ i je pouze horní index, nikoliv exponent

Buď $K \in \mathcal{T}_h$ libovolný element. Odhadneme stabilizační parametr $\tau|_K$. Na jednu stranu platí

$$\begin{aligned}
& \tau|_K(a_{K,j}) \stackrel{(1.10)}{=} \tau|_K \left((d+1)c_K - \sum_{k=1, k \neq j}^{d+1} a_{K,k} \right) \stackrel{(1.64)}{=} \\
& = \sum_{i=1}^d p_K^i \left((d+1)c_{K[i]} - \sum_{k=1, k \neq j}^{d+1} a_{K,k[i]} \right) + q_K = \\
& = (d+1) \left(\sum_{i=1}^d p_K^i c_{K[i]} + q_K \right) - \sum_{k=1, k \neq j}^{d+1} \left(\sum_{i=1}^d p_K^i a_{K,k[i]} + q_K \right) = \\
& = (d+1)\tau|_K(c_K) - \sum_{k=1, k \neq j}^{d+1} \tau|_K(a_{K,k}) \stackrel{\tau \geq 0}{\leq} (d+1)\tau(c_K) \quad \forall j \in \{1, 2, \dots, d+1\},
\end{aligned} \tag{1.65}$$

takže je

$$\tau|_K \leq \max_{x \in K} \tau|_K(x) \stackrel{(1.60)}{=} \max_{j \in \{1, 2, \dots, d+1\}} \tau|_K(a_{K,j}) \stackrel{(1.65)}{\leq} (d+1)\tau(c_K). \tag{1.66}$$

Na druhou stranu platí

$$\begin{aligned}
& \tau|_K(y_{K,j}) \stackrel{(1.58)}{=} \tau|_K \left(\frac{1}{2}a_{K,j} + \frac{1}{2}c_K \right) \stackrel{(1.64)}{=} \\
& = \sum_{i=1}^d p_K^i \left(\frac{1}{2}a_{K,j[i]} + \frac{1}{2}c_{K[i]} \right) + q_K = \\
& = \frac{1}{2} \left(\sum_{i=1}^d p_K^i a_{K,j[i]} + q_K \right) + \frac{1}{2} \left(\sum_{i=1}^d p_K^i c_{K[i]} + q_K \right) = \\
& = \frac{1}{2}\tau(y_{K,j}) + \frac{1}{2}\tau(c_K) \stackrel{\tau \geq 0}{\geq} \frac{1}{2}\tau(c_K),
\end{aligned} \tag{1.67}$$

takže je

$$\min_{x \in I(K)} \tau(x) \stackrel{(1.64)}{=} \min_{j \in \{1, 2, \dots, d+1\}} \tau(y_{K,j}) \stackrel{(1.67)}{\geq} \frac{1}{2}\tau(c_K). \tag{1.68}$$

Dále je

$$\begin{aligned}
& \|w\| \stackrel{(1.46)}{=} (\varepsilon |w|_{1,\Omega}^2 + \|\tau^{1/2}b \cdot \nabla w\|_{0,\Omega}^2)^{1/2} \stackrel{HN, (1.66)}{\leq} \\
& \leq \left(\sum_{K \in \mathcal{T}_h} (\varepsilon + (d+1)\tau(c_K)) \|b\|_{0,\infty,K}^2 |w|_{1,K}^2 \right)^{1/2} \leq \left(\sum_{K \in \mathcal{T}_h} \gamma_K |w|_{1,K}^2 \right)^{1/2} \leq \Xi_1.
\end{aligned} \tag{1.69}$$

Z definice formy a_{SUPG} (1.30) je

$$\begin{aligned} & a_{SUPG}(w, v_h) = \\ & = \underbrace{\varepsilon \int_{\Omega} \nabla w \cdot \nabla v_h + \int_{\Omega} b \cdot \nabla w \tau b \cdot \nabla v_h}_{\Theta_1} + \underbrace{\sum_{K \in \mathcal{T}_h} \int_{\text{int}(K)} -\varepsilon \Delta w \tau b \cdot \nabla v_h}_{\Theta_2} + \underbrace{\int_{\Omega} (b \cdot \nabla w) v_h}_{\Theta_3}. \end{aligned}$$

Odhadneme postupně výrazy Θ_1 , Θ_2 a Θ_3 .

Odhad výrazu Θ_1 :

$$\Theta_1 = \int_{\Omega} \varepsilon^{1/2} \nabla w \cdot \varepsilon^{1/2} \nabla v_h + \int_{\Omega} \tau^{1/2} b \cdot \nabla w \tau^{1/2} b \cdot \nabla v_h \stackrel{HN}{\leq} \|w\| \|v_h\| \stackrel{(1.69)}{\leq} \Xi_1 \|v_h\|. \quad (1.70)$$

Odhad výrazu Θ_2 :

$$\begin{aligned} \Theta_2 & \stackrel{HN}{\leq} \sum_{K \in \mathcal{T}_h} \varepsilon \|\Delta w\|_{0,K} \|\tau b \cdot \nabla v_h\|_{0,K} \stackrel{(1.45)}{\leq} \sum_{K \in \mathcal{T}_h} \frac{\varepsilon^{1/2} h_K}{\mu} \|\Delta w\|_{0,K} \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K} \stackrel{(1.46)}{\leq} \\ & \leq \frac{1}{\mu} \left(\sum_{K \in \mathcal{T}_h} \varepsilon h_K^2 \|\Delta w\|_{0,K}^2 \right)^{1/2} \|v_h\| \leq \frac{1}{\mu} \left(\sum_{K \in \mathcal{T}_h} \gamma_K h_K^2 \|\Delta w\|_{0,K}^2 \right)^{1/2} \|v_h\| \leq \frac{1}{\mu} \Xi_1 \|v_h\|. \end{aligned} \quad (1.71)$$

Odhad výrazu Θ_3 :

Nyní buď $K \in \mathcal{T}_h$ libovolný element. Je $v_h|_K \in P_2(K)$, takže s uvážením předpokladu (1.61) je $(b \cdot \nabla v_h)^2 \in P_{2k+2}(K)$, a tedy je

$$\begin{aligned} \tau(c_K) \|b \cdot \nabla v_h\|_{0,K}^2 & \stackrel{(1.59)}{\leq} C_E \tau(c_K) \|b \cdot \nabla v_h\|_{0,I(K)}^2 \leq \\ & \stackrel{(1.68)}{\leq} 2C_E \min_{x \in I(K)} \tau(x) \|b \cdot \nabla v_h\|_{0,I(K)}^2 = \\ & = 2C_E \left\| \left(\min_{x \in I(K)} \tau(x) \right)^{1/2} b \cdot \nabla v_h \right\|_{0,I(K)}^2 \leq \\ & \leq 2C_E \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2 \leq \\ & \stackrel{(1.46)}{\leq} 2C_E (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2). \end{aligned} \quad (1.72)$$

Odhad (1.72) vydělíme výrazem $2C_E$, odmocníme a vynásobíme výrazem $\|b\|_{0,\infty,K}$, a dostáváme tak odhad

$$\|b\|_{0,\infty,K} \left(\frac{\tau(c_K)}{2C_E} \right)^{1/2} \|b \cdot \nabla v_h\|_{0,K} \leq \|b\|_{0,\infty,K} (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2)^{1/2} \quad (1.73)$$

Dále je ale také

$$\begin{aligned}
\|b \cdot \nabla v_h\|_{0,K} &\stackrel{HN}{\leq} \|b\|_{0,\infty,K} |v_h|_{1,K} = \\
&= \varepsilon^{-1/2} \|b\|_{0,\infty,K} \varepsilon^{1/2} |v_h|_{1,K} \leq \\
&\stackrel{(1.46)}{\leq} \varepsilon^{-1/2} \|b\|_{0,\infty,K} (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2)^{1/2}.
\end{aligned} \tag{1.74}$$

Odhad (1.74) vynásobíme výrazem $\varepsilon^{1/2}$, a dostáváme tak odhad

$$\varepsilon^{1/2} \|b \cdot \nabla v_h\|_{0,K} \leq \|b\|_{0,\infty,K} (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2)^{1/2}. \tag{1.75}$$

Odhady (1.73) a (1.75) mají stejnou pravou stranu, takže z nich vyplývá odhad

$$\max \left\{ \varepsilon^{1/2}, \|b\|_{0,\infty,K} \left(\frac{\tau(c_K)}{2C_E} \right)^{1/2} \right\} \|b \cdot \nabla v_h\|_{0,K} \leq \|b\|_{0,\infty,K} (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2)^{1/2}. \tag{1.76}$$

Maximum na levé straně odhadu (1.76) je vždy kladné, takže jím můžeme tento odhad vydělit, a dostáváme tak odhad

$$\left(\max \left\{ \varepsilon^{1/2}, \|b\|_{0,\infty,K} \left(\frac{\tau(c_K)}{2C_E} \right)^{1/2} \right\} \right)^{-1} \|b \cdot \nabla v_h\|_{0,K} \leq \|b\|_{0,\infty,K} (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2)^{1/2}. \tag{1.77}$$

Element K byl volen libovolně, takže konečně je

$$\begin{aligned}
\Theta_3 &= \int_{\Omega} \sum_{i=1}^d b_{[i]} \frac{\partial w}{\partial x_{[i]}} v_h \stackrel{GV, (1.2)}{=} - \int_{\Omega} \sum_{i=1}^d b_{[i]} w \frac{\partial v_h}{\partial x_{[i]}} = - \sum_{K \in \mathcal{T}_h} \int_K (b \cdot \nabla v_h) w \stackrel{HN}{\leq} \\
&\stackrel{(1.77)}{\leq} \sum_{K \in \mathcal{T}_h} \|b \cdot \nabla v_h\|_{0,K} \|w\|_{0,K} \leq \\
&\leq \sum_{K \in \mathcal{T}_h} \left(\max \left\{ \varepsilon^{1/2}, \|b\|_{0,\infty,K} \left(\frac{\tau(c_K)}{2C_E} \right)^{1/2} \right\} \right)^{-1} \\
&\quad \cdot \|b\|_{0,\infty,K} (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2)^{1/2} \|w\|_{0,K} \stackrel{HN}{\leq} \\
&\leq \left(\sum_{K \in \mathcal{T}_h} \|b\|_{0,\infty,K}^2 \left(\max \left\{ \varepsilon, \|b\|_{0,\infty,K}^2 \frac{\tau(c_K)}{2C_E} \right\} \right)^{-1} \|w\|_{0,K}^2 \right)^{1/2} \quad (1.78) \\
&\quad \cdot \left(\sum_{K \in \mathcal{T}_h} (\varepsilon |v_h|_{1,K}^2 + \|\tau^{1/2} b \cdot \nabla v_h\|_{0,K}^2) \right)^{1/2} \stackrel{(1.46)}{=} \\
&= \left(\sum_{K \in \mathcal{T}_h} \|b\|_{0,\infty,K}^2 \left(\max \left\{ \varepsilon, \|b\|_{0,\infty,K}^2 \frac{\tau(c_K)}{2C_E} \right\} \right)^{-1} h_K^2 h_K^{-2} \|w\|_{0,K}^2 \right)^{1/2} \|v_h\| \stackrel{(1.62)}{\leq} \\
&\leq \left(\sum_{K \in \mathcal{T}_h} \gamma_K h_K^{-2} \|w\|_{0,K}^2 \right)^{1/2} \|v_h\| \leq \Xi_1 \|v_h\|.
\end{aligned}$$

Dokazovanou nerovnost (1.63) pak dostaneme sečtením nerovností (1.70), (1.71) a (1.78), vynásobením výsledné nerovnosti členem $2/\|v_h\|$ s předpokladem $\|v_h\| > 0$, přechodem k supremu přes $v_h \in V_h$ na levé straně výsledné nerovnosti, což lze, protože $v_h \in V_h$ bylo libovolné, a nakonec přičtením nerovnosti (1.69) k výsledné nerovnosti. ■

Definujeme význam operace \simeq .

$$x \simeq y \stackrel{\text{definice}}{\iff} \exists C_1, C_2 > 0 \text{ nezávislá na } h : C_1 y \leq x \leq C_2 y, \quad x, y \in \mathbb{R}. \quad (1.79)$$

Výrok $x \simeq y$ čteme „ x se chová jako y “. Dále definujeme význam operace \lesssim .

$$x \lesssim y \stackrel{\text{definice}}{\iff} \exists C > 0 \text{ nezávislá na } h : x \leq C y, \quad x, y \in \mathbb{R}.$$

Lemma 15: Existují $l \in \mathbb{N}$, $C_P > 0$ nezávislé na h taková, že interpolační operátor

$$\Pi_h \in \mathcal{L}(H^2(\Omega), U_h) \cap \mathcal{L}(H^2(\Omega) \cap H_0^1(\Omega), V_h)$$

splňuje

$$\begin{aligned} \forall K \in \mathcal{T}_h \quad \forall m \in \{1, 2, \dots, l\} \quad \forall v \in H^{m+1}(K) \\ h_K^{-1} \|v - \Pi_h v\|_{0,K} + |v - \Pi_h v|_{1,K} + h_K |v - \Pi_h v|_{2,K} \leq C_P h_K^m |v|_{m+1,K}. \end{aligned} \quad (1.80)$$

Důkaz: Viz [2].

■

Věta 16: Buď $k \in \{1, 2, \dots, l\}$ takové, že řešení u úlohy 2 splňuje $u \in H^{k+1}(\Omega)$, buď $\tilde{u}^B \in H^2(\Omega)$ rozšíření u^B a buď $\tilde{u}_h^B = \Pi_h \tilde{u}^B$, tj. \tilde{u}_h^B je funkce, jejíž stopa aproximuje funkci u^B . Pak existuje konstanta $C > 0$ nezávislá na h taková, že řešení u_h úlohy 7 splňuje odhad chyby

$$\| \|u - u_h\| \| \leq C h^k \left(\sum_{K \in \mathcal{T}_h} \gamma_K |u|_{k+1,K}^2 \right)^{1/2}. \quad (1.81)$$

Jestliže navíc stabilizační parametr splňuje relaci

$$\tau(c_K) \simeq \min \left\{ \frac{h_K}{\|b\|_{0,\infty,K}}, \frac{h_K^2}{\varepsilon} \right\} \quad \forall K \in \mathcal{T}_h, \quad (1.82)$$

pak existuje konstanta $C_c > 0$ nezávislá na h taková, že řešení u_h úlohy 7 splňuje odhad chyby

$$\| \|u - u_h\| \| \leq C_c (\varepsilon + h \|b\|_{0,\infty,\Omega})^{1/2} h^k |u|_{k+1,\Omega}. \quad (1.83)$$

Důkaz: Platí

$$\|\Delta v\|_{0,K}^2 = \int_K \left| \sum_{i=1}^d \frac{\partial^2 v}{\partial x_{[i]}^2} \right|^2 \leq \int_K \left(\sum_{i=1}^d \left| \frac{\partial^2 v}{\partial x_{[i]}^2} \right| \right)^2 \stackrel{HN}{\leq} d \int_K \sum_{i=1}^d \left| \frac{\partial^2 v}{\partial x_{[i]}^2} \right|^2 = d |v|_{2,K}^2. \quad (1.84)$$

Je

$$\begin{aligned}
& \underbrace{\|u - u_h\|}_{\substack{\text{lemma 13} \\ z_h := \Pi_h u}} \leq \|u - \Pi_h u\| + 2 \sup_{v_h \in V_h} \frac{a_{SUPG}(u - \Pi_h u, v_h)}{\|v_h\|} \underbrace{\|u - \Pi_h u\|}_{\substack{\text{lemma 14} \\ w := u - \Pi_h u}} \\
& \leq \left(5 + \frac{2}{\mu}\right) \left(\sum_{K \in \mathcal{T}_h} \gamma_K (h_K^{-2} \|u - \Pi_h u\|_{0,K}^2 + |u - \Pi_h u|_{1,K}^2 + h_K^2 \|\Delta(u - \Pi_h u)\|_{0,K}^2) \right)^{1/2} \underbrace{\leq}_{(1.84)} \\
& \leq \sqrt{d} \left(5 + \frac{2}{\mu}\right) \left(\sum_{K \in \mathcal{T}_h} \gamma_K (h_K^{-2} \|u - \Pi_h u\|_{0,K}^2 + |u - \Pi_h u|_{1,K}^2 + h_K^2 |u - \Pi_h u|_{2,K}^2) \right)^{1/2} \leq \\
& \leq \sqrt{d} \left(5 + \frac{2}{\mu}\right) \left(\sum_{K \in \mathcal{T}_h} \gamma_K (h_K^{-1} \|u - \Pi_h u\|_{0,K} + |u - \Pi_h u|_{1,K} + h_K |u - \Pi_h u|_{2,K})^2 \right)^{1/2} \underbrace{\leq}_{(1.80)} \\
& \leq C_P \sqrt{d} \left(5 + \frac{2}{\mu}\right) h_K^k \left(\sum_{K \in \mathcal{T}_h} \gamma_K |u|_{k+1,K}^2 \right)^{1/2} \underbrace{=}_{(1.9)} C h^k \left(\sum_{K \in \mathcal{T}_h} \gamma_K |u|_{k+1,K}^2 \right)^{1/2}.
\end{aligned}$$

Dále, je-li nejprve

$$\frac{h_K}{\|b\|_{0,\infty,K}} \leq \frac{h_K^2}{\varepsilon}, \quad (1.85)$$

pak z (1.82) plyne

$$\tau(c_K) \simeq \frac{h_K}{\|b\|_{0,\infty,K}}. \quad (1.86)$$

(1.85) je ekvivalentní také s

$$\varepsilon \leq \|b\|_{0,\infty,K} h_K. \quad (1.87)$$

Odhadneme nejdřív maximum v konstantě γ_K .

$$\begin{aligned}
& \max \left\{ \varepsilon, \|b\|_{0,\infty,K}^2 \frac{\tau(c_K)}{2C_E} \right\} \underbrace{\geq}_{(1.86), (1.79)} \max \left\{ \varepsilon, \frac{C_1 \|b\|_{0,\infty,K} h_K}{2C_E} \right\} \geq \\
& \geq \min \left\{ 1, \frac{C_1}{2C_E} \right\} \max \{ \varepsilon, \|b\|_{0,\infty,K} h_K \} \underbrace{=}_{(1.87)} \min \left\{ 1, \frac{C_1}{2C_E} \right\} \|b\|_{0,\infty,K} h_K.
\end{aligned} \quad (1.88)$$

Pak je

$$\begin{aligned}
& \underbrace{\gamma_K}_{(1.86), (1.88), (1.79)} \leq \varepsilon + (d+1)C_2 \|b\|_{0,\infty,K} h_K + \left(\min \left\{ 1, \frac{C_1}{2C_E} \right\} \right)^{-1} \|b\|_{0,\infty,K} h_K \leq \\
& \leq \max \left\{ 1, (d+1)C_2 + \left(\min \left\{ 1, \frac{C_1}{2C_E} \right\} \right)^{-1} \right\} (\varepsilon + \|b\|_{0,\infty,K} h_K)
\end{aligned} \quad (1.89)$$

Je-li nyní naopak

$$\frac{h_K}{\|b\|_{0,\infty,K}} > \frac{h_K^2}{\varepsilon}, \quad (1.90)$$

pak z (1.82) plyne

$$\tau(c_K) \simeq \frac{h_K^2}{\varepsilon}. \quad (1.91)$$

(1.90) je také ekvivalentní nyní s

$$\varepsilon > \|b\|_{0,\infty,K} h_K. \quad (1.92)$$

Odhadneme opět nejdřív maximum v konstantě γ_K .

$$\begin{aligned} & \max \left\{ \varepsilon, \|b\|_{0,\infty,K}^2 \frac{\tau(c_K)}{2C_E} \right\} \stackrel{(1.91), (1.79)}{\geq} \max \left\{ \varepsilon, \frac{C_1 \|b\|_{0,\infty,K}^2 h_K^2}{2C_E \varepsilon} \right\} \geq \\ & \geq \frac{1}{\varepsilon} \min \left\{ 1, \frac{C_1}{2C_E} \right\} \max \{ \varepsilon^2, \|b\|_{0,\infty,K}^2 h_K^2 \} \stackrel{(1.92)}{\geq} \frac{1}{\varepsilon} \min \left\{ 1, \frac{C_1}{2C_E} \right\} \|b\|_{0,\infty,K}^2 h_K^2. \end{aligned} \quad (1.93)$$

Pak je

$$\begin{aligned} \gamma_K & \stackrel{(1.79), (1.91), (1.93)}{\leq} \varepsilon + (d+1)C_2 \frac{\|b\|_{0,\infty,K}^2 h_K^2}{\varepsilon} + \left(\min \left\{ 1, \frac{C_1}{2C_E} \right\} \right)^{-1} \varepsilon \stackrel{(1.92)}{\leq} \\ & \leq \varepsilon + (d+1)C_2 \varepsilon + \left(\min \left\{ 1, \frac{C_1}{2C_E} \right\} \right)^{-1} \varepsilon \leq \\ & \leq \left(1 + (d+1)C_2 + \left(\min \left\{ 1, \frac{C_1}{2C_E} \right\} \right)^{-1} \right) \varepsilon. \end{aligned} \quad (1.94)$$

Z (1.9) a odhadů (1.89) a (1.94) plyne $\gamma_K \lesssim \varepsilon + \|b\|_{0,\infty,K} h_K$, což po dosazení do již dokázané nerovnosti (1.81), po uvážení vztahu L^∞ -normy a integračního oboru a po úpravách dává dokazovanou nerovnost (1.83).

■

Poznámka: Motivace pro to, aby stabilizační parametr splňoval právě relaci (1.82), je následující. Je

$$\gamma_K \simeq \varepsilon + \tau(c_K) \|b\|_{0,\infty,K}^2 + \frac{\|b\|_{0,\infty,K}^2 h_K^2}{\max\{\varepsilon, \tau(c_K) \|b\|_{0,\infty,K}^2\}}. \quad (1.95)$$

Je-li nejprve $\varepsilon \leq \tau(c_K) \|b\|_{0,\infty,K}^2$, plyne z (1.95)

$$\gamma_K \simeq \varepsilon + \tau(c_K) \|b\|_{0,\infty,K}^2 + \frac{h_K^2}{\tau(c_K)}. \quad (1.96)$$

V rovnici (1.96) požadujeme stejné chování členů obsahujících $\tau(c_K)$, tj. požadujeme

$$\tau(c_K)\|b\|_{0,\infty,K}^2 \simeq \frac{h_K^2}{\tau(c_K)} \iff \tau(c_K) \simeq \frac{h_K}{\|b\|_{0,\infty,K}}. \quad (1.97)$$

Je-li nyní $\varepsilon \geq \tau(c_K)\|b\|_{0,\infty,K}^2$, plyne z (1.95)

$$\gamma_K \simeq \varepsilon + \tau(c_K)\|b\|_{0,\infty,K}^2 + \frac{\|b\|_{0,\infty,K}^2 h_K^2}{\varepsilon}. \quad (1.98)$$

V rovnici (1.98) požadujeme stejné chování druhých dvou členů, tj. požadujeme

$$\tau(c_K)\|b\|_{0,\infty,K}^2 \simeq \frac{\|b\|_{0,\infty,K}^2 h_K^2}{\varepsilon} \iff \tau(c_K) \simeq \frac{h_K^2}{\varepsilon}. \quad (1.99)$$

Požadavky (1.97) a (1.99) pak sjednotíme do jediného požadavku (1.82).

2. Situace pro $d = 1$

V celé této kapitole budeme uvažovat $d = 1$. S ohledem na předpoklad (1.2) budeme dále v celé této kapitole uvažovat

$$b(x) = b \in \mathbb{R} \setminus \{0\}, \quad x \in \bar{\Omega}. \quad (2.1)$$

2.1 Modelový problém – klasická a Galerkinova formulace

Uvažujeme data

$$\begin{aligned} \Omega &= (0, 1), & b & \text{jako v (2.1),} \\ f(x) &= 0, & x & \in \Omega, \\ u^B(0) &= u^0 \in \mathbb{R}, & u^B(1) &= u^1 \in \mathbb{R}. \end{aligned} \quad (2.2)$$

a definujeme

$$Q := \frac{b}{\varepsilon}. \quad (2.3)$$

Řešení úlohy 1 s daty (2.2) s využitím definice (2.3) je

$$u(x) = \frac{u^0 \left(1 - \exp(Q(x-1))\right) - u^1 \left(\exp(-Q) - \exp(Q(x-1))\right)}{1 - \exp(-Q)}, \quad x \in \Omega.$$

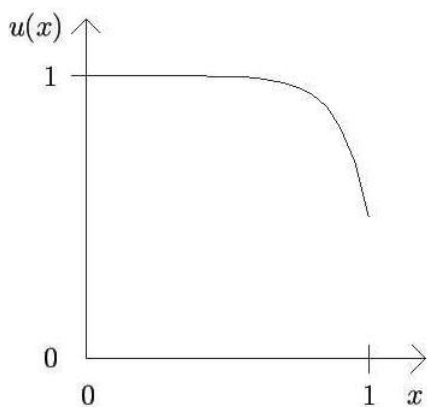
Uvažujeme data

$$\begin{aligned} \Omega &= (0, 1), & b &= 1, \\ f(x) &= 0, & x & \in \Omega, \\ u^B(0) &= 1, & u^B(1) &= 1/2, \end{aligned} \quad (2.4)$$

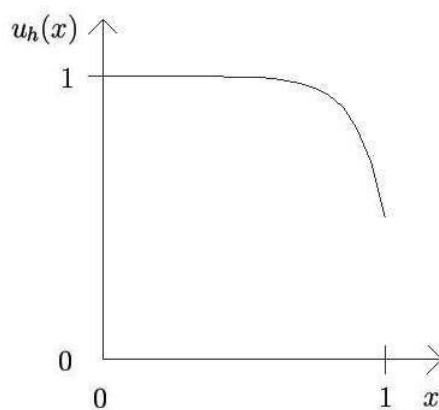
a jako triangulaci uvažujeme rovnoměrné dělení

$$\left. \begin{aligned} \text{card}(\mathcal{T}_h) &= M \in \mathbb{N}, \\ h &:= 1/\text{card}(\mathcal{T}_h) = 1/M, \\ N_T &= \{n_j \in [0, 1] : n_j = jh/2, j \in \{0, 2, 4, \dots, 2M\}\}, \\ N_S &= \{n_j \in [0, 1] : n_j = jh/2, j \in \{1, 3, 5, \dots, 2M-1\}\}. \\ &\implies \\ N_N &= \{n_j \in [0, 1] : n_j = jh/2, j \in \{0, 1, 2, \dots, 2M\}\}, \\ \mathcal{N}_h &= 2M + 1, \\ J &= \{0, 1, 2, \dots, 2M\}, \\ J(\Omega) &= \{1, 2, 3, \dots, 2M-1\}, \\ J(\partial\Omega) &= \{0, 2M\}, \\ \mathcal{T}_h &= \{[n_j, n_{j+2}], j \in \{0, 2, 4, \dots, 2M-2\}\}, \\ h_K &= h \forall K \in \mathcal{T}_h. \end{aligned} \right\} \quad (2.5)$$

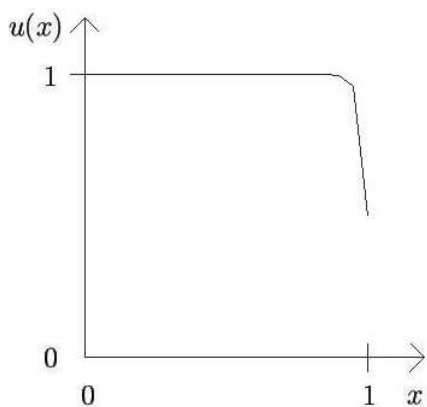
Řešení úlohy 1 a 8 s daty (2.4), různými ε a dělením (2.5) s $M = 10$ jsou vyobrazena na následujících obrázcích.



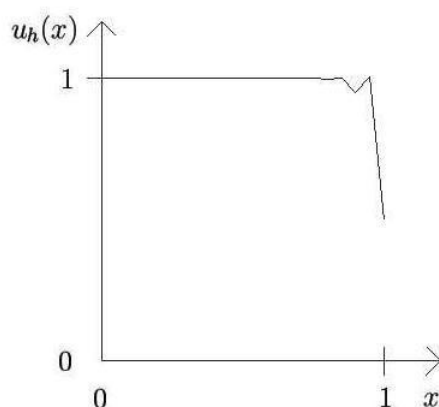
Obrázek 1: řešení úlohy 1 s daty (2.4) a $\varepsilon = 0,1$



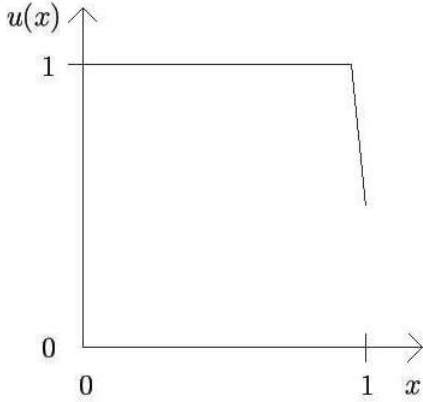
Obrázek 2: řešení úlohy 8 s daty (2.4), $\varepsilon = 0,1$ a dělením (2.5) s $M = 10$



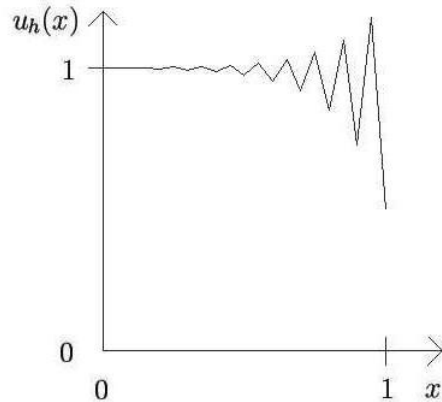
Obrázek 3: řešení úlohy 1 s daty (2.4) a $\varepsilon = 0,02$



Obrázek 4: řešení úlohy 8 s daty (2.4), $\varepsilon = 0,02$ a dělením (2.5) s $M = 10$



Obrázek 5: řešení úlohy 1 s daty (2.4) a $\varepsilon = 0,005$



Obrázek 6: řešení úlohy 8 s daty (2.4), $\varepsilon = 0,005$ a dělením (2.5) s $M = 10$

Pro dané ε má být řešení úlohy 8 přibližným řešením úlohy 1, z obrázků 1 až 6 je ale zřejmé, že v řešeních úlohy 8 se vyskytují nefyzikální oscilace, což je nežádoucí jev. Prozkoumáme další souvislosti týkající se těchto oscilací. Porovnáme obrázek 1 resp. 3 resp. 5 s obrázkem 2 resp. 4 resp. 6 a vidíme, že velikost absolutní hodnoty derivace řešení úlohy 1 odpovídá velikosti oscilací řešení úlohy 8 v tom smyslu, že čím větší absolutní hodnota derivace se v řešení úlohy 1 vyskytuje, tím větší jsou oscilace v řešení úlohy 8. Přitom oscilace v řešení úlohy 8 se nacházejí v okolí bodu $x = 1$, což je bod, ve kterém vektor b „směřuje ven z výpočetní oblasti“. Toto okolí je tím větší, čím menší je ε , a tedy čím více převládá konvekce nad difúzí.

Dále, kdyby bylo $u^0 = u^1$, bylo by řešení triviálně rovné u^0 v celém intervalu $[0, 1]$. Protože je ale $u^0 \neq u^1$ a konvekce je převládající jev, je řešení skoro v celém intervalu $[0, 1]$ přibližně rovno u^0 a jeho hodnota se mění na hodnotu u^1 až poblíž bodu $x = 1$. Hodnota u^1 svou odlišností od hodnoty u^0 ve spojení s převládající konvekcí tedy blízce souvisí s prudkými růsty či poklesy řešení, s čímž je, jak je uvedeno výše, spojen výskyt oscilací. Na tom je postavena myšlenka potlačení oscilací, která je uvedena níže.

Pojem „bod, ve kterém vektor b směřuje ven z výpočetní oblasti a ve kterém okrajová podmínka předepisuje hodnotu řešení odlišnou od hodnoty řešení vyskytující se ve většině bodů z intervalu $[0, 1]$,“ se jeví být důležitý, budeme tedy tento pojem precizovat obecnější definicí.

Výtoková hranice oblasti Ω je množina

$$\partial\Omega_L := \{x \in \partial\Omega : b(x) \cdot \nu(x) > 0\}. \quad (2.6)$$

V případě úlohy 8 s daty (2.4) a dělením (2.5) je podle definice (2.6)

$$\partial\Omega_L = \{1\}, \quad J(\partial\Omega_L) = \{2M\}. \quad (2.7)$$

Uvažujeme následující myšlenku z práce [10].

Myšlenka ★: Potlačení oscilací může být dosaženo co největším zmenšením vlivu okrajové podmínky ve výpočetních uzlech z výtokové hranice oblasti Ω na hodnoty řešení ve výpočetních uzlech z vnitřku oblasti.

Soustava rovnic (1.42) v úloze 8 po použití dat (2.2) a dělení (2.5) přejde na tvar

$$\sum_{j=1}^{2M-1} a_G(\varphi_j, \varphi_i) u_j = -u^0 a_G(\varphi_0, \varphi_i) - \underbrace{u^1 a_G(\varphi_{2M}, \varphi_i)}_{G1}, \quad i \in \{1, 2, \dots, 2M-1\}. \quad (2.8)$$

Zmíněný vliv z myšlenky ★ je v případě úlohy 8 v soustavě rovnic (2.8) realizován výrazem označeným $G1$ protože ten jako jediný v této soustavě obsahuje hodnotu u^1 , což je hodnota okrajové podmínky na výtokové hranici oblasti Ω . Výraz $G1$ je přitom nenulový jen pro ta i , pro která mají nosiče funkcí φ_{2M} a φ_i potažmo jejich gradientů neprázdný průnik. V takovém výrazu je pak totiž podle definice (1.6) obecně nenulová hodnota $a_G(\varphi_{2M}, \varphi_i)$. Zavedeme tedy obecnější definici. Definujeme *výtokovou hraniční vrstvu*,

$$L := \{K \in \mathcal{T}_h : \exists i \in J(\Omega), j \in J(\partial\Omega_L) : K = \text{supp}(\varphi_i) \cap \text{supp}(\varphi_j)\}. \quad (2.9)$$

V případě úlohy 8 s daty (2.4) a dělením (2.5) je podle definice (2.9)

$$L = \{[1-h, 1]\}, \quad J(L \cap \Omega) = \{2M-2, 2M-1\}. \quad (2.10)$$

Realizace myšlenky ★ se budeme snažit dosáhnout použitím metody SUPG s vhodnou volbou stabilizačního parametru.

2.2 Modelový problém – SUPG formulace, volba stabilizačního parametru

Ukážeme, že metoda SUPG dává rozdíl od metody Galerkinovy přibližné řešení s výrazně potlačenými nežádoucími oscilacemi.

Uvažujeme data

$$\begin{aligned} \Omega &= (0, 1), & b &> 0, \\ f &\in C((0, 1)), \\ u^B(0) &= u^0 \in \mathbb{R}, & u^B(1) &= u^1 \in \mathbb{R}. \end{aligned} \quad (2.11)$$

Po použití dat (2.11) a dělení (2.5) přejde rovnice (1.43) v úloze 9 na tvar

$$\begin{aligned} \sum_{j=1}^{2M-1} a_{SUPG}(\varphi_j, \varphi_i) u_j &= F_{SUPG}(\varphi_i) - u^0 a_{SUPG}(\varphi_0, \varphi_i) - \underbrace{u^1 a_{SUPG}(\varphi_{2M}, \varphi_i)}_{G2}, \\ i &\in \{1, 2, \dots, 2M-1\}. \end{aligned} \quad (2.12)$$

V případě úlohy 9 s daty (2.11) a dělením (2.5) podle definic (2.6) a (2.9) opět platí (2.7) a (2.10). Označme jediný element v L jako T ,

$$T := [1 - h, 1].$$

Výpočetním uzlům, které obsahuje, ponecháme ty indexy, které jsou jim přiřazeny v definici dělení (2.5), tj. levý krajní resp. střední resp. pravý krajní výpočetní uzel má index $2M - 2$ resp. $2M - 1$ resp. $2M$.

Myšlenka ★ nyní znamená právě to, že chceme, aby soustava (2.12) co nejméně závisela na hodnotě u^1 . Ta se objevuje na její pravé straně ve výrazu označeném $G2$. Takže požadujeme

$$a_{SUPG}(\varphi_{2M}, \varphi_i) \approx 0, \quad i \in J(L \cap \Omega). \quad (2.13)$$

Připomeňme, že $a_{SUPG}(\varphi_{2M}, \varphi_i) = 0 \forall i \in J(\Omega \setminus L)$.

Uvažujeme převládající konvekci, tedy $\varepsilon \ll |b|h$, takže v definicích (1.6) a (1.30) lze zanedbat členy obsahující ε , a je pak

$$a_{SUPG}(u, v) \approx \int_{\Omega} (bu')(v + \tau b \nabla v'),$$

takže požadavek (2.13) lze přeformulovat na požadavek

$$\int_T (b\varphi'_{2M})(\varphi_i + \tau b\varphi'_i) = 0, \quad i \in J(L \cap \Omega). \quad (2.14)$$

2.2.1 Volba stabilizačního parametru metodou C

Jde o volbu po elementech konstantního stabilizačního parametru, která vychází z práce [6]. Nejprve definujeme

$$\tau_C := \frac{h}{4|b|} \left(\coth(Pe) - \frac{1}{Pe} \right), \quad Pe = \frac{|b|h}{2\varepsilon}. \quad (2.15)$$

Veličina Pe je tzv. *Pécletovo číslo*. Pak položíme

$$\tau|_K = \tau_C, \quad K \in \mathcal{T}_h.$$

kde τ_C je definováno v (2.15).

2.2.2 Volba stabilizačního parametru metodou 1

V první metodě prostě vyřešíme soustavu rovnic (2.14),

$$\begin{aligned} \int_T b\varphi'_{2M}(\varphi_{2M-2} + \tau b\varphi'_{2M-2}) &= 0, \\ \int_T b\varphi'_{2M}(\varphi_{2M-1} + \tau b\varphi'_{2M-1}) &= 0. \end{aligned} \quad (2.16)$$

V soustavě (2.16) jsou celkem **dvě rovnice lineární v parametru** τ . Nabízí se tedy **volit** $\tau|_T$ jako **afinní** funkci, protože taková funkce má právě **dva parametry**. To je nový přístup oproti pracím [6] a [10], ve kterých se parametr τ volí vždy konstantní na každém elementu $K \in \mathcal{T}_h$.

Uvažujeme tedy $\tau|_T$ v následujícím tvaru vyjádřeném pomocí barycentrických souřadnic na elementu T :

$$\tau|_T = \tau_{T,2M-2}\lambda_{T,2M-2} + \tau_{T,2M}\lambda_{T,2M}, \quad \tau_{T,2M-2}, \tau_{T,2M} \in \mathbb{R}. \quad (2.17)$$

Do soustavy rovnic (2.16) dosadíme za bázové funkce z vyjádření (1.15) a za τ z vyjádření (2.17), roznásobíme integrandy, použijeme větu 5, zkrátíme, a dostaneme tak soustavu rovnic

$$\underbrace{\begin{pmatrix} b, & b \\ b, & 3b \end{pmatrix}}_{\mathbf{M}_1} \begin{pmatrix} \tau_{T,2M-2} \\ \tau_{T,2M} \end{pmatrix} = \begin{pmatrix} h \\ h \end{pmatrix}. \quad (2.18)$$

Pro determinant matice \mathbf{M}_1 platí

$$\det \mathbf{M}_1 = 2b^2 \neq 0.$$

To, že je determinant matice \mathbf{M}_1 nenulový, plyne z předpokladu kladeného na b v (2.4). Soustava (2.18) tedy má jediné řešení. Tím je

$$\tau_{T,2M-2} = \frac{h}{b}, \quad \tau_{T,2M} = 0. \quad (2.19)$$

Z rovnic (2.17) a (2.19) vzhledem k předpokladům $h > 0$ a (2.4) a vzhledem k definici barycentrických souřadnic (1.12) plyne, že τ je nezáporné na celém elementu T .

Na každém elementu $K \in \mathcal{T}_h \setminus L$ zvolíme $\tau|_K = \tau_C$, viz (2.15).

2.2.3 Volba stabilizačního parametru metodou 2

Teď pouze za účelem volby τ uvažujeme jinou bázi prostoru X_h , než je $\text{base}(X_h)$, a totiž

$$\text{base}_2(X_h) := \left\{ \psi_j \in X_h : \psi_j = \begin{cases} \lambda_{T,2M}, & j = 2M \\ \varphi_j, & j = 0, 1, \dots, 2M-1 \end{cases} \right\}. \quad (2.20)$$

Platí $\text{supp}(\varphi_j) = \text{supp}(\psi_j) \forall j \in \{0, 1, \dots, 2M\}$, takže výtoková hraniční vrstva je stejná jako ve výpočtech pro předchozí metodu, a tedy je třeba splnit opět požadavek (2.14), do kterého nyní dosadíme $\varphi_{2M} = \psi_{2M}$ a $\varphi_i = \psi_i$. Dostáváme soustavu rovnic

$$\begin{aligned} \int_{\Omega} b\psi'_{2M}(\psi_{2M-2} + \tau b\psi'_{2M-2}) &= 0, \\ \int_{\Omega} b\psi'_{2M}(\psi_{2M-1} + \tau b\psi'_{2M-1}) &= 0. \end{aligned} \quad (2.21)$$

Do soustavy (2.21) dosadíme za bázové funkce tentokrát funkce z množiny (2.20) a za τ dosadíme opět z vyjádření (2.17), roznásobíme integrandy, použijeme větu 5, zkrátíme, a dostaneme tak soustavu rovnic

$$\underbrace{\begin{pmatrix} 5b & b \\ b & -b \end{pmatrix}}_{\mathbf{M}_2} \begin{pmatrix} \tau_{T,2M-2} \\ \tau_{T,2M} \end{pmatrix} = \begin{pmatrix} h \\ -h \end{pmatrix}. \quad (2.22)$$

Pro determinant matice \mathbf{M}_2 platí

$$\det \mathbf{M}_2 = -6b^2 \neq 0.$$

To, že je determinant matice \mathbf{M}_2 nenulový, plyne opět z předpokladu kladeného na b v (2.4).

Soustava (2.22) tedy má jediné řešení. Tím je

$$\tau_{T,2M-2} = 0, \quad \tau_{T,2M} = \frac{h}{b}.$$

Z rovnic (2.17) a (2.19) vzhledem k předpokladům $h > 0$ a (2.4) a vzhledem k definici barycentrických souřadnic (1.12) opět plyne, že τ je nezáporné na celém elementu T .

Na každém elementu $K \in \mathcal{T}_h \setminus L$ zvolíme stejně jako v metodě 1 $\tau|_K = \tau_C$, viz (2.15).

2.2.4 Rozdíl mezi metodou 1 a 2

Metodou 1 i 2 je potlačena závislost řešení úlohy 9 na hodnotě u^1 . Souřadnice řešení u_j , $j \in \{1, 2, \dots, 2M - 1\}$, jsou počítány vzhledem k bázi $\text{base}(X_h)$, a to bez ohledu na použitou metodu volby stabilizačního parametru. V případě metody 1 je stabilizační parametr volen opět vzhledem k bázi $\text{base}(X_h)$, zatímco v případě metody 2 je stabilizační parametr volen vzhledem k bázi $\text{base}_2(X_h)$, která se od báze $\text{base}(X_h)$ liší. Pro bázi $\text{base}(X_h)$ totiž platí $\psi_{2M}(n_{2M-1}) = 0$, zatímco pro bázi $\text{base}_2(X_h)$ platí $\psi_{2M}(n_{2M-1}) = 1/2$, a tak pro hodnoty u_h a souřadnice řešení vzhledem k bázi $\text{base}(X_h)$ platí

$$u_h(n_j) = u_j, \quad j \in \{1, 2, \dots, 2M - 1\},$$

tj. souřadnice řešení jsou hodnotami u_h ve výpočetních uzlech, zatímco pro hodnoty u_h a souřadnice řešení vzhledem k bázi $\text{base}_2(X_h)$ platí

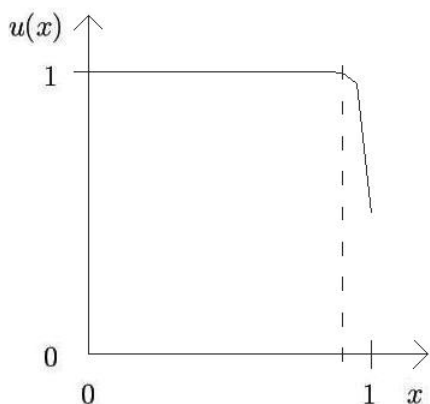
$$u_h(n_j) = u_j, \quad j \in \{1, 2, \dots, 2M - 2\},$$

$$u_h(n_{2M-1}) = u_{2M-1} + \frac{1}{2}u^1.$$

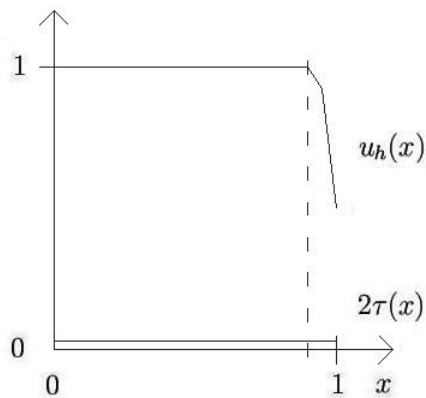
Na základě této odlišnosti báze pro výpočet řešení od báze pro volbu stabilizačního parametru v případě metody 2 proto lze očekávat, že hodnota $u_h(n_{2M-1})$ bude v případě metody 2 zatížena velkou chybou.

2.3 Modelový problém – numerické výsledky

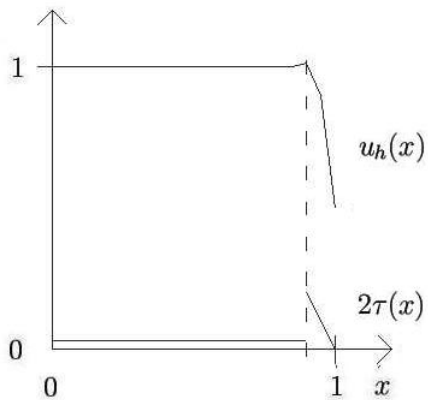
Řešení úlohy 1 a 9 s daty (2.4), různými ε , dělením (2.5) s $M = 10$ a τ získanými metodou C, 1 a 2 jsou vyobrazena na následujících obrázcích. Pro přehlednost je zobrazen dvojnásobek stabilizačního parametru, dále je přerušovanou čarou naznačen element T .



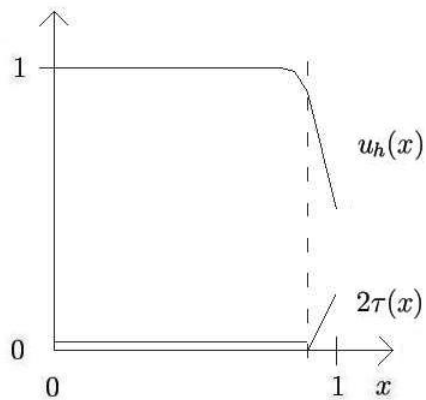
Obrázek 7: řešení úlohy 1 s daty (2.4) a $\varepsilon = 0,02$



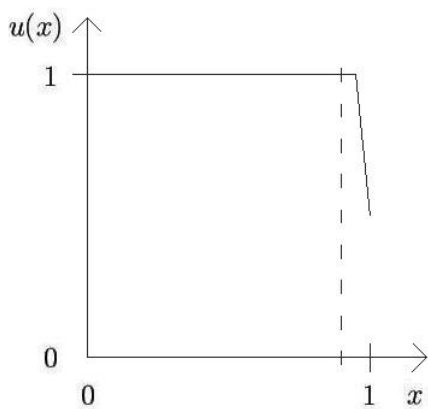
Obrázek 8: řešení úlohy 9 s daty (2.4), $\varepsilon = 0,02$ a dělením (2.5) s $M = 10$, τ voleno metodou C



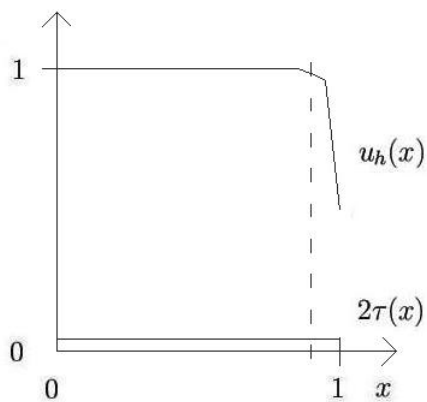
Obrázek 9: řešení úlohy 9 s daty (2.4), $\varepsilon = 0,02$ a dělením (2.5) s $M = 10$, τ voleno metodou 1



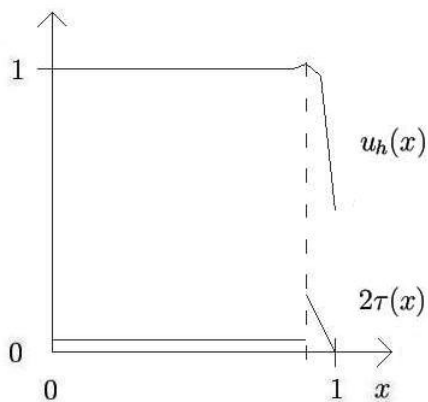
Obrázek 10: řešení úlohy 9 s daty (2.4), $\varepsilon = 0,02$ a dělením (2.5) s $M = 10$, τ voleno metodou 2



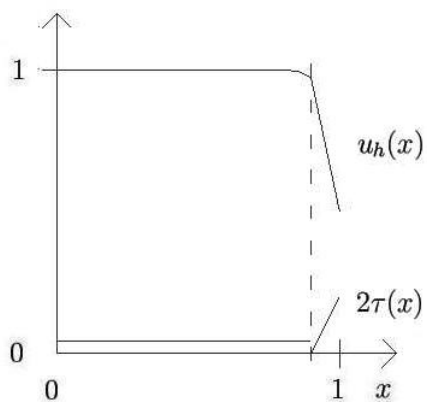
Obrázek 11: řešení úlohy 1 s daty (2.4) a $\varepsilon = 0,005$



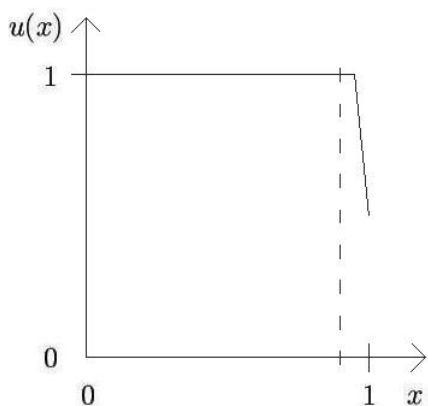
Obrázek 12: řešení úlohy 9 s daty (2.4), $\varepsilon = 0,005$ a dělením (2.5) s $M = 10$, τ voleno metodou C



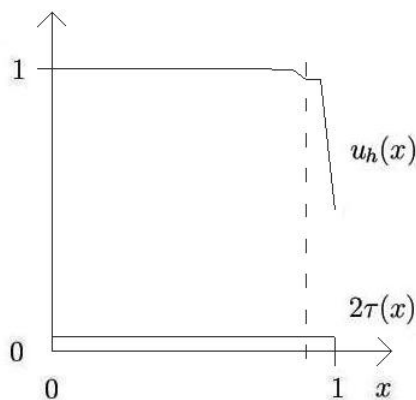
Obrázek 13: řešení úlohy 9 s daty (2.4), $\varepsilon = 0,005$ a dělením (2.5) s $M = 10$, τ voleno metodou 1



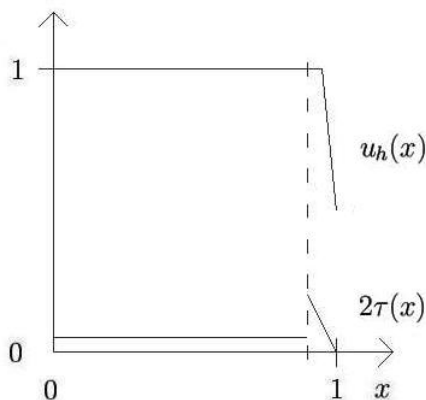
Obrázek 14: řešení úlohy 9 s daty (2.4), $\varepsilon = 0,005$ a dělením (2.5) s $M = 10$, τ voleno metodou 2



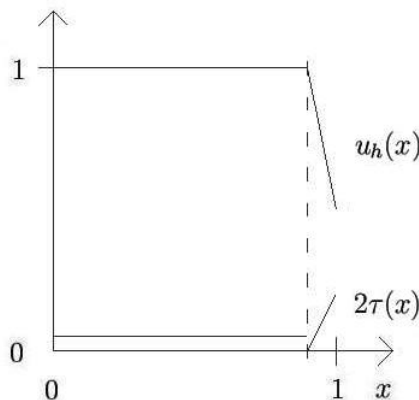
Obrázek 15: řešení úlohy 1 s daty (2.4) a $\varepsilon = 10^{-8}$



Obrázek 16: řešení úlohy 9 s daty (2.4), $\varepsilon = 10^{-8}$ a dělením (2.5) s $M = 10$, τ voleno metodou C



Obrázek 17: řešení úlohy 9 s daty (2.4), $\varepsilon = 10^{-8}$ a dělením (2.5) s $M = 10$, τ voleno metodou 1



Obrázek 18: řešení úlohy 9 s daty (2.4), $\varepsilon = 10^{-8}$ a dělením (2.5) s $M = 10$, τ voleno metodou 2

2.4 Obecný problém – volba stabilizačního parametru

Uvažujeme b jako v (2.1) a obecně nerovnoměrné dělení výpočetního intervalu, tj. $\exists K, T \in \mathcal{T}_h, K \neq T : h_K \neq h_T$. Zobecníme metodu C, 1 a 2.

volba po částech konstantního stabilizačního parametru metodou C (viz práci [6]):

- Uvažujeme

$$\tau|_K = \tau_{CK} := \frac{h_K}{4|b|} \left(\coth(Pe) - \frac{1}{Pe} \right), \quad Pe = \frac{|b|h_K}{2\varepsilon}, \quad K \in \mathcal{T}_h.$$

volba po částech afinního stabilizačního parametru metodou 1 a 2:

Pro daný element $K \in \mathcal{T}_h$ přiřadíme jeho levému resp. pravému krajnímu bodu index 1 resp. 2.

- Pro obě metody uvažujeme

$$\begin{aligned} \tau|_K &= \tau_{K,1}\lambda_{K,1} + \tau_{K,2}\lambda_{K,2}, & \tau_{K,1}, \tau_{K,2} &\in \mathbb{R}, & K &\in \mathcal{T}_h, \\ \tau_{K,1} &= \tau_{K,2} = \tau_{CK}, & K &\in \mathcal{T}_h \setminus L. \end{aligned}$$

- Pro metodu 1 uvažujeme

$$\begin{aligned} \tau_{K,1} &= \frac{h_K}{|b|}, & K &\in L \wedge a_{K,1} \in \Omega, \\ \tau_{K,2} &= 0, & K &\in L \wedge a_{K,2} \in \partial\Omega. \end{aligned}$$

- Pro metodu 2 uvažujeme

$$\begin{aligned} \tau_{K,1} &= 0, & K &\in L \wedge a_{K,1} \in \Omega, \\ \tau_{K,2} &= \frac{h_K}{|b|}, & K &\in L \wedge a_{K,2} \in \partial\Omega, \end{aligned}$$

2.5 Další numerické výsledky

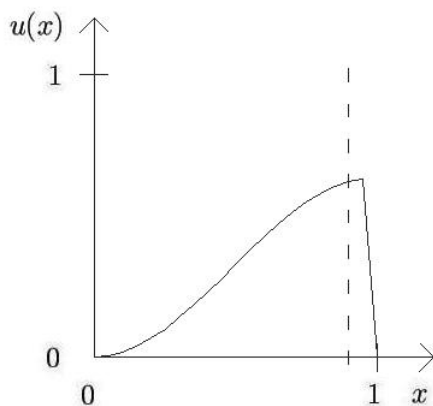
Uvažujeme data

$$\begin{aligned} \Omega &= (0, 1), & b &= 1, \\ f(x) &= \sin(\pi x), & x &\in \Omega, \\ u^B(0) &= u^B(1) = 0. \end{aligned} \tag{2.23}$$

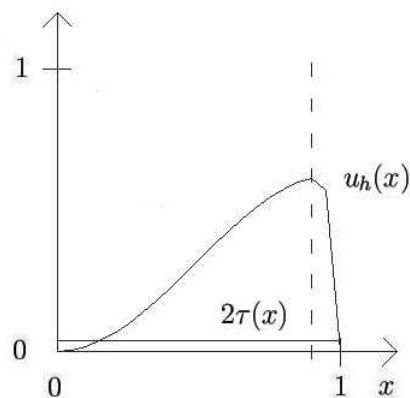
Řešení úlohy 1 s daty (2.23) s využitím definice (2.3) je

$$u(x) = \frac{1}{\varepsilon(Q^2 + \pi^2)} \left(\sin(\pi x) - \frac{Q}{\pi} \cos(\pi x) + \frac{Q(\exp(-Q) + 1 - 2\exp(Q(x-1)))}{\pi(1 - \exp(-Q))} \right).$$

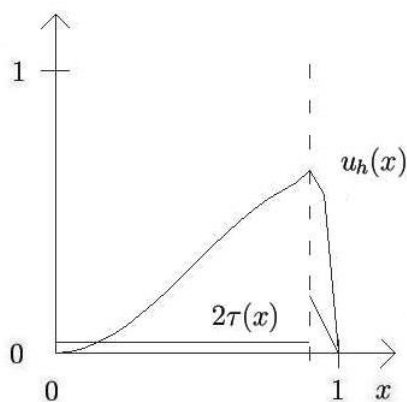
Řešení úlohy 1 a 8 s daty (2.23), různými ε a dělením (2.5) s $M = 10$ jsou vyobrazena na následujících obrázcích. Pro přehlednost je opět zobrazen dvojnásobek stabilizačního parametru a dále je opět přerušovanou čarou naznačen element T .



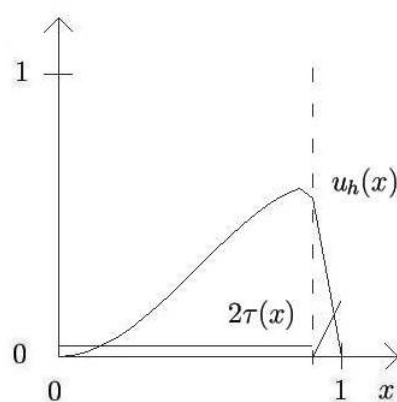
Obrázek 19: řešení úlohy 1 s daty (2.23) a $\varepsilon = 0,02$



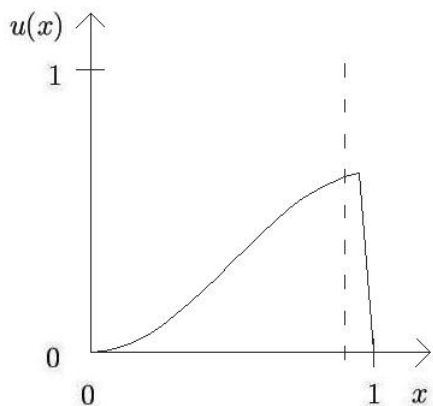
Obrázek 20: řešení úlohy 9 s daty (2.23), $\varepsilon = 0,02$ a dělením (2.5) s $M = 10$, τ voleno metodou C



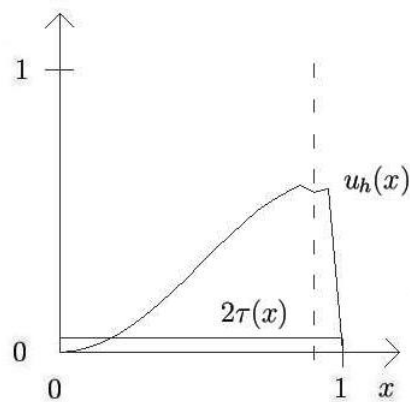
Obrázek 21: řešení úlohy 9 s daty (2.23), $\varepsilon = 0,02$ a dělením (2.5) s $M = 10$, τ voleno metodou 1



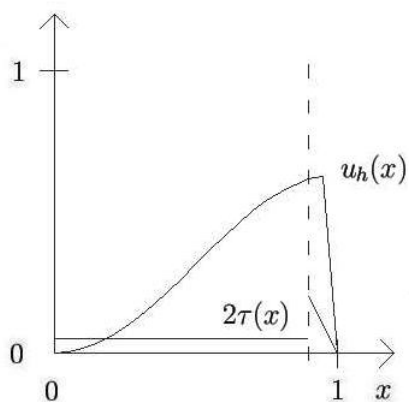
Obrázek 22: řešení úlohy 9 s daty (2.23), $\varepsilon = 0,02$ a dělením (2.5) s $M = 10$, τ voleno metodou 2



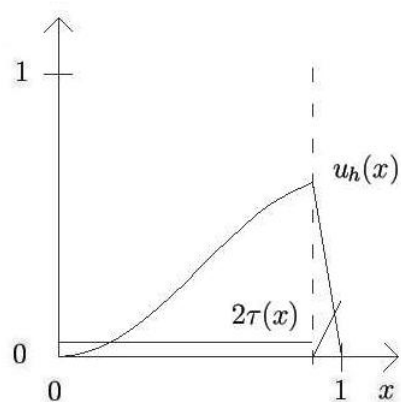
Obrázek 23: řešení úlohy 1 s daty (2.23) a $\varepsilon = 10^{-8}$



Obrázek 24: řešení úlohy 9 s daty (2.23), $\varepsilon = 10^{-8}$ a dělením (2.5) s $M = 10$, τ voleno metodou C



Obrázek 25: řešení úlohy 9 s daty (2.23), $\varepsilon = 10^{-8}$ a dělením (2.5) s $M = 10$, τ voleno metodou 1



Obrázek 26: řešení úlohy 9 s daty (2.23), $\varepsilon = 10^{-8}$ a dělením (2.5) s $M = 10$, τ voleno metodou 2

3. Situace pro $d = 2$

V celé této kapitole budeme uvažovat $d = 2$.

3.1 Modelový problém – klasická a Galerkinova formulace

Uvažujeme data

$$\begin{aligned} \Omega &= (0, 1) \times (0, 1), \\ b(x) &= (1, 2), \quad x \in \bar{\Omega}, \\ f(x) &= 0, \quad x \in \Omega, \\ u^B(x) &= 1 - \frac{1}{4} \left(\frac{1 - \exp(Q_1(x_{[1]} - 1))}{1 - \exp(-Q_1)} \right) \left(\frac{1 - \exp(Q_2(x_{[2]} - 1))}{1 - \exp(-Q_2)} \right), \\ & \quad x \in \partial\Omega, \end{aligned} \tag{3.1}$$

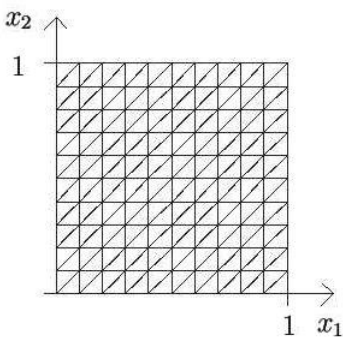
kde

$$Q_1 := \frac{b_{[1]}}{\varepsilon}, \quad Q_2 := \frac{b_{[2]}}{\varepsilon}. \tag{3.2}$$

Pro řešení úlohy 1 s daty (3.1) s využitím definice (3.2) platí

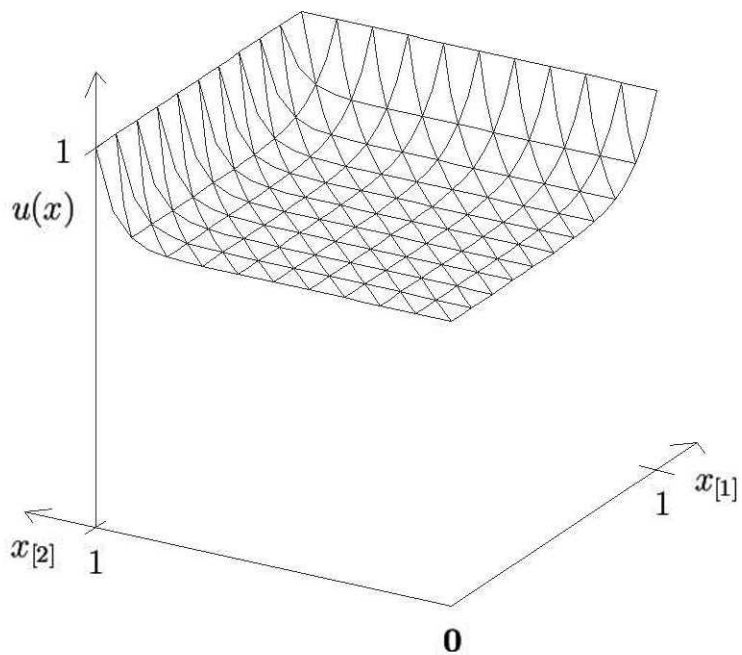
$$u(x) = u^B(x), \quad x \in \Omega.$$

Uvažujeme triangulaci vyobrazenou na následujícím obrázku.

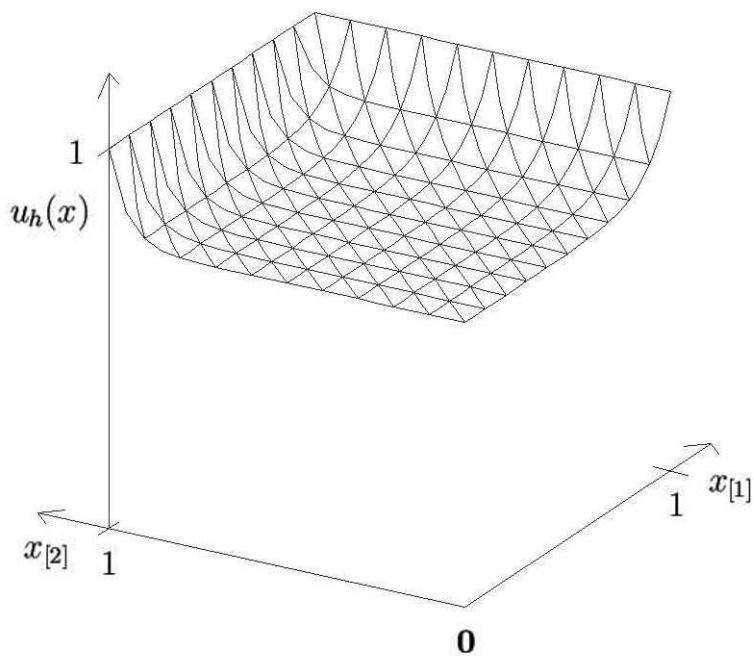


Obrázek 27: triangulace $\mathcal{T}_{h,1}$

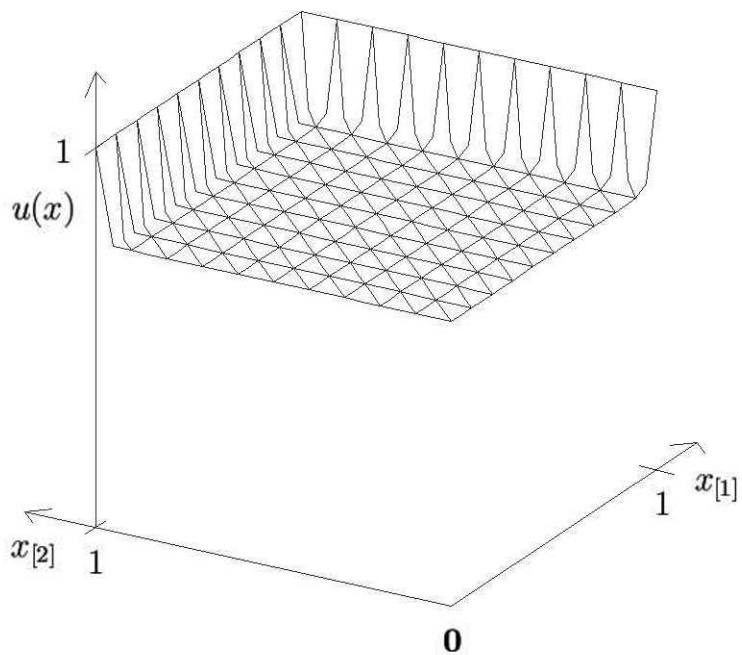
Triangulace $\mathcal{T}_{h,1}$ má 10×10 stejných dílčích čtverců, dále dělených svými diagonálami. Řešení úlohy 1 a 8 s daty (3.1), různými ε a triangulací $\mathcal{T}_{h,1}$ jsou vyobrazena na následujících obrázcích.



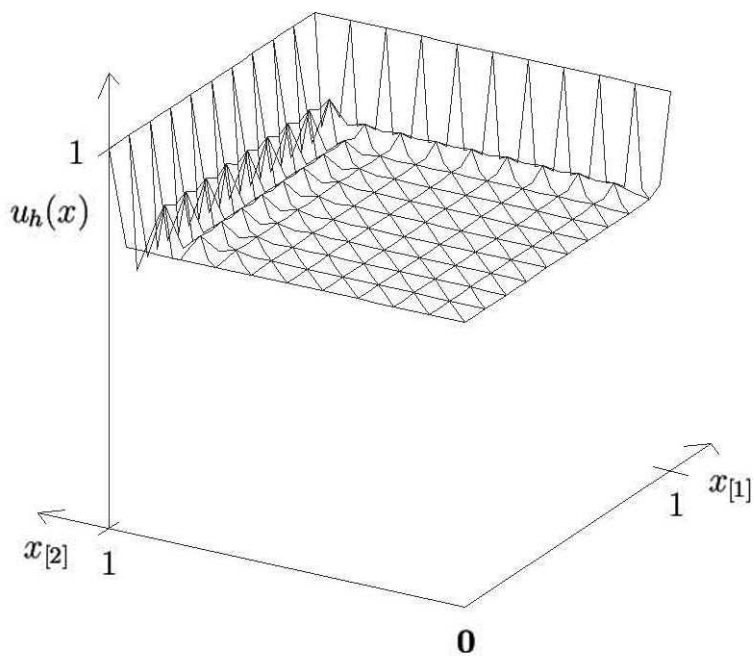
Obrázek 28: řešení úlohy 1 s daty (3.1), $\varepsilon = 0,1$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27



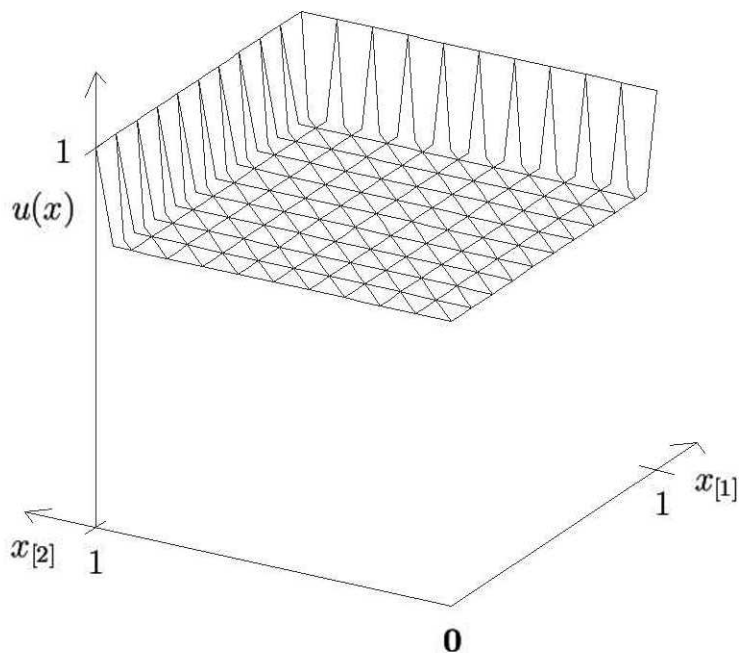
Obrázek 29: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,1$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27



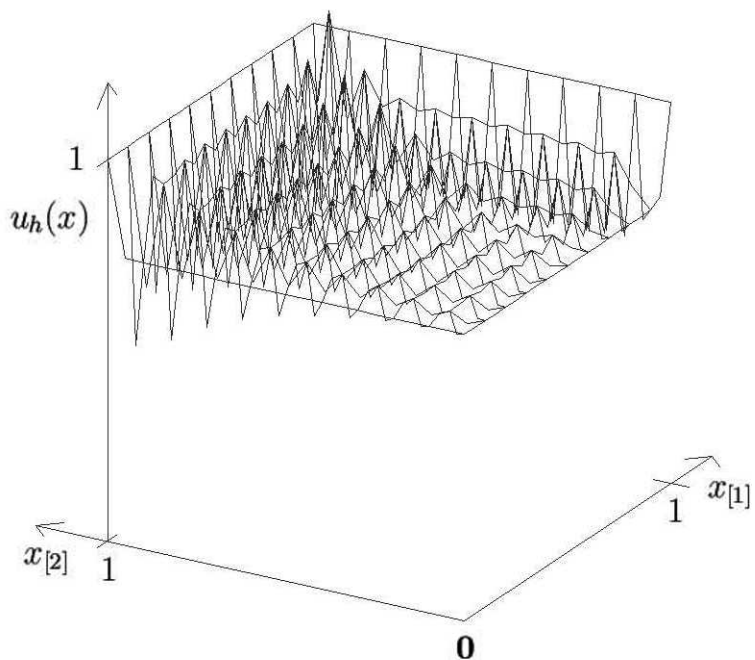
Obrázek 30: řešení úlohy 1 s daty (3.1), $\varepsilon = 0,02$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27



Obrázek 31: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,02$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27



Obrázek 32: řešení úlohy 1 s daty (3.1), $\varepsilon = 0,005$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27



Obrázek 33: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,005$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27

Pro dané ε a b má opět být řešení úlohy 8 přibližným řešením úlohy 1, z obrázků 28 až 33 je ale opět zřejmé, že v řešeních úlohy 8 se vyskytují nefyzikální oscilace, což je nežádoucí

jev. Porovnáme obrázek 28 resp. 30 resp. 32 s obrázkem 29 resp. 31 resp. 33 a vidíme, že velikost gradientu řešení úlohy 1 odpovídá velikosti oscilací řešení úlohy 8 v tom smyslu, že čím větší gradient se v řešení úlohy 1 vyskytuje, tím větší jsou oscilace v řešení úlohy 8. Přitom oscilace v řešení úlohy 8 se nacházejí v okolí části hranice $(\{1\} \times [0, 1]) \cup ([0, 1] \times \{1\})$, což je v této modelové úloze výtoková hranice oblasti Ω podle definice (2.6). Toto okolí je tím větší, čím menší je ε , a tedy čím více převládá konvekce nad difuzí.

3.2 Obecný problém – SUPG formulace, volba stabilizačního parametru

Ukážeme, že metoda SUPG dává rozdíly od metody Galerkinovy přibližné řešení opět s výrazně potlačenými nežádoucími oscilacemi.

Uvažujeme obecná data

$$\Omega = \text{mnohoúhelník}, \quad b \in L^\infty(\Omega), \quad f \in L^2(\Omega), \quad u^B \in C(\partial\Omega).$$

a připomeňme rovnici (1.43) z úlohy 9

$$\sum_{j \in J(\Omega)} a_{SUPG}(\varphi_j, \varphi_i) u_j = F_{SUPG}(\varphi_i) - \underbrace{\sum_{j \in J(\partial\Omega)} u^B(n_j) a_{SUPG}(\varphi_j, \varphi_i)}_{G3}, \quad i \in J(\Omega). \quad (3.3)$$

Myšlenka ★ nyní znamená právě to, že chceme, aby soustava (3.3) co nejméně závisela na hodnotách $u^B(n_j)$, $j \in J(\partial\Omega_L)$. Ty se objevují na její pravé straně ve výrazu označeném $G3$. Postup obdobný postupu z případu pro $d = 1$ by byl požadovat $G3 \approx 0$, $i \in J(\Omega)$, $j \in J(\partial\Omega_L)$. My však tento požadavek zúžíme jen na některé členy, ty budeme specifikovat dále. K tomu budeme potřebovat následující úvahy a pojmy.

V případě $d = 2$ existuje ve výtokové hraniční vrstvě několik typů elementů. Tyto typy se liší průnikem elementu s výtokovou hranicí oblasti Ω .

Mějme dán libovolný element $K \in L$. Následně pro účely dalších úvah a výpočtů přiřadíme indexy výpočetním uzlům náležejícím elementu K .

Objekt:	Označení:
výpočetní uzly, které jsou zároveň vrcholy elementu K , tedy prvky množiny $N_T \cap K$	n_1, n_2, n_3
výpočetní uzly, které jsou zároveň středy hran elementu K , tedy prvky množiny $N_S \cap K$	$n_{12}, n_{23}, n_{31},$ $n_{ij} = \text{střed hrany mezi uzly } n_i \text{ a } n_j,$ $i, j \in \{1, 2, 3\}$

(3.4)

Označení (3.4) budeme od teď dál uvažovat vždy, když bude dán nějaký element $K \in L$.

Přiřazení indexů výpočetním uzlům v (3.4) zatím nezohledňuje vztah elementu K a výtokové hranice oblasti Ω . Zvlášť se budeme zabývat dvěma typy elementů z výtokové hraniční vrstvy. Tyto typy definujeme následujícími dvěma definicemi.

- Řekneme, že element $K \in L$ je *typu 1*, jestliže má s výtokovou hranicí oblasti Ω společnou právě jednu hranu, v takovém případě dále uvažujeme

$$n_1 \in \Omega, \quad K \cap \partial\Omega_L = \{(1-c)n_2 + cn_3, c \in [0, 1]\}. \quad (3.5)$$

- Řekneme, že element $K \in L$ je *typu 2*, jestliže má s výtokovou hranicí oblasti Ω společný právě jeden vrchol, v takovém případě dále uvažujeme

$$n_1, n_2 \in \Omega, \quad K \cap \partial\Omega_L = \{n_3\}. \quad (3.6)$$

Množinu všech elementů $K \in L$, které jsou typu 1 nebo 2, budeme značit $L(1, 2)$.

Uvažujeme po elementech konstantní aproximaci vektoru b ,

$$\begin{aligned} b|_K &= b_K \in \mathbb{R}^2, & K \in \mathcal{T}_h, \\ \|b_K\| &> 0, & K \in L(1, 2) \end{aligned} \quad (3.7)$$

Definujeme normálu k výtokové hranici elementu typu 1 a 2,

$$\nu(K) := \begin{cases} \nu(n_{23}), & K \text{ je typu 1,} \\ \nu(n_3), & K \text{ je typu 2,} \end{cases}$$

(připomínáme definici normály ve vrcholu výpočetní oblasti (1.1))

a dále definujeme průmět vektoru b_K do směru $\nu(K)$,

$$b_K^\nu := (b_K \cdot \nu(K))\nu(K), \quad K \in L(1, 2).$$

Opět uvažujeme převládající konvekci, tedy $\varepsilon \ll \|b\|h$, takže v definicích (1.6) a (1.30) lze opět zanedbat členy obsahující ε . Platí

$$\begin{aligned} & a_{SUPG}(\varphi_j, \varphi_i) \underbrace{\varepsilon \ll \|b\|h}_{\approx} \sum_{K \in \mathcal{T}_h} \int_K (b_K \cdot \nabla \varphi_j)(\varphi_i + \tau b_K \cdot \nabla \varphi_i) = \\ &= \underbrace{\sum_{K \in L(1,2)} \int_K (b_K \cdot \nabla \varphi_j)(\varphi_i + \tau b_K \cdot \nabla \varphi_i)}_{G4_{ji}} + \underbrace{\sum_{K \in \mathcal{T}_h \setminus L(1,2)} \int_K (b_K \cdot \nabla \varphi_j)(\varphi_i + \tau b_K \cdot \nabla \varphi_i)}_{G5_{ji}}, \end{aligned} \quad (3.8)$$

kde

$$\begin{aligned} G4_{ji} &= \sum_{K \in L(1,2)} \int_K ((b_K^\nu + (b_K - b_K^\nu)) \cdot \nabla \varphi_j)(\varphi_i + \tau(b_K^\nu + (b_K - b_K^\nu)) \cdot \nabla \varphi_i) = \\ &= \underbrace{\sum_{K \in L(1,2)} \int_K (b_K^\nu \cdot \nabla \varphi_j)(\varphi_i + \tau b_K^\nu \cdot \nabla \varphi_i)}_{G6_{ji}} + \underbrace{\sum_{K \in L(1,2)} \int_K (b_K^\nu \cdot \nabla \varphi_j)(\tau(b_K - b_K^\nu) \cdot \nabla \varphi_i)}_{G7_{ji}} + \\ & \quad + \underbrace{\sum_{K \in L(1,2)} \int_K ((b_K - b_K^\nu) \cdot \nabla \varphi_j)(\varphi_i + \tau(b_K^\nu + (b_K - b_K^\nu)) \cdot \nabla \varphi_i)}_{G8_{ji}}, \end{aligned} \quad (3.9)$$

takže po dosazení přibližné rovnosti (3.8) a rovnosti (3.9) do $G3$ a použití linearitly sum ve výrazu s podvýrazem $G6_{ji}$ je

$$G3 \approx \underbrace{\sum_{K \in L(1,2)} \sum_{j \in J(\partial\Omega)} u^B(n_j) \int_K (b_K^\nu \cdot \nabla \varphi_j)(\varphi_i + \tau b_K^\nu \cdot \nabla \varphi_i)}_{G9} + \sum_{j \in J(\partial\Omega)} (G5_{ji} + G7_{ji} + G8_{ji}),$$

$$i \in J(\Omega).$$

Myšlenku \star použijeme na výraz $G9$, tedy s využitím definice (2.9) požadujeme

$$\sum_{K \in L(1,2)} \sum_{j \in J(\partial\Omega_L \cap K)} u^B(n_j) \int_K (b_K^\nu \cdot \nabla \varphi_j)(\varphi_i + \tau b_K^\nu \cdot \nabla \varphi_i) = 0, \quad i \in J(\Omega \cap L). \quad (3.10)$$

Jednou z postačujících, nikoliv však nutných podmínek pro splnění požadavku (3.10) je požadavek

$$\sum_{j \in J(\partial\Omega_L \cap K)} u^B(n_j) \int_K (b_K^\nu \cdot \nabla \varphi_j)(\varphi_i + \tau b_K^\nu \cdot \nabla \varphi_i) = 0 \quad (3.11)$$

$$\forall K \in L(1,2), \quad \forall i \in J(\Omega \cap K).$$

Z požadavku (3.11) budeme pro volbu stabilizačního parametru v této práci nadále vycházet. Při volbě stabilizačního parametru budeme typ daného elementu zohledňovat.

Ještě před popisem voleb stabilizačního parametru zavedeme značení

$$S_{T,i} := b_T^\nu \cdot \nabla \lambda_{T,i}, \quad T \in L(1,2), \quad i \in \{1,2,3\}. \quad (3.12)$$

Veličina $S_{T,i}$ má kromě zjednodušení zápisu jistých výrazů také další užitečné vlastnosti. Díky rovnici (1.14) a linearitě skalárního součinu platí

$$\sum_{i \in \{1,2,3\}} S_{T,i} = 0, \quad T \in L(1,2). \quad (3.13)$$

Dále, protože trojúhelník nemůže mít žádné své dvě různé strany rovnoběžné a protože gradient barycentrické souřadnice definované na trojúhelníku je kolmý ke straně protilehlé vrcholu, ke kterému je tato barycentrická souřadnice přidružena, nemůžou být rovnoběžné ani žádné dva gradienty příslušné dvěma různým barycentrickým souřadnicím. Takže uvažujeme-li element $T \in L(1,2)$, pak platí následující.

$$\begin{aligned} &\text{Je-li pro nějaké } i \in \{1,2,3\} \text{ splněno } S_{T,i} = 0, \\ &\text{pak z nenulovosti } b_T \text{ musí být } S_{T,j} \neq 0, j \in \{1,2,3\} \setminus \{i\}, \\ &\text{takže z (3.13) pak plyne } S_{T,j} = -S_{T,k} \neq 0, j, k \in \{1,2,3\}, j \neq i, k \neq i. \end{aligned} \quad (3.14)$$

3.2.1 Volba stabilizačního parametru metodou C

Jde o volbu opět po elementech konstantního stabilizačního parametru, která je obdobou volby metodou C z případu $d = 1$. Nejprve definujeme

$$\tau_{C2K} := \frac{h_K^b}{4\|b_K\|} \left(\coth(Pe) - \frac{1}{Pe} \right), \quad Pe = \frac{\|b_K\| h_K^b}{2\varepsilon}. \quad (3.15)$$

Pak položíme

$$\tau|_K = \tau_{C2K}, \quad K \in \mathcal{T}_h.$$

kde τ_{C2K} je definováno v (3.15).

3.2.2 Volba stabilizačního parametru metodou 1

Mějme dán libovolný element $T \in L(1, 2)$.

Případ, kdy element T je typu 1

Uvažujeme značení (3.4) a navíc to, že je splněno (3.5). Pak z požadavku (3.11) dostaneme soustavu rovnic

$$\sum_{j \in \{2, 23, 3\}} u^B(n_j) \int_T (b_T^\nu \cdot \nabla \varphi_j)(\varphi_i + \tau b_T^\nu \cdot \nabla \varphi_i) = 0, \quad i \in \{1, 12, 13\}. \quad (3.16)$$

Uvažujeme-li speciálně okrajovou podmínku

$$u^B(n_2) = u^B(n_{23}) = u^B(n_3), \quad (3.17)$$

lze $u^B(n_j)$ za sumy v soustavě (3.16) vytknout, a použijeme-li dále linearitu integrálu, skalárního součinu a gradientu, je pro platnost soustavy (3.16) postačující platnost soustavy

$$\int_T (b_T^\nu \cdot \nabla \left(\sum_{j \in \{2, 23, 3\}} \varphi_j \right)) (\varphi_i + \tau b_T^\nu \cdot \nabla \varphi_i) = 0, \quad i \in \{1, 12, 13\}. \quad (3.18)$$

V soustavě (3.18) jsou celkem **tři rovnice lineární v parametru τ** . Nabízí se tedy **volit $\tau|_T$ opět jako afinní funkci**, protože taková funkce má právě **tři parametry**.

Uvažujeme tedy $\tau|_T$ v následujícím tvaru vyjádřeném pomocí barycentrických souřadnic na elementu T :

$$\tau|_T = \tau_{T,1} \lambda_{T,1} + \tau_{T,2} \lambda_{T,2} + \tau_{T,3} \lambda_{T,3}, \quad \tau_{T,1}, \tau_{T,2}, \tau_{T,3} \in \mathbb{R}. \quad (3.19)$$

Do soustavy rovnic (3.18) dosadíme za báze funkce z vyjádření (1.15). Při dosazování je pro usnadnění výpočtu vhodné použít úpravy

$$\begin{aligned} \nabla \left(\sum_{j \in \{2,23,3\}} \varphi_j \right) &= \nabla \left(2(\lambda_{T,2} + \lambda_{T,3})^2 - (\lambda_{T,2} + \lambda_{T,3}) \right) = \\ &= \underbrace{\left(4(\lambda_{T,2} + \lambda_{T,3}) - 1 \right)}_{G10} \nabla \underbrace{(\lambda_{T,2} + \lambda_{T,3})}_{G10} = (4\lambda_{T,1} - 3) \nabla \lambda_{T,1}. \\ &G10 \stackrel{(1.13)}{=} 1 - \lambda_{T,1} \end{aligned}$$

Dále do doustavy (3.18) dosadíme za τ z vyjádření (3.19), roznásobíme integrandy, použijeme značení (3.12) a větu 5 a provedeme ekvivalentní úpravy. Dostaneme tak soustavu

$$S_{T,1} \left[\underbrace{\begin{pmatrix} S_{T,1}, & -3S_{T,1}, & -3S_{T,1}, \\ 7S_{T,1} + 6S_{T,2}, & 22S_{T,1} + 7S_{T,2}, & 11S_{T,1} + 7S_{T,2}, \\ 7S_{T,1} + 6S_{T,3}, & 11S_{T,1} + 7S_{T,3}, & 22S_{T,1} + 7S_{T,3}, \end{pmatrix}}_{=: \mathbf{M}_3} \begin{pmatrix} \tau_{T,1} \\ \tau_{T,2} \\ \tau_{T,3} \end{pmatrix} + \begin{pmatrix} -2 \\ 7 \\ 7 \end{pmatrix} \right] = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (3.20)$$

Soustava (3.20) je ve tvaru „součin = 0“, stačí tedy vyřešit soustavu

$$\mathbf{M}_3 \begin{pmatrix} \tau_{T,1} \\ \tau_{T,2} \\ \tau_{T,3} \end{pmatrix} = \begin{pmatrix} 2 \\ -7 \\ -7 \end{pmatrix} \quad (3.21)$$

Je

$$\begin{aligned} \det \mathbf{M}_3 &= S_{T,1}(825S_{T,1}^2 + 275S_{T,1}S_{T,2} + 275S_{T,1}S_{T,3}) = \\ &= S_{T,1}(550S_{T,1}^2 + 275S_{T,1} \sum_{i \in \{1,2,3\}} S_{T,i}) \stackrel{(3.13)}{=} 550S_{T,1}^3 \neq 0 \end{aligned}$$

To, že je determinant matice \mathbf{M}_3 nenulový, plyne z definice výtokové hranice oblasti Ω (2.6) a z toho, že

$$\exists c > 0 : \nabla \lambda_{T,1} = -c\nu(n_{23}).$$

Soustava (3.21) tedy má jediné řešení. Vypočítáme ho Cramerovým pravidlem.

Je

$$\begin{aligned}
\det \mathbf{M}_{31} &= 2S_{T,1}(132S_{T,1} + 77S_{T,2} + 77S_{T,3}) = \\
&= 2S_{T,1}(55S_{T,1} + 77 \sum_{i \in \{1,2,3\}} S_{T,i}) \stackrel{(3.13)}{=} 110S_{T,1}^2, \\
\det \mathbf{M}_{32} &= S_{T,1}(-231S_{T,1} + 9S_{T,2} - 141S_{T,3}) = \\
&= S_{T,1}(-231 \sum_{i \in \{1,2,3\}} S_{T,i} + 240S_{T,2} + 90S_{T,3}) \stackrel{(3.13)}{=} S_{T,1}(240S_{T,2} + 90S_{T,3}), \\
\det \mathbf{M}_{33} &= S_{T,1}(-231S_{T,1} - 141S_{T,2} + 9S_{T,3}) = \\
&= S_{T,1}(-231 \sum_{i \in \{1,2,3\}} S_{T,i} + 90S_{T,2} + 240S_{T,3}) \stackrel{(3.13)}{=} S_{T,1}(90S_{T,2} + 240S_{T,3}).
\end{aligned}$$

Jediné řešení soustavy (3.21) pak je

$$\tau_{T,1} = \frac{1}{5S_{T,1}}, \quad \tau_{T,2} = \frac{24S_{T,2} + 9S_{T,3}}{55S_{T,1}^2}, \quad \tau_{T,3} = \frac{9S_{T,2} + 24S_{T,3}}{55S_{T,1}^2}.$$

Případ, kdy element T je typu 2

Uvažujeme opět značení (3.4) a navíc to, že je splněno tentokrát (3.6). Pak z požadavku (3.11) dostaneme soustavu rovnic

$$u^B(n_3) \int_T (b_T^\nu \cdot \nabla \varphi_3)(\varphi_i + \tau b_T^\nu \cdot \nabla \varphi_i) = 0, \quad i \in \{1, 12, 2, 13, 23\}. \quad (3.22)$$

Volíme $\tau|_T$ opět jako afinní funkci, tedy ve tvaru (3.19). Do soustavy rovnic (3.22) dosadíme za báze funkce z vyjádření (1.15) a za τ z vyjádření (3.19), roznásobíme integrandy, použijeme značení (3.12) a větu 5 a provedeme ekvivalentní úpravy. Dostaneme tak soustavu

$$S_{T,1} \left[\begin{pmatrix} 2S_{T,1} & S_{T,1} & 2S_{T,1} \\ S_{T,2} & 2S_{T,2} & 2S_{T,2} \\ S_{T,1} + 2S_{T,2} & 2S_{T,1} + S_{T,2} & -3S_{T,1} - 3S_{T,2} \\ 3S_{T,1} - 2S_{T,3} & 3S_{T,1} - S_{T,3} & 14S_{T,1} + 3S_{T,3} \\ 3S_{T,2} - S_{T,3} & 3S_{T,2} - 2S_{T,3} & 14S_{T,2} + 3S_{T,3} \end{pmatrix} \begin{pmatrix} \tau_{T,1} \\ \tau_{T,2} \\ \tau_{T,3} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \\ 3 \end{pmatrix} \right] = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (3.23)$$

Soustava (3.23) je ve tvaru „součin = 0“, stačí tedy vyřešit soustavu

$$\begin{array}{l} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \underbrace{\begin{pmatrix} 2S_{T,1}, & S_{T,1}, & 2S_{T,1} \\ S_{T,2}, & 2S_{T,2}, & 2S_{T,2} \\ S_{T,1} + 2S_{T,2}, & 2S_{T,1} + S_{T,2}, & -3S_{T,1} - 3S_{T,2} \\ 3S_{T,1} - 2S_{T,3}, & 3S_{T,1} - S_{T,3}, & 14S_{T,1} + 3S_{T,3} \\ 3S_{T,2} - S_{T,3}, & 3S_{T,2} - 2S_{T,3}, & 14S_{T,2} + 3S_{T,3} \end{pmatrix}}_{\mathbf{M}_4} \underbrace{\begin{pmatrix} \tau_{T,1} \\ \tau_{T,2} \\ \tau_{T,3} \end{pmatrix}}_r = \underbrace{\begin{pmatrix} -1 \\ -1 \\ -1 \\ -3 \\ -3 \end{pmatrix}}_r. \quad (3.24)$$

Jako $\mathbf{M}_{4,i}$, $i \in \{1, 2, 3, 4, 5\}$, označíme i -tý řádek matice \mathbf{M}_4 .

Soustava (3.24) je přeuredená, je tedy třeba nějak zeslabit požadavky na řešení. Jednou z možností je metoda nejmenších čtverců. Vyšetříme lineární nezávislost sloupců matice \mathbf{M}_4 . Při tom využijeme vlastnosti (3.14).

$$\bullet S_{T,1} = 0 \stackrel{(3.14)}{\implies} S_{T,3} = -S_{T,2} \neq 0 \implies \det \begin{pmatrix} \mathbf{M}_{4,2} \\ \mathbf{M}_{4,3} \\ \mathbf{M}_{4,5} \end{pmatrix} = -30S_{T,2}^3 \neq 0.$$

$$\bullet S_{T,2} = 0 \stackrel{(3.14)}{\implies} S_{T,3} = -S_{T,1} \neq 0 \implies \det \begin{pmatrix} \mathbf{M}_{4,1} \\ \mathbf{M}_{4,3} \\ \mathbf{M}_{4,4} \end{pmatrix} = 30S_{T,1}^3 \neq 0.$$

$$\bullet S_{T,3} = 0 \stackrel{(3.14)}{\implies} S_{T,2} = -S_{T,1} \neq 0 \implies \det \begin{pmatrix} \mathbf{M}_{4,1} \\ \mathbf{M}_{4,2} \\ \mathbf{M}_{4,4} \end{pmatrix} = -30S_{T,1}^3 \neq 0.$$

$$\bullet (S_{T,1} \neq 0 \wedge S_{T,2} \neq 0 \wedge S_{T,3} \neq 0) \implies \det \begin{pmatrix} \mathbf{M}_{4,1} \\ \mathbf{M}_{4,2} \\ \mathbf{M}_{4,3} \end{pmatrix} = 15S_{T,1}S_{T,2}S_{T,3} \neq 0.$$

V každém případě tedy existuje čtvercová podmatice matice \mathbf{M}_4 stupně 3, jejíž determinant je nenulový, takže matice \mathbf{M}_4 má lineárně nezávislé sloupce. Můžeme tedy najít jednoznačně určené přibližné řešení metodou nejmenších čtverců vyřešením soustavy normálních rovnic

$$\mathbf{M}_4^T \mathbf{M}_4 \begin{pmatrix} \tau_{T,1} \\ \tau_{T,2} \\ \tau_{T,3} \end{pmatrix} = \mathbf{M}_4^T r. \quad (3.25)$$

Dokončení metody 1

Nebudeme analyzovat, kdy je stabilizační parametr získaný metodou 1 nezáporný a shora omezený. Místo toho navrhuje postup, který jeho nezápornost a omezenost shora v každém případě zajistí.

Zvolíme vhodné $\tau_{MAX} > 0$ a definujeme

$$\tau_{T2,i} := \tau_{T,i} - \left(\min_{j \in \{1,2,3\}} \tau_{T,j} \right)^-, \quad i \in \{1, 2, 3\}, \quad (3.26)$$

$$\tau_{T3,i} := \min\{\tau_{MAX}, \tau_{T2,i}\}, \quad i \in \{1, 2, 3\},$$

a nakonec předefinujeme

$$\tau_{T,i} := \tau_{T3,i}, \quad i \in \{1, 2, 3\}.$$

Na každém elementu $K \in \mathcal{T}_h \setminus L(1, 2)$ zvolíme $\tau|_K = \tau_{C2K}$, viz (3.15).

3.2.3 Volba stabilizačního parametru metodou 2

Bude se jednat o heuristické zobecnění metody 2 ze situace pro $d = 1$.

Uvažujeme $\tau|_T$ ve tvaru (3.19).

Mějme dán libovolný element $T \in L(1, 2)$.

Případ, kdy element T je typu 1

Uvažujeme značení (3.4) a navíc to, že je splněno (3.5). Položíme

$$\tau_{T,1} = 0, \quad \tau_{T,2} = \tau_{T,3} = \frac{h_T^{\nu(n_{23})}}{\|b_T^{\nu(n_{23})}\|} \quad (3.27)$$

Případ, kdy element T je typu 2

Uvažujeme opět značení (3.4) a navíc to, že je splněno tentokrát (3.6). Položíme

$$\tau_{T,1} = \tau_{T,2} = 0, \quad \tau_{T,3} = \frac{h_T^{\nu(n_3)}}{\|b_T^{\nu(n_3)}\|}, \quad (3.28)$$

kde $\nu(n_3)$ je definováno v (1.1).

Dokončení metody 2

Jestliže se b blíží vektoru tečnému k $\partial\Omega_L$, pak z (3.27) a (3.28) plyne, že $\tau \rightarrow \infty$ někde na $\partial\Omega_L$. Vzhledem k předpokladu (1.45) je tedy potřeba nějak omezit shora všechna $\tau_{T,i}$, $i \in \{1, 2, 3\}$. Navrhujeme tedy volit vhodné $\tau_{MAX} > 0$, položit

$$\tau_{T2,i} := \min\{\tau_{MAX}, \tau_{T,i}\}, \quad i \in \{1, 2, 3\},$$

a předefinovat

$$\tau_{T,i} := \tau_{T2,i}, \quad i \in \{1, 2, 3\}.$$

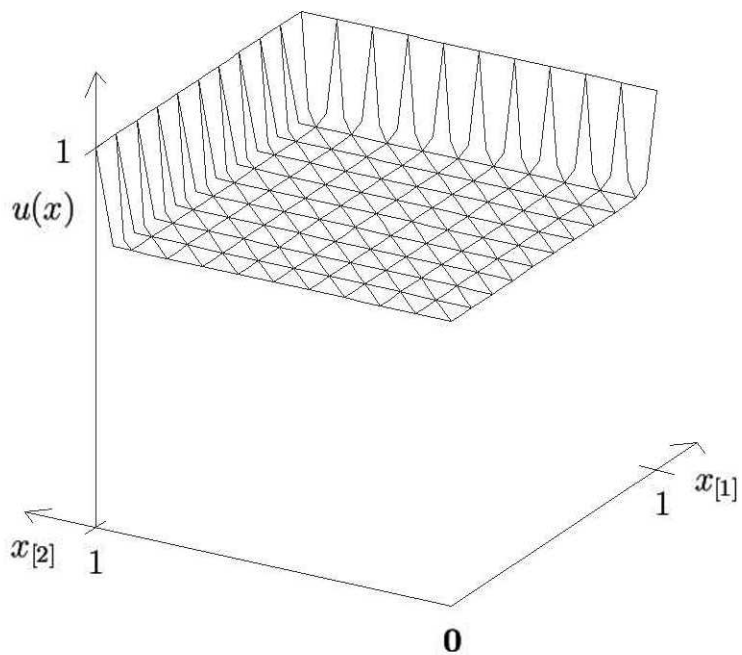
Na každém elementu $K \in \mathcal{T}_h \setminus L(1, 2)$ zvolíme $\tau|_K = \tau_{C2K}$, viz (3.15).

Poznámka: Heuristické zobecnění metody 1 z případu $d = 1$ obdobné tomu, které bylo právě popsáno, nedávalo dobré výsledky.

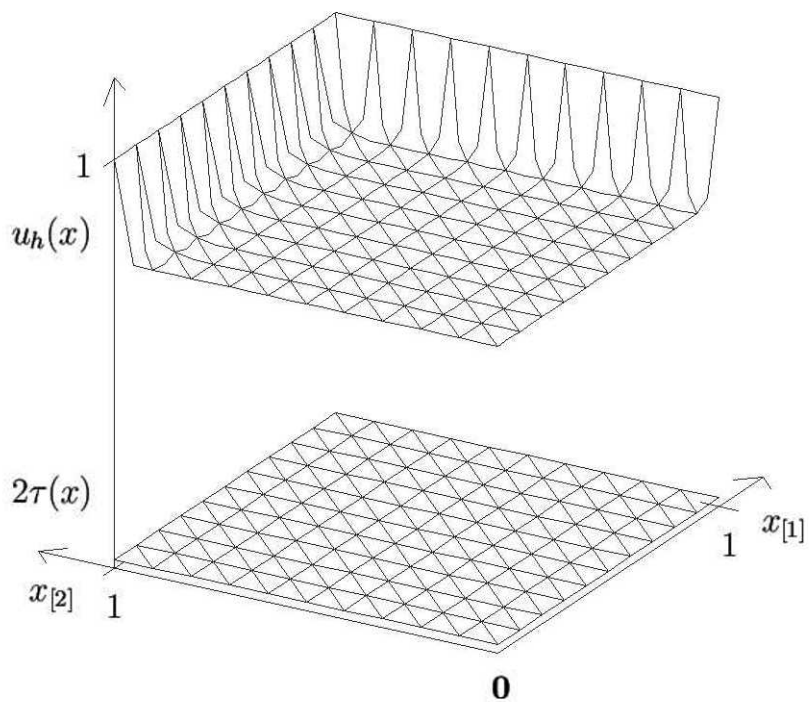
3.3 Modelový problém – numerické výsledky

Řešení úlohy 1 a 9 s daty (3.1), různými ε , triangulací $\mathcal{T}_{h,1}$ a τ získanými metodou C, 1 a 2 jsou vyobrazena na následujících obrázcích. Pro přehlednost je zobrazen dvojnásobek stabilizačního parametru. Pro výpočet je voleno

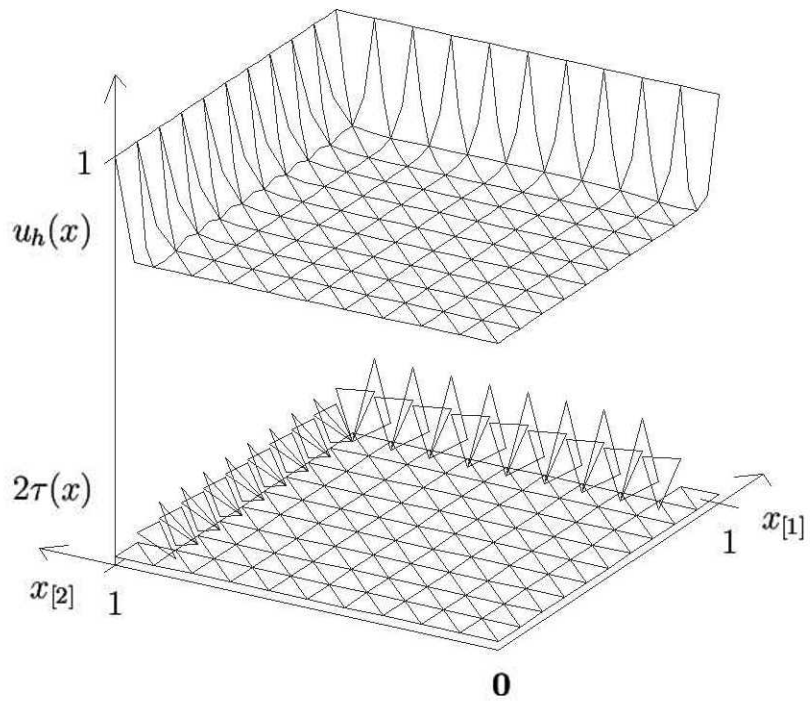
$$\begin{aligned} \tau_{MAX} &= 1, \\ b|_K &= b(c_K), \quad K \in \mathcal{T}_h. \end{aligned} \tag{3.29}$$



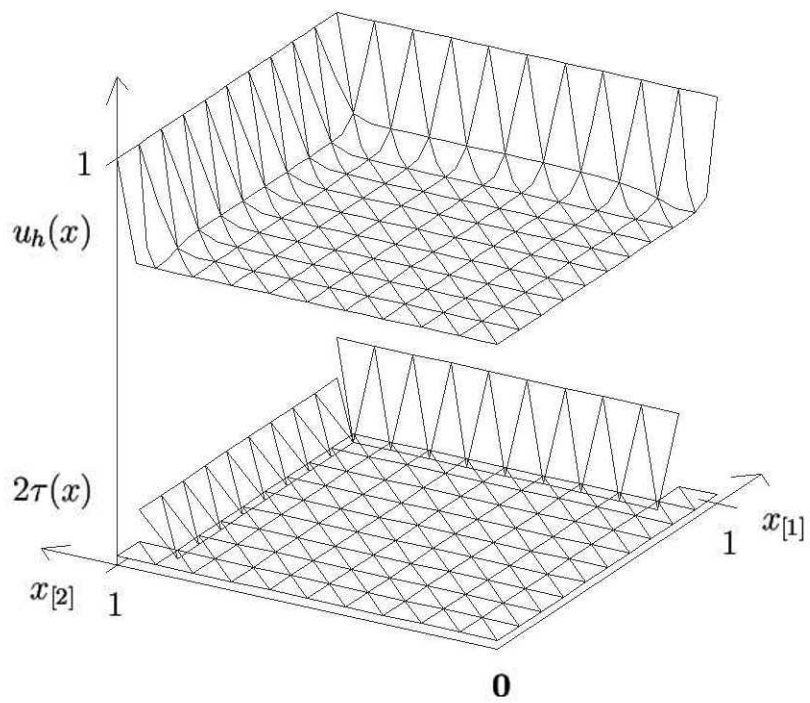
Obrázek 34: řešení úlohy 1 s daty (3.1), $\varepsilon = 0,02$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27



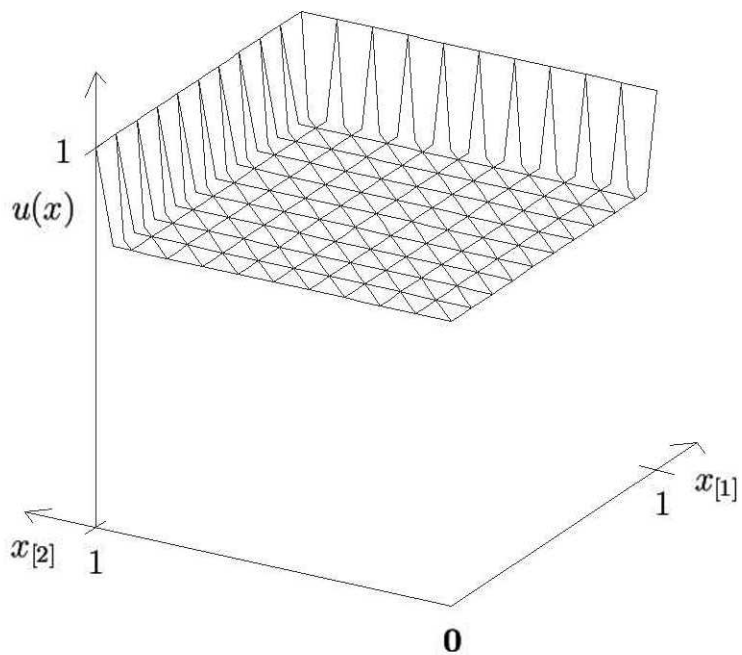
Obrázek 35: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,02$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou C



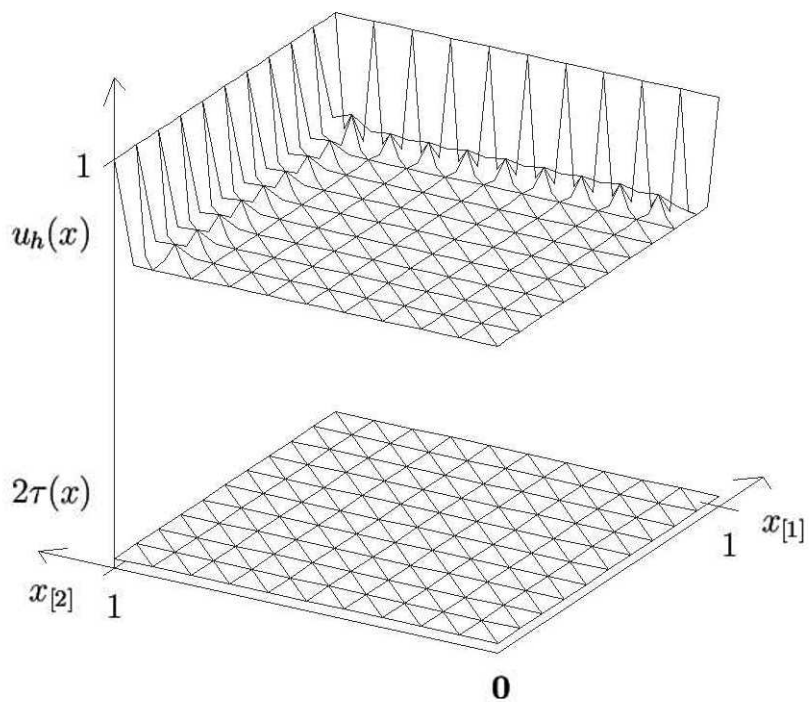
Obrázek 36: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,02$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou 1



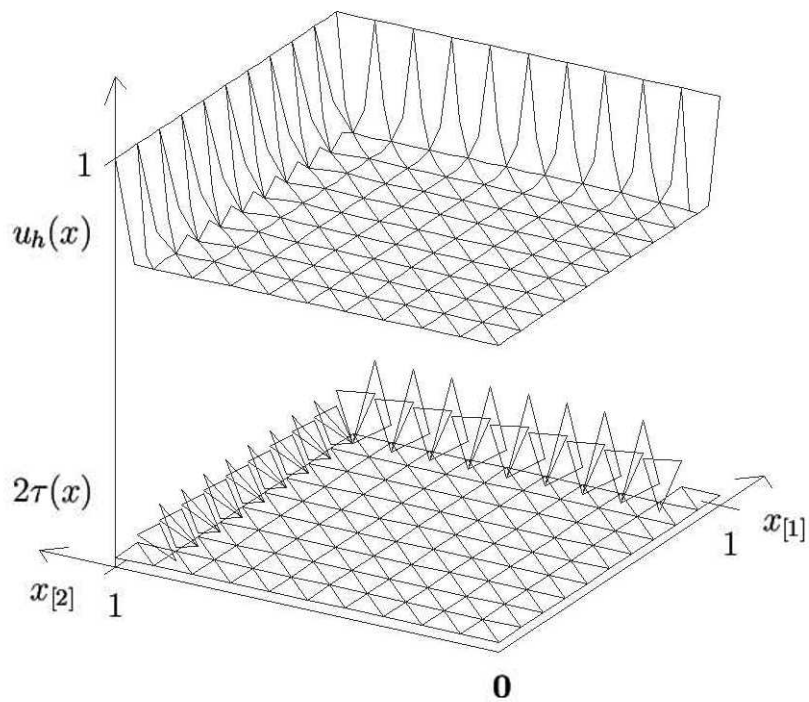
Obrázek 37: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,02$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou 2



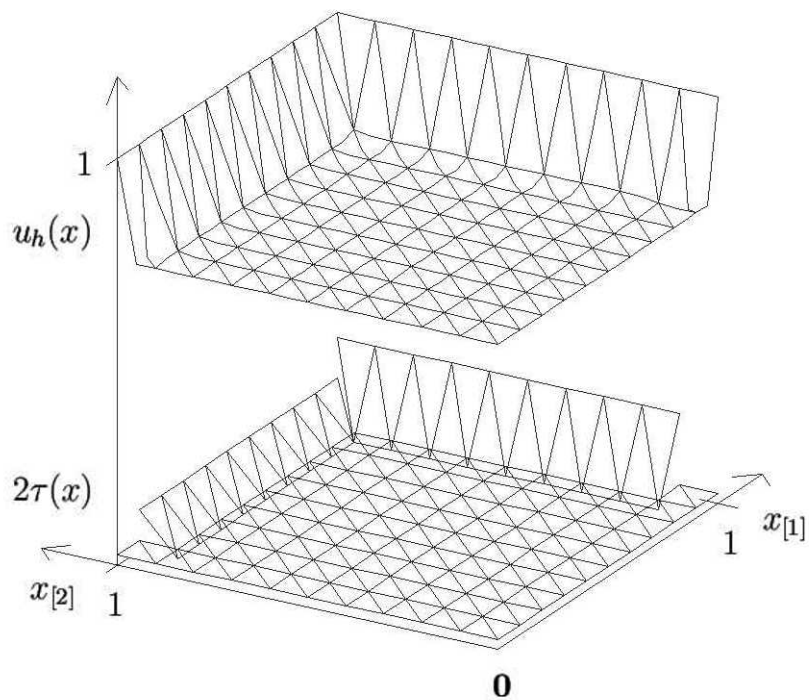
Obrázek 38: řešení úlohy 1 s daty (3.1), $\varepsilon = 0,005$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27



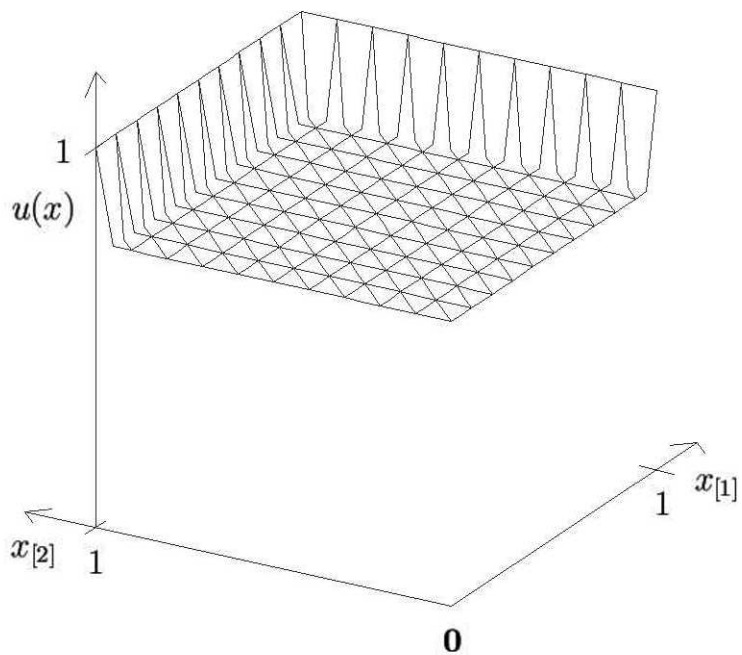
Obrázek 39: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,005$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou C



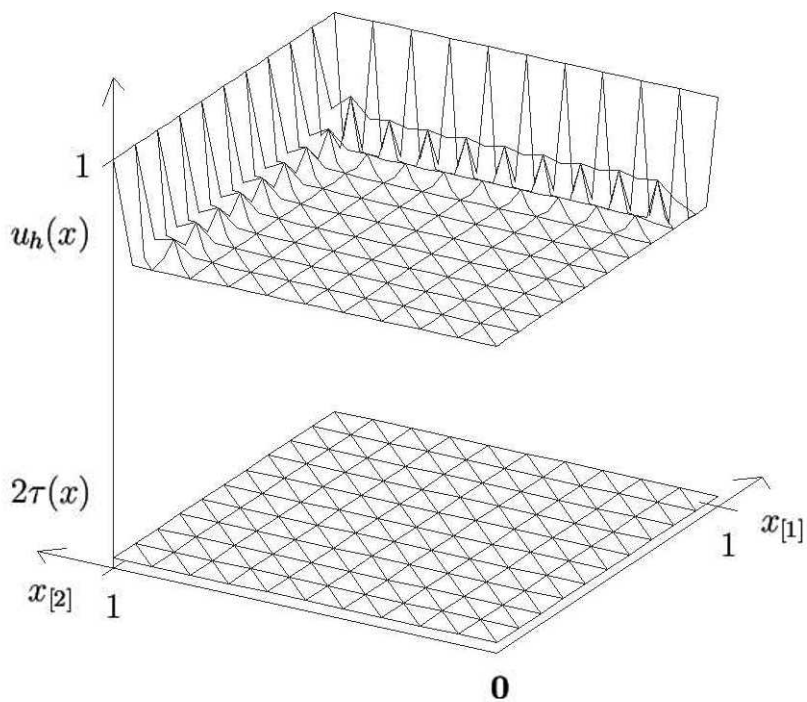
Obrázek 40: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,005$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou 1



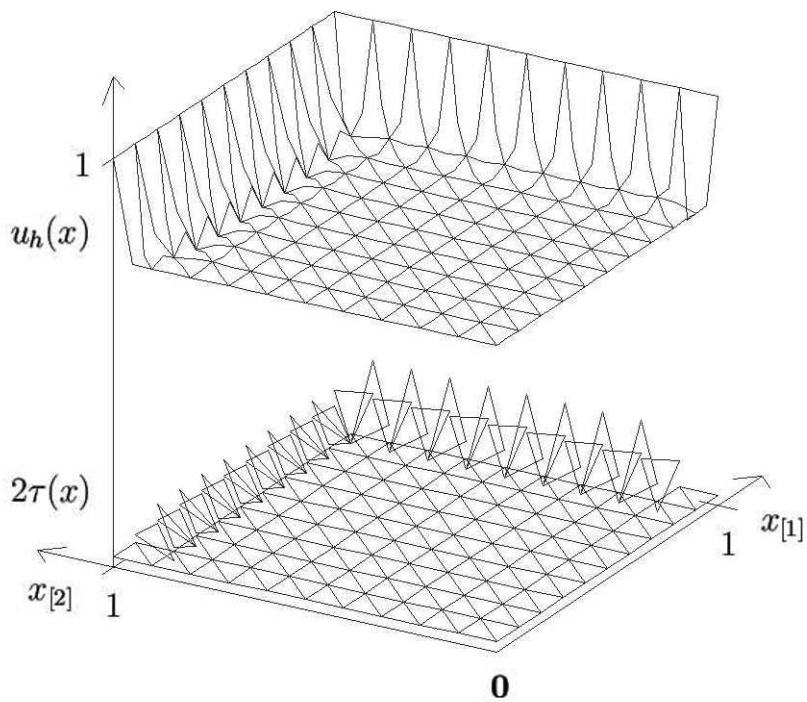
Obrázek 41: řešení úlohy 9 s daty (3.1), $\varepsilon = 0,005$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou 2



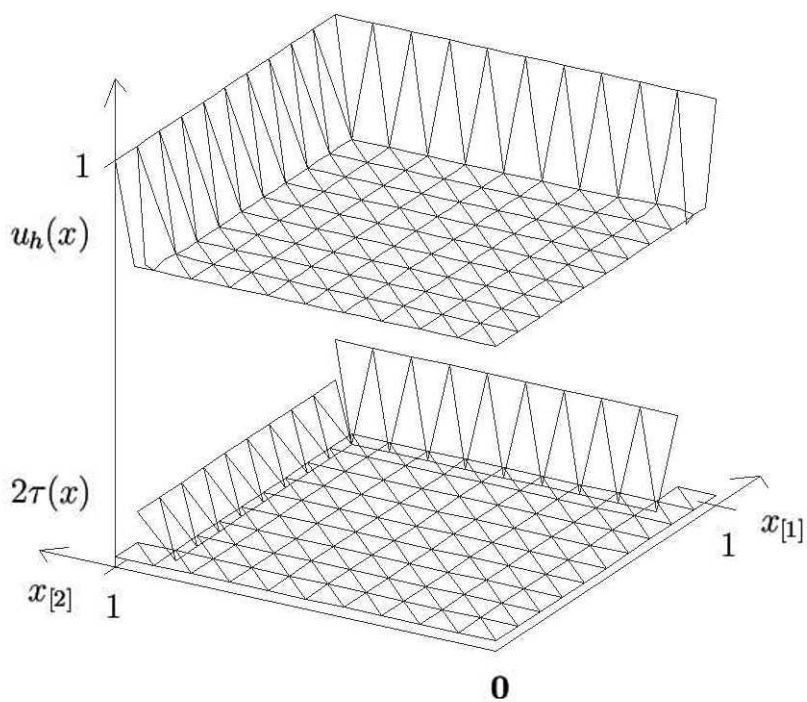
Obrázek 42: řešení úlohy 1 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27



Obrázek 43: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou C

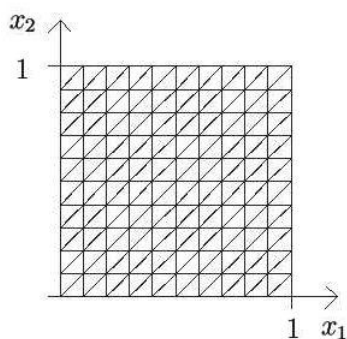


Obrázek 44: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou 1

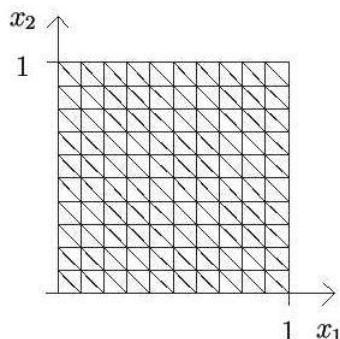


Obrázek 45: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,1}$ z obrázku 27, τ voleno metodou 2

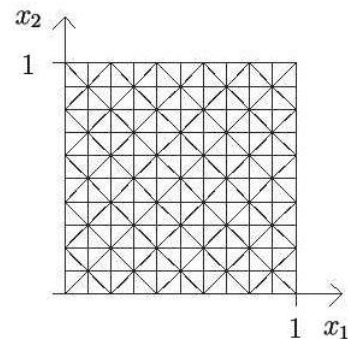
Uvažujeme dvě další triangulace $\mathcal{T}_{h,2}$ a $\mathcal{T}_{h,3}$, které se od triangulace $\mathcal{T}_{h,1}$ liší pouze orientací dílčích diagonál. Všechny tři triangulace jsou vyobrazeny na následujících obrázcích.



Obrázek 46: triangulace $\mathcal{T}_{h,1}$

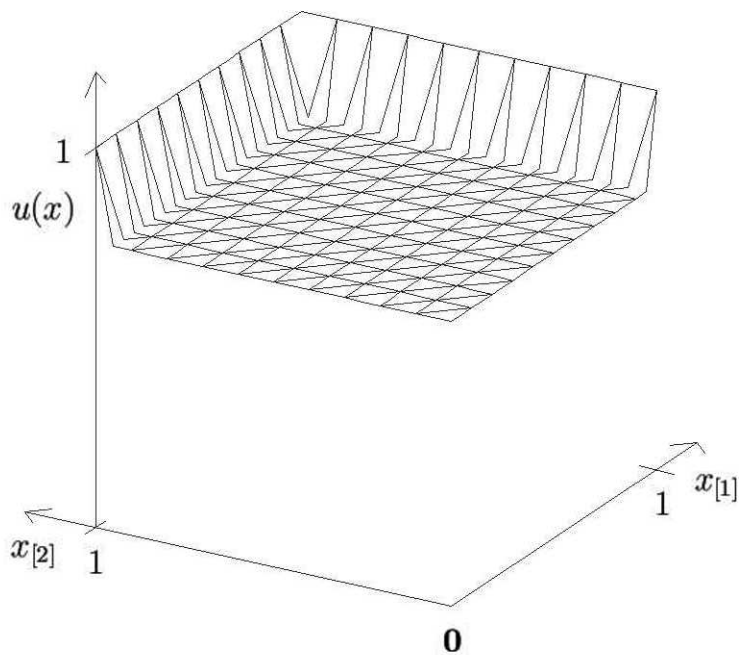


Obrázek 47: triangulace $\mathcal{T}_{h,2}$

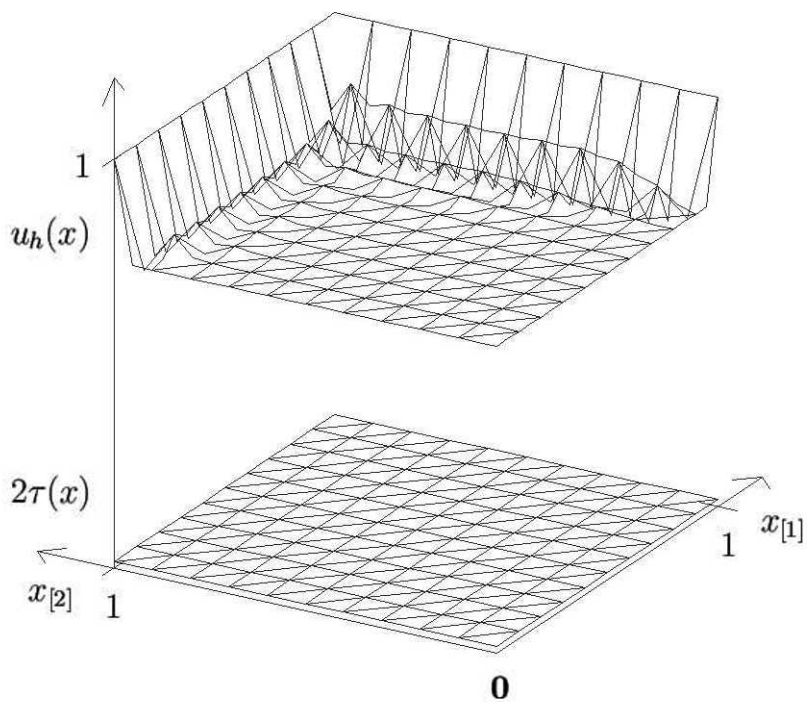


Obrázek 48: triangulace $\mathcal{T}_{h,3}$

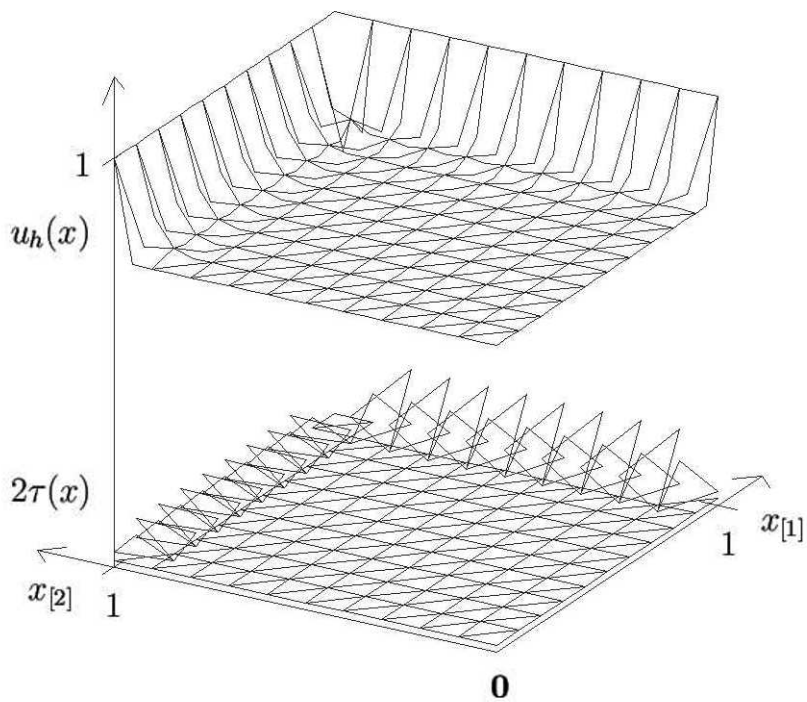
Řešení úloh 1 a 9 s daty (3.1), $\varepsilon = 10^{-8}$, triangulacemi $\mathcal{T}_{h,2}$ a $\mathcal{T}_{h,3}$ a τ získanými metodou C, 1 a 2 jsou vyobrazena na následujících obrázcích.



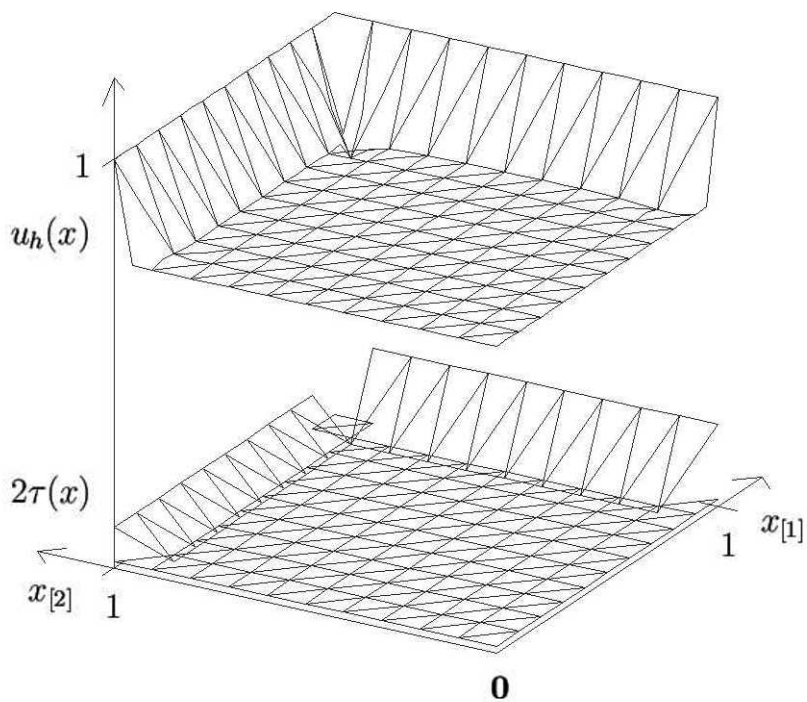
Obrázek 49: řešení úlohy 1 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,2}$ z obrázku 47



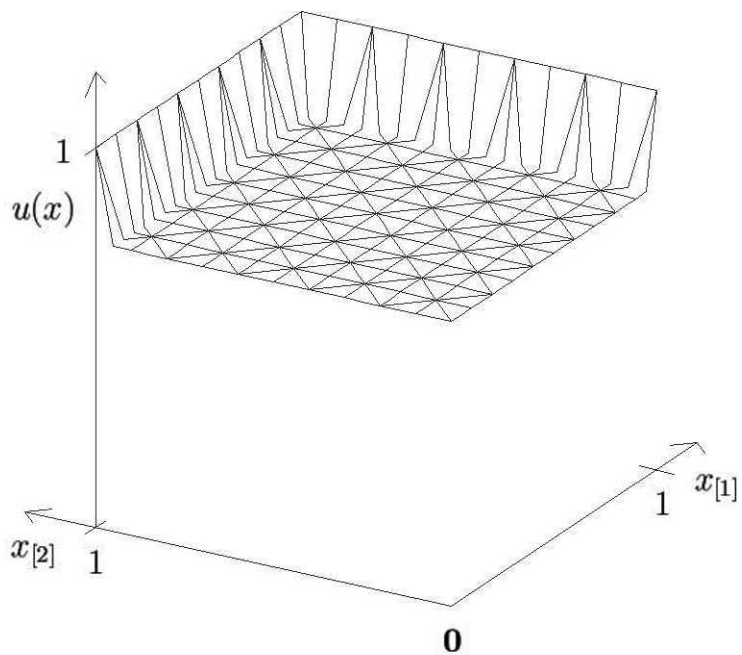
Obrázek 50: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,2}$ z obrázku 47, τ voleno metodou C



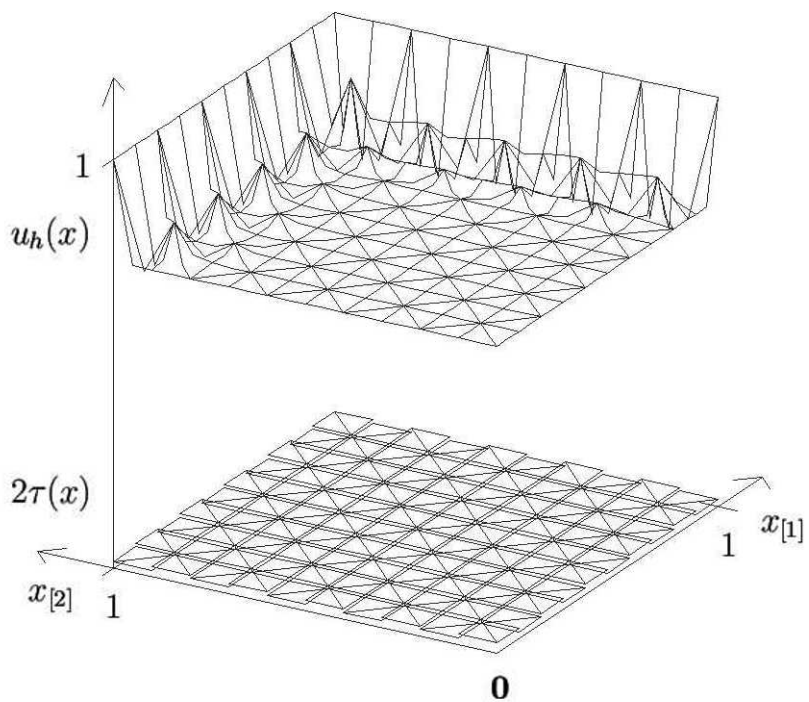
Obrázek 51: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,2}$ z obrázku 47, τ voleno metodou 1



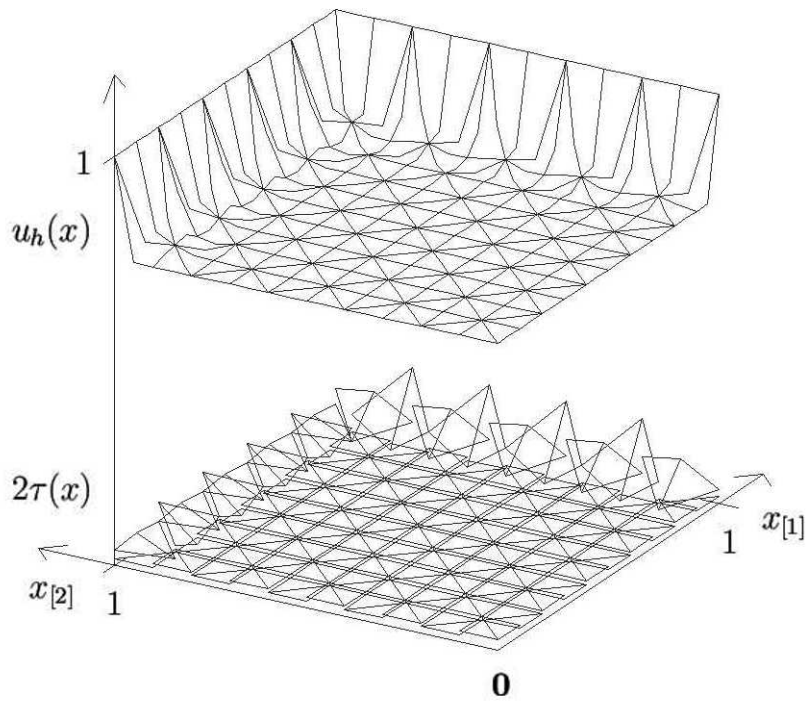
Obrázek 52: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,2}$ z obrázku 47, τ voleno metodou 2



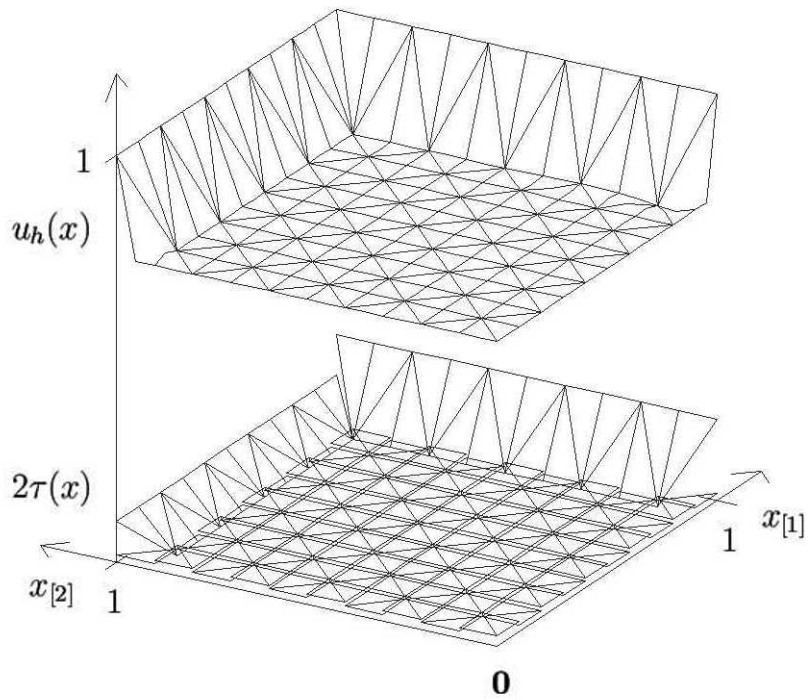
Obrázek 53: řešení úlohy 1 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,3}$ z obrázku 48



Obrázek 54: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,3}$ z obrázku 48, τ voleno metodou C



Obrázek 55: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,3}$ z obrázku 48, τ voleno metodou 1



Obrázek 56: řešení úlohy 9 s daty (3.1), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,3}$ z obrázku 48, τ voleno metodou 2

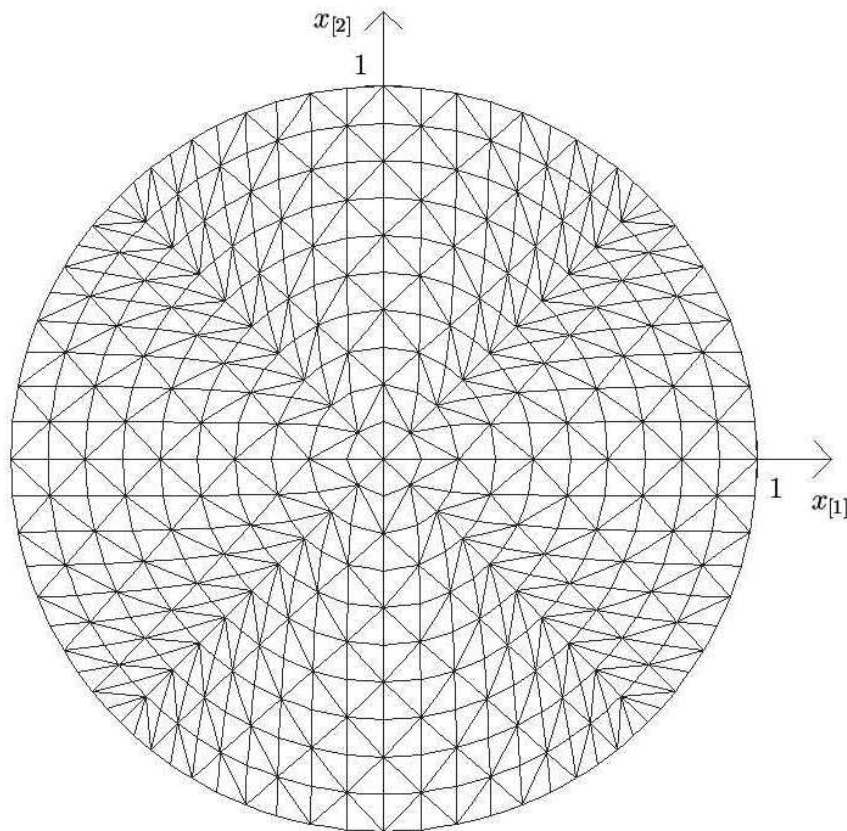
3.4 Další numerické výsledky

Uvažujeme data

$\Omega =$ vnitřek konvexního mnohoúhelníka s opsanou kružnicí splňující rovnici $x_{[1]}^2 + x_{[2]}^2 = 1$,

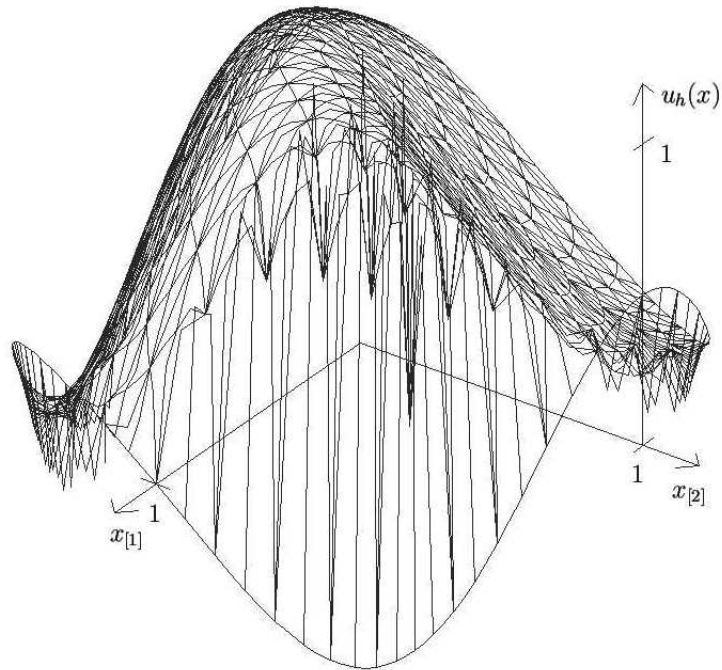
$$\begin{aligned} b(x) &= \frac{1}{\|\omega(x)\|}(-\omega_{[2]}(x), \omega_{[1]}(x)), & \omega(x) &= x - \left(-\frac{3}{2}, 2\right), & x &\in \overline{\Omega}, \\ f(x) &= \frac{1}{3}(-x_{[1]} - x_{[2]} + 1), & x &\in \Omega, \\ u^B(x) &= -\frac{1}{2}\sin\left(\frac{\pi}{2}(2x_{[1]} - x_{[2]})\right), & x &\in \partial\Omega, \end{aligned} \tag{3.30}$$

a triangulaci vyobrazenou na následujícím obrázku.

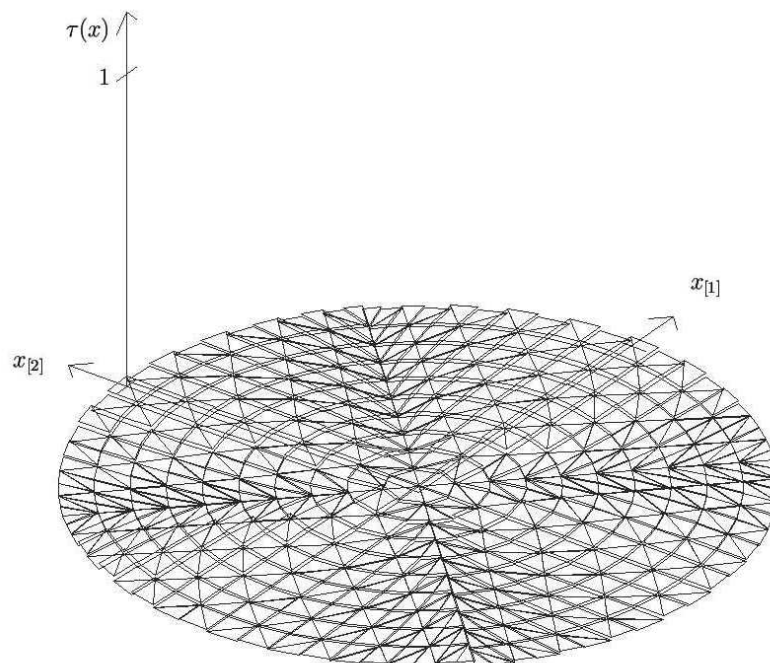


Obrázek 57: triangulace $\mathcal{T}_{h,4}$

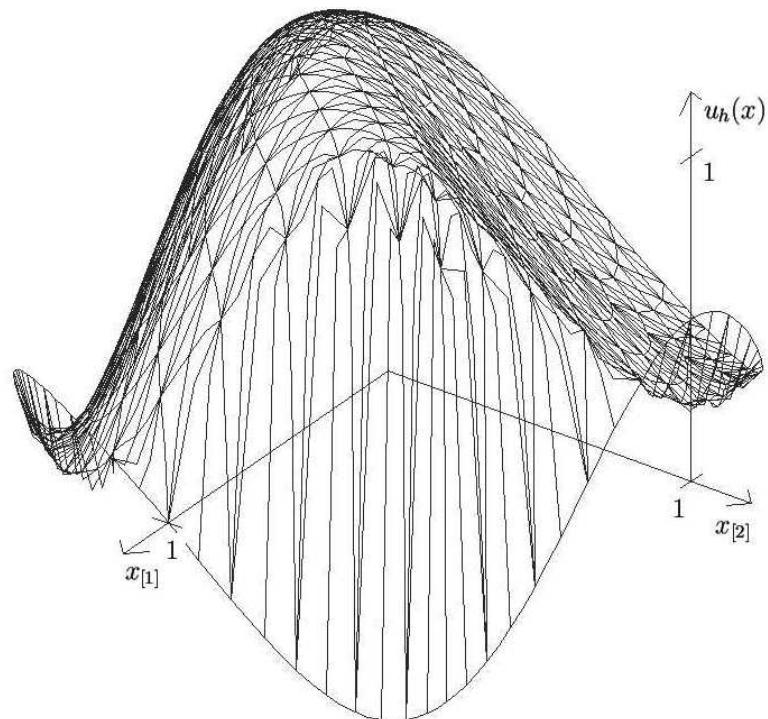
Řešení úlohy 9 s daty (3.30), $\varepsilon = 10^{-8}$, triangulací $\mathcal{T}_{h,4}$ a τ získanými metodou C, 1 a 2 jsou vyobrazena na následujících obrázcích. Data (3.30) jsou volena tak, aby nebyly splněny předpoklady (3.7) a (3.17), a mohly se tak projevit případné nepřesnosti řešení získaného metodou 1. Pro výpočet je voleno opět (3.29).



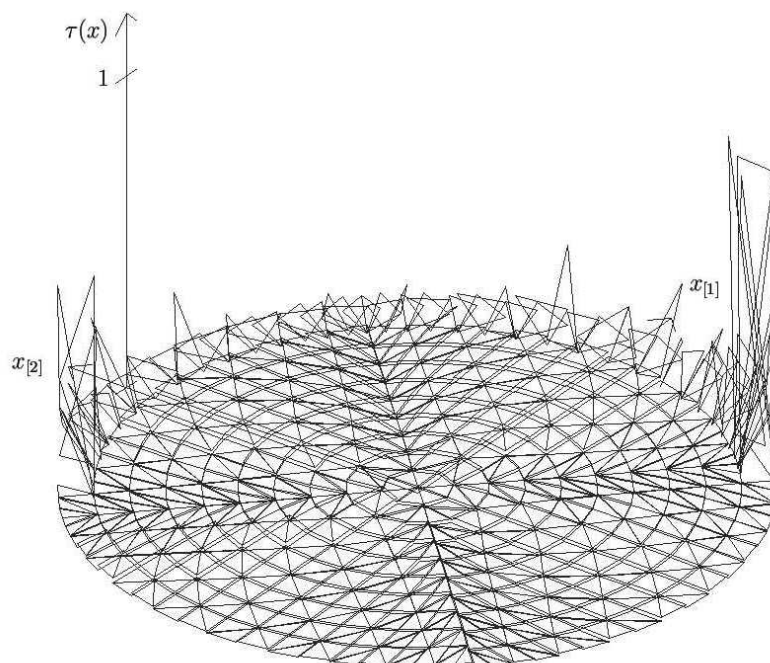
Obrázek 58: řešení úlohy 9 s daty (3.30), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,2}$ z obrázku 47, τ voleno metodou C



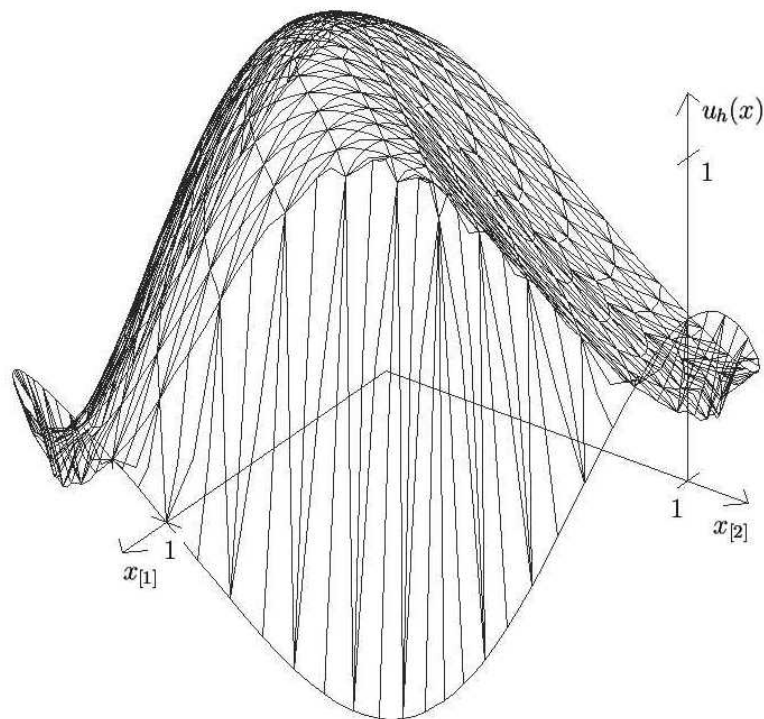
Obrázek 59: τ příslušné k řešení z obrázku 58



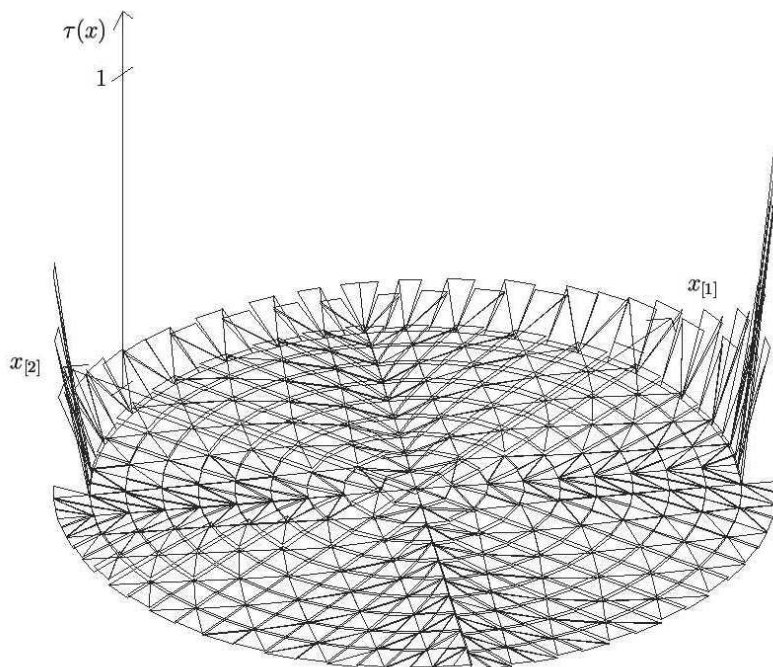
Obrázek 60: řešení úlohy 9 s daty (3.30), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,2}$ z obrázku 47, τ voleno metodou 1



Obrázek 61: τ příslušné k řešení z obrázku 60



Obrázek 62: řešení úlohy 9 s daty (3.30), $\varepsilon = 10^{-8}$ a triangulací $\mathcal{T}_{h,2}$ z obrázku 47, τ voleno metodou 2



Obrázek 63: τ příslušné k řešení z obrázku 62

Závěr

Volbou stabilizačního parametru afinního po elementech výtokové hraniční vrstvy navrhovanými metodami 1 a 2 bylo dosaženo zpřesnění přibližných řešení oproti standardní volbě stabilizačního parametru konstantního po elementech metodou C.

V případě $d = 1$ vyšla metoda 1 jako lepší než metoda 2. Metoda 1 dává velmi dobré výsledky – řešení je velmi přesné ve všech uzlech, zatímco metoda 2 dává výsledky nepřesné v uzlu z výpočetní oblasti nejbližším k výtokové hranici, jak se očekávalo v oddílu 2.2.4. V ostatních uzlech však i metoda 2 dává výsledky velmi přesné. Zmíněná nepřesnost metody 2 příliš nenarušuje celkový charakter řešení.

V případě $d = 2$ vyšla jako lepší naopak metoda 2 než metoda 1, ačkoliv metoda 2 je v případě $d = 2$ pouhým heuristickým zobecněním metody 2 z případu $d = 1$, zatímco metoda 1 je v případě $d = 2$ odvozena pečlivěji než metoda 2. Metoda 2 se v případě $d = 2$ chová podobně jako v případě $d = 1$ – dává celkem přesné řešení až na případ uzlů z výpočetní oblasti nejbližších výtokové hranici, nepřesnost v těchto uzlech však opět příliš nenarušuje celkový charakter řešení. Naproti tomu metoda 1 dává řešení, jehož podstatné nepřesnosti na jednu stranu nejsou tak velké jako zmíněná nepřesnost řešení získaného metodou 2, na druhou stranu se tyto nepřesnosti vyskytují v širší vrstvě uzlů podél výtokové hranice oproti metodě 2. Navíc mají tyto nepřesnosti lehce oscilační charakter.

Nepřesnosti řešení získaných metodou 1 v případě $d = 2$ přisuzujeme tomu, že v (3.8) a (3.9) byl vyšetřován jen člen vesměs s projekcí vektoru konvekce do směru vnější normály a ostatní členy vyšetřovány nebyly. Odhadujeme, že zlepšení metody 1 by se dalo dosáhnout vyšetřením i těchto ostatních členů. Dalším problémem je to, že metoda 1 využívá po elementech konstantní aproximaci vektoru konvekce, viz (3.7), předpoklad konstantnosti okrajové podmínky v uzlech výtokové hranice, viz (3.17), metodu nejmenších čtverců, viz (3.25), a elevaci hodnot stabilizačního parametru za účelem dosažení jeho nezápornosti, viz (3.26). Odhadujeme, že dalšího zlepšení metody 1 by se dalo dosáhnout použitím jiných předpokladů a postupů, než jsou posledně zmíněné. Například by se dalo využít metody nezáporných nejmenších čtverců.

Seznam použité literatury

- [1] ROOS, H. G., STYNES, M., TOBISKA, L.:
Robust Numerical Methods for Singularly Perturbed Differential Equations,
Convection-Diffusion-Reaction and Flow Problems,
second edition.
Springer-Verlag, Berlin - Heidelberg 2008,
ISBN 978-3-540-34467-4
- [2] CIARLET, P. G.: The Finite Element Method for Elliptic Problems.
North-Holland Publishing Company, Amsterdam - New York - Oxford 1978
- [3] REUSKEN, A.: Numerical methods for elliptic partial differential equations.
poznámky k přednáškám,
<http://www.igpm.rwth-aachen.de/Download/ws1112/numanaIV/Numa4.pdf>
- [4] ZIENKIEWICZ, O. C., TAYLOR, R. L.: The Finite Element Method Volume 3: Fluid
Dynamics, Fifth edition.
Butterworth-Heinemann, Oxford 2000
- [5] BROOKS, A. N., HUGHES, T. J. R.: Streamline upwind/Petrov-Galerkin formulations
for convection dominated flows with particular emphasis on the incompressible Navier-
Stokes equations.
Comput. Methods Appl. Mech. Engrg. 32, 199-259 (1982)
- [6] CODINA, R., OÑATE, E, CERVERA, M.: The intrinsic time for the streamline
upwind/Petrov-Galerkin formulation using quadratic elements.
Comput. Methods Appl. Mech. Engrg. 94, 239-262 (1992)
- [7] FRANCA, L. P., HAUKE, G., MASUD, A.: Stabilized finite element methods.
Finite element methods: 1970's and beyond, CIMNE, Barcelona, Spain 2003
- [8] FRIES, T.-P., MATTHIES, H. G.: A Review of Petrov-Galerkin Stabilization
Approaches and an Extension to Meshfree Methods, (2004).
- [9] HARARI, I., FRANCA, L. P., OLIVIERA, S. P.: Streamline Design of Stability
Parameters For Advection-diffusion Problems.
- [10] KNOBLOCH, P.: On the choice of the SUPG parameter at outflow boundary layers.
Adv. Comput. Math. 31, 369-389 (2009)
- [11] MIZUKAMI, A.: An implementation of the streamline upwind/Petrov-Galerkin
method for linear triangular elements.
Comput. Methods Appl. Mech. Engrg. 49, 357-364 (1985)

- [12] TEZDUYAR, T. E., OSAWA, Y.: Finite element stabilization parameters computed from element matrices and vectors.
Comput. Methods Appl. Mech. Engrg. 190, 411-430 (2000)
- [13] JOHN, V., KNOBLOCH, P.: On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations: Part I - A review.
Comput. Methods Appl. Mech. Engrg. 196, 2197-2215 (2007)