UNIVERSITATIS PALACKIANAE OLOMUCENSIS
FACULTAS RERUM NATURALIUM

Department of Mathematical Analysis
and Applications of Mathematics

ODAM
2003

Editors: Jiří V. Horák & Miloslav Závodný

OBSAH — CONTENTS


Univ. Palacki. Olomuc., Fac. rer. nat.,
Dept of Math. Anal. and Appl. of Math.
ODAM (2003) 4–31

Modelování interaktivní soustavy pro kotvení stropů: I. Základní numerické experimenty *

Jiří V. HORÁK, Ivona SVOBODOVÁ

Katedra matematické analýzy a aplikací matematiky
přírodovědecká fakulta Univerzity Palackého
Tomkova 40, 779 00 Olomouc
e-mail: jhorak@risc.upol.cz

Abstrakt

V následujícím textu jsou stručně uvedeny problémy související s hledáním hodného výpočtového modelu, který by odpovídal pasivnímu kotevním systému tvořenému interaktivní soustavou pružných těles ve vzájemném kontaktu se třením. K numerické realizaci je užito SW systému ANSYS, který k diskretizaci používá metodu konečných prvků.

Při ponechání odpovídajících stupňů volnosti je výsledná matice tuhosti semikoercivní. Je proto nutno zajistit podmínky řešitelnosti. Vybrané aspekty této problematky ilustrujeme na vzorových numerických příkladech pro zjednodušený výpočtový model. Chování na hranici těles je aproxinmováno různým typem okrajových podmínek a celková efektivnost navrženého modelu je testována lineární i nelineární teorií matematické pružnosti. Úlohy jsou formulovány jako variační nerovnice druhého druhu pro jednostranný Signoriniho problém s Coulombovským třením.

*Předložená práce byla realizována v rámci výzkumného záměru katedry MAaAM PřF UP Olomouc, č. J14/98: 1531 000 11.

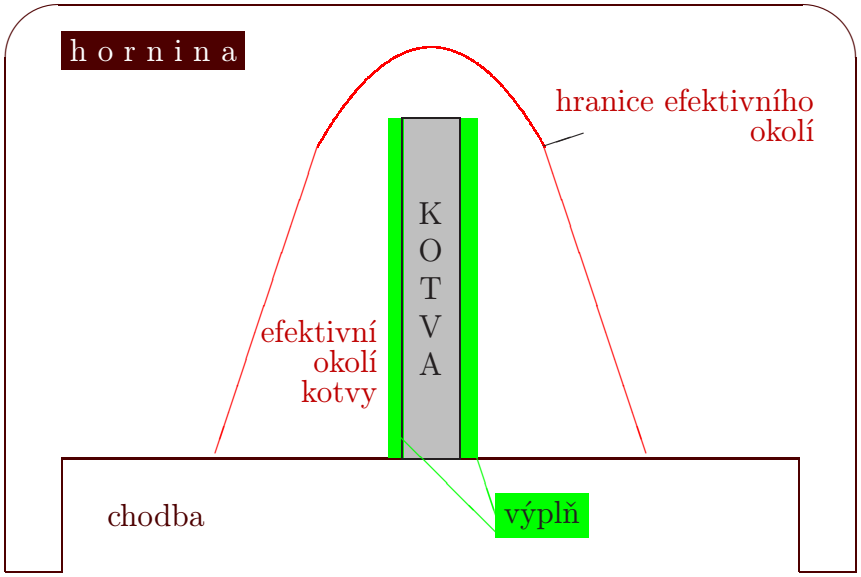
1 Úvod

Jako výchozí situaci uveďme problém, který nazýváme *problémem nestabilnosti stropů* a ke kterému například dochází při stavbě tunelů či chodeb v dolech. Otázkou zůstává jak stropy efektivně *vyztužit* a zabránit tak nežádoucímu sesuvu. Jedno z možných řešení je „ukotvení“ pomocí tzv. *pasivního systému kotev*.

Prakticky to znamená, že se do horniny tvořící strop navrtají kotvy (dlouhé kovové tyče s poměrem $\frac{\text{délka}}{\text{průměr}}$ například $\frac{100}{4}$, $\frac{200}{3}$, atp.) Okolo každé kotvy je vrstva výplně (v příkladech užíváme materiálových konstant epoxidu) o tloušťce odpovídající zhruba poloměru kotvy. Všechna tři tělesa (hornina, výplň i kotva) jsou ve vzájemném *kontaktu*. Vznikající *tření* na konaktních plochách brání uvolnění kotvy, která jinak není nijak uchycena. Výplň mez kotvou a horninou toto tření ještě zvětšuje.

Navržený systém kotev by měl vytvářet „dostatečně velké“ tření na „přiměřeně velkých“ kontaktních plochách, aby dostatečně zpevnil nestabilní strop. Při dostatečném tření se totiž na kotvy přenáší tahová napětí způsobující porušení materiálu a případně i zhroucení stropu.

Úkolem je tedy nalézt výpočtový model popsané interaktivní soustavy těles, který by vyhovujícím způsobem modeloval chování pasivního systému výztuže. Jako první krok při hledání tohoto modelu jsme zvolili numerické experimenty s navrženými velmi zjednodušenými variantami, které reprezentují některé aspekty lokálního modelu *výztuže s jedinou kotvou („single bolt model“)*, viz kapitola 2 na straně 6. V následující části 3 na straně 7 popisujeme varianty zjednodušeného výpočtového modelu a v poslední části 4, strana 20, uvádíme výsledky provedených experimentů.

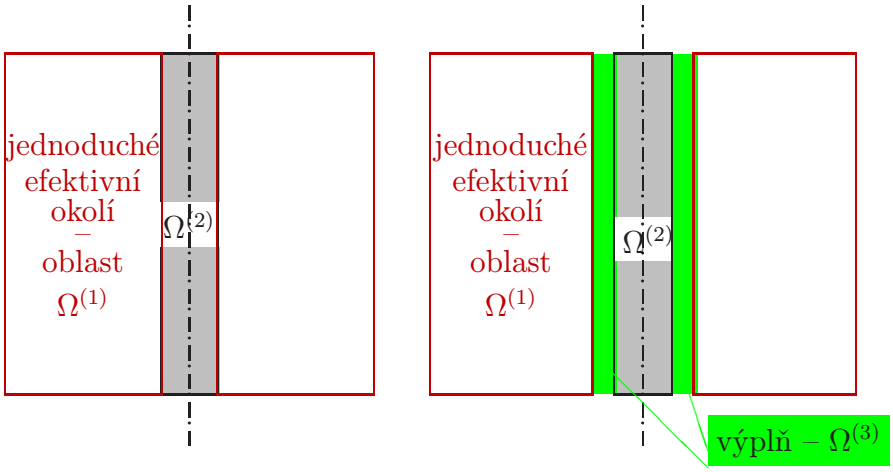


Obrázek 1. 2D průřez popisující lokální model výztuže s jedinou kotvou.

2 Zjednodušení cílového modelu

Nejprve uveďme schématický nákres (2D průřez), který popisuje lokální model výztuže s jedinou kotvou.

Jak jsme se již v úvodu zmínili, numerické experimenty jsme prováděli s výpočtovými modely (dále jen modely), jimiž jsme situaci z obrázku 1 aproximovali. Uvažovali jsme oblasti reprezentující řezy jednolivými tělesy tak, aby měly jednoduchou geometrii. Proto jsme do aproximujících modelů nezahrnuli svrchní část efektivního okolí, tj. část nad myšleným horizontálním předělem procházejícím horním ukončením kotvy. Navíc jsme zbývající část efektivního okolí zjednodušili na těleso s obdélníkovým průřezem. Další zjednodušení vyplynou z předepisovaných okrajových podmínek a předpokladů uvedených v následující kapitole (kapitola 3, strana 7).



(a) zjednodušený model (b) model kotvy

Obrázek 2. výpočtové modely

Z obrázku 2. je zřejmé, že vzhledem ke zjednodušením, která jsou do výpočtových modelů (dále jen modelů) zahrnuta, budou provedené numerické experimenty sledovat pouze některé aspekty chování cílového numerického modelu.

Při hledání vyhovujícího modelu, který by dostatečně přesně aproximoval lokální model z obrázku 1., dojdeme hned k několika „otevřeným“ problémům. Například neznáme

- velikost „efektivního okolí“ kotvy, tj. oblasti, jejíž pole napětí a deformace je tvarem kotvy významně ovlivněno. Tento problém zde neřešíme, pouze jej částečně aproximujeme škálou příkladů s *různým poměrem velikostí* jednotlivých těles;

- chování na hranici těles, přesněji řečeno mluvíme o hranici efektivního okolí kotvy a okolního prostředí. Aproximaci tohoto chování zachycujeme volbou *různých typů okrajových podmínek*.

Chování navrženého modelu je formulováno a následně vypočítáno pomocí

- *lineární* i *nelineární matematické teorie pružnosti.*

Nakonec uveďme, že při diskretizaci spojitého modelu etodou konečných prvků používáme

- *hrubší* a *jemnější diskretizační síť.*

Při zkoumání jsme se v tomto příspěvku zaměřili především na zjednodušený model uvedený na obrázku 2.(a) a pro něj uvedeme i tabulkový přehled výsledků.

Výpočtové modely jednotlivých úloh jsou sestaveny a řešeny pomocí komerčního SW systému ANSYS, který byl na katedře MAaAM k dispozici v univerzitní verzi (University high option RELEASE 6.1). SW ANSYS při diskretizaci užívá metodu konečných prvků.

Pro model se třemi oblastmi z obrázku 2.(b), který jsme nazvali modelem kotvy, nebudeme formulovat spojité úlohy, pouze uvedeme grafická řešení pro jednotlivé typy okrajových podmínek, viz kapitola 4.3.

3 Řešené úlohy

Jak jsme se v předchozí části zmínili, nabízí se několik možných variant volby vstupních dat či způsobu formulace úlohy, které mohou při výpočtu ovlivnit výslednou funkčnost ukotvení, tedy schopnost systému dostatečně vyztužit danou oblast v hornině, aby pomocí tření na kontaktních plochách zabránil zhroucení stropu.

Numerické experimenty jsme prováděli hlavně se zjednodušeným modelem z obrázku 2. (a) a proto se veškeré další předpoklady a formulace budou vztahovat k tomuto modelu (bez výplně).

3.1 Předpoklady společné všem výpočtovým modelům

Systém je tvořen dvěma (případně třemi) pružnými tělesy ve vzájemném kontaktu. Jestliže kotvě (případně výplni) ponecháme některé stupně volnosti, bude model semikoercivní. Proto musíme formulovat podmínky řešitelnosti. Zvolili jsme tedy parametrickou třídu modelů, na které jsme zkoumali vliv změny vybraných parametrů. Třída je určena skupinou předpokladů uvedených v následujícím výčtu. Jednotlivé parametry jsou popsané v oddílu 3.2 a odpovídají výběru typu geometrie, okrajové podmínky na zvolené části hranice, hodnoty součinitele smykového tření, metody pro spojitou formulaci a hustoty diskretizace.

I. Pro jednoduchost předpokládejme, že všechna tělesa jsou *homogenní* a *izotropická*. Materiálové koeficienty odpovídající jednotlivým tělesům tedy budou konstantní.

II. Formulujme spojitou úlohu jako problém *rovinné deformace*[1], který vychází z následujících předpokladů[2].

1. Nechť je v souřadném systému $\{0, \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3\}$ dán nekonečně dlouhý válec s osou rovnoběžnou se směrem $\mathbf{i}_3$ a s konstantním řezem vedeným rovinou kolmou k ose $\mathbf{i}_3$. Reprezentujme 3D těleso tímto válcem a jeho řez obecně označme jako oblast

$$\Omega \subset \mathbf{R}^2 \,.$$

2. Nechť následující veličiny nezávisí na třetí proměnné:

- vektor objemových sil, tj. $\mathbf{F} = \mathbf{F}(x_1, x_2)$,
- a vektor povrchových napětí, $\mathbf{T} = \mathbf{T}(x_1, x_2)$.

Ve výčtu by následovaly materiálové koeficienty a počáteční posunutí, $\lambda = \lambda(x_1, x_2)$, $\mu = \mu(x_1, x_2)$ a $\mathbf{u}_0 = \mathbf{u}_0(x_1, x_2)$. Tyto veličiny uvádíme jen pro úplnost, protože počáteční posunutí nebudeme nadále v „našem modelu“ potřebovat a protože λ a μ budou konstantami díky předpokladu v **I.** odstavci.

3. Dále předpokládáme, že veškerá zatížení působí v rovině oblasti Ω. Pro vektory objemových i povrchových sil $\mathbf{F}$ a $\mathbf{T}$ tedy platí, že

$$F_3 = 0 \text{ a } T_3 = 0 \,.$$

Na základě výše uvedených předpokladů můžeme i o hledaném vektoru posunutí $\mathbf{u}$ tvrdit, že $u_{1,3} = u_{2,3} = u_3 = 0$. Tedy pro ostatní složky vektoru posunutí platí, že

$$u_i = u_i(x_1, x_2) \quad \text{pro } \; i = 1, 2 \,.$$

Potom i složky tenzoru deformací s alespoň jedním indexem odpovídajícím třetí proměnné jsou nulové, tj. $e_{33} = e_{13} = e_{23} = 0$. Odtud název rovinná deformace.

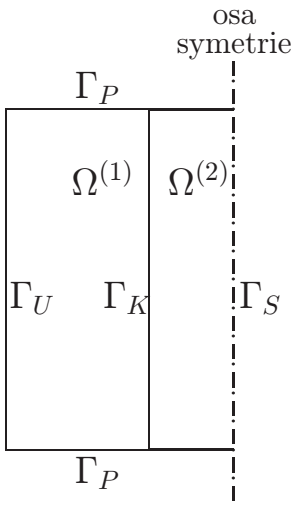
III. Jako zatížení uvažujme tíhovou sílu $\mathbf{F} = -g \cdot \rho \cdot \mathbf{i}_2$ odpovídající tíhovému zrychlení $g = 9.81 {}^{m}\!/_{s^2}$ a hustotě ρ zvoleného materiálu. Protože jsme v **II.** části převedli 3D úlohu na 2D, bude tíhová síla $\mathbf{F}$ působit ve směru $-\mathbf{i}_2$.

IV. Okrajové podmínky společné pro všechny modelové úlohy odpovídají následujícímu označení jednotlivých částí hranice.

---

[1] Viz naříklad [Nečas, Hlaváček, strana 149 až 177].

[2] Naším záměrem je zkoumat vliv poměrů velikostí řezů jednotlivými tělesy, viz dále v kapitole 3.2.1 na straně 9. Proto vycházíme z předpokladů *rovinné deformace*, ačkoli tím silně omezujeme aproximativnost modelu výstuže s jedinou kotvou, který by jinak byl výstižněji popsán jako problém *rotační symetrie*.



Obrázek 3. označení částí hranice pro zjednodušený výpočtový model

Na hranicích $\partial\Omega^{(1)}$ a $\partial\Omega^{(2)}$ předepisujeme následující okrajové podmínky:

a) podmínka *symetrie* na hranici Γ_S, tj. pro posunutí bodů na hranici Γ_S oblasti $\Omega^{(2)}$ ve směru vnější normály $\mathbf{n}$ platí, že $u_n^{(2)} = 0$, kde $u_n^{(2)} = u_i^{(2)} n_i$;

b) podmínka *kontaktu* na hranici Γ_K, tj. součet posunutí ve směru normály bodů hranice $\partial\Omega^{(1)}$ a bodů hranice $\partial\Omega^{(2)}$ $u_n^{(1)} - u_n^{(2)} \leq 0$;

c) na Γ_P předepisujeme *nulové zatížení*, tj. pro složky tenzoru napětí τ_{ij} platí rovnost $\tau_{ij}(\mathbf{u})n_j = 0$, a nakonec

d) typ podmínky předepisované na hranici Γ_U se mění podle zvoleného typu podpory. Blíže tuto podmínku specifikujeme v části 3.2.2 popisující varianty předepsaných okrajových podmínek.

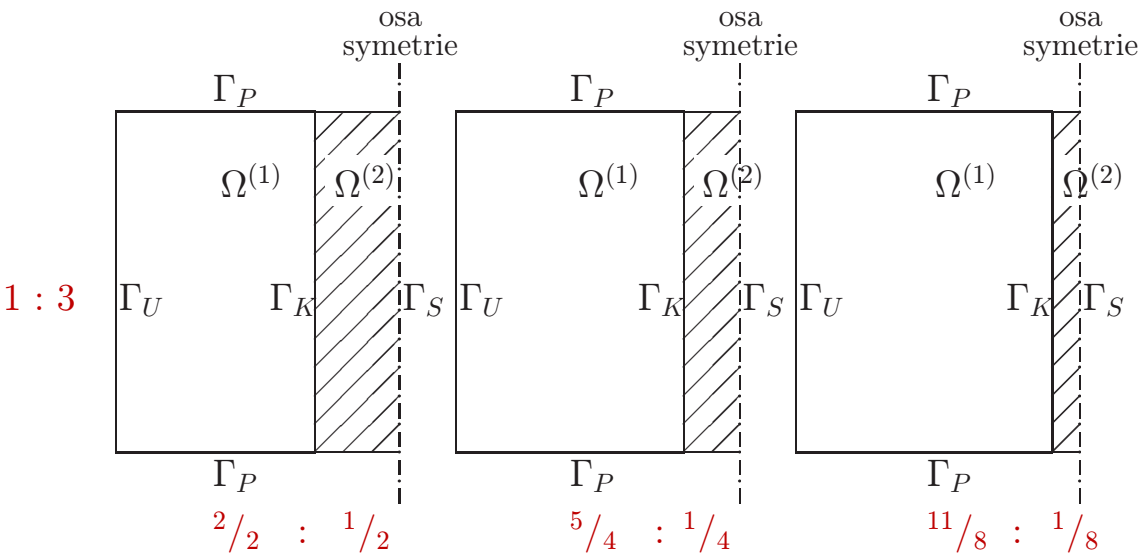
3.2 Přehled volených variant vstupních dat a dalších předpokladů

V předchozích odstavcích máme shrnuty předpoklady, konstantně se vyskytujících u všech modelů. Jak jsme již předeslali, numerickými experimenty jsme testovali vliv změny některého z následujících pěti předpokladů či vstupních dat na zvolený model, tj. zjednodušený výpočtový model.

3.2.1 Varianty geometrie

Protože efektivní okolí (oblast $\Omega^{(2)}$) kotvy není přesně známo, nabízí se možnost zkoumat alespoň přibližně vliv jeho dimenze (a typu okrajových podmínek, viz následující oddíl 3.2.2) pomocí volby různých variant poměru tloušťky kotvy a velikosti neznámé efektivní oblasti.

Obrázek zachycuje varianty poměru šířky efektivní oblasti a poloměru kotvy, předpokládáme-li, že šířka modelu je 3krát větší než výška.



Obrázek 4. 2D schémata ilustrující jednu skupinu z uvažovaných variant geometrie, kde 1 : 3 je poměr šířky k výšce celého modelu. Pomocí $^2/_2$: $^1/_2$, $^5/_4$: $^1/_4$ a $^{11}/_8$: $^1/_8$ charakterizujeme rozdělení poloviční šířky modelu mezi oblastmi $\Omega^{(1)}$ a $\Omega^{(2)}$, tj. podíl šířky efektivní oblasti kotvy a poloměru kotvy.

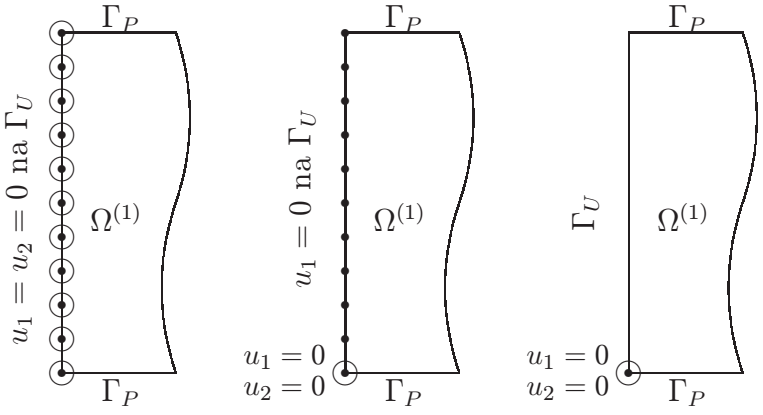
Celkový přehled zvolených geometrických vztahů je uveden v následující tabulce.

výška : šířka (pro celý model)	šířka efektivní oblasti : poloměr kotvy (pro poloviční model)		
1 : 3	$^2/_2$: $^1/_2$	$^5/_4$: $^1/_4$	$^{11}/_8$: $^1/_8$
1 : 2	$^1/_2$: $^1/_2$	$^3/_4$: $^1/_4$	$^7/_8$: $^1/_8$
1 : 1		$^1/_4$: $^1/_4$	$^3/_8$: $^1/_8$

Tabulka č.1 : přehled uvažovaných variant geometrie

3.2.2 Varianty předepsaných okrajových podmínek

Chování na části Γ_U hranice $\partial\Omega^{(1)}$, tj. chování na hranici mezi efektivní oblastí a okolní horninou, aproximujeme jednou z následujících možností podmínek v uzlech diskretizovaného modelu:



a) podmínka typu 1 b) podmínka typu 2 c) podmínka typu 3

Obrázek 5. varianty okrajových podmínek

3.2.3 Varianty hodnot součinitelů smykového tření

Kotva má za úkol zpevnit nadloží tím, že díky vznikající třecí síle na kontaktních plochách převezme část tahového napětí z horniny. Užitý *Coulombovský model tření* vychází ze vztahu normálového napětí a napětí ve směru tangenty

$$\begin{cases} ||\mathbf{T}_t|| \leq \mathcal{F}\,|T_n|\,; \\ \text{jestliže} \quad ||\mathbf{T}_t|| < \mathcal{F}\,|T_n|\,, \quad \text{potom} \quad \mathbf{u}_t = \mathbf{0}\,, \\ \text{jestliže} \quad ||\mathbf{T}_t|| = \mathcal{F}\,|T_n|\,, \quad \text{potom} \quad \exists \chi \geq 0 \ \text{ tak, že } \ \mathbf{u}_t = -\chi \mathbf{T}_t\,, \end{cases}$$

kde pro vnější normálu $\mathbf{n}(\mathbf{x}) = \{n_i(\mathbf{x})\}_i$ je $T_n = \tau_{ij}(\mathbf{u}) n_i n_j$ normálové napětí, $\mathbf{T}_t = \tau_{ij}(\mathbf{u}) n_j - T_n \mathbf{n}$ vektor tangenciálních napětí. Analogicky uvažujeme značení $u_n (= u_i n_i)$ a $\mathbf{u}_t = \mathbf{u} - u_n \mathbf{n}$. Symbol $\mathcal{F}$ je součinitelem smykového tření.

Budeme-li experimentálně zvyšovat hodnotu $\mathcal{F}$, poroste i velikost třecí síly. Tím dojde i k lepšímu zpevnění nadloží. Nicméně součinitel smykového tření je fyzikální veličina získávaná měřením. Jde o tangens úhlu sklonu podložky, pod kterým se těleso z daného materiálu dává po podložce do pohybu. Díky tomu je $\mathcal{F} \in \langle 0, 1 \rangle$. Hodnoty $\mathcal{F}$ jsou vždy tabelovány za jasně daných podmínek, mimo jiné i pro tělesa z konkrétních materiálů. Je tedy zřejmé, že součinitel smykového tření stejně jako materálové konstanty se vztahují ke zvolenému materiálu. Při získávání prezentovaných výsledků jsme ale zvyšovali jen hodnotu $\mathcal{F}$ a jednalo se tedy pouze o teoretickou úvahu. Navíc vzhledem ke způsobu měření není příliš běžné, aby se hodnoty $\mathcal{F}$ vyskytovali v blízkosti jedničky. To jsou důvody, proč u takto získaných výsledků pouze konstatujeme, že po zvýšení hodnoty součinitele $\mathcal{F}$ program nabízí řešení, k němuž došel na základě zvolené výpočtové metody.

3.2.4 Varianty užitých metod pro spojitou formulaci

Na zvolených výpočtových modelech jsme testovali, jak se projeví předpoklad o chování materiálů, ze kterých jsou tělesa složena. Jak jsme dříve uvedli, pro jednoduchost uvažujeme jen homogenní tělesa z izotropických materiálů. V následujících dvou odstavcích shrneme uvažované varianty o závislosti přetvoření těles na vektoru posunutí $\mathbf{u}$.

I. Předpokládejme, že výslednou deformaci tělesa způsobenou vnějším zatížením můžeme dostatečně přesně popsat vztahy, které jsou lineární vzhledem k vektoru posunutí $\mathbf{u} = \{u_j\}_1^3$ libovolného bodu $\mathbf{x} = \{x_i\}_1^3$ daného tělesa. Dochází tedy pouze k tzv. *malým deformacím*. Tenzor malých deformací $\mathbf{e} = \{e_{ij}\}_{i,j}$ je tvaru

$$\mathbf{e}(\mathbf{u}) = \frac{1}{2}\left(\nabla\mathbf{u} + [\nabla\mathbf{u}]^T\right). \tag{1}$$

Rovnice rovnováhy budou formulovány pomocí *Cauchyho tenzoru napětí* $\mathbf{t} = \{\tau_{ij}\}_{i,j}$.

II. Výslednou deformaci tělesa způsobenou vnějším zatížením popíšeme vztahy, které jsou nelineární vzhledem k vektoru posunutí $\mathbf{u}$, tj. když dochází k tzv. *velkým deformacím*. Tenzor velkých deformací označený $\mathbf{d} = \{d_{ij}\}_{i,j}$ je tvaru

$$\mathbf{d}(\mathbf{u}) = \frac{1}{2}\left(\nabla\mathbf{u} + [\nabla\mathbf{u}]^T + [\nabla\mathbf{u}]^T \cdot \nabla\mathbf{u}\right). \tag{2}$$

V lineární teorii se vlastně veškerá zatížení předepisují na deformované oblasti $f(\hat{\Omega}) = \Omega$, stejně tak se i všechny veličiny, jejichž hodnota popisuje vlastnosti v bodě tělesa, vztahují k pozici x materiálového bodu $\hat{x}$ v deformované oblasti Ω. To ale s sebou nese předpoklad, že Jakobián $det[\nabla f] = \frac{vol\ f(\hat{\Omega})}{vol\ \hat{\Omega}}$ deformace f je předem znám. V nelineární teorii předepisujeme veškerá zatížení na referenční (nedeformované) oblasti $\hat{\Omega}$. Rovnice rovnováhy tedy formulujeme pomocí *Piola–Kirchhoffova tenzoru napětí* $\hat{\mathbf{s}} = \{\hat{\sigma}_j^i\}_{i,j}$.

3.2.5 Varianty hustoty diskretizace

Chování diskretizovaného modelu se v průběhu řešení obecně jeví jako tužší v případě volby hrubší sítě. Proto jsme každý model řešili dvakrát. Poprvé jsme provedli diskretizaci jednotlivých oblastí „hrubší“ sítí a potom znovu sítí „jemnější“ (zhruba dvojnásobnou).

3.3 Formulace matematických modelů řešených úloh

3.3.1 Rovnice rovnováhy

Mějme dánu oblast $\Omega \subset \mathbf{R}^2$ s hranicí $\partial\Omega$. Na základě předpokladů uvedených v části 3.2.4 budou i rovnice rovnováhy dvojího typu v závislosti na zvoleném tenzoru deformace.

Lineární chování deformovaného tělesa popisujeme tenzorem malých deformací, $\mathbf{e}(\mathbf{u}) = \frac{1}{2}\,(\nabla\mathbf{u} + \nabla^T\mathbf{u}\,)$.

Rovnice rovnováhy (pro stacionární případ) odpovídající předpokladu o rovinné deformaci mají tvar

$$\begin{array}{l} \operatorname{div} \mathrm{t} + \mathbf{F} = 0 \;\text{ v }\; \Omega\,, \;\text{ tj.} \\ \frac{\partial \tau_{ij}}{\partial x_j} + F_i = 0 \;\; (i,j = 1,2)\,, \end{array} \tag{3}$$

kde $\mathbf{F} = \{F_i(x_1, x_2)\}_i$ $(F_3 = 0)$ jsou složky vektoru objemových sil a $\mathrm{t} = \{\tau_{ij}\}_{i,j}$ je Cauchyho tenzor napětí, pro nejž platí *zobecněný Hookeův zákon* ve tvaru

$$\tau_{ij} = \lambda\,\delta_{ij} e_{kk} + 2\mu\, e_{ij}\,. \tag{4}$$

Součinitelé λ a μ jsou Lamého koeficienty, $\forall \mathbf{x} \in \Omega$ platí $\lambda(\mathbf{x}) = \lambda > 0$ a $\mu(\mathbf{x}) = \mu > 0$.

Nelineární tenzor deformace $\mathbf{d}(\mathbf{u})$ popisuje konečná přetvoření v tělese a platí, že

$$\mathbf{d}(\mathbf{u}) = \frac{1}{2}\,(\nabla\mathbf{u} + \nabla^T\mathbf{u} + \nabla^T\mathbf{u}\cdot\nabla\mathbf{u}\,).$$

Uvažujme referenční oblast $\hat{\Omega}$. Nechť symbol $\hat{\sigma}(\hat{\mathbf{x}}, \hat{\mathbf{n}}(\hat{\mathbf{x}})\,)$ označuje vektor povrchového napětí ve směru $\hat{\mathbf{n}}(\hat{\mathbf{x}}) = \{\hat{n}_i(\hat{\mathbf{x}})\,\}_i$ k nedeformovanému povrchu. Potom platí, že $\hat{\sigma}_i(\hat{\mathbf{x}}, \hat{\mathbf{n}}(\hat{\mathbf{x}})) = \hat{\sigma}^i_j(\hat{\mathbf{x}})\,\hat{n}_j(\hat{\mathbf{x}})$, kde $\hat{\mathbf{s}} = \{\hat{\sigma}^i_j\}_{i,j}$ je *I. Piola–Kirchhoffův tenzor napětí*. Tento tenzor je obecně nesymetrický. Splňuje pouze rovnost $\hat{\mathrm{s}}\,[\nabla f]^T = [\nabla f]\,\hat{\mathrm{s}}^T$.

Rovnice rovnováhy (opět pro stacionární případ) mají pro rovinnou deformaci tvar

$$\begin{array}{l} \widehat{\operatorname{div}}\,\hat{\mathbf{s}} + \hat{\mathbf{F}} = \mathbf{0} \quad \text{v } \hat{\Omega} \quad \text{tj.} \\ \frac{\partial \hat{\sigma}^i_j}{\partial \hat{x}_j} + \hat{F}_i = 0 \;\; (i,j = 1,2)\,, \end{array} \tag{5}$$

kde $\hat{\mathbf{F}} = \{\hat{F}_i\}_i$ je vektor objemových sil působících na těleso před deformací. Jestliže deformace f přiřadí bodu $\hat{\mathbf{x}} \in \hat{\Omega}$ polohu $\mathbf{x} = f(\hat{\mathbf{x}})$ v Ω, pak vztahy mezi vektory z (5) ve tvaru před deformací a po deformaci a je dán předpisem

$$\hat{\mathbf{F}}(\hat{\mathbf{x}}) = det[\nabla f(\hat{\mathbf{x}})] \cdot \mathbf{F}(\mathbf{x}) \quad \text{a} \quad \widehat{\operatorname{div}}\,\hat{\mathbf{s}}(\hat{\mathbf{x}}) = det[\nabla f(\hat{\mathbf{x}})] \cdot \operatorname{div} s(\mathbf{x}) \tag{6}$$

který je odvozený z transformačního zákona[3].

Předpokládáme-li , že vektory $\frac{\partial f(\hat{\mathbf{x}})}{\partial \hat{x}_j}$ jsou lineárně nezávislé[4], pak se tenzor označený $\hat{\mathbf{t}} = \{\hat{t}_{ij}\}_{i,j}$ a splňující

$$\hat{\mathbf{t}}(\hat{\mathbf{x}}) = [\nabla f(\hat{\mathbf{x}})]^{-1}\hat{\mathbf{s}}(\hat{\mathbf{x}}) = det[\nabla f(\hat{\mathbf{x}})] \cdot [\nabla f(\hat{\mathbf{x}})\,]^{-1}\mathbf{s}(\mathbf{x})[\nabla f(\hat{\mathbf{x}})\,]^{-T} \qquad (7)$$

nazývá *II. Piola–Kirchhoffův tenzor napětí*. Poslední výraz v (7) je opět výsledkem transformačních vztahů tentokrát pro hraniční členy a platí, že $\mathbf{x} = f(\hat{\mathbf{x}})$ pro $\hat{\mathbf{x}} \in \hat{\Omega}$.

Elastické těleso nazveme *hyperelastickým*, jestliže II. Piola–Kirchhoffův tenzor napětí $\hat{\mathbf{t}}$ ze vztahu (7) můžeme vyjádřit ve tvaru Hookeova zákona

$$\hat{\mathbf{t}}(\hat{\mathbf{x}}) = \frac{\partial \widehat{W}(\hat{\mathbf{x}}, \mathbf{d})}{\partial \mathbf{d}} = det[\nabla f(\hat{\mathbf{x}})] \cdot \frac{\partial W(\mathbf{x}, \mathbf{d})}{\partial \mathbf{d}}[\nabla f(\hat{\mathbf{x}})\,]^{-T}\,, \qquad (8)$$

kde W je dostatečně hladká funkce, která je invariantní vzhledem k záměně d_{ji} za d_{ij} a která se nazývá *potenciál deformační energie*. Dosazením II. Piola–Kirchhoffova tenzoru podle (7) a Hookeova zákona (8) do rovnic rovnováhy (5) a díky transformačním vztahům (6) získáme rovnice rovnováhy ve tvaru

$$\frac{\partial}{\partial x_i}\left(\frac{\partial W}{\partial d_{ij}} + \frac{\partial W}{\partial d_{ik}}\frac{\partial u_j}{\partial \hat{x}_k}(\hat{\mathbf{x}})\right) + F_j(\mathbf{x}) = 0 \;\; (i,j = 1,2) \;\; \text{v} \;\; \Omega \qquad (9)$$

pro $\mathbf{x} = f(\hat{\mathbf{x}})$, $\hat{\mathbf{x}} \in \hat{\Omega}$.

Nechť $\mathcal{I}_1$, $\mathcal{I}_2$ a $\mathcal{I}_3$ jsou hlavní invarianty Cauchy–Greenova tenzoru deformace $\mathbf{C} = [\nabla f]^T \nabla f = \mathbf{I} + 2 \cdot \mathbf{d}(\mathbf{u})$, kde $\mathbf{I}$ je jednotkovým tenzorem. Pro výpočtové modely jsme zvolili tzv. *Neo-Hookeanův potenciál deformační energie*, který je ve tvaru

$$W(d_{ij}) = W(\mathcal{I}_1, [\nabla_f]\,) = \frac{\mu}{2}(\,[\nabla_f]^{-\frac{1}{3}}\mathcal{I}_1 - 3) + \frac{1}{2}(\lambda + \frac{2}{3}\mu) \cdot (\,[\nabla_f] - 1)^2\,,$$

kde $[\nabla_f]$ je Jakobián deformace f, $[\nabla_f] = det[\nabla f]$,
λ a μ Laméovy materiálové konstanty.

3.3.2 Okrajové podmínky

Nechť Γ_P, Γ_S, Γ_K a Γ_U jsou otevřené disjunktní části hranice jedné z oblastí $\Omega^{(1)}$ případně $\Omega^{(2)}$ tak, že $\overline{\Gamma}_P \cup \overline{\Gamma}_K \cup \overline{\Gamma}_U = \partial\Omega^{(1)}$ a $\overline{\Gamma}_P \cup \overline{\Gamma}_K \cup \overline{\Gamma}_S = \partial\Omega^{(2)}$. Část Γ_U hranice $\partial\Omega^{(2)}$, kde budeme předepisovat různé typy okrajových podmínek, navíc

[3]Viz např. [Gurtin, str. 51].

[4]Lineární nezávislost sloupců ∇f nám zaručí, že objem $f(\hat{\Omega})$ je nenulový.

rozdělme na dvě části Γ_U^a a Γ_U^b, $\overline{\Gamma}_U = \overline{\Gamma}_U^{\,a} \cup \overline{\Gamma}_U^{\,b}$. Předepišme okrajové podmínky dle dříve uvedených předpokladů

$$\begin{aligned}
&\text{na } \Gamma_S && u_n^{(2)} = 0\,;\\
&\text{na } \Gamma_K && u_n^{(1)} - u_n^{(2)} \le 0\,,\\
& && T_n^{(1)} = T_n^{(2)} = T_n \le 0 \quad \text{a}\\
& && (\,u_n^{(1)} + u_n^{(2)}\,)T_n = 0\,,\\
& && ||\mathbf{T}_t(\mathbf{u})|| \le \mathcal{F}|T_n(\mathbf{u})| \quad \text{a}\\
& && \mathbf{T}_t(\,\mathbf{u}_t\,) + \mathcal{F}|T_n(\mathbf{u})|\,||\,\mathbf{u}_t\,|| = 0\\
& && \text{pro } \mathbf{u}_t = \mathbf{u}_t^{(1)} - \mathbf{u}_t^{(2)}\,;\\
&\text{na } \Gamma_U && u_1 = u_2 = 0 \quad (\,\text{tedy } \mu(\Gamma_U^b) = 0\,)\,,\\
&\text{nebo} && u_1 = u_2 = 0 \quad \text{na části } \Gamma_U^a \quad \text{a}\\
& && u_1 = 0 \quad \text{na zbytku } \Gamma_U^b = \Gamma_U - \Gamma_U^a\,,\\
&\text{nebo} && u_1 = u_2 = 0 \quad \text{pouze na části } \Gamma_U^a\\
& && \text{a navíc je } \Gamma_U^b \subset \Gamma_P
\end{aligned}$$

pro daný součinitel smykového tření $\mathcal{F}$. V podmínkách předepsaných na hranicích Γ_S a Γ_K jsme navíc potřebovali rozlišit posunutí tak, aby odpovídala posunutí bodů jednotlivých oblastí $\Omega^{(r)}$. Proto jsme zavedli označení $u_n^{(r)}$ a analogicky $T_n^{(r)} = T_n(\mathbf{u}^{(r)})$. Nakonec jsme užili rovnosti $\mathbf{n} = \mathbf{n}^{(1)} = -\mathbf{n}^{(2)}$ na Γ_K, kde $\mathbf{n}^{(r)}$ je vnější jednotková normála k části hranice $\partial\Omega^{(r)}$ $(r = 1, 2)$. Toto rozlišení nezdůrazňujeme v případě, kdy se celý předpis vztahuje pouze k jediné oblasti.

Poznámka 1 Podmínka kontaktu předepisovaná na části Γ_K je v SW systému ANSYS realizována výběrem speciálních prvků přiřazených uzlům diskretizační sítě, které leží na kontaktních plochách. Viz kapitola 3.4 na straně 18.

3.3.3 Variační formulace

Nechť $\mu^{(r)}, \lambda^{(r)}$ jsou konstanty, $\mu^{(r)} > 0$, $\lambda^{(r)} > 0$. Nechť $V^{(r)}$ jsou prostory virtuálních posunutí

$$\begin{aligned}
V^{(r)} = \{\mathbf{v}^{(r)} \in H^1(\Omega^{(r)}) \times H^1(\Omega^{(r)}) \,|\, &\mathbf{v}^{(r)} = 0 \ \text{ na } \ \Gamma_U^a\,,\\
&v_1^{(r)} = 0 \text{ na } \Gamma_U^b \quad \text{a } v_n^{(r)} = 0 \quad \Gamma_S\}\,.
\end{aligned}$$

Připomeňme, že při formulacích uvažujeme případ zjednodušeného výpočtového modelu bez výplně.

Nejdříve definujme jednotlivé funkcionály tak, jak odpovídají okrajovým podmínkám (**A.**), lineární (**B.**) a nelineární (**C.**) teorii. Nakonec uvedeme definici variačního řešení

A. Nechť $d\Gamma$ je elementem hranice oblasti. Definujme funkcionál $j(\ ,\ )$ vztahem

$$j(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}) := \int_{\Gamma_K} \mathcal{F}|T_n(\mathbf{u})|\,||\mathbf{u}_t^{(1)} - \mathbf{u}_t^{(2)}||\,d\Gamma\,.$$

Tímto nediferencovatelným konvexním fukcionálem přispějí do variační nerovnic okrajové podmínky.

Poznámka 2 S takto definovaným funkcionálem $j(\mathbf{u}^{(1)}, \mathbf{u}^{(2)})$ je úloha komplikovaná, proto se meznámá hodnota $|T_n(\mathbf{u})|$ nahradí známým funkconálem g a variačním řešením problému s Coulombovským třením rozumíme takové kinematicky přípustné posunutí, pro nějž je $-T_n(\mathbf{u})$ pevným bodem zobrazení $\mathcal{G} : H^+ \mapsto H^+$, kde

$$H^+ = \{g \in H^{-\frac{1}{2}}(\Gamma_K) \mid \langle g; v\rangle \geq 0 \ \ \forall v, \ \ v \geq 0 \ \ \text{na}\ \Gamma_K\}\,.$$

B. Lineárnímu tenzoru deformace (1) a příslušné rovnici rovnováhy (3) spolu s Hookeovým zákonem (4) odpovídá funkcionál energie $\mathcal{J}^{(e)}$ definovaný na $K^{(e,1)} \times K^{(e,2)}$, kde $K^{(e,r)} = V^{(r)}$, předpisem

$$\mathcal{J}^{(e)}(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}) := \mathcal{J}_0^{(e,1)}(\mathbf{u}^{(1)}) + \mathcal{J}_0^{(e,2)}(\mathbf{u}^{(2)}) + j(\mathbf{u}^{(1)}, \mathbf{u}^{(2)})\,, \ \ \text{kde}$$

$$\mathcal{J}_0^{(e,r)}(\mathbf{u}^{(1)}) := \int_{\Omega^{(r)}} [\,\frac{1}{2}\lambda^{(r)}\varepsilon_{kk}(\mathbf{u}^{(r)}) + \mu^{(r)}\varepsilon_{ij}(\mathbf{u}^{(r)}) \cdot \varepsilon_{ij}(\mathbf{u}^{(r)})\,]\,d\Omega^{(r)}$$

$$- \int_{\Omega^{(r)}} F_i^{(r)} u_i^{(r)}\,d\Omega^{(r)}\,.$$

C. Nechť $K^{(h,r)}$ je množina všech kinematicky přípustných posunutí, která je pro $r = 1, 2$ definována

$$K^{(h,r)} := \{(\mathbf{v}^{(r)} \in V^{(r)} \,|\, det[\mathrm{Id} + \nabla \mathbf{v}^{(r)}] > 0 \ \ \text{s.v.} \ \ \text{v} \ \ \Omega^{(r)}\}\,.$$

Nelineární tenzor deformace (2) a rovnice rovnováhy (9) spolu s Hookeovým zákonem (8) určují funkcionál energie $\mathcal{J}^{(h)}$ definovaný na $K^{(h,1)} \times K^{(h,2)}$ předpisem

$$\mathcal{J}^{(h)}(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}) := \mathcal{J}_0^{(h,1)}(\mathbf{u}^{(1)}) + \mathcal{J}_0^{(h,2)}(\mathbf{u}^{(2)}) + j(\mathbf{u}^{(1)}, \mathbf{u}^{(2)})\,, \ \ \text{kde}$$

$$\mathcal{J}_0^{(h,r)}(\mathbf{u}^{(r)}) := \int_{\Omega^{(r)}} W(d_{ij}(\mathbf{u}^{(r)})\,d\Omega^{(r)} - \int_{\Omega^{(r)}} F_i^{(r)} u_i^{(r)}\,d\Omega^{(r)}\,.$$

Definice 1 V závislosti na volbě lineární či nelineární teorie, tj. pro zvolené $\alpha \in \{e, h\}$, definujme konvexní množinu

$$K^{(\alpha)} := \{(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}) \in K^{(\alpha,1)} \times K^{(\alpha,1)} \,|\, v_n^{(1)} + v_n^{(2)} \leq 0 \;\text{ na }\; \Gamma_K\}\,.$$

Variačním řešením nazveme takovou dvojici kinematicky přípustných posunutí $(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}) \in K^{(\alpha)}$, řeší minimalizační úlohu

$$\mathcal{J}^{(\alpha)}(\mathbf{u}^{(1)}, \mathbf{u}^{(2)}) \leq \mathcal{J}^{(\alpha)}(\mathbf{v}^{(1)}, \mathbf{v}^{(2)})$$

na množině všech kinematicky přípustných posunutí $(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}) \in K^{(\alpha)}$.

Poznámka 3 Připoměňme, že ve formulaci $j(\mathbf{u}^{(1)}, \mathbf{u}^{(2)})$ pro hyperelastický materiál počítáme s příslušnou transformací.

Protože vektor povrchových sil odpovídající I. Piola–Kirchhoffovu tenzoru napětí je ve tvaru

$$\hat{\mathbf{s}}(\hat{\mathbf{x}})\hat{\mathbf{n}} = \hat{\mathbf{s}}_{\mathbf{n}}(\hat{\mathbf{x}}) \cdot \hat{\mathbf{n}}(\hat{\mathbf{x}}) + \hat{\mathbf{s}}(\hat{\mathbf{x}}) \;\; \hat{\mathbf{x}} \in \partial\hat{\Omega}\,,$$

má tato rovnost pro II. Piola–Kirchhoffův tenzor napětí tvar

$$\nabla f\, \hat{\mathbf{t}}(\hat{\mathbf{x}})\hat{\mathbf{n}} = \nabla f\, \hat{\mathbf{t}}_{\mathbf{n}}(\hat{\mathbf{x}}) \cdot \hat{\mathbf{n}}(\hat{\mathbf{x}}) + \nabla f\, \hat{\mathbf{t}}(\hat{\mathbf{x}}) \;\; \hat{\mathbf{x}} \in \partial\hat{\Omega}\,.$$

Potom po transformaci do deformovaného $f(\partial\hat{\Omega}) = \partial\Omega$ a po dosazení podmínek na Γ_K získáme vztah

$$\nabla f\; \hat{\mathbf{t}}(\hat{\mathbf{x}})\, \hat{\mathbf{n}}(\hat{\mathbf{x}})\; d\hat{\Gamma} = det[\nabla f\,] \cdot \nabla f\; \mathbf{t}(\mathbf{x})\, [\nabla f\,]^{-T} \mathbf{n}(\mathbf{x})\; d\Gamma\,. \qquad \square$$

Poznámka 4 SW systém ANSYS používá jiného způsobu formulace. Nabízí možnost užití penalizačního funkcionálu, popřípadě funkconálu rozšířených Lagrangiánů („Penalty function + Lagrange multiplier“). My jsme zvolili verzi rozšířených Lagrangiánů, viz následující oddíl 3.4 va straně 18.

3.4 Přehled výběru elementů a výchozích nastavení

Software: ANSYS/University high option RELEASE 6.1

Spojitý model: *geometrie* konkrétního modelu viz schémata v kapitole 3.2.1
zatížení vlastní vahou s tíhovým zrychlením $g = 9.81^{m}/_{s^2}$
okraj. podm.: viz schémata v kapitole 3.2.2

Diskretizace:

Elementy:

PLANE42	2D element pro pevné matetriály (4 uzly), lineární teorie,
PLANE182	2D element pro pevné mat. (4 uzly), hyperelasticita - nelineární teorie,
CONTAC48	2D element modelující kontakt „bodu na povrch“ (3 uzly).

Nastavené klíč. vlast.:

PLANE42	keyopt(3) „Plane strain (Z strain = 0.0)“,
PLANE182	keyopt(3) „Plane strain (Z strain = 0.0)“,
CONTAC48	keyopt(2) „Penalty function + Lagrange multiplier“, keyopt(3) „elastic coulomb“.

Dále viz ANSYS 6.1 Documentation [ANSYS Elemet].

Volba parametrů přiřazených elementu CONTAC48:

$$Toln = 10^{-5}m.$$

Dále viz ANSYS 6.1 Documentation [ANSYS Contact].

Materiály:

Lineární: pevný, homogenní, izomorfní, elastický
Nelineární: hyperelastický materiál s Neo-Hookeanovým tvarem potenciálu v Hookeově zákonu
Dále ANSYS 6.1 Documentation viz [ANSYS Nonlinear] a [ANSYS Static].

Mat. konst.:

materiál	Young. mod. E (MPa)	Poiss. č. σ	hustota ρ $({}^{kg}/_{m^3})$
kov	200 000	0.30	78 500
epoxid	10 000	0.25	2 300
hornina	2 000	0.35	2 500

3.4.1 Poznámka ke zvolenému kontaktnímu elementu

Element nazvaný CONTAC48 je 2D kontaktní 3-uzlový element s uzly

K – kontaktní uzel na „slave surface“ a

I,J – krajní uzly cílové úsečky na „master surface“,

kde *master–slave* vztah mezi uzly je dán *podmínkou nepronikání* na kontaktní hranici, tj. že „podřízené“ uzly kontaktního tělesa nesmí proniknout do tělesa cílového, jehož povrch je dán „řídícími“ uzly.

Jestliže tedy během řešení získáme $u_n(K) > 0$, tj. dojde k penetraci kontaktního uzlu K cílovou úsečkou $|IJ|$, pak nalezneme normálovou sílu F_n, která působí ve směru vnější normály k cílovému povrchu prezentovanému cílovou úsečkou $|IJ|$. F_n je při použití *metody rozšířených lagrangiánů* počítána dle vztahu

$$F_n(K) = min\{0; k_n u_n(K) + \lambda_{i+1}\}\,, \quad \text{kde}$$

k_n je tzv. kontaktní tuhost (penaltový parametr v metodě rozšířených lagrangiánů),

$u_n(K)$ je velikost penetrace kontaktního uzlu K za hranici cílové oblasti a

λ_{i+1} je $i+1$-ní iterace lagrangeova multiplikátoru upravující hodnotu normálové síly F_n.

Pro λ_{i+1} platí, že

$$\lambda_{i+1} = \begin{cases} \lambda_i + \alpha k_n u_n(K) & |u_n(K)| \geq Toln \\ \lambda_i & \text{jinak}\,, \end{cases}$$

kde $Toln$ je volený parametr a $\alpha < 1$.

Poznámka 5 Elementy CONTAC48 jsou vytvořeny příkazem

gcgen,kontaktni_uzly,cilove_uzly,2 .

Číslo *2* znamená, že k jedné cílové úsečce tvořené cílovými uzly T_1 a T_2 jsou přiřazeny dva kontaktní uzly, tj. cílová úsečka figuruje ve dvou kontaktních elementech a to:

$$\begin{array}{lcl} E_1 & : & T_1, T_2 \quad \text{a} \quad K_1, \\ E_2 & : & T_1, T_2 \quad \text{a} \quad K_2, \end{array}$$

pro kontaktní uzly K_1 a K_2. Dále viz ANSYS 6.1 Documentation [ANSYS Elemet].

4 Výsledky

4.1 Výsledky rozdělené dle zvoleného typu okrajové podmínky na hranici Γ_U

Legenda k tabulkám:

A varianta zkonvergovala, tj. software ANSYS nalezl řešení;

× model je nestabilní, výslednice tíhových sil převýšila výslednici sil třecích, které vznikají na kontaktních plochách;

[**1**] posloupnost diskrétních úloh podle ANSYSu zkonvergovala až po zvýšení součinitele smykového tření f.

Okraj. podm. typu 1	*lineární model*		*nelineární model*	
	hrubší síť	*jemnější síť*	*hrubší síť*	*jemnější síť*
1 : 3 $^{11}/_8 : {}^1/_8$	A	A	A	A
1 : 3 $^5/_4 : {}^1/_4$	[1]	[1]	[1]	[1]
1 : 3 $^2/_2 : {}^1/_2$	×	×	×	×
1 : 2 $^7/_8 : {}^1/_8$	×	×	A	[1]
1 : 2 $^3/_4 : {}^1/_4$	×	×	A	
1 : 2 $^1/_2 : {}^1/_2$	×	×	×	×
1 : 1 $^3/_8 : {}^1/_8$	×	×	×	×
1 : 1 $^1/_4 : {}^1/_4$	×	×	×	×

Tabulka 2. Výsledky pro příklad s homogenní Dirichletovu podmínkou podél celé Γ_U.

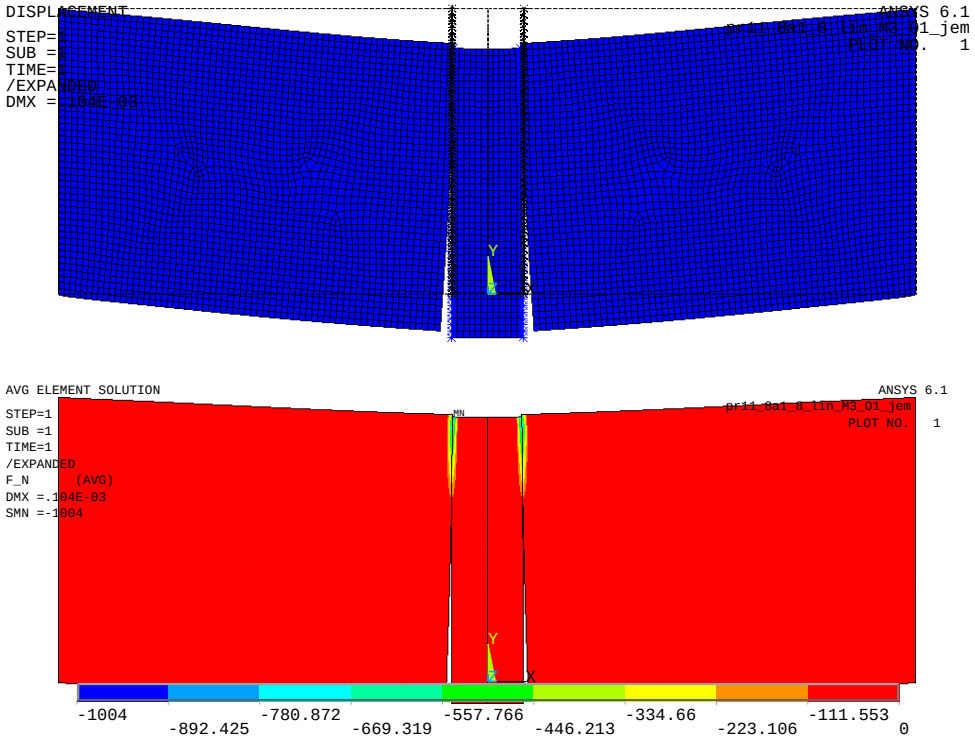
Okraj. podm. typu 2	*lineární model*		*nelineární model*	
	hrubší síť	*jemnější síť*	*hrubší síť*	*jemnější síť*
1 : 3 $^{11}/_8 : {}^1/_8$	A	A	A	A
1 : 3 $^5/_4 : {}^1/_4$	[1]	[1]	[1]	[1]
1 : 3 $^2/_2 : {}^1/_2$	×	×	×	×
1 : 2 $^7/_8 : {}^1/_8$	A		A	
1 : 2 $^3/_4 : {}^1/_4$	×	×	×	×
1 : 2 $^1/_2 : {}^1/_2$	×	×	×	×
1 : 1 $^3/_8 : {}^1/_8$	×	×	×	×
1 : 1 $^1/_4 : {}^1/_4$	×	×	×	×

Tabulka 3. Výsledky pro příklad s podminkou $u_1 = 0$ podél celé Γ_U a homogenní Dirichletovou podmínkou ($u_1 = u_2 = 0$) ve spodním uzlu.

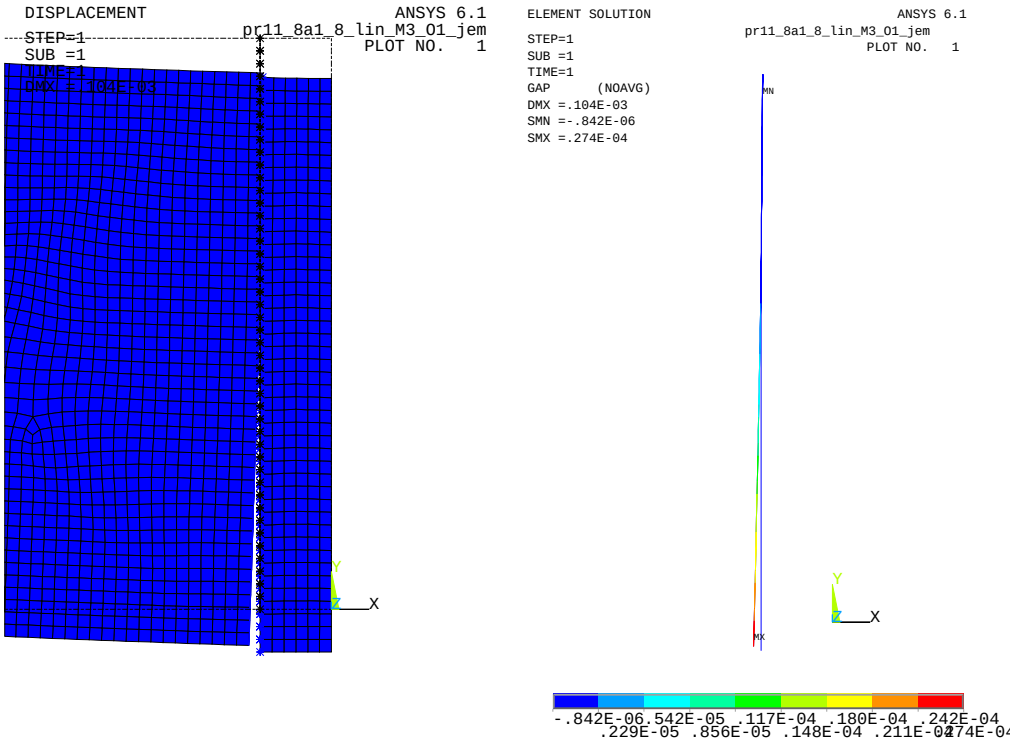
Okraj. podm. typu 3	*lineární model*		*nelineární model*	
	hrubší síť	*jemnější síť*	*hrubší síť*	*jemnější síť*
1 : 3 $^{11}/_{8} : {}^{1}/_{8}$	A	A	A	A
1 : 3 $^{5}/_{4} : {}^{1}/_{4}$	A	A	A	
1 : 3 $^{2}/_{2} : {}^{1}/_{2}$	A	A	[1]	[1]
1 : 2 $^{7}/_{8} : {}^{1}/_{8}$	A		A	
1 : 2 $^{3}/_{4} : {}^{1}/_{4}$	A	×	×	×
1 : 2 $^{1}/_{2} : {}^{1}/_{2}$	×	×	×	×
1 : 1 $^{3}/_{8} : {}^{1}/_{8}$	×	×	[1]	×
1 : 1 $^{1}/_{4} : {}^{1}/_{4}$	×	×	×	×

Tabulka 4. Výsledky pro příklad s homogenní Dirichletovou podmínkou ve spodním uzlu.

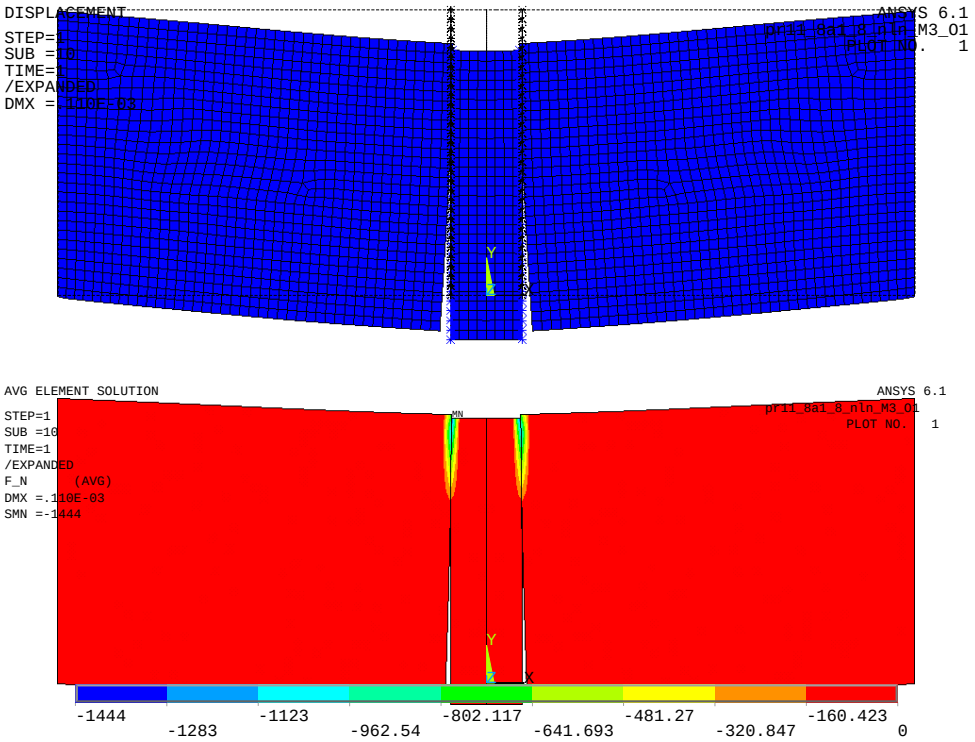
4.2 Výsledná vyobrazení pro zjednodušený model bez výplně



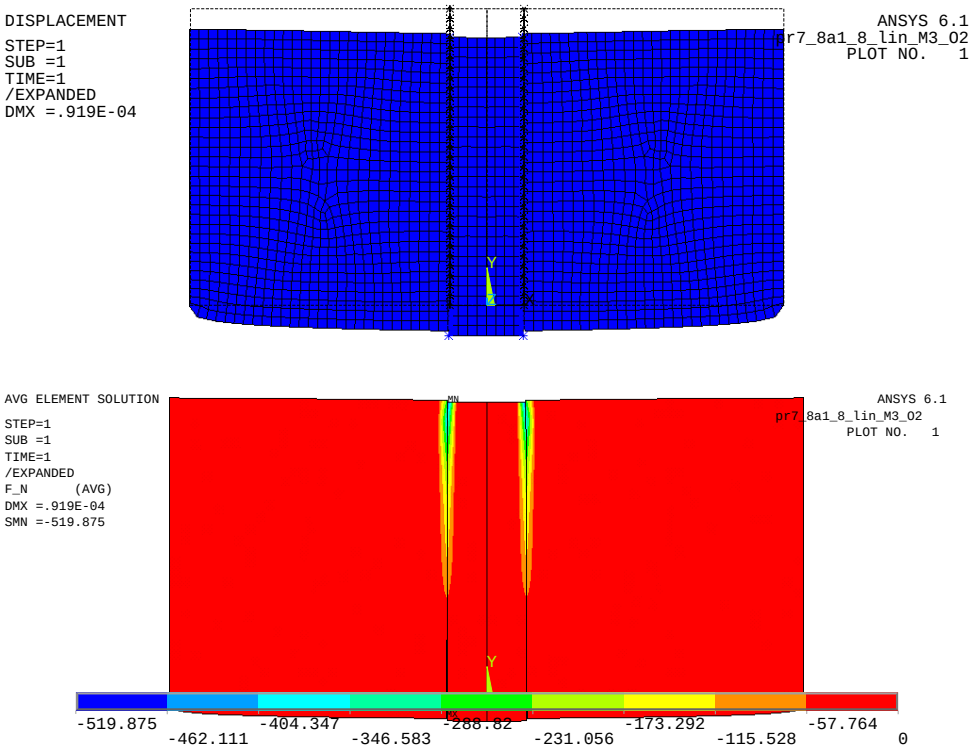
Obrázek 6. Výsledek příkladu s *prvním typem okrajových podmínek* a s poměrem velikostí 1:3 $^{11}/_{8} : {}^{1}/_{8}$ s *lineárním tenzorem deformace*; vykresleny jsou oblasti po deformaci a normálová napětí v okolí kontaktu.



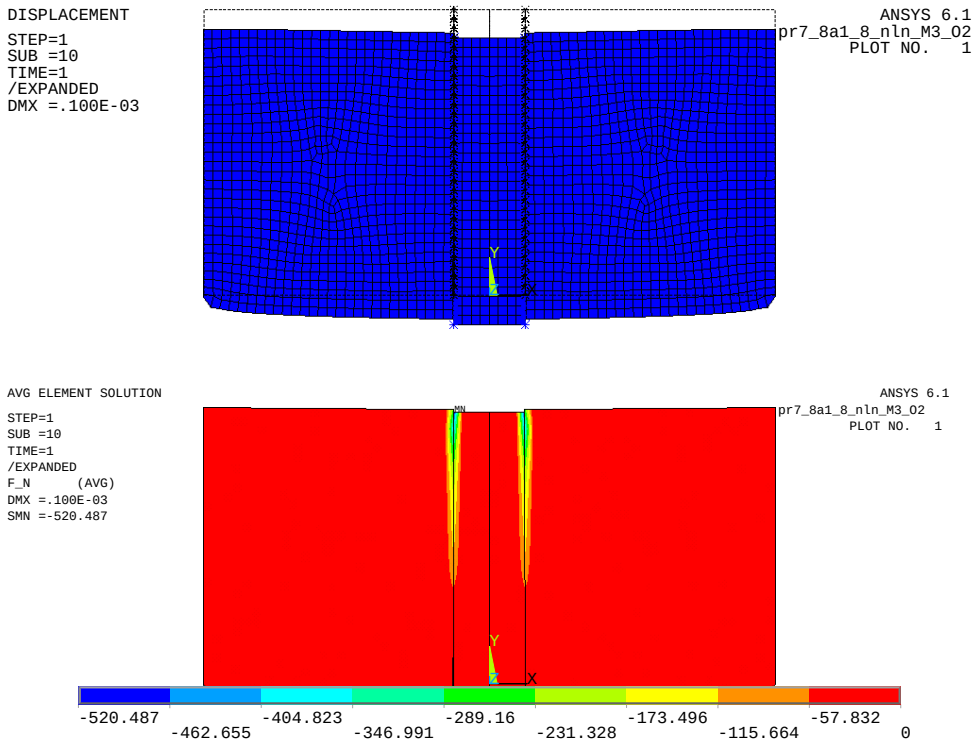
Obrázek 7. Výsledek příkladu s *prvním typem okrajových podmínek* a s poměrem velikostí 1:3 $^{11}/_{8} : {}^{1}/_{8}$ s *lineárním tenzorem deformace*; vykreslena je část řešené oblasti (půlka modelu) a zvlášť vynesen průběh otvírání na kontaktu.



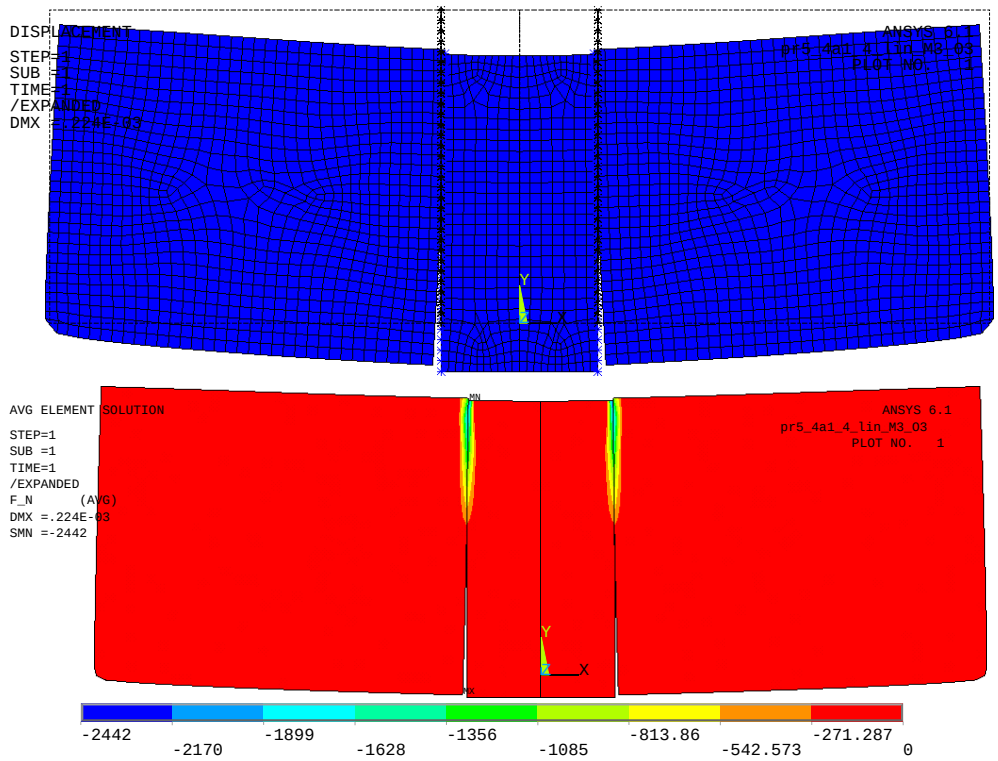
Obrázek 8. Výsledek příkladu s *prvním typem okrajových podmínek* a s poměrem velikostí 1:3 $^{11}/_{8}$: $^{1}/_{8}$ s *nelineárním tenzorem deformace*; vykresleny jsou oblasti po deformaci a normálová napětí v okolí kontaktu.



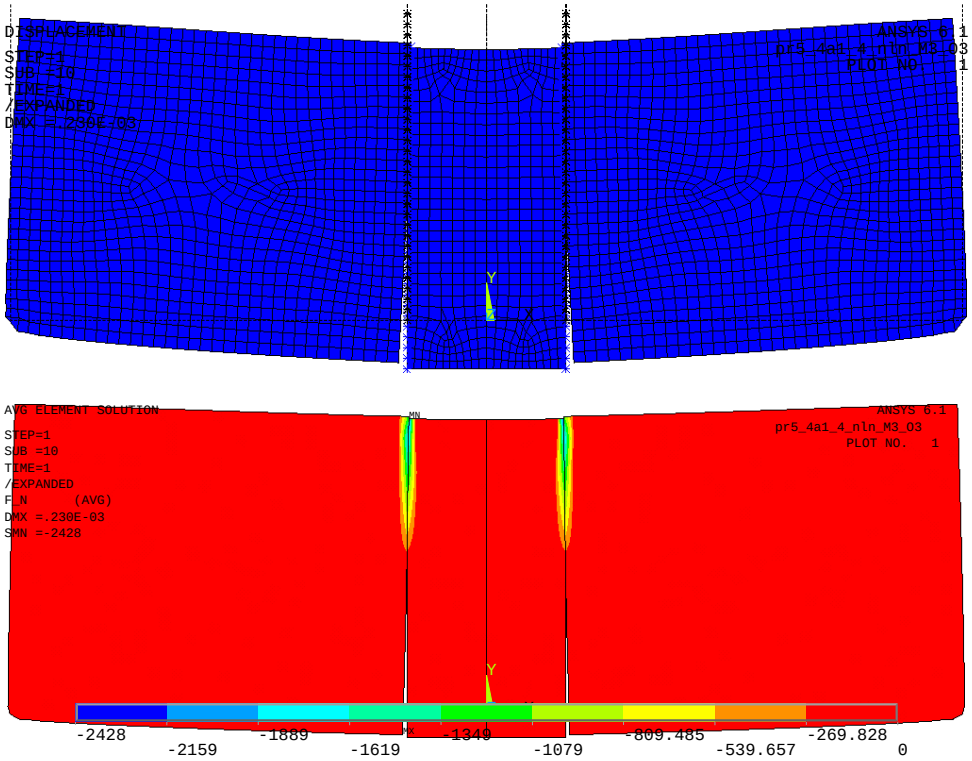
Obrázek 9. Výsledek příkladu s *druhým typem okrajových podmínek* a s poměrem velikostí 1:2 $^{7}/_{8}$: $^{1}/_{8}$ s *lineárním tenzorem deformace*; vykresleny jsou oblasti po deformaci a normálová napětí v okolí kontaktu.



Obrázek 10. Výsledek příkladu s *druhým typem okrajových podmínek* a s poměrem velikostí 1:2 $^{7}/_{8}$: $^{1}/_{8}$ s *nelineárním tenzorem deformace*; vykresleny jsou oblasti po deformaci a normálová napětí v okolí kontaktu.

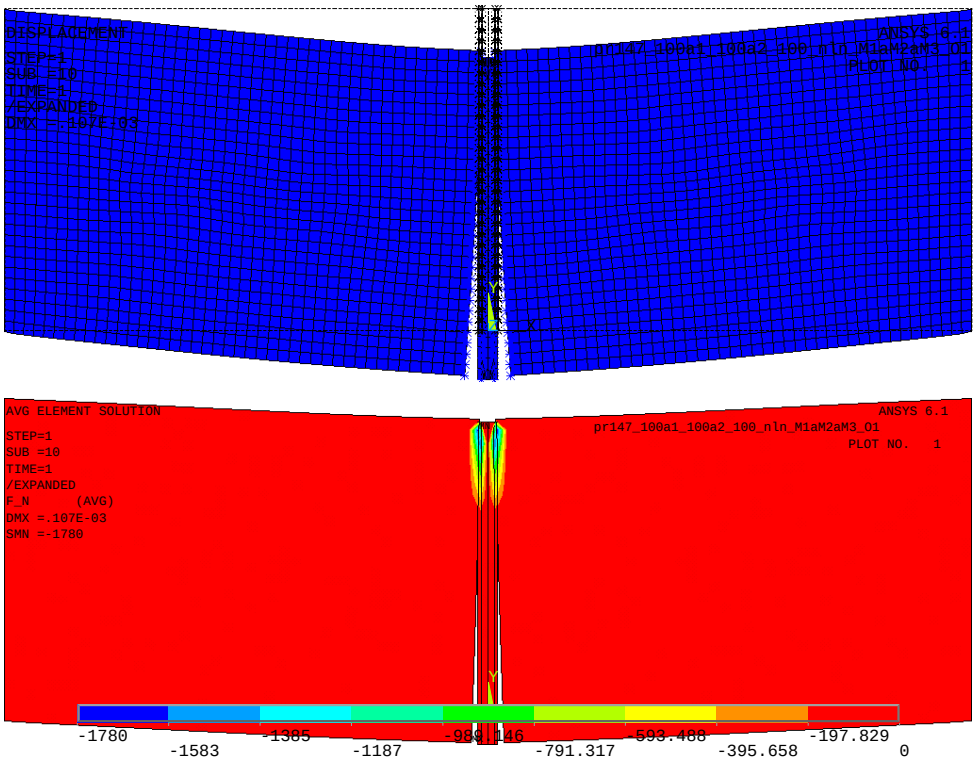


Obrázek 11. Výsledek příkladu s *třetím typem okrajových podmínek* a s poměrem velikostí 1:3 $^{5}/_{4} : {}^{1}/_{4}$ s *lineárním tenzorem deformace*; vykresleny jsou oblasti po deformaci a normálová napětí v okolí kontaktu.

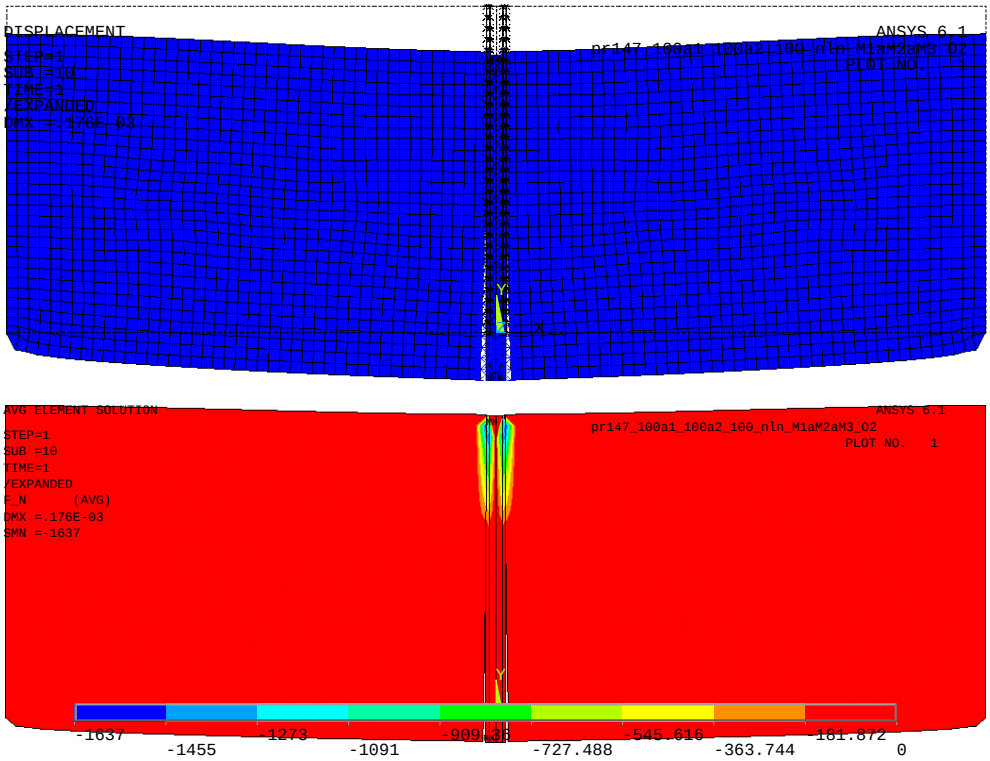


Obrázek 12. Výsledek příkladu s *třetím typem okrajových podmínek* a s poměrem velikostí 1:3 $^{5}/_{4} : {}^{1}/_{4}$ s *nelineárním tenzorem deformace*; vykresleny jsou oblasti po deformaci a normálová napětí v okolí kontaktu.

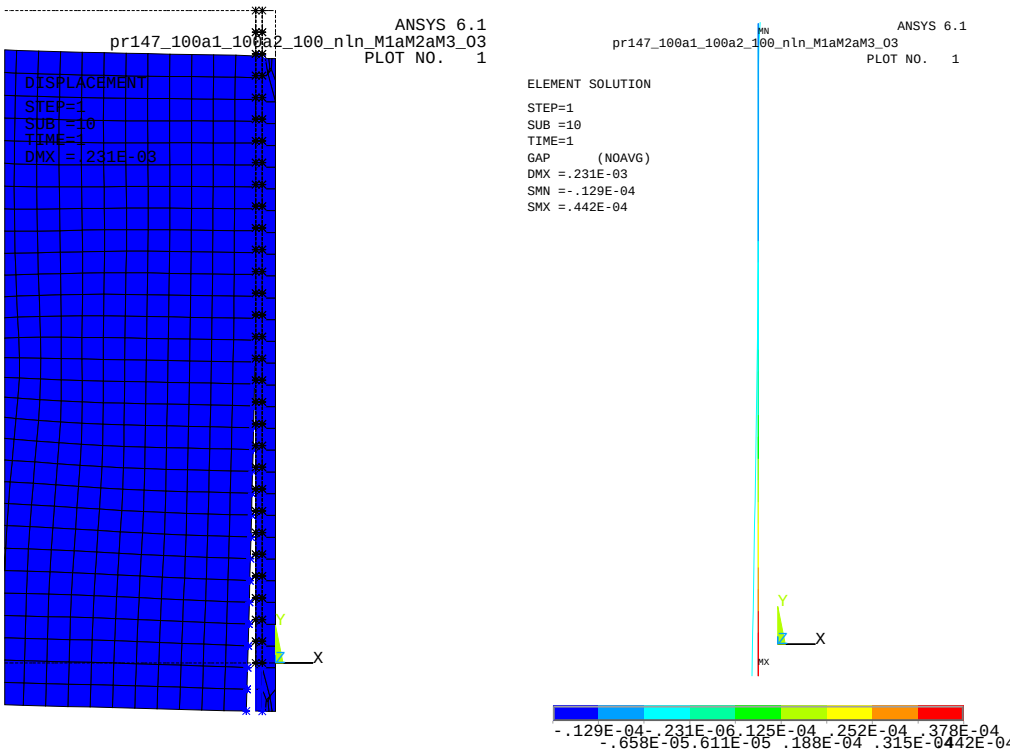
4.3 Výsledná vyobrazení pro model kotvy



Obrázek 13. Výsledek příkladu pro model kotvy (3 odlasti) s *nelineárním tenzorem deformace* pro *okrajovou podmínku typu 1*; vykresleny jsou oblasti po deformaci a normálová napětí v okolí kontaktu.



Obrázek 14. Výsledek příkladu pro model kotvy (3 odlasti) s *nelineárním tenzorem deformace* pro *okrajovou podmínku typu 2*; vykresleny jsou oblasti po deformaci a normálová napětí v okolí kontaktu.



Obrázek 15. Výsledek příkladu pro model kotvy (3 odlasti) s *nelineárním tenzorem deformace* pro *okrajovou podmínku typu 3*; vykreslena je část řešené oblasti (půlka modelu) a zvlášť vynesen průběh otvírání na kontaktu.

5 Shrnutí

Tento text je prvním krokem k hledání a numerickému ověřování podmínek řešitelnosti interaktivní soustavy pružných těles, která jsou ve vzájemném kontaktu s uvažovaným vlivem tření na kontaktních plochách.

Zvolili jsme model, zakreslený na obrázku 2. na straně 6, zjednodušující příklad lokální výztuže jedinou kotvou, který je zachycen na obrázku 1. na straně 5. Uvádíme výsledky numerického testování citlivosti navrženého modelu na změně předem vybraných předpokladů.

Z výsledků je patrný vliv především zvolené geometrie a typu okrajové podmínky, kterou jsme se snažili popsat chování na rozhraní mezi efektivním okolím kotvy a okolní horninou. Změna předpokladu lineárního chování materiálu na nelineární se efektivnost modelu výrazně nezměnila; nejvíce snad u prvního typu okrajové podmínky, tabulka č.2 strana 20. Ještě menší vliv na sledovanou stabilnost návrhu měla volba hustoty diskretizace.

Z výsledných obrázků na stranách 22 až 30 je také zřejmé, že při použití jak lineární tak nelineární teorie dochází k „otvírání“ spáry v dolní části převážně s tahovým napětím v blízkosti kontaktních ploch a tedy v případě kotevního systému k možnému snížení účinnosti kotvy. K nalezení uspokojivého návrhu výpočtového modelu bude tedy třeba se nadále zabývat hledáním podmínek řešitelnosti pro spojitý i diskrétní model odpovídající nelineární teorii.

Poděkování. Autoři článku děkují doktoru H. Netukovi z katedry matematické analýzy a aplikací matematiky přírodovědecké fakulty Univerzity Palackého v Olomouci za poskytnuté konzultace a podnětné připomínky v průběhu práce a dále děkují Radě vlády ČR pro výzkum a vývoj za finanční podporu výzkumné činnosti.

Reference

[Gurtin] Gurtin, M. E.: *An Introduction to Continuum Mechanics.* Academic Press, 1981.

[Nečas, Hlaváček] Nečas, J., Hlaváček, I.: *Mathematical Theory of Elastic and Elastico-Plastic Bodies: An Introduction.* Elsevier Scientific Publishing Company, 1981.

[Glowinski, Tallec] Glowinski, R., Le Tallec, P.: *Augmented Lagrangian and Operator–Splitting Methods.* SIAM, 1989.

[ANSYS Contact] *ANSYS 6.1 Documantation, Structural Analysys Guide, Contact, Performing Node-to-Surface Contact Analysys.*

[ANSYS Elemet] *ANSYS 6.1 Documantation, ANSYS Element Reference.*

[ANSYS Nonlinear] *ANSYS 6.1 Documantation, Nonlinear Structural Analysys.*

[ANSYS Static] *ANSYS 6.1 Documantation, Structural Static Analysys.*


Univ. Palacki. Olomuc., Fac. rer. nat.,
Dept of Math. Anal. and Appl. of Math.
ODAM (2003) 33–55

Quadratic Polynomials and Splines Interpolating 1D Mean Values on Simplest Triangulations*

Jiří KOBZA

Department of Mathematical Analysis and Applications of Mathematics,
Faculty of Science, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kobza@inf.upol.cz

Abstract

The problem of the one-dimensional mean value interpolation with quadratic splines on some simple triangulations is discussed. The function values in the vertices, edge midpoints and one-dimensional mean values are used for spline local representations and in continuity conditions. With prescribed mean values for all edges an interpolant exists in very special cases only (as consequence of Euler's rule). We have to search for the set of edges with unique solution of such problem (the Lagrange Interpolation Set) or for some mixed set of such local parameters—the space dimension is not equal in all cases to the number of 1D mean values we can prescribe.

1 Statement of the problem

In the FEM and FVM theories, in interpolation and optimal recovery theories we can find many results concerning interpolation of function values or mean values on triangulated domains in 2D (see [1]–[15]) with higher degree polynomials and

*Supported by the Council of Czech Government, J 14/153100011.

another special types of functions (radial basis functions, thin plate splines). We will discuss the problems connected with 1D mean values interpolation (MVI) on simple triangulations with quadratic C^1-splines. The author's results for function values interpolation with quadratic splines on such triangulations can be found in [6], for 2D MVI in [8]; more details about interpolation with quadratic polynomials in [7].

We can find in the literature the technique of the BB control set (see [4]) and technique using function values and gradients (see [5],[14],[15]) for the analysis of C^1-continuity conditions. We will use function values (FV) in the vertices and edge midpoints and 1D mean values along edges for spline local representation and for writing C^1-continuity conditions (CC) for triangles with common edge or vertice. The aim of this contribution is also to compare the results for MVI, FVI with quadratic splines and polynomials and also to present algorithms for computing local parameters of resulting splines.

Let us have the regular triangulation $\Delta = \bigcup T_i$ of some polygonal domain D into T triangles with V vertices and E edges. The Euler's rule states that $V - E + T = 1, \quad V : T : E \sim 1 : 2 : 3$ for large triangulations. The quadratic C^1-spline on Δ is the function, which is a quadratic polynomial (with the total degree used) over each triangle and with continuous first derivatives over the whole domain D. When for each edge V_iV_j with the length $|V_iV_j|$ (or for some subset of edges only) we know the one-dimensional mean value $m_{ij} \in R$, we can state the following

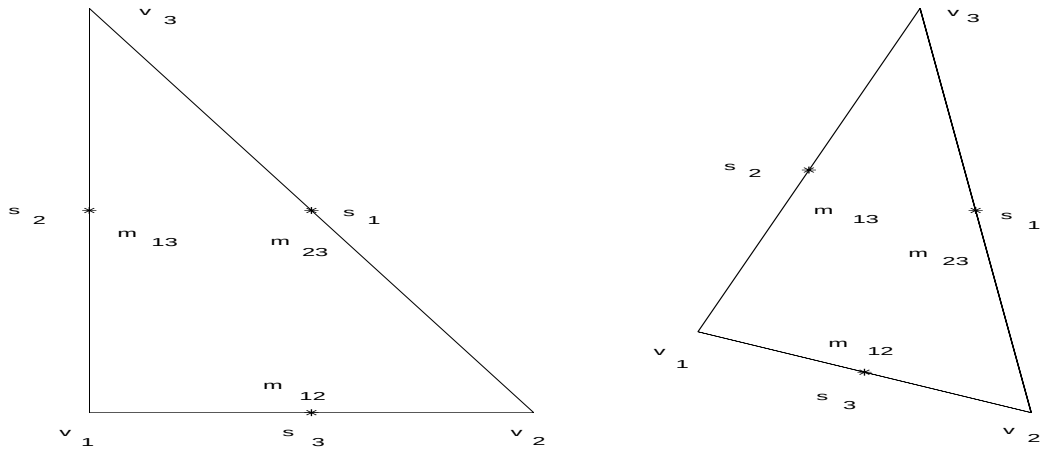
Problem 1D MVI: *With given 1D mean values*

$$m_{ij} = \frac{1}{|V_iV_j|} \int_{V_iV_j} s(x,y)\, dl \quad \textit{along some edges } V_iV_j, \qquad (CI)$$

does there exist a quadratic spline or polynomial with such mean values?

Problem LIS: *Find a subset (all such subsets) of edges (or edges, vertices, edge midpoints) on given triangulation (the Lagrange Interpolation Set) for which the 1D MVI problem (or problem with given mean and function values) has the unique solution with any given 1D mean (function) values on such subset.*

In the following we will discuss solvability of such problems using notation T_i, V_j, S_k for triangles, vertices and edge midpoints, $s(x,y),\ v_j = s(V_j),\ s_i = s(S_i)$ for the quadratic spline function values (see Figs 1a, 1b—the numbering will be precised in special cases). As a consequence of the mean value theorem of the integral calculus (m_{ij} is the function value in some point on the edge) the results give an answer also to some problems of function values interpolation on triangulations (when sample points are not specified or chaotic).



Figs 1a, 1b

2 Spline local representations on triangle

The quadratic spline over any triangle is a quadratic polynomial and we can describe it using some local polynomial coefficients or various local parameters as v_i, s_k, m_{ij}. There are simple connections between different spline local parameters which help us to use parameters we are interested in. The quadrature formula valid for any edge of triangle T with vertices V_j and local parameters s_i, v_j, $i, j = 1 : 3$ as in Figs 1a, 1b

$$m_{13} = \frac{1}{|V_1V_3|}\int_{V_1V_3} s(x,y)\,dl = \frac{1}{6}(v_1 + 4s_2 + v_3)$$

is exact for quadratic functions and allows us to write down the relation between mean values m_{ij} and function values s_i, v_j in any triangle T as e.g.

$$v_1 + 4s_3 + v_2 = 6m_{13}. \tag{1}$$

The values s_i we can compute from values m_{ij}, v_j as

$$4s_1 = 6m_{23} - v_2 - v_3,\ 4s_2 = 6m_{13} - v_1 - v_3,\ 4s_3 = 6m_{12} - v_1 - v_2. \tag{2}$$

With given parameters m_{ij} there is one-to-one correspondence between parameters s_i, v_j.

2.1 Cartesian coordinates on reference triangle

A quadratic polynomial on reference triangle T_0 with vertices $V_1 = [0,0]$, $V_2 = [1,0]$, $V_3 = [0,1]$ and edge midpoints S_i (as used in FEM—see Fig. 1a for numbering) we can write in *Taylor's representation*

$$s(x,y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 \tag{3}$$

Six coefficients a_{ij} can be uniquelly determined from six conditions of interpolation (CI) with local parameters v_i, m_{ij} and we obtain local representation

$$\begin{aligned} s(x,y) = {} & [1-4(x+y)+3(x+y)^2]v_1 + x(3x-2)v_2 + y(3y-2)v_3 \\ & + 6[xym_{23} + (1-x-y)(ym_{13}+xm_{12})]. \end{aligned} \tag{4}$$

For the triangle in a general position the coefficients a_{ij} can be also computed from CI and used in the corresponding local representation.

In the FEM interpolation theory (see e.g. [15]) the transformation of any triangle to some reference triangle is used. We shall try to use local spline parameters only without such transformations. *The dimension of the space of quadratic polynomials is six in all cases considered.*

2.2 Spline representations in barycentric coordinates

The barycentric coordinates $\mathbf{t} = [t_1, t_2, t_3]$ of the point $P = [x, y]$ with respect to the triangle $V_1V_2V_3$ in general position allow us to write the quadratic polynomial with local parameters v_i, s_i (see [4]) in (**v,s**)-representation

$$s(\mathbf{t}) = t_1(2t_1-1)v_1 + t_2(2t_2-1)v_2 + t_3(2t_3-1)v_3 + 4(t_1t_2s_3 + t_1t_3s_2 + t_2t_3s_1) \tag{5}$$

or—after substitution from (4)—with parameters v_i, m_{ij} in (**v,m**)-representation

$$s(\mathbf{t}) = t_1(3t_1-2)v_1 + t_2(3t_2-2)v_2 + t_3(3t_3-2)v_3 + 6(t_1t_2m_{12} + t_1t_3m_{13} + t_2t_3m_{23}) \tag{6}$$

Substitution for v_i from (7) gives us the spline in (**s,m**)-representation

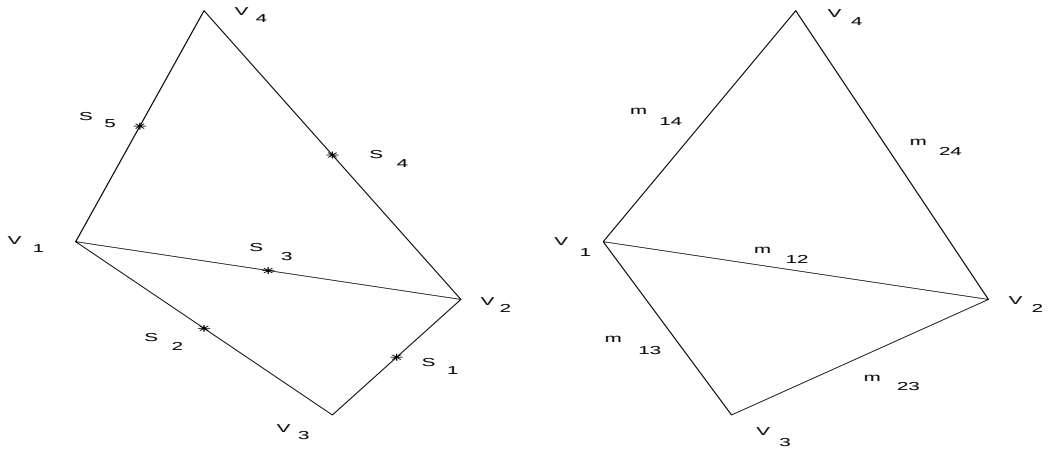
$$\begin{aligned} s(\mathbf{t}) = {} & 3(1-2t_3-4t_1t_2)m_{12} + 3(1-2t_2-4t_1t_3)m_{13} + 3(1-2t_1-4t_2t_3)m_{23} \\ & + 2(-1+2t_1+6t_2t_3)s_1 + 2(-1+2t_2+6t_1t_3)s_2 + 2(-1+2t_3+6t_1t_2)s_3 \end{aligned} \tag{7}$$

with local parameters m_{ij}, s_i. Let us mention that the segment of the boundary curve depends now on boundary mean values on another segments (we have only two parameters on such a segment).

3 Interpolation and continuity conditions

Let us consider the 1D MVI problem for two triangles $V_1V_2V_3$, $V_1V_2V_4$ with common edge V_1V_2 and with barycentric coordinates $\mathbf{t} = [t_1, t_2, t_3]$ of V_4 with respect to the triangle $V_1V_2V_3$ (see Fig. 2 for notation). Let the five 1D MV m_{ij} are known—we search for MVI quadratic polynomial and quadratic spline.

In the *MVI problem with quadratic polynomial* we have to find the corresponding six parameters only (in the LR chosen) from values prescribed in CI. When we use (**v,m**)-LR and compute the FV s_4, s_5, v_4 as functions of $v_1, v_2, v_3, m_{12}, m_{13}, m_{23}$



Figs 2a, 2b

and when we use the CI for m_{14}, m_{24}, we obtain three equations relating nine parameters v_j, m_{ij}. The (3,4)-matrix related to parameters v_j is

$$\begin{bmatrix} t_1(3t_1-2) & t_2(3t_2-2) & t_3(3t_3-2) & -1 \\ t_1(3t_1+2) & t_2(3t_2-4) & t_3(3t_3-4) & 1 \\ t_1(3t_1-4) & t_2(3t_2+2) & t_3(3t_3-4) & 1 \end{bmatrix}. \tag{8}$$

The four determinants corresponding to the free parameters v_j are equal $\pm 36t_1t_2t_3$, $36t_1(1-t_3)t_3$, $36t_2(t_3-1)t_3$. From the position of V_4 ($t_j \neq 0, t_3 < 0$) follows now that they are all different from zero in case $t_2 \neq 0$.
When we use (**s,m**)-LR for computing s_4, s_5, v_4 and substitute the results into CI for m_{14}, m_{15}, we obtain two relations for parameters s_1, s_2, s_3 and m_{ij} with simple matrix [1,-1,-1; -1,1,1] corresponding to columns for s_i. We can summarize the results in the following

Statement 1: *1D MVI problem for two triangles with the common edge and five given m_{ij} (see Fig.2 for the notation) in the class of quadratic polynomials (with the dimension equal to six)*
– has the unique solution with one free parameter v_j, $j \in \{1:4\}$ if $t_2 \neq 0$;
– has the unique solution with $t_2 = 0$ and free parameters $v_2 + 5m_{ij}$;
– has the unique solution with one free parameter $s_i, i \neq 3$ (FV in midpoint of the common edge).

To solve our problem *with quadratic splines* we have to find the spline local parameters for each patch of the spline under C^1-continuity (smoothing) conditions (CC) and CI. We can find e.g. in [3],[5],[9] the C^1-smoothing conditions for the common edge expressed with local parameters FV and gradients in the vertices. We will show that we can use also for this purpose the FV in the vertices and edge midpoints only or FV in the vertices and MV along the edges. The dimension of the spline space will be growing up with the number of triangles considered.

The C^1-continuity conditions can be expressed in various local parameters used

for the spline representation. With common FV v_1, v_2, s_3 (see Fig. 2a for numbering) these patches have common boundary curve over the connecting edge. The continuity conditions for the first derivatives along this edge can be expressed as common directional derivatives along edges V_1V_4, V_2V_4 (see [6]–[8]) with the local parameters $\mathbf{s} = [s_1, s_2, s_3, s_4, s_5]^T$, $\mathbf{v} = [v_1, v_2, v_3, v_4]^T$ and (8) as

$$4\begin{bmatrix} 0 & t_3 & t_2 & 0 & -1 \\ t_3 & 0 & t_1 & -1 & 0 \end{bmatrix}\mathbf{s} = \begin{bmatrix} -3t_1 & t_2 & t_3 & -1 \\ t_1 & -3t_2 & t_3 & -1 \end{bmatrix}\mathbf{v}. \tag{9}$$

We can substitute for s_i from (2) to obtain the CC in (**v,m**)-LR or to use this LR directly—we obtain then the CC conditions with the local parameters $\mathbf{v} = [v_1, v_2, v_3, v_4]$, $\mathbf{m} = [m_{12}, m_{13}, m_{23}, m_{14}, m_{24}]^T$ as

$$\begin{bmatrix} 2t_1 & -t_2 & -t_3 & 1 & 3t_2 & 3t_3 & 0 & -3 & 0 \\ -t_1 & 2t_2 & -t_3 & 1 & 3t_1 & 0 & 3t_3 & 0 & -3 \end{bmatrix}\begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0} \tag{10}$$

We can use the CC for solving 1D MVI problem with given all five values m_{ij} and two parameters v_j. The only variant with no solution in general is now with given v_1, v_2.

We can try to obtain the CC expressed also with the local parameters s_k, m_{ij}. When we compare the expressions for the derivatives along edges V_1V_4, V_2V_4 using local representation $(\mathbf{s}, \mathbf{m})$, we obtain the C^1-continuity conditions

$$\begin{aligned} 2[(t_1-1)s_1 + (t_2+3t_3)s_2 + (3t_2+t_3-1)s_3 + s_4 - 3s_5] = \\ = 3[(2t_2+t_3-1)m_{12} + (t_2+2t_3)m_{13} - 2m_{14} + (t_1-1)m_{23} + m_{24}], \\ 2[(t_1+3t_3)s_1 + (t_2-1)s_2 + (3t_1+t_3-1)s_3 - 3s_4 + s_5] = \\ = 3[(2t_1+t_3-1)]m_{12} + (t_2-1)m_{13} + m_{14} + (t_1+2t_3)m_{23} - 2m_{24}]. \end{aligned} \tag{11}$$

But the local representation (**s,m**) does not ensure us now the C-continuity of the neighbouring patches! The condition of equal function values in both common vertices V_1, V_2 gives us one complementary condition

$$2(s_1 - s_2 - s_4 + s_5) = 3(-m_{13} + m_{14} + m_{23} - m_{24}), \tag{12}$$

which we have to add to foregoing C^1-conditions.

With vectors $\mathbf{s} = [s_i,\ i = 1:5]$, $\mathbf{m} = [m_{12}, m_{13}, m_{14}, m_{23}, m_{24}]^T$ and matrices

$$\mathbf{A_s} = \begin{bmatrix} t_1-1 & t_2+3t_3 & 3t_2+t_3-1 & 1 & -3 \\ t_1+3t_3 & t_2-1 & 3t_1+t_3-1 & -3 & 1 \\ 1 & -1 & 0 & -1 & 1 \end{bmatrix},$$

$$\mathbf{A_m} = \begin{bmatrix} 2t_2+t_3-1 & t_2+2t_3 & -2 & t_1-1 & 1 \\ 2t_1+t_3-1 & t_2-1 & 1 & t_1+2t_3 & -2 \\ 0 & -1 & 1 & 1 & -1 \end{bmatrix}$$

we can write the C^1 conditions in this LR as

$$2\mathbf{A_s s} = 3\mathbf{A_m m}. \tag{13}$$

Through detailed discussion of determinants from columns of the matrix $\mathbf{A_s}$ we can search for the solvability conditions.

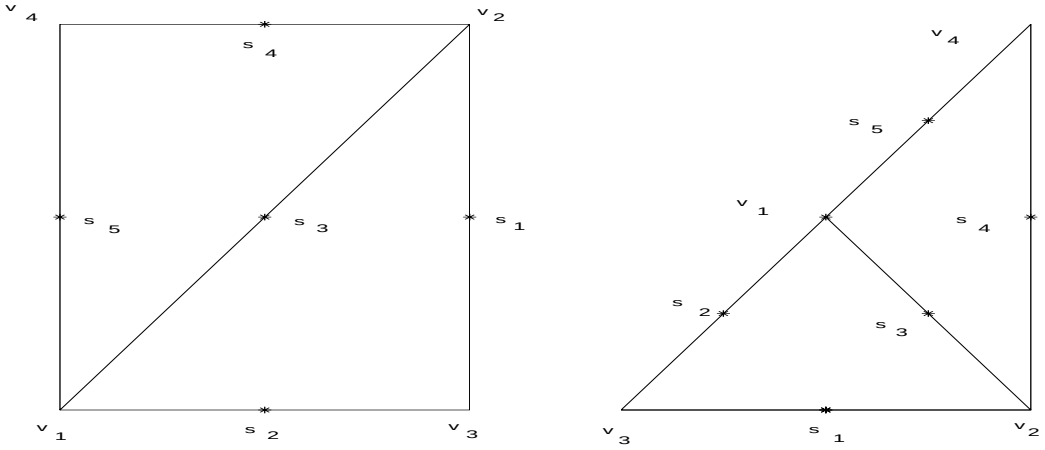
The results of both cases of LP v_j, s_i we can summarize in the following

Statement 2: *The dimension of solutions of 1D MVI problem for two triangles with one common edge and with given five values m_{ij} in the space of quadratic splines is seven.*

Such a problem has a unique solution with any two free parameters v_j with the exception of the couple v_1, v_2.

With two free parameters s_i there is a unique solution with the exceptions of cases

– *s_1, s_4, when the vertex V_4 is located on the line V_2V_3;*
– *s_4, s_5 and vertex V_4 on the line V_1V_2 or on the axis of the angle $V_1V_3V_2$;*
– *s_1, s_2 and vertex V_4 on the axis of $\angle V_1V_3V_2$ and in symmetric cases.*



Figs 3a, 3b

3.1 Special cases of CC + CI

We can frequently meet the special cases of triangulations, where some symmetry results in repeating simple coefficients in CC. We find here some part of results proved for the general cases in the foregoing subsection.

3.1.1 Quadrilateral

In case that *two connected triangles form a quadrilateral* (see Fig. 3a), we have $V_4 = V_1 + V_2 - V_3$, with barycentric coordinates $\mathbf{t} = [1, 1, -1]$.

When we use the (**s,v**)-LR to express the values s_4, s_5, v_4 as FV of *one quadratic polynomial* and eliminate the parameters s_i from five CI for m_{ij}, we obtain for vectors $\mathbf{v} = [v_1, v_2, v_3, v_4]$, $\mathbf{m} = [m_{12}, m_{13}, m_{14}, m_{23}, m_{24}]$ the relations

$$\begin{bmatrix} -1 & 5 & 7 & 1 & 12 & -6 & 0 & -12 & -6 \\ 5 & -1 & 7 & 1 & -12 & -6 & -6 & 0 & \\ 1 & 1 & 5 & -1 & 6 & -6 & 0 & -6 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0}. \tag{14}$$

All four (3,3)-determinants from first 4 columns are nonzero—*the 1D MVI problem with prescribed 5 MV m_{ij} + any v_j has a unique solution in the class of quadratic polynomials.*

For *quadratic splines* the CC + CI conditions with parameters (**v,s**) are

$$\begin{bmatrix} 3 & -1 & 1 & 1 & 0 & -4 & 4 & 0 & -4 \\ -1 & 3 & 1 & 1 & -4 & 0 & 4 & -4 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 4 \\ 0 & 1 & 1 & 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 4 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{s} \end{bmatrix} = 6 \begin{bmatrix} 0 \\ 0 \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{23} \\ m_{24} \end{bmatrix}. \tag{15}$$

We have 7 equations for $4 + 5 = 9$ parameters; with 2 free parameters and 5 MV m_{ij} we have confirmed again that the dimension is equal to seven.

With the local parameters $(\mathbf{v}, \mathbf{m})$ we can write more simply the corresponding CC as

$$\begin{bmatrix} 2 & -1 & 1 & 1 & 3 & -3 & 0 & -3 & 0 \\ -1 & 2 & 1 & 1 & 3 & 0 & -3 & 0 & -3 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0}. \tag{16}$$

We can now easily to see that *the only case with no solution with prescribed 5 m_{ij} + 2 v_j is with free parameters v_1, v_2.*

In the (**s,m**)-representation we can write the CC now as

$$2 \begin{bmatrix} 0 & -2 & 1 & 1 & -3 \\ -2 & 0 & 1 & -3 & 1 \\ 1 & -1 & 0 & -1 & 1 \end{bmatrix} \mathbf{s} = 3 \begin{bmatrix} 0 & -1 & -2 & 0 & 1 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & -1 & 1 & 1 & -1 \end{bmatrix} \mathbf{m} \tag{17}$$

We find that the computed parameters s_i do not depend on the value m_{12} now and that *our MVI problem has no solution in general in two variants with free parameters—s_1, s_2 and s_4, s_5 only.*

3.1.2 Half of quadrilateral

When *two connected triangles form one half of the quadrilateral* (as in the Δ^2-triangulation of the rectangular mesh— see Fig. 3b and Sect. 6.1), then $V_4 = 2V_1 - V_3$ and $\mathbf{t} = [2, 0, -1]$. We can use the Statement 1 to find the unique MVI polynomial to the data $5m_{ij} + 1v_2$ only.

For MVI quadratic spline the system of equations for CC with the local parameters $(\mathbf{v}, \mathbf{m})$ has now simple form

$$\begin{bmatrix} 4 & 0 & 1 & 1 & 0 & -3 & 0 & -3 & 0 \\ -2 & 0 & 1 & 1 & 6 & 0 & -3 & 0 & -3 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0}. \tag{18}$$

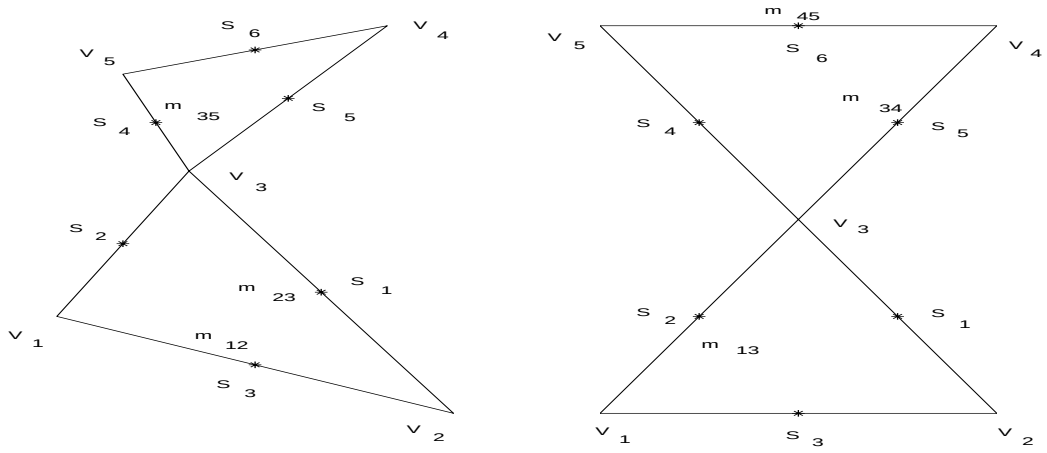
With given values m_{ij} *we find the solvabillity of our MVI problem in general with the free parameters* v_2, v_3 *and* v_2, v_4 *only.*

With parameters (**s,m**) the CC (13) can be written now as

$$\begin{bmatrix} 2 & -6 & -4 & 2 & -6 & 6 & 6 & 6 & -3 & -3 \\ -2 & -2 & 8 & -6 & 2 & -6 & 3 & -3 & 0 & 6 \\ 2 & -2 & 0 & -2 & 2 & 0 & 3 & -3 & -3 & 3 \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ \mathbf{m} \end{bmatrix} = \mathbf{0}. \tag{19}$$

We have now 3 equations with 5 parameters s_i and given m_{ij}—two free parameters s_i. We find that *from ten variants of two free parameters the only couple with no solution in general is* s_2, s_5.

3.2 Triangles with common vertex only



Figs 4a, 4b

Let us discuss the case of two triangles with the common one vertex only as in Figs 4a, 4b and 1D MV m_{ij} along all six edges.

With quadratic polynomial and (**v,m**)-LR we can compute the parameters v_4, v_5, s_4, s_5, s_6 and substitute for s_4, s_5, s_6 in the CI for m_{34}, m_{35}, m_{45}.

We obtain so five equations relating eleven parameters $\mathbf{v} = [v_1, v_2, v_3, v_4, v_5]$, $\mathbf{m} = [m_{12}, m_{13}, m_{23}, m_{34}, m_{35}, m_{45}]$. We find the unique solution with given six values m_{ij} in case of nonzero determinant (general case). We find e.g. with $t_1^4 = -1$, $t_2^5 = -1$ and free parameters t_2^4, t_1^5 the determinant to be equal $216(-4t_2^4 - 4t_1^5 + 4t_2^4 t_1^5 + 5(t_2^4)^2 t_1^5 + 5t_2^4 (t_1^5)^2 - 5(t_2^4)^2 (t_1^5)^2 - (t_2^4)^3 (t_1^5)^2 - (t_2^4)^2 (t_1^5)^3 + (t_2^4)^3 (t_1^5)^3$

and we can find the unique solution of our problem e.g. with $\mathbf{t^4} = [-1, 3, -1]$, $\mathbf{t^5} = [-2, -1, 4]$.

In the special case with $V_4 = [-1, 0, 2]$, $V_5 = [0, -1, 2]$ (V_3 is center of edges V_1V_4, V_2V_5—as in Fig. 4b and frequently used triangulations) we obtain the system of relations

$$\begin{bmatrix} 5 & 0 & 8 & -1 & 0 & 0 & -12 & 0 & 0 & 0 & 0 \\ 0 & 5 & 8 & 0 & -1 & 0 & 0 & -12 & 0 & 0 & 0 \\ 0 & 7 & 16 & 0 & 1 & 0 & 0 & -18 & 0 & -6 & 0 \\ 7 & 0 & 16 & 1 & 0 & 0 & -18 & 0 & -6 & 0 & 0 \\ 7 & 7 & 32 & 1 & 1 & 6 & -24 & -24 & 0 & 0 & -6 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0}.$$

We find the determinant with columns corresponding to parameters v_j equal to zero—*the problem with 1D MV prescribed for all six edges has not a unique solution with quadratic polynomial in general.* The problem has a solution under condition $m_{12} - m_{13} - m_{23} = m_{45} - m_{34} - m_{35}$ only (mistake in [7]), or with given 5 values m_{ij} and one from parameters s_i, v_j.

Let us discuss *the solution of such problem with quadratic splines.* With vectors $\mathbf{v} = [v_i,\ i = 1 : 5]$, $\mathbf{s} = [s_1, s_2, s_4, s_5]$ and barycentric coordinates $\mathbf{t}^4 = [t_1^4, t_2^4, t_3^4]$, $\mathbf{t}^5 = [t_1^5, t_2^5, t_3^5]$ of vertices V_4, V_5 related to the triangle $V_1V_2V_3$ we can write the spline CC as

$$\begin{bmatrix} -t_1^4 & -t_2^4 & 3t_3^4 & 1 & 0 \\ -t_1^5 & -t_2^5 & 3t_3^5 & 0 & 1 \end{bmatrix} \mathbf{v} + 4 \begin{bmatrix} t_2^4 & t_1^4 & 0 & -1 \\ t_2^5 & t_1^5 & -1 & 0 \end{bmatrix} \mathbf{s} = \mathbf{0}. \tag{20}$$

Let us mention that the CC are independent on parameters s_3, s_6—they appear in CI only.

The CC in special case with $\mathbf{t}^4 = [-1, 0, 2],\ \mathbf{t}^5 = [0, -1, 2]$ we can write as

$$\begin{bmatrix} 1 & 0 & 6 & 1 & 0 \\ 0 & 1 & 6 & 0 & 1 \end{bmatrix} \mathbf{v} = 4 \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \mathbf{s}. \tag{21}$$

When we add 6 CI with given 1D MV m_{ij}, we obtain the system with full rank (8,11)-matrix. With six m_{ij} we can choose three free parameters from s_i, v_j – *the dimension of the corresponding spline space is equal to nine !* The discussion of the solvability of that system gives the result that *we obtain the unique solution with each from all 20 triples of values s_i and with 8 triples of values v_j which are not located in the vertices on one line.*

In the (**v,m**)-LR we can write the CC with vectors $\mathbf{v} = [v_1, v_2, v_3, v_4, v_5]$, $\mathbf{m} = [m_{13}, m_{23}, m_{34}, m_{35}]$ (independence on m_{12}, m_{45}) in the general case as

$$\begin{bmatrix} -t_1^4 & -t_2^4 & 2t_3^4 & 1 & 0 \\ -t_1^5 & -t_2^5 & 2t_3^5 & 0 & 1 \end{bmatrix} \mathbf{v} + 3 \begin{bmatrix} t_1^4 & t_2^4 & -1 & 0 \\ t_1^5 & t_2^5 & 0 & -1 \end{bmatrix} \mathbf{m} = \mathbf{0}. \tag{22}$$

With given six m_{ij} we have again three free parameters v_j—*the dimension is equal to nine.* We can see that we do not obtain unique solution in cases when

prescribed parameters v_j correspond to three vertices which are located on one line.

Statement 3. *For two triangles with common vertice only and given six 1D MV m_{ij} there exist quadratic MVI splines. The dimension of such spline space is nine. We obtain the unique solution e.g. with triples of free parameters v_j for vertices not located on one line and for any triple of free parameters s_i.*

In all solvable cases mentioned above we can choose free parameters according to our needs (known FV or MV, boundary conditions) or to use them to optimization purposes - we usually minimize some functional or norm of local parameters. In case of minimization of 2-norms we can easily use pseudoinverse solution of CC+CI.

Remark: Quadratic spline—one quadratic polynomial?
The space of quadratic polynomials is a subspace of the space of quadratic splines with the dimension equal to six—so we can expect the existence of MVI quadratic polynomial for our problem with six CI. In the case of two triangles with the common edge we have found the dimension of the spline space equal to seven. To ensure that both quadratic patches belong to the same quadratic polynomial it is enough to add one CC—e.g. equal derivative in the direction V_3V_4 in the common edge midpoint S_3. Another possibility is to prescribe sixth CI—1D MV along V_3V_4 (see [1]).

3.3 Adding neighbouring triangular patch

Let us have given the quadratic patch over the triangle $V_1V_2V_3$ and the neighbouring triangle $V_1V_2V_4$ with 3 parameters v_4, m_{14}, m_{24}. *Can we find a quadratic MVI patch over triangle $V_1V_2V_4$ such that both patches form a quadratic spline?* Similar problem we can state for the case of triangles with common vertex only and 3 free parameters.

When we consider in the first problem the CC (10) with free parameter v_4, we can compute uniquely any from couples $[v_4, m_{14}], [v_4, m_{24}], [m_{14}, m_{24}]$ (nonzero determinants) – but we cannot compute v_4 only from given m_{14}, m_{24}. In the second case with CC (22) and parameters v_4, v_5, m_{14}, m_{24} we can compute uniquely both $[s_4, s_5]$ or $[m_{34}, m_{35}]$ from the remaining given seven parameters (solvability does not depend on free parameters m_{12}, m_{45}.)

Statement 4: *Given quadratic patch we can extend as quadratic C^1-spline to the neighbouring triangle*
– uniquelly in case of common vertex only from given values m_{34}, m_{35} (v_4, v_5 computed from CC);
– uniquelly in case of common edge and given value v_4 (m_{14}, m_{24} computed from CC); no solution is here with given m_{14}, m_{24} in general.

Remark: The above result gives an example of the set of triangles with prescribed 1D MV along edges for which exists MVI quadratic spline interpolant—

any set of triangles generated from the "mother triangle" with "blossoming" through the vertices without contact with the neighbours (the dimension is growing with each new triangle about one). More difficult situation is with the set of triangles generated from mother triangle with "flaps" along the vertices.

4 Simple triangular complexes

We can add to any couple of two triangles discussed above some another triangles with common edges or vertices and with given 1D MV along some vertices and to discuss the solvability conditions of the 1D MVI problem with quadratic polynomials or splines on such or more general complexes of triangles.

With quadratic polynomials we cannot expect solution for complexes with m_{ij} given in more than six edges in general. Yet for three triangles from Fig. 5a with 7 edges we obtain the solution e.g. with m_{12} computed from another six CI, but we cannot compute so m_{25} in general. The more general result for higher degree polynomials (depending on the number of vertices) can be found in [1].

Using quadratic splines we have also to obey the C^1-smoothness of the spline patches over such complexes. For simple triangulations with one interior vertex—called "cells" in [12]—we can deduce from the results published here the expression for the dimension of the corresponding spline space S_2^1 as

$$dim(S_2^1) = 3 + E_i + (3 - e)_+ \tag{23}$$

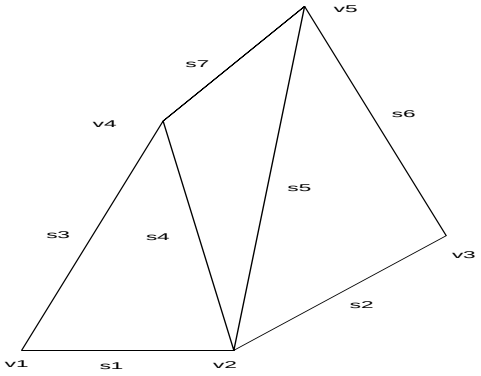
with E_i—the number of interior edges, e—the number of of interior edges with different slopes attached to interior vertex. For the triangulated regular polyhedrons H_n with $T = n$, $V = n + 1$, $E = 2n$ we can find then $\dim S_2^1(H_n) = n + 3$ for odd n, $\dim S_2^1(H_n) = n + 3 + (3 - k)_+$ for even $n = 2k$. So we have the result $\dim S_2^1(H_n) = n + 3$ for $n \geq 3$.

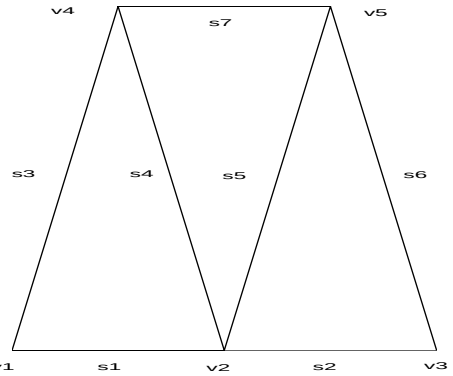
4.1 Three triangles—general and special cases

Let us consider the case of three triangles with two common edges as in Fig. 5a. We can use the foregoing results and write six CC for the connections of two pairs of triangles with the common edge and one with the common vertice.

For the local parameters (see Fig. 5a for numbering) $\mathbf{v} = [v_1, v_2, v_3, v_4, v_5]$, $\mathbf{m} = [m_{12}, m_{14}, m_{23}, m_{24}, m_{25}, m_{35}, m_{45}]$ we find dependent two CC corresponding to triangles with common vertex only (for the geometric interpretation see [4]—common derivatives along the edges V_2V_3, V_2V_5) and we obtain so four independent CC

$$\begin{bmatrix} -t_2^5 & -t_3^5 & 0 & 2t_1^5 & 1 & 0 & 3t_2^5 & 0 & 3t_3^5 & 0 & 0 & -3 \\ t_2^5 & 2t_3^5 & 0 & -t_1^5 & 1 & 3t_2^5 & 0 & 0 & 3t_1^5 & -3 & 0 & 0 \\ 0 & -t_2^3 & 1 & -t_3^3 & 2t_1^3 & 0 & 0 & 0 & 0 & 3t_2^3 & -3 & 3t_3^3 \\ 0 & 2t_2^3 & 1 & -t_3^3 & -t_1^3 & 0 & 0 & -3 & 3t_3^3 & 3t_1^3 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0} \tag{24}$$





Figs 5a, 5b

with the components of vector $\mathbf{t^5}$ of barycentric coordinates of the vertex V_5 with respect to triangle $V_4V_1V_2$ and $\mathbf{t^3}$ for V_3 in triangle $V_5V_2V_4$. We have 4 equations with 12 parameters —*the dimension of the MVI spline space is eight in general.*

To obtain the solution with given seven m_{ij} and one v_j in general, we need to have nonzero determinant from remaining four columns. It is valid for configurations

– $j = 1, V_5$ not located on lines V_1V_2, V_1V_4;
– $j = 2, V_5$ not on lines $V_1V_2, V_2V_4,\ V_3$ not on V_2V_4;
– $j = 3, V_5$ not on $V_2V_4,\ V_3$ not on $V_2V_4,\ t_1^5/t_2^5 = t_3^3/t_2^3$;
– $j = 4, V_5$ not on $V_1V_4, V_2V_4;\ V_3$ not on V_2V_4;
– $j = 5, V_5$ not on $V_1V_2, V_2V_4;\ V_3$ not on V_4V_5 .

Similar configuration of three triangles we obtain when we delete one triangle from *the general triangulated quadrangle with $V_1V_2V_3V_4$ with the internal vertex V_5 not located on lines V_1V_3, V_2V_4* .

In the case that the triangles with the common edges form a quadrilaterals (see Fig. 5b—half of the regular hexagon) we have $\mathbf{t^5} = [1, 1, -1] = \mathbf{t^3}$, $\mathbf{t^2} = [0, -1, 2]$ and the CC form the system with four independent CC only—

$$\begin{bmatrix} 1 & -1 & 0 & 2 & 1 & 0 & -3 & 0 & 3 & 0 & 0 & -3 \\ 1 & 2 & 0 & -1 & 1 & -3 & 0 & 0 & 3 & -3 & 0 & 0 \\ 0 & -1 & 1 & 1 & 2 & 0 & 0 & 0 & 0 & 3 & -3 & -3 \\ 0 & 2 & 1 & 1 & -1 & 0 & 0 & -3 & -3 & 3 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0}. \tag{25}$$

We find all five (4,4)-subdeterminants from the first five columns to be nonzero. *The dimension of the spline space is eight again.* We can prescribe all 7 values m_{ij} and any v_j.

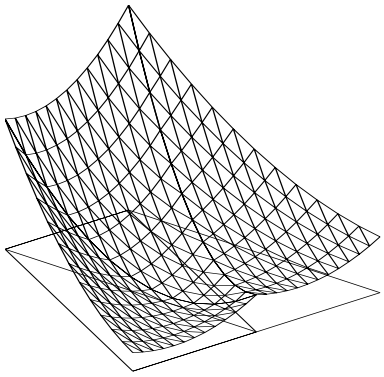
Similar special case we obtain when we delete from the quadrilateral with two diagonals one of four triangles (see Fig. 6b). The previous result and detailed analysis give us the result, that *there is not solution with 7 m_{ij} and any v_j, $j \in \{1 : 5\}$ in general*—the solvability condition is now $-m_{12} + m_{15} + m_{25} = -m_{34} + m_{35} + m_{45}$.

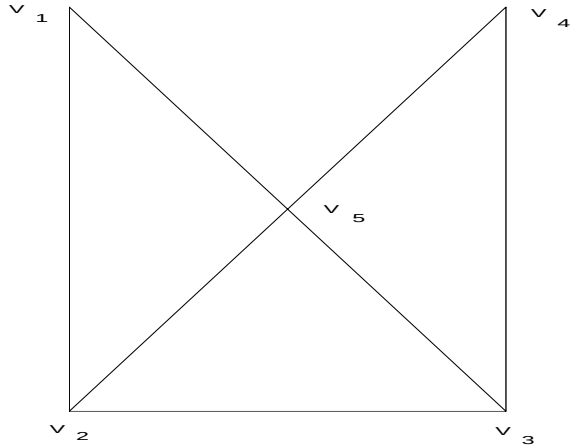
But when we use the CC (13) in (**s,m**)-LR, we find that *there is the unique solution with* $7m_{ij}$ *and any free parameter* $s_i,\ i \in \{1:7\}$.

The foregoing example is a special case with the edges on the common lines (called as *degenerate case* in [12], [13], [10]).

Statement 5: *The dimension of our spline space is equal to eight for vertices in general position. Such a 1D MVI problem has the unique solution with prescribed 7 values* m_{ij} *+ some* $v_j,\ j \in \{1:5\}$ *under conditions discussed above. In special cases (degenerate—one triangle deleted from triangulated equilateral, one half of hexagon) the dimension of the spline space is also equal to eight and we can find a solution with seven* m_{ij} *and one proper free parameter* v_j *or* s_i.

An example of such spline interpolating data $v_2 = 2,\ \mathbf{m} = [1,2,3,2,4,6,8]$ is plotted in Fig. 6a.





Figs 6a, 6b)

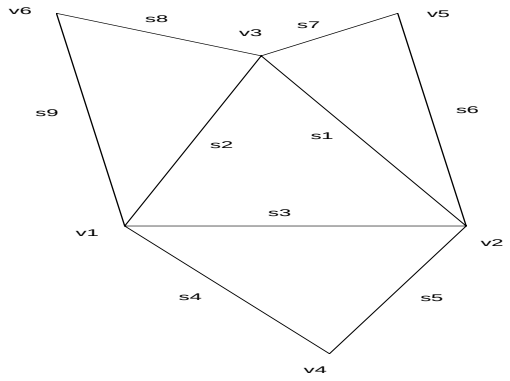
In all examples with three triangles discussed till now the dimension of the spline spaces was eight—we have obtained different patches over different triangles in general.

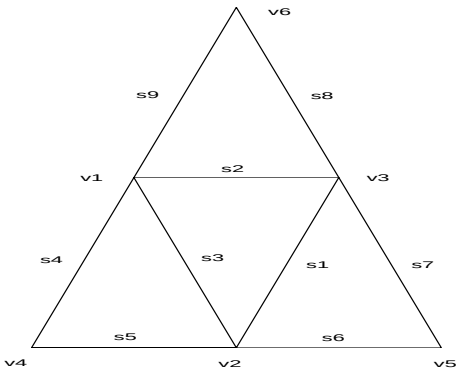
4.2 Four triangles

Let us consider the case of the triangle $V_1V_2V_3$ with flaps to each of three edges as in Fig. 7a. With nine edges now we cannot expect to find the quadratic polynomial interpolating 1D MV along all edges. We can try to search for MVI polynomial with six CI (e.g. along the boundary—some special cases will be mentioned in the following).

For the discussion of such MVI problem with splines, we can expect with two new CC and three new parameters the dimension of the spline space to be nine.

Let us denote $\mathbf{t^4}, \mathbf{t^5}, \mathbf{t^6}$ the vectors of barycentric coordinates of the vertices V_4, V_5, V_6 with respect to the triangles $V_2V_1V_3,\ V_3V_2V_1,\ V_1V_3V_2$. We can write down three pairs of CC (10) for the triangles with a common edge and three





Figs 7a, 7b

pairs of CC (20) for triangles with common vertex (with adequate changes of indexes corresponding to numbering in Fig. 7a). When we discuss in detail these last three pairs of CC, we find them again dependent on the first three pairs of CC (as the technique of BB-nets helps us to understand and forecast it). So we have 6 CC with 6 parameters v_j and 9 parameters m_{ij}. The determinant of the CC matrix with the columns corresponding to six LP v_j is equal to $27(t_2^4t_2^5t_2^6 - t_1^4t_1^5t_1^6)$. We can deduce from it the following

Statement 6: *The MVI problem for four triangles connected along edges as in Fig. 7a and with prescribed mean values along all nine edges has the unique solution under condition* $t_2^4t_2^5t_2^6 \neq t_1^4t_1^5t_1^6$ *only (with b.c.* $\mathbf{t^4}, \mathbf{t^5}, \mathbf{t^6}$ *w.r to triangle* $V_1V_2V_3$ *this condition is* $t_2^4t_3^5t_1^6 \neq t_1^4t_2^5t_3^6$*). So to each couple of vertices* V_4, V_5 *we can find the vertex* V_6 *such that the MVI problem has not solution in general. The dimension of the spline space is equal to nine.*

The above solvability condition is not fulfilled in the most cases with some symmetry—let us to see some frequently used cases.

For the special case of *the uniform triangulation* $\Delta(2)$ *of the triangle* (see Fig. 7b) we have $\mathbf{t^4} = [1, 1, -1] = \mathbf{t^5} = \mathbf{t^6}$. When we write six CI with LP $m_{12}, m_{13}, m_{23}, m_{45}, m_{46}, m_{56}$ (3 on the boundary, 3 inside) for *the quadratic polynomial*, we find the solution only under two conditions

$$m_{23} + m_{45} = m_{13} + m_{46},\ m_{23} + m_{45} = m_{12} + m_{56}\,.$$

For given six 1D MV along the boundary we find the solvability condition

$$m_{14} + m_{26} + m_{35} = m_{15} + m_{24} + m_{36}.$$

We can find quadratic polynomial interpolating four m_{ij} and two v_j in 144 from the total of 225 cases (e.g. with two internal m_{ij}, no e.g. with two MV's along parallel edges). There is no quadratic polynomial interpolating any six or five such values m_{ij} in general.

The LIS for the 1D MVI polynomial problem on the $\Delta(2)$*-triangulation consists from 74 groups of six edges (chosen from all 9 edges), where are not omitted*

3 m_{ij}:
a) from edges on parallel lines (3 variants),
b) in one triangle (4 variants),
c) m_{15}, m_{56}, m_{26} and symmetric two cases.

We mentioned the dimension of *the spline space* to be nine. When we use the (**v,m**)-LR and write six CC with 6 parameters v_j and 9 parameters m_{ij}, we obtain with $\mathbf{v} = [v_1, \cdots, v_6]$, $\mathbf{m} = [m_{12}, m_{13}, m_{14}, m_{16}, m_{23}, m_{24}, m_{25}, m_{35}, m_{36}]$ and matrix $\mathbf{A_{vm}}$

$$\begin{bmatrix} 2 & -1 & 1 & 1 & 0 & 0 & 3 & -3 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & 1 & 1 & 0 & 0 & 3 & 0 & 0 & 0 & -3 & -3 & 0 & 0 & 0 \\ 1 & -1 & 2 & 0 & 1 & 0 & 0 & -3 & 0 & 0 & 3 & 0 & 0 & -3 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 & -3 & 0 & 0 & 0 & 3 & 0 & -3 & 0 & 0 \\ 2 & 1 & -1 & 0 & 0 & 1 & 0 & 3 & 0 & 0 & -3 & 0 & 0 & 0 & -3 \end{bmatrix}$$

the system $\mathbf{A_{vm}}[\mathbf{v}, \mathbf{m}]^T = 0$. We find the matrix corresponding to parameters v_j to be singular and the MVI problem with given 9 parameters m_{ij} to be solvable in general only under condition

$$m_{14} + m_{25} + m_{36} = m_{16} + m_{24} + m_{35}.$$

But we can find the unique solution of the problem with given
a) 3 internal and any 5 boundary values m_{ij} + any v_j ;
b) 6 boundary parameters m_{ij} and 3 free parameters $s_2, s_3; v_j$ (but no with s_1, s_2, s_3 or v_1, v_2, v_3).

In both general and special cases we have obtained the spline space dimension equal to nine—we could use free parameters for some another (e.g. optimization) purposes. The results obtained in the discussion of foregoing examples allow us to claim the summary in the following

Statement 7: *The dimension of quadratic splines space over the uniform $\Delta(2)$-triangulation is nine. But there is not a unique solution of the 1D MVI problem in general with given nine 1D MV's along all 9 edges—one m_{ij} has to be computed from the solvability condition mentioned.*

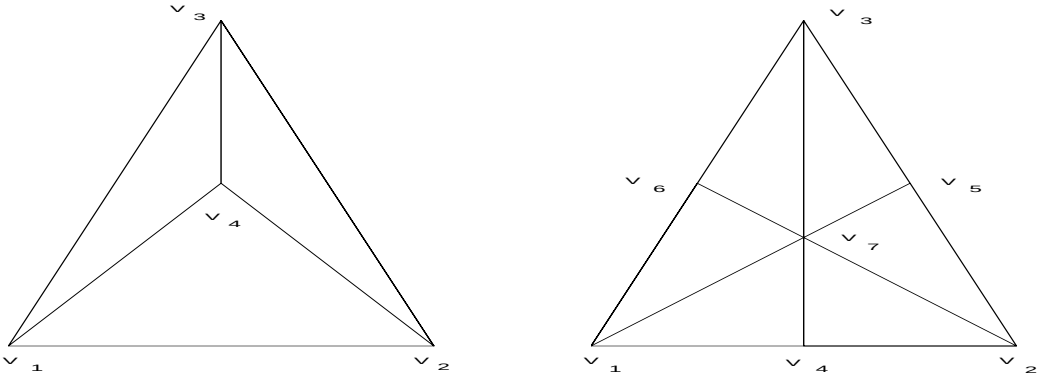
There is the unique solution with given eight m_{ij} (5 on the boundary + 3 inside) + one v_j.

In the problem with six MV's given along 3 inner and 3 outer edges we can find solution with 3 additional free parameters s_i, v_j (no e.g. s_1, s_2, s_3 or v_1, v_2, v_3).

5 Clough–Tocher triangle

Let us have the triangle $V_1V_2V_3$ divided into three triangles with internal vertex V_4 as in CT element known in FEM method (see Fig. 8 a) for parameters numbering). Frequently the special case of V_4 as the center is considered. In the discussion of

CC+CI we will use now the results from the foregoing sections and [4],[6]-[8]. We have learned that as for the polynomials and for the splines the space dimension is equal to six now.



Figs 8a, 8b

5.1 One quadratic MVI polynomial

With barycentric coordinates $V_4 = [t_1, t_2, t_3]$ in the triangle $V_1V_2V_3$ and LR (**v,s**) we can write the CI (after some eliminations) as the system of six equations for parameters $v_1, v_1, v_3, s_1, s_2, s_3$ with regular matrix.

Statement 8. *The MVI problem on CT split of triangle with prescribed 1D MV along all six edges has unique solution in the class of quadratic polynomials.*

Remark: This result follows also from the more general Theorem 12.1 in [1].

5.2 MVI quadratic spline

Let us consider the case when for the MVI interpolating spline on CT triangle the values v_i and the one-dimensional mean values m_{ij} over edges are used as local parameters. Using relations (4) we can write the four independent CC now with vectors $\mathbf{v} = [v_1, v_2, v_3, v_4], \mathbf{m} = [m_{12}, m_{13}, m_{14}, m_{23}, m_{24}, m_{34}]^T$ and vectors $\mathbf{t^1}, \mathbf{t^2}, \mathbf{t^3}$ of barycentric coordinates of the vertices V_3, V_1, V_2 with respect to triangles $V_4V_2V_1, V_4V_3V_2, V_4V_3V_1$ as

$$\begin{bmatrix} -t_3^1 & -t_2^1 & 1 & 2t_1^1 & 0 & 0 & 3t_3^1 & 0 & 3t_2^1 & -3 \\ -t_3^1 & 2t_2^1 & 1 & -t_1^1 & 3t_3^1 & 0 & 0 & -3 & 3t_1^1 & 0 \\ 1 & -t_3^2 & 2t_2^2 & -t_1^2 & 0 & -3 & 0 & 3t_3^2 & 0 & 3t_1^2 \\ 2t_2^3 & 1 & -t_3^3 & -t_1^3 & -3 & 3t_3^3 & 3t_1^3 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0}. \tag{26}$$

With the barycentric coordinates $\mathbf{t} = [t_1, t_2, t_3]$ of the vertice V_4 with respect to the triangle $V_1V_2V_3$ we can write the matrix of CC (26) as

$$\begin{bmatrix} t_1 & t_2 & t_3 & 2 & 0 & 0 & -3t_1 & 0 & -3t_2 & -3t_3 \\ t_1 & -2t_2 & t_3 & -1 & -3t_1 & 0 & 0 & -3t_3 & 3 & 0 \\ t_1 & t_2 & -2t_3 & -1 & 0 & -3t_1 & 0 & -3t_2 & 0 & 3 \\ -2t_1 & t_2 & t_3 & -1 & -3t_2 & -3t_3 & 3 & 0 & 0 & 0 \end{bmatrix}. \tag{27}$$

In the special case with V_4 as centroid we have $t_1 = t_2 = t_3 = 1/3$ and the CC have simple form

$$\begin{bmatrix} 1 & 1 & 1 & 6 & 0 & 0 & -3 & 0 & -3 & -3 \\ 1 & -2 & 1 & -3 & -3 & 0 & 0 & -3 & 9 & 0 \\ 1 & 1 & -2 & -3 & 0 & -3 & 0 & -3 & 0 & 9 \\ -2 & 1 & 1 & -3 & -3 & -3 & 9 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{m} \end{bmatrix} = \mathbf{0} \tag{28}$$

We can consider as prescribed parameters all six 1D mean values $m_{i,j}$ and to compute from CC (27) (with determinant equal to $-27t_1t_2t_3$)

$$\begin{aligned} v_1 &= m_{12} + m_{13} + m_{14} + [(1-t_3)(m_{23} - m_{34}) - (1-t_2)m_{24}]/t_1, \\ v_2 &= m_{12} + m_{23} + m_{24} + [(1-t_2)m_{13} - (t_2+t_3)m_{14} + t_3m_{34}]/t_2, \\ v_3 &= m_{13} + m_{23} + m_{34} + [(1-t_3)m_{12} - (t_2+t_3)m_{14} - (1-t_2)m_{24}]/t_3, \\ v_4 &= m_{14} + m_{24} + m_{34} - (1-t_3)m_{12} - (1-t_2)m_{13} + (t_2-t_3)m_{23}. \end{aligned}$$

Now we can use $(\mathbf{v}, \mathbf{m})$- LR for each quadratic patch separately (the LP s_i can be computed from v_j, m_{ij} if needed).

Statement 9: *There is a unique quadratic spline with prescribed six values m_{ij} on CT triangle. The dimension of the corresponding spline space is six.*

Remark. In case of $V_4 = [1,1,1]/3$ (barycenter) we obtain specially

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 2 & -2 & -2 \\ 1 & 2 & -2 & 1 & 1 & -2 \\ 2 & 1 & -2 & 1 & -2 & 1 \\ -2/3 & -2/3 & 1 & -2/3 & 1 & 1 \end{bmatrix} [m_{12}, m_{13}, m_{14}, m_{23}, m_{24}, m_{34}]^T. \tag{29}$$

Also the mixed problem with prescribed parameters $v_1, v_2, v_3, m_{12}, m_{13}, m_{23}$ is uniquelly solvable - with $t_i = 1/3$ we obtain for the remaining parameters

$$\begin{bmatrix} v_4 \\ m_{14} \\ m_{24} \\ m_{34} \end{bmatrix} = \frac{1}{9} \begin{bmatrix} -3 & -3 & -3 & 6 & 6 & 6 \\ 1 & -2 & -2 & 5 & 5 & 2 \\ -2 & 1 & -2 & 5 & 2 & 5 \\ -2 & -2 & 1 & 2 & 5 & 5 \end{bmatrix} [v_1, v_2, v_3, m_{12}, m_{13}, m_{23}]^T. \tag{30}$$

5.3 Quadratic MVI spline over CT triangulation is quadratic polynomial

We have mentioned in the foregoing subsection that (also as in (23) and 2D MVI or FVI problems—see [6]–[8]) in all cases on CT triangulations we have obtained the dimension of the spline space equal to six—the same as for unique quadratic polynomial.

Statement 10. *All three patches of the MVI quadratic spline on the CT triangulation (with 1D or 2D mean values prescribed) belong to the same quadratic polynomial over triangle* $V_1V_2V_3$.

Proof: For the quadratic MVI polynomial with parameters $v_1 = v_2 = v_3 = s_1 = s_4 = s_6 = 0$ and quadratic spline with all six $m_{ij} = 0$ we find in both cases only trivial solution—and from it follows now $P_2 \equiv S_2^1$.

Another way of the proof can be the following: When we use (**v,s**)-local representation of the quadratic polynomial and quadratic spline and compare the the results e.g. for the values v_4, s_2, s_5 (with the relations between parameters mentioned in the foregiong subsections), we obtain identical results in six points for both objects. Similarly in (**v,m**)- representation we obtain the same results in computing values $v_4, m_{12}, m_{13}, m_{23}$ with the data from the polynomial or the spline. The data $v_1, v_2, v_3, s_1, s_4, s_6$ uniquelly determine both unique quadratic polynomial and mentioned values m_{ij} and also quadratic spline—both object have to be identical.

For the 1D MVI spline we can use the above relation to compare the expressions for $v_4, m_{14}, m_{24}, m_{34}$ in quadratic spline and quadratic polynomial—we will find them identical.

Remark: The result mentioned in Statement 10 is stated in [4]. In FEM interpolation theory the different CT element is used in problem with cubic polynomial patches: different cubic patches form here C^1-element.

5.4 Powell–Sabin split

Let us consider the in FEM well-known Powell–Sabin split of the triangle $V_1V_2V_3$ into six triangles (the vertices connected with the edge midpoints through the center—see Fig. 8b). When we prescribe the 2D MV for each subtriangle, there is not quadratic MVI polynomial—we have to prescribe one another local parameter to obtain the solution (see [7]) .

For the MVI polynomial problem with given six 1D MV along six edges the positive answer is not given in [[1], Th. 12.1], because the vertices are not "in the general position" now. When we compute the FV in edge midpoints of V_1V_5, V_2V_5, V_3V_4 as combinations of FV v_j, we obtain the following system of

equations

$$\begin{bmatrix} 1 & 1 & 0 & 4 & 0 & 0 \\ 0 & 1 & 1 & 0 & 4 & 0 \\ 1 & 0 & 1 & 0 & 0 & 4 \\ 1 & -1/2 & -1/2 & 2 & 2 & 2 \\ -1/2 & 1 & -1/2 & 2 & 2 & 2 \\ -1/2 & -1/2 & 1 & 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix} = 6 \begin{bmatrix} m_{12} \\ m_{23} \\ m_{13} \\ m_{15} \\ m_{26} \\ m_{34} \end{bmatrix} \tag{31}$$

with the regular matrix. This proves the following

Statement 11. *The 1D MVI problem on the PS split with prescribed values m_{ij} along six edges $m_{12}, m_{23}, m_{13}, m_{15}, m_{26}, m_{34}$ has unique solution in the class of quadratic polynomials.*

For the analysis of such MVI problem with quadratic splines we know from (23) that the space dimension is nine. We have now $V = 7$, $T = 6$, $E = 12$ and we can write 10 independent CC (4 around the center V_7, 6 at the vertices V_j, $j = 1 : 6$) in (**v**,**m**)-LR, $\mathbf{v} = [v_1, \cdots, v_7], \mathbf{m} = [m_1, \cdots, m_{12}]$ (numbering from Fig. 8b) as

$$\mathbf{A_v v} + \mathbf{A_m m} = \mathbf{0} \tag{32}$$

with matrices

$$\mathbf{A_v} = \begin{bmatrix} 1 & 1 & 0 & 4 & 0 & 0 & 0 \\ 1 & 1 & 0 & -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 2 & 2 & 0 & 6 \\ 0 & 2 & 0 & 2 & 2 & 0 & -3 \\ 0 & 1 & 1 & 0 & 4 & 0 & 0 \\ 0 & 1 & 1 & 0 & -2 & 0 & 0 \\ 0 & 0 & -1 & 0 & 2 & 2 & 6 \\ 0 & 0 & 2 & 0 & 2 & 2 & -3 \\ 1 & 0 & 1 & 0 & 0 & 4 & 0 \\ 2 & 0 & 0 & 2 & 0 & 2 & -3 \end{bmatrix},$$

$$\mathbf{A_m} = \begin{bmatrix} -3 & -3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 6 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -6 & 3 & 0 & 0 & -6 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 & 0 & 9 & -6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & -3 & 0 & 0 & 6 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -6 & -6 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -6 & 9 & -6 \\ 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 0 & -3 & 0 & 0 \\ -6 & 0 & -6 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We can confirm now that the dimension of the spline space is equal to nine and we can search for the "Lagrange interpolation set" (LIS) of nine from twelve) edges,

for which the MV's m_i determine uniquely the remaining LP. Some computations of determinants with columns from matrices $\mathbf{A_v}, \mathbf{A_m}$ result in the following

Statement 12. *The MVI problem with prescribed nine one-dimensdional mean values along edges of the Powell–Sabin split of the triangle has the unique solution in 211 from the total 220 cases.*

There is no solution in cases when

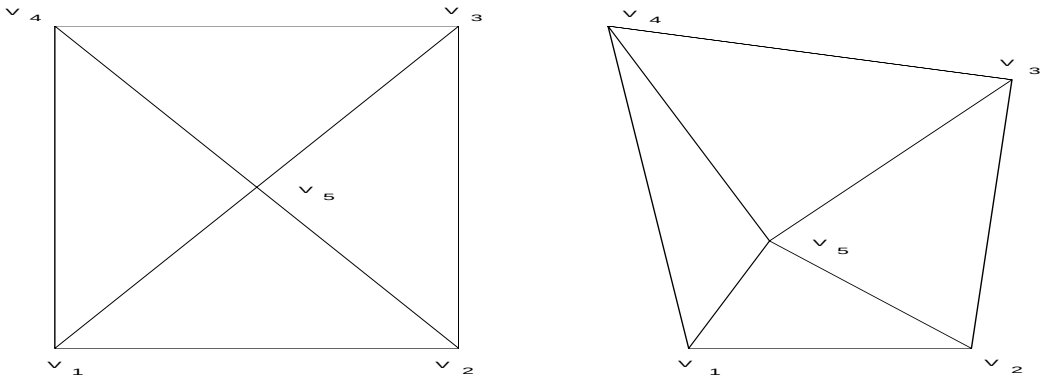
– *we omit three mean values in one subtriangle (6 variants),*

– *we omit m_{16}, m_{47}, m_{25} and symmetric cases (3 variants).*

There is the unique solution with given v_1, v_2, v_3 and six values m_{ij} along the boundary.

There is no solution in general with given v_1, v_2, v_3 and six values m_{ij} along internal edges.

6 Triangulated quadrangle



Figs 9a, 9b

The general quadrangle we can triangulate into four triangles with some internal vertex (see Fig. 9b). We have mentioned the known formula (39) for the dimension of the spline space, depending on the geometry of edges.

6.1 Triangulated rectangle (quadrilateral)

Let us have the rectangle (or quadrilateral) $V_1V_2V_3V_4$ divided with two diagonals into four triangles with the common vertex V_5, and denote s_i, $i = 1 : 8$, v_j, $j = 1 : 5$ the (unknown) function values in edge midpoints and vertices (for the numbering see Fig. 9a).

With given six 1D MV along four boundary edges and two diagonals there exists *the unique MVI quadratic polynomial* (consequence with Th. 12.1 in [1]). We can find its parameters $\mathbf{v} = [v_1, v_2, v_3, v_4, v_5], \mathbf{s} = [s_1, s_4, s_5, s_8]$ e.g. from the

regular system

$$\begin{bmatrix} 1&1&0&0&0&4&0&0&0\\ 1&0&1&0&4&0&0&0&0\\ 1&0&0&1&0&1&4&0&0\\ 0&1&1&0&0&0&0&6&0\\ 0&1&0&1&4&0&0&0&0\\ 0&0&1&1&0&0&0&0&4\\ 1&3&1&-1&4&-4&0&-4&0\\ 1&1&0&0&2&-2&-1&-1&0\\ 0&1&1&0&2&-1&0&-2&-1 \end{bmatrix} \begin{bmatrix} \mathbf{v}\\ \mathbf{s} \end{bmatrix} = 6 \begin{bmatrix} m_{12}\\ m_{13}\\ m_{14}\\ m_{23}\\ m_{24}\\ m_{34}\\ 0\\ 0\\ 0 \end{bmatrix}. \tag{33}$$

We can also obtain the unique solution with given $m_{12}, m_{23}, m_{34}, m_{14}, m_{15}, m_{25}$ and in symmetric cases (which correspond more exactly to our original problem).

Statement 13: *There exist unique quadratic MVI polynomial over triangulated rectangle with prescribed six m_{ij}*
– along four boudary edges and two diagonals,
– along four boundary edges and two neighbouring edges V_jV_5.

To discuss solvability of such problem *in the space of quadratic splines* with the dimension equal to eight we can write six independent CC with parameters $\mathbf{v},\ \mathbf{m} = [m_{12}, m_{15}, m_{25}, m_{14}, m_{23}, m_{45}, m_{35}, m_{34}]$ as

$$\begin{bmatrix} 1&0&1&0&4&0&-3&0&0&0&-3&0\\ 1&0&1&0&-2&-3&0&6&0&-3&0&0\\ 0&1&0&1&4&0&0&-3&0&0&-3&0\\ 1&0&1&0&-2&0&0&0&-3&0&6&-3\\ 0&1&0&1&-2&-3&6&0&-3&0&0&0 \end{bmatrix} \begin{bmatrix} \mathbf{v}\\ \mathbf{m} \end{bmatrix} = \mathbf{0}. \tag{34}$$

Although the dimension of our spline space is equal to eight, the analysis of this system gives us the following results:

Statement 14: *The 1D MVI problem on the triangulated rectangle (quadrilateral) has dimension equal to eight;*
– it has no solution with prescribed all 8 m_{ij} in general,
– it has no solution with prescribed 7 $m_{ij} + 1$ v_j,
– it has the unique solution with 6 $m_{ij} + 2$ v_j in 96 cases (e.g. with no four data on two lines, two v_j on the diagonal).

6.2 General triangulation of quadrangle

On *the general triangulated quadrangle* as in Fig. 9b with nondegenerate vertex V_5 and barycentric coordinates of the vertices $V_3 = [t_1^3, t_2^3, t_3^3],\ V_4 = [t_1^4, t_2^4, t_3^4]$ with respect to the triangle $V_5V_2V_1$, (V_5 in the general position inside quadrangle—no on common line) we find now six linearly independent CC (two around the center, four around vertices) with 13 parameters $v_j,\ m_{ij}$. So we have only 7 free

parameters in general—we confirmend the known result (see (39), [12]) that *the dimension of such spline space is equal to seven!*

Reference

[1] Bojanov, B. D., Hakopian, H. A., Sahakian, A. A.: *Spline Functions and Multivariate Interpolation.* Kluwer A. P., 1993.

[2] Chui, Ch. K.: *Multivariate Splines. Theory and Applications.* Texas AM Univ., 1987, 1991.

[3] Chui, Ch. K., Tian-Xiao, He: *Bivariate C^1 quadratic finite elements and vertex splines.* Math. Comp., **54**, 189 (1990), 169–187.

[4] Farin, G. : *Triangular Bernstein-Bezier patches.* CAGD **3** (1986), 83–127.

[5] Heindl, G. : *Interpolation and approximation by piecewise quadratic C^1-functions of two variables.* Multivariate Approx. Theory **2** (1979).

[6] Kobza, J.: *Quadratic interpolatory splines on triangulations.* In: Proc. PANM'02 (MI AS CR), 65–99.

[7] Kobza, J.: *Quadratic polynomials interpolating function or mean values on simple triangulations.* In: Proc. SANM'03 (to appear).

[8] Kobza, J.: *Quadratic splines interpolating mean values on simple triangulations.* In: Proc. WS, St. Petersburg 2003 (to appear).

[9] Gmelig Meyling, Pfluger, P. R.: *On the dimension of the space $S_2^1(\Delta)$ in special cases.* Multivariate Approx. Theory **3** (1985),180–190.

[10] Micula Gh., Micula, S.: *Handbook of Splines.* Kluwer, 1999.

[11] Nuernberger, G., Zeilfelder, F.: *Development in spline interpolation.* Jour. Comp. Appl. Math. **121** (2000), 125–152.

[12] Schumaker, L. L.: *On the dimension of spaces of piecewise polynomials in two variables.* Multivariate Approx. Theory **2** (1979), 396–412.

[13] Schumaker, L. L.: *Bounds on the dimension of spaces of multivariate piecewise polynomials.* Rocky Mountains J. Math. **14**, 1 (1984), 251–264.

[14] Wang, R. H.: *The dimension and basis of spaces of multivariate splines.* J. Comp. Appl. Math. 12–13 (1985),163–177

[15] Ženíšek, A.: *Nonlinear elliptic and evolution problems and their FE approximations.* Academic Press 1990


Univ. Palacki. Olomuc., Fac. rer. nat.,
Dept of Math. Anal. and Appl. of Math.
ODAM (2003) 56–75

A Fast Method for Solving Saddle-Point Systems with Singular Blocks Arising in Wavelet-Galerkin Discretizations of PDEs*

Radek KUČERA

Department of Mathematics and Descriptive Geometry
VŠB–Technical University of Ostrava, 17. listopadu 15
CZ-708 33 Ostrava-Poruba, Czech Republic
e-mail: radek.kucera@vsb.cz

Abstract

The paper deals with fast solving of large saddle-point systems arising in wavelet-Galerkin discretizations of separable elliptic PDEs. A special structure of matrices make possible to use the fast Fourier transform that determines the complexity of the algorithm. Numerical experiments confirm theoretical results.

Key words: wavelet-Galerkin discretization, saddle-point systems, conjugate gradient method, circulant matrices, fast Fourier transform, Kronecker product.

1991 Mathematics Subject Classification: 65T60, 65F10, 65T50, 65N30

*Supported by grant MSM 272400019 and HPRNT-CT-2002-00286.

1 Introduction

In this paper, we shall propose a fast method for finding a pair $(u, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m$ that solves the linear system of algebraic equations called the *saddle-point* system:

$$\begin{pmatrix} A & B^\top \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}, \tag{1}$$

where the $n \times n$ matrix A is symmetric positive semi-definite, the $m \times n$ matrix B has full row-rank and the vectors f, g are of the order n, m, respectively. We shall be interested especially in systems (1) with n large, A singular, B sparse and m much smaller than n. Moreover we shall assume that the defect of A, i.e. $l = n - rankA$, is much smaller than m. Systems of this type arise e.g. when we want to solve quadratic programming problems with equality constraints [2], mixed formulations of second-order elliptic problems [1] or if we use the fictitious domain method to Dirichlet problems [4].

This paper is motivated by a class of saddle-point systems arising in wavelet-Galerkin discretizations of separable elliptic PDEs. It is well-known that the orthogonal compactly supported wavelets are defined on a bounded interval (or a rectangular domain) via periodization [3]. Therefore they are a natural tool for solving problems with the periodic boundary conditions. Other types of boundary conditions (e.g. Dirichlet or Neumann) can be treated by means of the *fictitious domain method* [7]. In this case, it is necessary to solve the saddle-point system (1), where the diagonal block A represents the PDE and the off-diagonal block B describes the geometry of the domain. More precisely, A is the stiffness matrix on a new fictitious domain, where the periodic boundary conditions are considered. On the one hand, it can happen that the matrix A is singular, on the other hand, the periodic boundary conditions lead to the block circulant structure of A. Since the circulant matrices are diagonalizabel by the *discrete Fourier transform* (DFT), one can evaluate eigenvalues of A by using the efficient algorithm called the *fast Fourier transform* (FFT). This one does the situation much more easy because it is possible to treat the singularity of A without greate computational costs.

There are several basic approaches used for solving the saddle-point systems (1). We turn our attention to the class of methods called *primar* (the Schur complement methods, the range space methods or static condensation). The key idea is based on eliminating the first unknown u. If A is non-singular, we obtain the linear system in terms of the second unknown λ with the positive definite matrix. Then the *conjugate gradient method* (CGM) can be used for computation of the solution.

The situation is not so easy if A is singular because the first unknown u can not be eliminated completly from (1). We obtain the (second) linear system in terms of λ and a new unknown, say α, that describes the part of u corresponding to the null-space of A. The new linear system has again the saddle-point structure and

its diagonal block is non-singular in many practical situations. Therefore we can repeatedly eliminate the first unknown, now it is λ, and we obtain the (third) linear system in terms of α with the positive definite matrix. The resulting linear system can be solved easily, e.g. by a direct method, since it is of the small order. Let us point out that the CGM can be used during the elimination process in order to compute the matrix and the right-hand side vector for the third linear system without necessity to compute and store the diagonal block of the second linear system. Although this method seems to be cumbersome, we shall propose its efficient implementation with small memory requirements. In advance, we shall show that even the fast implementation is possible, e.g. for the saddle-point systems arising in the wavelet-Galerkin discretizations of PDEs as mentioned above.

The paper is organized as follows. In Section 2, we summarize theoretical results concerning existency and uniqueness of the solution to the saddle-point systems (1) and elimination of the first unknown in the case of singular A. A general form of the algorithm for solving the saddle-point system (1) is proposed in Section 3. A fast implementation based on the use of the FFT and the Kronecker product is described in Section 4. Finally, Section 5 presents results of numerical experiments.

2 Preliminaries

Let us denote the null-space of B by

$$\mathcal{N}(B) = \{v \in \mathbb{R}^n : Bv = o\}$$

and the range-space of B by

$$\mathcal{R}(B) = \{\mu \in \mathbb{R}^m : \mu = Bv\}.$$

Lemma 1 *Let A be symmetric positive semi-definite. Then $v \in \mathcal{N}(A)$ iff $v\top Av = 0$.*

Proof. The proof follows from the fact that $\mathcal{R}(A)$ is the orthogonal complement to $\mathcal{N}(A)$. □

Theorem 1 *The saddle-point system (1) has a unique solution iff*

$$\mathcal{N}(A) \cap \mathcal{N}(B) = \{o\}. \tag{2}$$

Proof. There is a unique solution to (1) iff the homogennous saddle-point system

$$Au + B^\top \lambda = o, \tag{3}$$
$$Bu = o, \tag{4}$$

has only the trivial solution $(u, \lambda) = (o, o)$. Let us suppose contradictorily to (2) that $v \in \mathcal{N}(A) \cap \mathcal{N}(B)$, $v \neq o$. Then the pair (v, o) is the non-trivial solution to (3), (4) so that the assumption (2) have to be neccessarily satisfied. In order to prove sufficientness of (2), we multiply (3), (4) from the left by $u^\top$, $\lambda^\top$, respectively. A simple manipulation gives $u^\top A u = 0$ so that $u \in \mathcal{N}(A)$ by Lemma 1. In view of (4), we have $u \in \mathcal{N}(A) \cap \mathcal{N}(B)$ and (2) implies $u = o$. This one reduces (3) to $B^\top \lambda = o$ and, because of full-row rank of B, we obtain $\lambda = o$. □

We shall assume that A in the saddle-point system (1) is singular with $l = \dim \mathcal{N}(A)$, $l \geq 1$. Consider a $l \times p$ matrix N whose columns span the null-space $\mathcal{N}(A)$ and denote by $A^\dagger$ a generalized inverse to A that satisfies

$$A = AA^\dagger A. \tag{5}$$

Let us point out that $A^\dagger$ is not determined by (5) uniquely. The following remark shows that we can easily find symmetric positive semi-definite $A^\dagger$.

Remark 1 Any symmetric positive semi-definite matrix A can be factored into a product $LDL^\top$ with non-singular lower tri-diagonal L and diagonal $D = diag(d_1, \ldots, d_n)$, see [5]. Let us define $D^\dagger = diag(d_1^\dagger, \ldots, d_n^\dagger)$, where $d_i^\dagger = 1/d_i$ if $d_i \neq 0$ and $d_i^\dagger = 0$ if $d_i = 0$. It may be verifed directly that $A^\dagger = (L^\top)^{-1} D^\dagger L^{-1}$ is symmetric positive semi-definite and satisfies (5).

Theorem 2 *Let us assume that (2) is satisfied and $A^\dagger$ is symmetric positive semi-definite. The second component λ of the solution to (1) is the first component of the solution to the linear system*

$$\begin{pmatrix} BA^\dagger B^\top & -BN \\ -N^\top B^\top & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \alpha \end{pmatrix} = \begin{pmatrix} BA^\dagger f - g \\ -N^\top f \end{pmatrix}. \tag{6}$$

The first component u of the solution to (1) is given by the formulea

$$u = A^\dagger (f - B^\top \lambda) + N\alpha. \tag{7}$$

Proof. First we shall prove that there is a unique solution to (6). Under our assumptions, it is easy to show that $BA^\dagger B^\top$ is symmetric positive semi-definite and $-N^\top B^\top$ has full row-rank. Therefore (6) is again the saddle-point system of the type (1). Let us have $\mu \in \mathcal{N}(BA^\dagger B^\top) \cap \mathcal{N}(-N^\top B^\top)$ and denote $v = B^\top \mu$. Because of $N^\top v = o$, v belongs to $\mathcal{R}(A)$ so that $v = Aw$. Furthermore $\mu^\top BA^\dagger B^\top \mu = 0$ yields

$$0 = v^\top A^\dagger v = w^\top AA^\dagger Aw = w^\top Aw.$$

According to Lemma 1, we obtain $w \in \mathcal{N}(A)$ so that $v = o$ or equivalently $B^\top \mu = o$. The last equivality gives $\mu = o$ because B has full-row rank. Hence

$\mathcal{N}(BA^\dagger B^\top) \cap \mathcal{N}(-N^\top B^\top) = \{o\}$ and Theorem 1 implies that there is a unique solution to the saddle-point system (6). It remains to prove that the pair (u, λ) satisfying (6) and (7) is the solution to the saddle-point system (1). It may be directly verified by substituting (7) and then (6) into (1). □

The algorithm proposed in the next section is based on the previous theorem. We shall confine ourselves to situations in which $BA^\dagger B^\top$ is non-singular. The sufficient condition guaranting this property is proved in the following theorem.

Theorem 3 *(i) The matrix $BA^\dagger B^\top$ in (6) is symmetric positive definite if*

$$\mathcal{N}(A^\dagger) \cap \mathcal{R}(B^\top) = \{o\}. \tag{8}$$

(ii) If $A^\dagger$ is the Moore-Penrose pseudoinverse to A then (8) is equivalent to

$$\mathcal{N}(A) \cap \mathcal{R}(B^\top) = \{o\}. \tag{9}$$

Proof. Denote $v = B^\top \mu$ for $\mu \neq o$. Because of $v \neq o$, the relation (8) yields $v \notin \mathcal{N}(A^\dagger)$. Using Lemma 1 for $A^\dagger$, we obtain

$$\mu^\top B A^\dagger B^\top \mu = v^\top A^\dagger v > 0$$

so that the statment *(i)* holds. The statment *(ii)* follows from the fact that if $A^\dagger$ is the Moore-Penrose pseudoinverse to A then $\mathcal{N}(A) = \mathcal{N}(A^\dagger)$. □

3 Algorithms

In this section, we shall propose an algorithm for solving the saddle-point system (1) with singular A. First we recall the well-known algorithm for the case of non-singular A that specifies a bound for efficiency of our algorithm. We shall asses computational costs both algorithms by means of floating point operations (flops) that are perfomed in matrix-vector multiplications.

3.1 Algorithm for the non-singular case

Let us consider the saddle-point system (1), i.e.

$$Au + B^\top \lambda = f, \tag{10}$$

$$Bu = g, \tag{11}$$

where A is non-singular. We can express u from (10),

$$u = A^{-1}(f - B^\top \lambda), \tag{12}$$

and substitute it into (11). We obtain the linear system in terms of λ,

$$C\lambda = p, \tag{13}$$

where $C = BA^{-1}B^\top$ is positive definite and $p = BA^{-1}f - g$. The algorithm for solving (10), (11) can be divided into three steps:

Algorithm I

(a) Compute p.

(b) Solve (13) by mens of the CGM.

(c) Compute u from (12).

Denote $n_{A^{-1}}$ the number of flops to evaluate the matrix-vector product $A^{-1}v$ and assume that $n_{A^{-1}} \geq n$. Using sparsity of B, the matrix-vector products Bv and $B^\top\mu$ can be evaluated by means of $m_B = \mathcal{O}(m)$ flops. Then the computational costs of the steps (a) and (c) are $2(n_{A^{-1}} + m_B)$ flops. The main reason for the use of the CGM in the step (b) is the fact that it requires one matrix-vector product $C\mu$ per iteration which can be evaluated by the following procedure:

$$v := B^\top\mu, \quad v := A^{-1}v, \quad C\mu := Bv. \tag{14}$$

Since (14) requires $n_{A^{-1}} + 2m_B$ flops and the CGM terminates at most after m iterations, the realization of the step (b) requires at most $m(n_{A^{-1}} + 2m_B)$ flops.

Lemma 2 *Algorithm I requires $\mathcal{O}((m+2)n_{A^{-1}})$ flops.*

Proof. The previous discussion shows that *Algorithm I* involves

$$(m+2)n_{A^{-1}} + 2(m+1)m_B$$

flops. The proof follows from the fact that m is much smaller than n, $m_B = \mathcal{O}(m)$ and from the assumption $n_{A^{-1}} \geq n$. □

3.2 Algorithm for the singular case

Let us consider the saddle-point system (1) with singular A and let us assume that (2) is satisfied, i.e. there is a unique solution (u, λ). The key idea of our algorithm consists of the use of Theorem 2. At first we compute the pair (λ, α) solving the linear system (6) and then we evaluate u by means of (7). In order to simplify the presentation, we introduce the following notations

$$\begin{aligned} C &= BA^\dagger B^\top, \quad & p &= BA^\dagger f - g, \\ D &= -N^\top B^\top, \quad & q &= -N^\top f, \end{aligned}$$

where C is the $m \times m$ matrix, D is the $l \times m$ matrix, p is the m vector and q is the l vector. Using the new notations, the linear system (6) reads as follows

$$\begin{pmatrix} C & D^\top \\ D & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \alpha \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}. \tag{15}$$

Moreover, we shall assume that (8) is satisfied. Then C is non-singular so that we can eliminate the first unknown λ from (15). We obtain

$$\lambda = C^{-1}(p - D^\top \alpha) \tag{16}$$

and

$$E\alpha = r, \tag{17}$$

where $E = DC^{-1}D^\top$ is the positive definite $l \times l$ matrix and $r = DC^{-1}p - q$ is the l vector.

Let us point out that, under our assumptions, D, E, p, q and r are relatively small and, on the other hand, C is large and non-sparse. Therefore we shall propose the algoritm in such a way that its not neccesary to store C.

Algorithm II

(a.1) Compute D, p, q.

(a.2) Compute E, r.

(a.3) Compute α solving the linear system (17).

(b) Compute λ from (16).

(c) Compute u from (7), i.e. $u = A^\dagger(f - B^\top\lambda) + N\alpha$.

An efficient implementation of this algorithm is based on the following three assumptions:

(A1) We shall assume that each of the matrix-vector products $A^\dagger v$, $N\alpha$ and $N^\top v$ can be evaluated by means of $n_{A^\dagger}$ flops, $n_{A^\dagger} \geq n$. We shall discuss such multiplying procedures in Section 4.

(A2) Using sparsity of B, the matrix-vector products Bv and $B^\top\mu$ involve $m_B = \mathcal{O}(m)$ flops.

(A3) The matrix-vector product $C\mu$ can be evaluated by the following procedure:

$$v := B^\top\mu, \quad v := A^\dagger v, \quad C\mu := Bv. \tag{18}$$

Using **(A1)** and **(A2)**, it requires $n_{A^\dagger} + 2m_B$ flops.

Now we can asses computational costs of the individual steps of *Algorithm II*:

Step (a.1): The computations of D, p and q based on **(A1)** and **(A2)** require $mn_{A^\dagger}$, $n_{A^\dagger} + m_B$ and $n_{A^\dagger}$ flops, respectively.

Step (a.2): It is convenient to divide the computations of E and r into further two steps:

(a.2.1) Solve $CX = D^\top$ and $Cx = p$ by means of the CGM.

(a.2.2) Compute $E = DX$ and $r = Dx - p$.

The $m \times l$ matrix X and the m vector x can be stored because they are relatively small. Let us point out that the CGM is executed $(l+1)$-times in the step (a.2.1) and terminates at most after m iterations. Since each iteration requires one matrix-vector product $C\mu$ evaluated by (18), the step (a.2.1) requires $(l+1)m(n_{A^\dagger} + 2m_B)$ flops. The computational costs of the step (a.2.2) are not significant.

Step (a.3): The computational costs of this step are not significant. The linear system (17) can be solved e.g. by a direct method.

Step (b): The formulae (16) can be rewritten as

$$\lambda = x - X\alpha,$$

where x, X are the results of the step (a.2.1). Then the computational costs are not significant.

Step (c): This step requires $2n_{A^\dagger} + m_B$ flops.

Lemma 3 *Algorithm II requires* $\mathcal{O}((l+2)mn_{A^\dagger})$ *flops.*

Proof. Summarizing the previous discussion, we obtain that *Algorithm II* involves

$$(ml + 2m + 4)n_{A^\dagger} + 2(ml + m + 1)m_B$$

flops. The proof follows from the fact that m is much smaller than n, l is much smaller than m, $m_B = \mathcal{O}(m)$ and from the assumption $n_{A^\dagger} \geq n$. □

In this section, we shall show how to evaluate efficiently the matrix-vector products $A^\dagger v$, $N\alpha$ and $N^\top v$ in *Algorithm II* provided

$$A = A_x \otimes I_y + I_x \otimes A_y, \tag{19}$$

where A_x, A_y are circulant symmetric positive semi-definite matrices, I_x, I_y are identity matrices and $\otimes$ stands for the Kronecker product. The subscript x, y denotes that the corresponding matrix is of the order n_x, n_y, respectively. Since we shall use the FFT, we assume that n_x, n_y are powers of two.

Before we give multiplying procedures for A of the form (19), we introduce them for the case of the circulant matrix.

3.3 Circulant matrices and the DFT

The matrix A is called the *circulant matrix* if

$$A = \begin{pmatrix} a_1 & a_n & \dots & a_2 \\ a_2 & a_1 & \dots & a_3 \\ a_3 & a_2 & \dots & a_4 \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n-1} & \dots & a_1 \end{pmatrix},$$

i.e. each column of A is a cyclic shift of the column above to the bottom. Let us denote the first column of A by a, i.e. $a = (a_1, a_2, \dots, a_n)^\top$.

There are important conections between the circulant matrices and the DFT. The *DFT matrix* is defined by $F = (\omega^{(k-1)(l-1)})_{k,l=1}^n$, where $\omega = e^{-i2\pi/n}$ is an nth root of unity, i.e. $\omega^n = 1$. If v is the n vector, then its DFT is

$$\widehat{v} = Fv. \tag{20}$$

If n is a power of two, then it is possible to evaluate (20) by the FFT with $\mathcal{O}(n \log_2 n)$ flops [5]. In this case, we shall use instead of (20) the notation

$$\widehat{v} := \texttt{fft}(v).$$

Let us point out that F is symmetric and fulfils $F\overline{F}^\top = nI$ so that

$$v = \frac{1}{n} F \overline{\widehat{v}}. \tag{21}$$

In view of (21), the inverse DFT can be evaluated again by the FFT with $\mathcal{O}(n \log_2 n)$ flops that shall be denoted by

$$v := \texttt{ifft}(\widehat{v}).$$

A relationship between the circulant matrix and the DFT is given in the lemma.

Lemma 4 *Let A be the circulant matrix. Then*

$$A = F^{-1} \Lambda F, \tag{22}$$

where $\Lambda = diag(\widehat{a})$.

Proof. It is well-known that the Fourier transform changes translation operators onto modulation ones [8]. The equality of columns in $FA = \Lambda F$ represents this property. □

Let us point out that diagonal entries of Λ are eigenvalues of A and columns of F^{-1} are corresponding eigenvectors. Lemma 4 proves that eigenvalues of the circulant matrix can be computed very cheaply by the FFT of its first column a and, moreover, eigenvectors need not be computed at all because they are known apriori.

Lemma 5 *Let A be the circulant matrix, $\Lambda = diag(\widehat{a})$ and let $\Lambda^\dagger$ be defined so that the non-zero entries of Λ are inverted. Then*

$$A^\dagger = F^{-1}\Lambda^\dagger F \tag{23}$$

is the Moore–Penrose pseudoinverse to A.

Proof. It is easy to verify that $A^\dagger$ fulfils relations defining the Moore–Penrose pseudoinverse, see e.g. [5]. □

Using the lemma, we can propose the multiplying procedure which makes possible to evaluate efficiently the matrix-vector product $A^\dagger v$:

`Aplus_v` (for the circulant matrix)

Input: $\widehat{a}, v$

1° $v := \mathtt{fft}(v)$

2° $v := \Lambda^\dagger v$

3° $A^\dagger v := \mathtt{ifft}(v)$

Output: $A^\dagger v$

Since $\Lambda^\dagger$ is diagonal, the computational costs of `Aplus_v` are $\mathcal{O}(2n\log_2 n + n)$ flops.

Recall that we denoted by N a matrix whose columns span the null space of A. If A is the circulant matrix, we can suppose that N is composed by eigenvectors from F^{-1} coresponding to vanishing eigenvaluse of A, i.e. to vanishing entries of $\widehat{a}$. In order to determine the positions of desirable columns of F^{-1}, we intorduce the operation $\mathtt{ind}_{\widehat{a}}$,

$$\alpha \in \mathbb{R}^l \quad \Longleftrightarrow \quad v_\alpha := \mathtt{ind}_{\widehat{a}}(\alpha) \in \mathbb{R}^n,$$

so that the entries of α are put in v_α onto the positions of zeros in $\widehat{a}$ and the remaining entries of v_α vanish. Let us denote by $\mathtt{ind}_{\widehat{a}}^{-1}$ the reverse operation to $\mathtt{ind}_{\widehat{a}}$, i.e.

$$\alpha := \mathtt{ind}_{\widehat{a}}^{-1}(v_\alpha).$$

It is easy to verify that

$$N\alpha = F^{-1}\mathtt{ind}_{\widehat{\mathtt{a}}}(\alpha), \tag{24}$$

$$N^\top v = \mathtt{ind}_{\widehat{\mathtt{a}}}^{-1}(F^{-1}v). \tag{25}$$

The multiplying procedures for computations of $N\alpha$ and $N^\top v$ based on (24) and (25) read as follows:

`N_alpha` (for the circulant matrix)

Input: $\widehat{a}, \alpha$

1° $v_\alpha := \texttt{ind}_{\widehat{a}}(\alpha)$

2° $N\alpha := \texttt{ifft}(v_\alpha)$

Output: $N\alpha$

`Ntranspose_v` (for the circulant matrix)

Input: $\widehat{a}, v$

1° $v := \texttt{ifft}(v)$

2° $N^\top v := \texttt{ind}_{\widehat{a}}^{-1}(v)$

Output: $N^\top v$

The computational costs are $\mathcal{O}(n \log_2 n)$ flops for both `N_alpha` and `Ntranspose_v` because the operations $\texttt{ind}_{\widehat{\texttt{a}}}$ and $\texttt{ind}_{\widehat{a}}^{-1}$ do not require any flops.

3.4 Kronecker product

Let us consider the matrices A_x and A_y, respectively. Their *Kronecker product* (or tensor product) is defined by

$$A_x \otimes A_y = \begin{pmatrix} a^y_{11}A_x & \dots & a^y_{1n_y}A_x \\ \vdots & \ddots & \vdots \\ a^y_{n_y1}A_x & \dots & a^y_{n_yn_y}A_x \end{pmatrix},$$

where a^y_{kl} are entries of A_y. In other words, $A_x \otimes A_y$ is the $n \times n$ matrix (with $n = n_x n_y$) whose the (k,l)th block is $a^y_{kl}A_x$. A relationship between the Kronecker product and the matrix-matrix product is described by the following equality [5]:

$$(A_x \otimes A_y)(B_x \otimes B_y) = A_xB_x \otimes A_yB_y. \tag{26}$$

Let us point out that (26) implies

$$(A_x \otimes A_y)^{-1} = A_x^{-1} \otimes A_y^{-1} \tag{27}$$

provided non-singular A_x, A_y, respectively.

It is a very favourable feature of the Kronecker product that the matrix-vector product $(A_x \otimes A_y)v$ can be split into multiplications with the particular matrices A_x and A_y, respectively. To this end, we introduce the operation `vec`,

$$V = (v_1, \dots, v_{n_y}) \in \mathbb{R}^{n_x \times n_y} \quad \Longleftrightarrow \quad \texttt{vec}(V) = \begin{pmatrix} v_1 \\ \vdots \\ v_{n_y} \end{pmatrix} \in \mathbb{R}^{n_x n_y}.$$

Denote by $\mathtt{vec}^{-1}$ the reverse operation to $\mathtt{vec}$, i.e. if $v = \mathtt{vec}(V)$, then $V = \mathtt{vec}^{-1}(v)$.

Lemma 6 *It holds*

$$(A_x \otimes A_y)v = \mathtt{vec}(A_x V A_y^\top),$$

where $V = \mathtt{vec}^{-1}(v)$.

Proof. The proof follows comparing the right-hand sides of

$$A_x V A_y^\top = (a^y_{11}A_x v_1 + \ldots + a^y_{1n_y}A_x v_{n_y}, \ldots, a^y_{n_y1}A_x v_1 + \ldots + a^y_{n_yn_y}A_x v_{n_y})$$

and

$$(A_x \otimes A_y)v = \begin{pmatrix} a^y_{11}A_x v_1 + & \ldots & +a^y_{1n_y}A_x v_{n_y} \\ & \ldots & \\ a^y_{n_y1}A_x v_1 + & \ldots & +a^y_{n_yn_y}A_x v_{n_y} \end{pmatrix}.$$

□

The lemma yields that $(A_x \otimes A_y)v$ can be evaluated by two steps:

`AxtimesAy_v`

Input: $A_x, A_y, V := \mathtt{vec}^{-1}(v)$

1° $V := A_x V$

2° $V := V A_y^\top$

Output: $(A_x \otimes A_y)v := \mathtt{vec}(V)$

Denote n_{A_x}, n_{A_y} the number of flops to evaluate the matrix-vector products $A_x v_x$, $A_y v_y$, respectively. Since $\mathtt{vec}$ and $\mathtt{vec}^{-1}$ do not require any flops, the computational costs of `AxtimesAy_v` are $n_y n_{A_x} + n_x n_{A_y}$ flops.

3.5 Multiplying procedures

We shall combine the techniques from the previous two sections in order to obtain fast multiplying procedures for the matrix (19). We shall use the following notations: F_x, F_y denotes the DFT matrix of the order n_x, n_y, respectively; a_x, a_y denotes the first columns of the circulant matrices A_x, A_y, respectively.

Lemma 7 *Let* A_x, A_y *be the circulant matrices and* $A = A_x \otimes I_y + I_x \otimes A_y$. *Then*

$$A = F^{-1}\Lambda F, \tag{28}$$

where $F = F_x \otimes F_y$, $\Lambda = \Lambda_x \otimes I_y + I_x \otimes \Lambda_y$, $\Lambda_x = diag(\widehat{a}_x)$ *and* $\Lambda_y = diag(\widehat{a}_y)$.

Proof. Using $A_x = F_x^{-1}\Lambda_x F_x$, $A_y = F_y^{-1}\Lambda_y F_y$ (see Lemma 4) and (26), we obtain

$$A = F_x^{-1}\Lambda_x F_x \otimes F_y^{-1}F_y + F_x^{-1}F_x \otimes F_y^{-1}\Lambda_y F_y \; = \\ = (F_x^{-1} \otimes F_y^{-1})(\Lambda_x \otimes I_y + I_x \otimes \Lambda_y)(F_x \otimes F_y).$$

Then (27) proves the lemma. □

Let us point out that Lemma 7 is formally the same as Lemma 4. Therefore the lemma analogous to Lemma 5 holds.

Lemma 8 *Let A_x, A_y be the circulant matrices, $\Lambda_x = diag(\widehat{a}_x)$, $\Lambda_y = diag(\widehat{a}_y)$, $F = F_x \otimes F_y$, and let $\Lambda^\dagger$ be defined so that the non-zero entries of $\Lambda = \Lambda_x \otimes I_y + I_x \otimes \Lambda_y$ are inverted. Then*

$$A^\dagger = F^{-1}\Lambda^\dagger F, \tag{29}$$

is the Moore–Penrose pseudoinverse to $A = A_x \otimes I_y + I_x \otimes A_y$.

Using the lemma, we can propose the multiplying procedure to evaluate the matrix-vector product $A^\dagger v$ that is analogous to `Aplus_v` for the circulant matrix. Combining it with `AxtimesAy_v`, we obtain:

`Aplus_v`

Input: $\widehat{a}_x, \widehat{a}_y, V := \mathtt{vec}^{-1}(v)$

1° $V := \mathtt{fft}(V)$

2° $V := \mathtt{fft}(V^\top)^\top$

3° $V := \mathtt{vec}^{-1}(\Lambda^\dagger \mathtt{vec}(V))$

4° $V := \mathtt{ifft}(V)$

5° $V := \mathtt{ifft}(V^\top)^\top$

Output: $A^\dagger v := \mathtt{vec}(V)$

Here, we suppose that `fft` and `ifft` are independently performed for the individual columns of the matrices V or $V^\top$, respectively.

Lemma 9 *The multiplying procedure* `Aplus_v` *requires $\mathcal{O}(2n \log_2 n + n)$ flops, where $n = n_x n_y$.*

Proof. Recall that V is the $n_x \times n_y$ matrix. The steps 1° and 2° involve

$$n_y \mathcal{O}(n_x \log_2 n_x) + n_x \mathcal{O}(n_y \log_2 n_y) = \mathcal{O}(n \log_2 n) \tag{30}$$

flops. The same flops are required by the steps 4° and 5°. Since $\Lambda^\dagger$ is the diagonal matrix, the step 3° requires n flops. □

Let us turn our attention to the matrix N whose columns span the the null space of A. We start with an auxiliary lemma that makes possible to characterize N by means of the Kronecker product.

Lemma 10 *Let A_x, A_y be the symmetric positive semi-definite matrices with $l_x = \dim \mathcal{N}(A_x)$, $l_y = \dim \mathcal{N}(A_y)$ and $l_x \geq 1, l_y \geq 1$, respectively. Let N_x, N_y be the matrices whose columns span the null spaces of A_x, A_y, respectively. Then the columns of*

$$N = N_x \otimes N_y \tag{31}$$

span the null space of $A = A_x \otimes I_y + I_x \otimes A_y$.

Proof. Using (26), we obtain

$$AN = A_x N_x \otimes N_y + N_x \otimes A_y N_y = 0 \otimes N_y + N_x \otimes 0 = 0.$$

The lemma follows from the fact that the number of columns $l = l_x l_y$ of N is the same as $\dim \mathcal{N}(A)$. □

If A_x, A_y are the cirulant matrices, we can identify N_x, N_y with the columns of F_x^{-1}, F_y^{-1} corresponding to the positions of vanishing entries of $\widehat{a}_x$, $\widehat{a}_y$, respectively. To this end, we introduce the operation $\mathtt{Ind}_{\widehat{a}_x,\widehat{a}_y}$,

$$\alpha \in \mathbb{R}^{l_x l_y} \quad \Longleftrightarrow \quad V_\alpha := \mathtt{Ind}_{\widehat{a}_x,\widehat{a}_y}(\alpha) \in \mathbb{R}^{n_x \times n_y},$$

so that the entries of α are taken as the entries of V_α with the row indeces corresponding to the positions of zeros in $\widehat{a}_x$ and with the column indeces correponding to the positions of zeros in $\widehat{a}_y$. The remaining entries of V_α vanish. Let us denote by $\mathtt{Ind}^{-1}_{\widehat{a}_x,\widehat{a}_y}$ the reverse operation to $\mathtt{Ind}_{\widehat{a}_x,\widehat{a}_y}$, i.e.

$$\alpha := \mathtt{Ind}^{-1}_{\widehat{a}_x,\widehat{a}_y}(V_\alpha).$$

Using (31) and Lemma 6, it is easy to verify that

$$N\alpha = \mathtt{vec}(F_x^{-1}\mathtt{Ind}_{\widehat{a}_x,\widehat{a}_y}(\alpha)F_y^{-1}), \tag{32}$$

$$N^\top v = \mathtt{Ind}^{-1}_{\widehat{\mathtt{a}}_\mathtt{x},\widehat{\mathtt{a}}_\mathtt{y}}(F_x^{-1}\mathtt{vec}^{-1}(v)F_y^{-1}). \tag{33}$$

The multiplying procedures for computations of $N\alpha$ and $N^\top v$ based on (32) and (33) read as follows:

N_alpha

Input: $\widehat{a}_x, \widehat{a}_y, \alpha$

1° $V_\alpha := \mathtt{Ind}_{\widehat{a}_x,\widehat{a}_y}(\alpha)$

2° $V_\alpha := \mathtt{ifft}(V_\alpha)$

3° $V_\alpha := \mathtt{ifft}(V_\alpha^\top)^\top$

Output: $N\alpha := \mathtt{vec}(V_\alpha)$

Ntranspose_v

Input: $\widehat{a}_x, \widehat{a}_y, V := \mathtt{vec}^{-1}(v)$

1° $V := \mathtt{ifft}(V)$

2° $V := \mathtt{ifft}(V^\top)^\top$

3° $N^\top v := \mathtt{Ind}^{-1}_{\widehat{a}_x,\widehat{a}_y}(V)$

Output: $N^\top v$

Lemma 11 *The multiplying procedures* N_alpha *and* Ntranspose_v *require* $\mathcal{O}(n \log_2 n)$ *flops, where* $n = n_x n_y$.

Proof. The computational costs are the same as (30) because the operations $\mathtt{Ind}_{\widehat{a}_x,\widehat{a}_y}$ and $\mathtt{Ind}^{-1}_{\widehat{a}_x,\widehat{a}_y}$ do not require any flops. □

3.6 Complexity of algorithms

Now we can asses the computational costs of *Agorithm I* and *Algorithm II* for solving saddle-point system (1) provided A of the form (19). Using Lemma 9 and Lemma 11, we can suppose that both $n_{A^{-1}}$ from Lemma 2 and $n_{A^\dagger}$ from Lemma 3 equal to $\mathcal{O}(n \log_2 n)$. We obtain:

Theorem 4 *Algorithm I for solving (1) with non-singular A requires* $\mathcal{O}((m + 2)n \log_2 n)$ *flops.*

Theorem 5 *Algorithm II for solving (1) with singular A requires* $\mathcal{O}((l+2)mn \log_2 n)$ *flops.*

Let us point out that faster implementations are possible. Further accelerations can be done in the general form of the algorithm as well as in its fast implementation.

Accelerations in the general form of *Algorithm II*:

- The CGM is performed $(l + 1)$-times in the step (a.2.1). If it is done parallel then the computational costs of *Algorithm II* can be reduced to $\mathcal{O}(2mn \log_2 n)$ flops.
- The same acceleration can be achieved by using the CGM with multiple right-hand sides in the step (a.2.1).
- If an "ideal" preconditioner in the CGM is used, then m iterations of the CGM can be reduced onto $\mathcal{O}(1)$ iteratons. This one reduces the computational costs of *Algorithm II* to $\mathcal{O}(2n \log_2 n)$ flops.

Accelerations in the fast implementation of *Algorithm II*:

- The multiplying procedures `Aplus_v`, `N_alpha` and `Ntranspose_v` involve to carry out many FFT that can be done parallel. If we have to our disposal k CP units, $k \leq \max\{n_x, n_y\}$, then the computational costs of *Algorithm II* can be reduced to $\mathcal{O}(2k^{-1}n \log_2 n)$ flops.

4 Numerical experiments

4.1 Model problem

Let us consider the following PDEs problem:

$$-\Delta u + cu = f \quad \text{in } \omega \tag{34}$$

$$u = g \quad \text{on } \partial\omega, \tag{35}$$

where f, g are sufficiently smooth functions, c is a non-negative value and ω is a domain with sufficiently smooth boundary. The numerical experiments presented below are performed with $\omega = \{(x, y) \in \mathbb{R}^2 : (x/0.2)^2 + (y/0.3)^2 \leq 1\}$, $f(x, y) = 1$ on $\langle -0.5, 0.5\rangle \times \langle -0.5, 0.5\rangle$, $f(x, y) = 0$ elsewhere and $g \equiv 0$.

We solve the problem (34), (35) by means of the fictitious domain method with boundary Lagrange multipliers [4]. We plug ω into the new rectangular domain $\Omega = \langle -1, 1\rangle \times \langle -1, 1\rangle$, where we rewrite (34), (35) into a weak formulation. Then we discretize the problem by means of the periodized orthonormal compactly supported wavelets of the tensor product type [7]. This discretization leads to the saddle-point system (1) with the stiffness matrix A of the form (19). If $c = 0$ then A is singular with $l = 1$ $(= \dim \mathcal{N}(A))$. If $c \neq 0$ then A is non-singular.

All numerical experiments are carried out by Matlab. The tables bellow contains the following informations:

n	... the order of A,
m	... the number of the rows in B,
time	... the CPU time in seconds,
CG_step	... the number of the conjugate gradient steps,
err_u	... the relative norm of the residual to the first part of (1),
err_λ	... the relative norm of the residual to the second part of (1).

The CGM has been terminated if the relative norm of the residual in the solved linear system fell under the terminating tolerance $tol = 10^{-4}$. The dependence of the CPU time onto n, m and l is in agreement with the statements of Theorem 4 and Theorem 5.

n	m	time	CG_step	err_u	err_λ
1024	64	0.03	13	4.5e-10	1.7e-5
2048	88	0.05	17	1.6e-10	1.1e-5
4096	128	0.06	17	2.2e-10	8.1e-6
8192	180	0.17	23	6.9e-10	4.4e-6
16384	256	0.28	21	9.8e-10	3.6e-6
32768	360	0.67	27	2.9e-10	1.9e-6
65536	512	2.27	27	4.5e-9	1.0e-6
131072	716	7.22	35	1.3e-8	6.9e-7
262144	1024	14.72	34	1.9e-8	3.9e-7
524288	1432	35.70	45	5.3e-8	2.4e-7
1048576	2048	65.56	41	7.8e-8	1.4e-7
2097152	2868	173.53	54	2.2e-8	9.1e-8
4194304	4096	337.95	48	3.4e-8	4.2e-8

Tabulka 1: The reults of *Algorithm I*, i.e. $c = 1$.

n	m	time	CG_step	err_u	err_λ
1024	64	0.06	10+16	1.8e-10	3.4e-5
2048	88	0.08	14+21	5.5e-10	1.1e-5
4096	128	0.13	13+21	1.0e-10	1.0e-6
8192	180	0.30	20+29	2.8e-10	6.6e-6
16384	256	0.48	18+25	4.3e-10	2.6e-6
32768	360	1.20	25+33	1.1e-10	2.1e-6
65536	512	5.58	25+33	1.8e-10	1.4e-6
131072	716	15.17	33+43	5.0e-9	7.9e-7
262144	1024	29.33	29+38	7.6e-8	4.0e-7
524288	1432	73.41	43+53	2.0e-8	2.5e-7
1048576	2048	133.75	38+49	3.2e-8	1.5e-7
2097152	2868	347.41	50+62	8.0e-8	8.0e-8
4194304	4096	655.00	42+57	2.7e-8	3.1e-8

Tabulka 2: The reults of *Algorithm II*, i.e. $c = 0$.

4.2 Preconditioning

We can use the precondioned CGM [5] in *Algorithm II* and *Algorithm I*. Let us point out that the linear systems with the matrix

$$C = BA^{\dagger}B^{\top}$$

are solved, where that the matrix C is not assembled. Moreover, B is sparse and A has the special structure. These facts give restrictions on an appropriate selection of an efficient preconditioner M.

We have tested three preconditioners:

- Preconditioner I:
$$M^{-1} = (B^{\dagger})^{\top}AB^{\dagger}$$
- Preconditioner II:
$$M^{-1} = BAB^{\top}$$
- Preconditioner III:
$$M^{-1} = \widetilde{B}A\widetilde{B}^{\top},$$
where only entry $\tilde{b}_{ij}$ is non-vanishing in the i-th row of $\widetilde{B}$ so that
$$j = round\,(\sum_{k=1}^{n} kb_{ik} / \sum_{k=1}^{n} b_{ik}),$$
$$\tilde{b}_{ij} = \sum_{k=1}^{n} b_{ik}.$$

		$c=1$		$c=0$	
n	m	time(sec.)	CG steps	time(sec.)	CG steps
1024	64	0.13	7	0.20	7+10
2048	88	0.28	9	0.55	8+12
4096	128	1.09	9	2.16	8+12
8192	180	3.31	12	6.61	10+15
16384	256	20.02	13	39.83	10+16

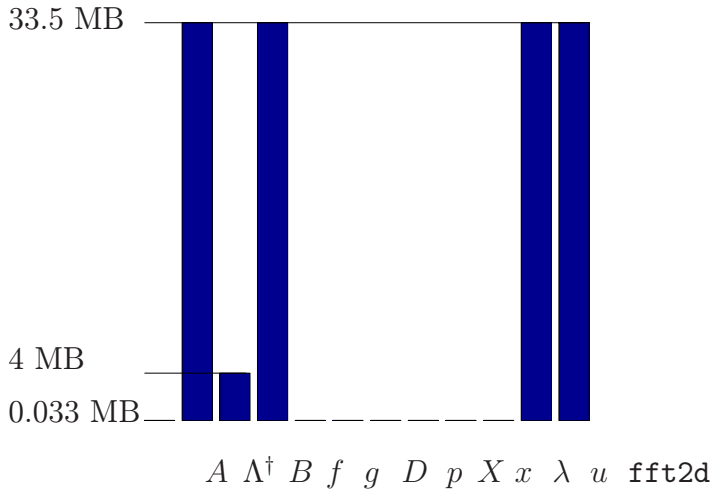
Tabulka 3: *Preconditioner I.*

		$c = 1$		$c = 0$	
n	m	time(sec.)	CG steps	time(sec.)	CG steps
1024	64	0.06	7	0.09	7+11
2048	88	0.08	9	0.11	8+12
4096	128	0.09	10	0.16	9+13
8192	180	0.19	12	0.30	10+15
16384	256	0.31	14	0.53	11+16
32768	360	0.64	14	1.19	12+18
65536	512	3.64	18	6.45	14+19
131072	716	7.75	19	15.94	16+23
262144	1024	17.63	21	32.56	18+23
524288	1432	37.11	24	73.06	20+28
1048576	2048	75.83	24	149.89	21+28
2097152	2868	184.80	29	351.07	25+31
4194304	4096	406.88	29	754.59	25+32

Tabulka 4: *Preconditioner II.*

		$c = 1$		$c = 0$	
n	m	time(sec.)	CG steps	time(sec.)	CG steps
1024	64	0.05	8	0.06	12+11
2048	88	0.06	12	0.11	16+16
4096	128	0.11	11	0.20	15+15
8192	180	0.22	15	0.36	18+18
16384	256	0.39	15	0.78	18+18
32768	360	0.81	19	1.73	21+22
65536	512	4.30	19	10.05	21+23
131072	716	11.28	22	24.06	26+28
262144	1024	23.16	23	50.05	25+28
524288	1432	52.52	25	114.55	29+29
1048576	2048	120.08	26	255.72	29+30
2097152	2868	296.05	29	627.92	33+33
4194304	4096	717.33	31	1591.95	35+35

Tabulka 5: *Preconditioner III.*



Obrázek 1: Memory requirements.

4.3 Memory requirements

The memory requirements for the problem with $n = 4194304$ and $m = 4096$ are shown in the graph of Fig. 1.

Reference

[1] Brezzi, F., Fortin, M.: *Mixed and hybrid finite element methods.* Springer, New York, 1991.

[2] Fortin, M., Glowinski, R.: *Augmented Lagrangian methods: Applications to the numerical solution of boundary-value problems.* North-Holland, New York, 1983.

[3] Daubechies, I.: *Ten lectures on wavelets.* SIAM, Philadelphia, 1992.

[4] Glowinski, R., Pan, T., Periaux, J.: *A fictitious domain method for Dirichlet problem and applications.* J. Appl. Mech. and Eng. **111** (1994), 283–303.

[5] Golub, G. H., Van Loan, C. F.: *Matrix computation.* The Johns Hopkins University Press, Baltimore, 1996, 3rd ed.

[6] Gould, N. I. M.: *On practical conditions for the existence and uniqueness of solutions to the general equality quadratic programming problem.* Mathematical Programming **32** (1985), 90–99.

[7] Kučera, R.: *Wavelet solution of elliptic PDEs.* In: Proc. Matematyka v Naukach Technicznych i Przyrodniczych (2000), Krynica, AGH Krakow, 55–62.

[8] Rudin, W.: *Real and complex analysis.* McGraw–Hill, New York, 1987, 3rd ed.

Univ. Palacki. Olomuc., Fac. rer. nat.,
Dept of Math. Anal. and Appl. of Math.
ODAM (2003) 77–83

Úloha s oboustranným kontaktem a nemonotonním třením

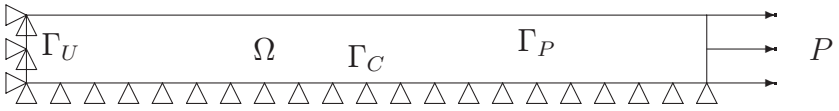
Zuzana MORÁVKOVÁ

Katedra matematiky a deskriptivní geometrie
VŠB TU Ostrava, 17. listopadu 15
708 33 Ostrava–Poruba,
e-mail: zuzana.moravkova@vsb.cz

1 Klasická formulace

Úloha s oboustranným kontaktem a nemonotonním třením je úloha rovinné napjatosti s klasickými okrajovými podmínkami, s podmínkou oboustranného kontaktu a dále s nemonotonním třením.

Oblast $\Omega = 100\,\text{mm} \times 10\,\text{mm}$ a rozdělení hranice $\partial\Omega$ na části Γ_U, Γ_P a Γ_C vyplývá z obr. 1.1.



obr. 1.1

Těleso je z homogenního izotropního materiálu, jehož deformace vede na úlohu rovinné napjatosti s modulem pružnosti $E = 2.1\,10^5\ \text{N/mm}^2$, Poissonovou konstantou $\sigma = 0.3$ a tloušťkou $t = 5\,\text{mm}$. Těleso je upevněno na části hranice Γ_U, t.j. jsou zde předepsána nulová posunutí v obou směrech:

$$u_i = 0 \quad \text{na } \Gamma_U,\ i = 1, 2\,. \tag{1}$$

Na části Γ_P působí zatížení $T = (P, 0)$, kde $P \in (L^2(\Gamma_P))^2,\ P \geq 0$ s.v. na Γ_P (viz obr. 1.1):

$$T_1 = P \quad \text{na } \Gamma_P\,. \tag{2}$$

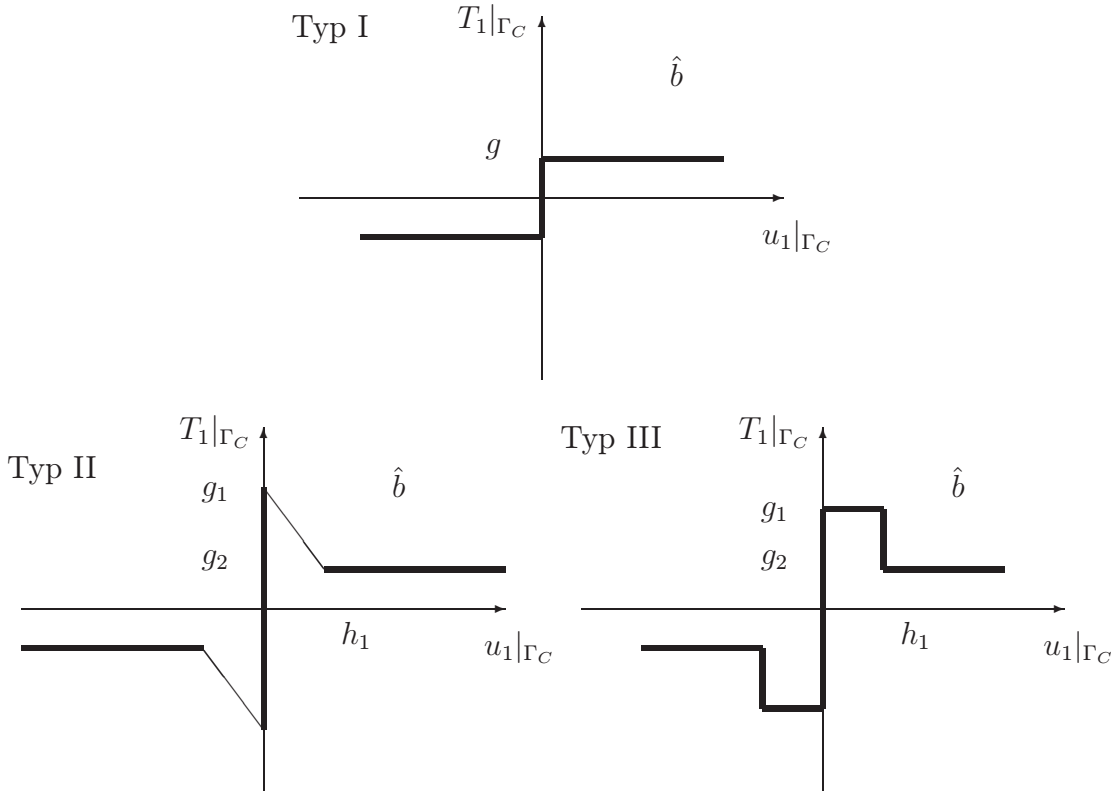
Předpokládejme, že objemové síly jsou nulové.

Na části Γ_C leží na tuhé podložce. Budeme uvažovat *oboustrannou* kontaktní podmínku:

$$u_2 = 0 \quad \text{na } \Gamma_C\,. \tag{3}$$

Dále na Γ_C předepíšeme podmínku *nemonotonního tření*, budeme uvažovat tři typy tření (viz obr. 1.2):

$$-T_1(x) \in \hat{b}(u_1(x)) \qquad x \in \Gamma_C\,. \tag{4}$$



obr. 1.2

Okrajové podmínky (1)–(4) ještě doplníme o systém rovnic rovnováhy:

$$\frac{\partial \tau_{ij}(u)}{\partial x_j} = 0 \quad \text{v } \Omega, \qquad i = 1, 2\,. \tag{5}$$

Tenzor napětí $\{\tau_{ij}(u)\}_{i,j=1}^2$ je svázán s tenzorem malých posunutí $\{\varepsilon_{ij}(u)\}_{i,j=1}^2$ pomocí lineárního Hookeova zákona:

$$\tau_{ij}(u) = \frac{E\sigma}{1-\sigma^2}\delta_{ij}\vartheta + \frac{E}{1+\sigma}\varepsilon_{ij}(u), \quad i, j = 1, 2\,. \tag{6}$$

Zde $\vartheta := \varepsilon_{ii}(u)$ je stopa tenzoru $\{\varepsilon_{ij}(u)\}_{i,j=1}^2$ a δ_{ij} je Kronekerův symbol.

Klasickým řešením úlohy s nemonotonním třením a oboustranným kontaktem nazveme pole posunutí $u = (u_1, u_2)$ splňující okrajové podmínky (1), (2), (3), (4), rovnice rovnováhy (5) a lineární Hookeův zákon (6).

2 Slabá formulace

V tomto odstavci zformulujeme úlohu pomocí hemivariační rovnice.

Zaveďme nejprve tato označení:

$$\begin{aligned} V &= \{v \in (H^1(\Omega))^2 |\ v = 0 \text{ na } \Gamma_U, \quad v_2 = 0 \text{ na } \Gamma_C\} \\ a(u,v) &= \int_\Omega \tau_{ij}(u)\varepsilon_{ij}(v)\,dx\,, \\ L(v) &= \int_{\Gamma_P} Pv\,ds\,. \end{aligned}$$

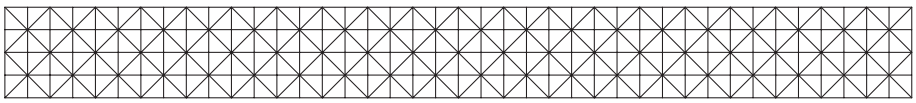
Slabá formulace úlohy s *nemonotonním třením* a *oboustranným kontaktem* je dána následující hemivariační rovnicí:

$$\begin{cases} \textit{Nalézt } (u,\Xi) \in V \times L^2(\Gamma_C) \textit{ takové, že} \\ a(u,v) + \int_{\Gamma_C} \Xi v_1 dx_1 = L(v) \quad \forall v \in V \\ \Xi(x) \in \hat{b}(u_1(x)) \quad \textit{pro s.v. } x \in \Gamma_C\,. \end{cases} \tag{7}$$

3 Diskretizace

Nyní popíšeme konstrukci dělení Ω a Γ_C a volby konečněprvkových prostorů, ve kterých poté hledáme řešení aproximované hemivariační nerovnice. Dále popíšeme aproximaci operátoru P a diskretizaci úlohy (7).

Nechť $\{\mathcal{D}_h\}$, $h \to 0+$ je systém *regulárních* triangulací oblasti $\overline{\Omega}$ (viz obr. 3.1).



obr. 3.1

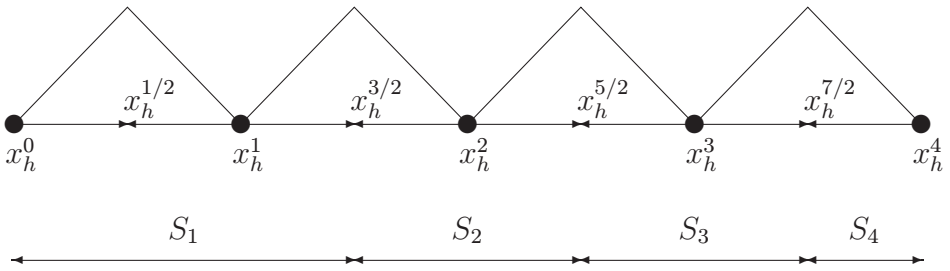
Označme V_h prostor všech spojitých, po částech lineárních vektorových funkcí nad dělením $\mathcal{D}_h$, který je aproximací prostoru V:

$$\begin{aligned} V_h = \{v_h = (v_{h1}, v_{h2})) \in (C(\overline{\Omega}))^2 |\ v_h|_T \in (P_1(T))^2 \ \ \forall T \in \mathcal{D}_h; \\ v_h = 0 \text{ na } \Gamma_U, \quad v_{h2} = 0 \text{ na } \Gamma_C\}\,. \end{aligned}$$

Označme $\{x_h^i\}_{i=1}^m$ množinu všech uzlů dělení $\mathcal{D}_h$ na části hranice $\overline{\Gamma}_C \backslash \overline{\Gamma}_U$. Pak zbývá popsat konstrukci prostoru Y_h. Nechť $x_h^{i+1/2}$ je bod ve středu intervalu

$[x_h^i, x_h^{i+1}]$, $i = 0, \ldots, m-1$. Dělení $\mathcal{T}_h$ části hranice $\overline{\Gamma}_C$ je tvořeno úsečkami S_i spojující body $x_h^{i-1/2}$, $x_h^{i+1/2}$, $i = 2, \ldots, m-1$ s následující úpravou pro krajní prvky S_1 a S_m (viz obr. 3.2):

$$S_1 = [x_h^0, x_h^{3/2}], \quad S_m = [x_h^{m-1/2}, x_h^m]\,.$$



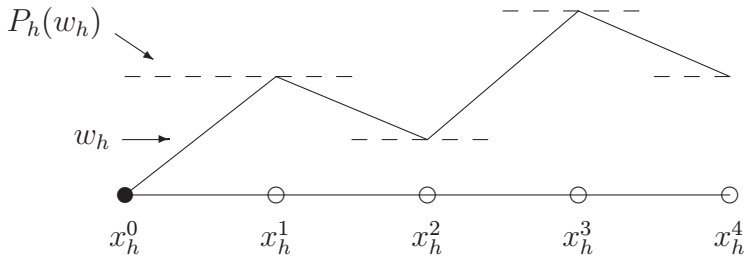
obr. 3.2

Nad každým takovým dělením $\mathcal{T}_h$ definujeme prostor Y_h všech po částech konstantních funkcí, přičemž hodnoty v bodech $\{x_h^i\}_{i=1}^m$ jsou stupně volnosti. Prostor $W_h := \Pi V_h$ je tvořen spojitými po částech lineární skalárními funkcemi nad dělením $\{x_h^i\}_{i=0}^m$, jež nabývají hodnoty 0 v prvním uzlu $x_h^0 \in \overline{\Gamma}_U$. Díky definici Y_h je také zřejmé, že $\dim W_h = \dim Y_h$.

Zavedem zobrazení $P_h:\ W_h \to Y_h$ P_h následujícím způsobem:

$$P_h(w_h) = \sum_{i=1}^{m} w_h(x_h^i)\chi_{S_i}(x_1) \quad w_h \in W_h\,,$$

kde χ_{S_i} je charakteristická funkce vnitřku S_i. Toto zobrazení nahradí funkci $w_h \in W_h$ její po částech konstantní Lagrangovou interpolaci nad $\mathcal{T}_h$ (viz obr. 3.3).



obr. 3.3

Pak tato *diskrétní hemivariační rovnice* aproximuje úlohu (7):

$$\begin{cases} \textit{Nalézt } (u_h, \Xi_h) \in V_h \times Y_h \textit{ takové, že} \\ a(u_h, v_h) + \int_{\Gamma_C} \Xi_h P_h v_{h1} dx_1 = L(v_h) \qquad \forall v_h = (v_{h1}, v_{h2}) \in V_h \\ \Xi_h(x_h^i) \in \hat{b}(P_h(u_{h1})(x_h^i)) \quad \forall i = 1, \ldots, m\,. \end{cases} \tag{8}$$

4 Algebraická reprezentace

Nyní krátce uvedeme algebraickou podobu problému (8). Předpokládáme, že body $\{x_h^i\}_{i=0}^m$ tvoří ekvidistantní dělení $\overline{\Gamma}_C$. Označme $\dim V_h = n$, $\dim Y_h = m$.

Nadefinujeme-li $\Xi_i := c_i \Xi_i$, kde $c_1 = 3/2h$, $c_2 = \cdots = c_{m-1} = h$, $c_m = h/2$, pak lze problém (8) zapsat v následovně:

$$\begin{cases} \textit{Nalézt } (\boldsymbol{u}, \boldsymbol{\Xi}) \in \mathbb{R}^n \times \mathbb{R}^m \textit{ takové, že} \\ (\boldsymbol{A}\boldsymbol{u}, \boldsymbol{v})_{\mathbb{R}^n} + (\boldsymbol{\Xi}, \boldsymbol{\Lambda}(\boldsymbol{v}))_{\mathbb{R}^m} = (\boldsymbol{f}, \boldsymbol{v})_{\mathbb{R}^n} \quad \forall \boldsymbol{v} \in \mathbb{R}^n \\ \Xi_i \in c_i \hat{b}((\boldsymbol{\Lambda}\boldsymbol{u})_i) \quad \forall i = 1, \ldots, m\,. \end{cases} \tag{9}$$

Protože bilineární forma a je symetrická, zkonstruujeme diskrétní superpotencionál $\mathcal{L}$, který je součtem kvadratické části a lipschitzovsky spojité perturbace Ψ, definované pomocí obdélníkové formule:

$$\Phi(v_h) \approx \sum_{i=1}^{m} c_i \int_0^{P_h(\Pi v_h)(x_h^i)} b(t)dt := \Psi(\boldsymbol{v}). \tag{10}$$

Nechť Φ je primitivní funkce k funkci b:

$$\Phi(x) = \int_0^x b(t)\, dt\,.$$

Pak

$$\Psi(\boldsymbol{v}) = \sum_i c_i \Phi((\boldsymbol{\Lambda}\boldsymbol{v})_i)\,.$$

Matice $\boldsymbol{\Pi}$ reprezentující operátor $\Pi_{|_{V_h}}$ má následující tvar:

$$(\boldsymbol{O}_{(m\times(n-m))} \quad |\boldsymbol{I}_{(m\times m)})\,.$$

pokud x_1-složky vektoru posunutí $\boldsymbol{v}$ v bodech x_h^i, $i = 1, \ldots, m$ jsou řazeny na konci vektoru $\boldsymbol{v}$. Symbol $\boldsymbol{I}$ značí jednotkovou a $\boldsymbol{O}$ nulovou matici uvedených rozměrů. Pak $(\boldsymbol{\Lambda}\boldsymbol{v})_i$:= x_1-složka vektoru posunutí $\boldsymbol{v}$ v bodech x_h^i, $i = 1, \ldots, m$.

Diskrétní superpotencionál $\mathcal{L}$ odpovídající algebraické hemivariační rovnici má tvar:

$$\mathcal{L}(\boldsymbol{v}) = \frac{1}{2}(\boldsymbol{A}\boldsymbol{v}, \boldsymbol{v})_{\mathbb{R}^n} - (\boldsymbol{f}, \boldsymbol{v})_{\mathbb{R}^n} + \Psi(\boldsymbol{v}), \quad \boldsymbol{v} \in \mathbb{R}^n\,. \tag{11}$$

Namísto úlohy (9) pak řešíme tento problém:

$$\begin{cases} \textit{Nalézt } \boldsymbol{u} \in \mathbb{R}^n \textit{ takové, že} \\ \boldsymbol{0} \in \overline{\partial}\mathcal{L}(\boldsymbol{u})\,. \end{cases} \tag{12}$$

kde $\overline{\partial}$ je zobecněný Clarkův gradient.

Pokud zobrazení P_h zobrazuje W_h *na* Y_h, pak oba problémy (9) a (12) jsou ekvivalentní, za předpokladu existence jednostranných limit $b(\xi\pm)$ pro každé $\xi \in \mathbb{R}$. Problém (12) budeme řešit nehladkou variantou Newtonovy metody.

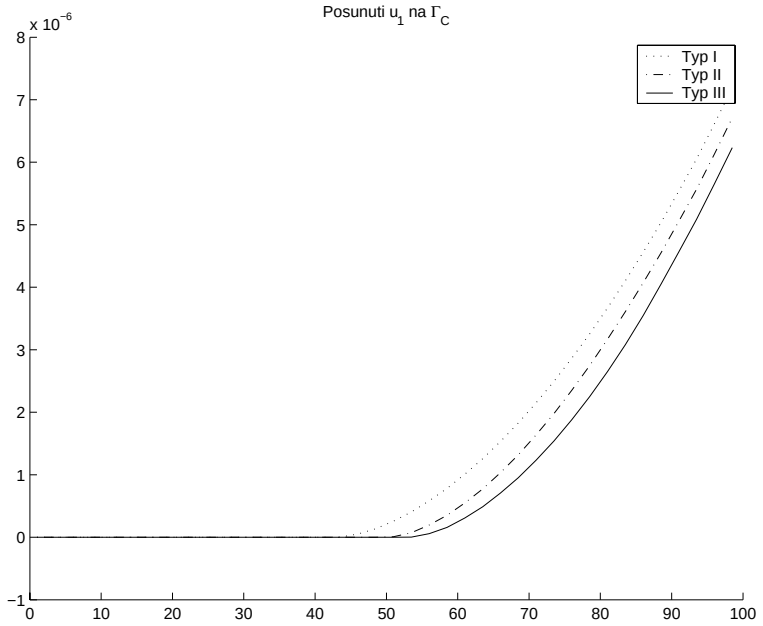
Z (12) obdržíme vektor posunutí $\boldsymbol{u}$, pak vektor $\boldsymbol{\Xi}$ vypočteme ze vztahu:

$$\boldsymbol{\Lambda}^T\boldsymbol{\Xi} = \boldsymbol{f} - \boldsymbol{A}\boldsymbol{u}\,.$$

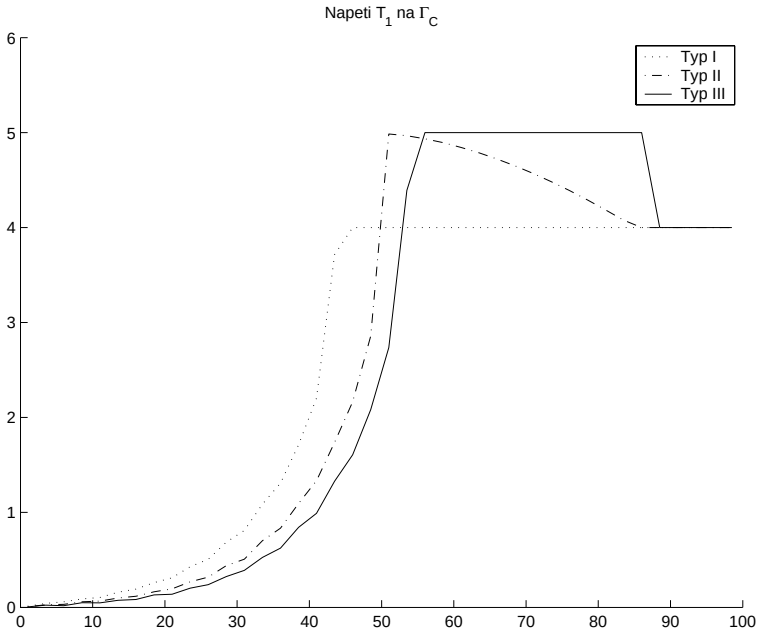
Tímto způsobem získáme obě složky řešení (9).

5 Numerické výsledky

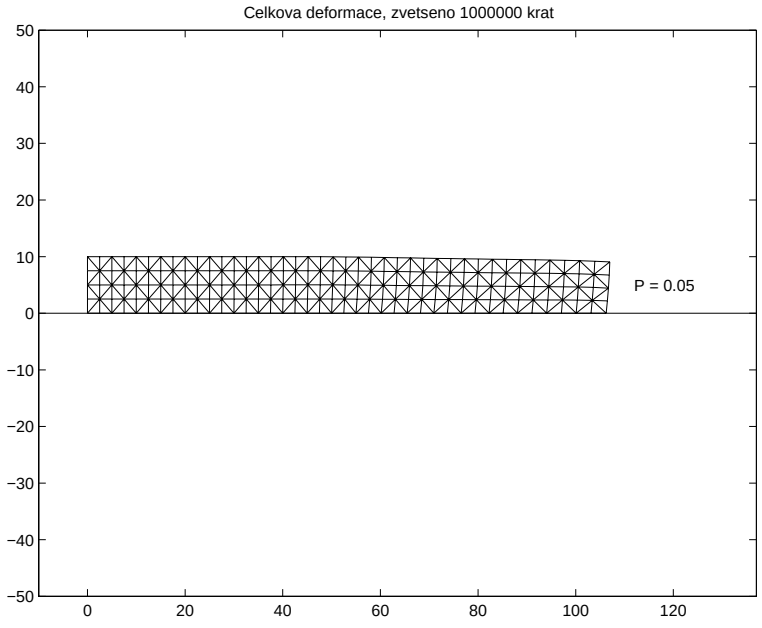
Budeme uvažovat dvoudimenzionální těleso z homogenního izotropního materiálu, jehož deformace vede na úlohu zobecněné rovinné napjatosti. Rozměry tělesa, jeho tloušťka a materiálové konstanty jsou uvedeny v odstavci 1, hodnotu zatížení P zvolíme 0.05 N/mm^2 a hodnoty parametrů v diagramech v obr. 1.1 takto: $h_1 = 4 \cdot 10^{-6}$ $g_1 = 5$ $g_2 = 4$.



obr. 5.1



obr. 5.2



obr. 5.3

Reference

[1] Haslinger, J., Miettinnen, M., Panagiotopoulos, P. D.: *Finite Element Method for Hemivariational Inequalities.* Nonconvex Optimization and its Applications 35, Kluwer Academic Publishers,1999.

[2] Lukšan, L., Vlček, J.: *A bundle–Newton method for nonsmooth unconstrain minimization.* Math. Prog. **83** (1998) 373–391.

[3] Lukšan, L., Vlček, J.: *A Bundle–Type Algorithms for Nonsmooth Optimization*, Tech. Report No. V-718, Prague, 1997.