

VĚDECKÉ SPISY VYSOKÉHO UČENÍ TECHNICKÉHO V BRNĚ

Edice Habilitační a inaugurační spisy, sv. 541

ISSN 1213-418X

Radek Kučera

OPTIMALIZAČNÍ ALGORITMY V KONTAKTNÍCH ÚLOHÁCH

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
Fakulta strojního inženýrství

doc. RNDr. Radek Kučera, Ph.D.

OPTIMALIZAČNÍ ALGORITMY V KONTAKTNÍCH ÚLOHÁCH

OPTIMIZATION ALGORITHMS
IN CONTACT PROBLEMS

TEZE PŘEDNÁŠKY
K PROFESORSKÉMU JMENOVACÍMU ŘÍZENÍ
V OBORU
APLIKOVANÁ MATEMATIKA



BRNO 2016

KLÍČOVÁ SLOVA

optimalizační algoritmy, metoda aktivních množin, metoda vnitřních bodů, semi-hladká Newtonova metoda, kontaktní úlohy, tření, úlohy proudění, skluzové podmínky

KEYWORDS

optimization algorithms, active set method, interior point method, semi-smooth Newton method, contact problems, friction, flow problems, stick-slip conditions

© Radek Kučera, 2016

ISBN 978-80-214-5388-3

ISSN 1213-418X

Obsah

| | |
|--|-----------|
| Představení autora | 4 |
| 1 Úvod | 5 |
| 2 Formulace úloh | 7 |
| 2.1 Kontaktní úlohy se třením | 7 |
| 2.1.1 Spojitá formulace | 7 |
| 2.1.2 Aproximace metodou konečných prvků | 9 |
| 2.1.3 Duální algebraické úlohy | 9 |
| 2.2 Úlohy proudění s podmínkou skluzu | 11 |
| 2.2.1 Spojitá formulace | 11 |
| 2.2.2 Aproximace metodou smíšených konečných prvků | 12 |
| 2.2.3 Duální algebraická úloha | 13 |
| 3 Optimalizační algoritmy | 13 |
| 3.1 Metoda aktivních množin | 14 |
| 3.1.1 Podmínky optimality | 14 |
| 3.1.2 Algoritmus KPRGP | 16 |
| 3.1.3 Projekce | 18 |
| 3.1.4 Numerické experimenty | 19 |
| 3.2 Metoda vnitřních bodů | 20 |
| 3.2.1 Podmínky optimality | 20 |
| 3.2.2 Algoritmus PF | 22 |
| 3.2.3 Předpodmínění | 23 |
| 3.2.4 Implementace | 25 |
| 3.2.5 Numerické experimenty | 26 |
| 3.3 Semi-hladká Newtonova metoda | 27 |
| 3.3.1 Podmínky optimality | 28 |
| 3.3.2 Diferencovatelnost „slantly“ | 28 |
| 3.3.3 Duální implementace | 31 |
| 3.3.4 Globalizace | 33 |
| 3.3.5 Numerické experimenty | 34 |
| 4 Závěr | 35 |
| Použitá literatura | 36 |
| Abstract | 40 |

Představení autora

Radek Kučera se narodil 13. 3. 1968 v Opavě. V roce 1986 maturoval na gymnáziu v Hlučíně. V letech 1986-1994 studoval na Přírodovědecké fakultě Univerzity Palackého v Olomouci nejprve obor *Přibližné a numerické metody* (RNDr.) a poté doktorské studium *Numerická matematika* (Ph.D.). V roce 2001 se habilitoval na Fakultě elektrotechniky a informatiky Vysoké školy báňské – Technické univerzity Ostrava (VŠB – TUO) v oboru *Aplikovaná matematika* (doc.). Od roku 1994 pracuje na Katedře matematiky a deskriptivní geometrie VŠB – TUO, kde v letech 2010-2015 zastával funkci vedoucího oddělení pro Hornicko-geologickou fakultu a od roku 2012 zastává funkci vedoucího katedry. Od roku 2011 je také zaměstnancem Národního superpočítačového centra IT4Innovations VŠB – TUO. Po jedno volební období byl členem Akademického senátu VŠB – TUO, v současnosti je členem Vědecké rady VŠB – TUO, členem Vědecké rady Hornicko-geologické fakulty a členem Interní akreditační komise VŠB – TUO. Je také členem Jednoty českých matematiků a fyziků, členem České matematické společnosti a členem Zastupitelstva obce Bohuslavice. V roce 1994 studoval čtyři měsíce na Technické univerzitě v Drážďanech, v roce 2016 absolvoval na základě pozvání měsíční stáž na Univerzitě Blaise Pascala v Clermont-Ferrand a od roku 2005 se pravidelně účastní týdenních pobytů na Univerzitě v Caen, Normandie. Je autorem řady odborných prací, z nichž 24 bylo publikováno v časopisech s impaktním faktorem. V databázi WoS eviduje 164 citací bez autocitací a jeho h-index má hodnotu 9.



Na VŠB – TUO vedl výuku na dvou fakultách a pro celouniverzitní studijní obory. Garantuje a přednáší předměty *Numerické metody a statistika*, *Vybrané kapitoly z matematiky a Maticová analýza a variační počet*. Je spoluautorem skript *Numerická matematika* a studijního textu *Numerické metody - testy*. V rámci katedry se podílí na organizování semináře o výuce matematiky. Od roku 2015 je garantem studijního programu *Aplikované vědy a technologie*. Působí jako vedoucí bakalářských, diplomových a disertačních prací. Podílel se na řešení řady projektů GAČR, MŠMT a EU. Pracuje jako recenzent pro Zentralblatt für Mathematik a pro odborné časopisy. V letech 2013-2014 byl předsedou organizačního výboru česko-polského semináře *Moderní matematické metody v inženýrství*, který byl podpořen z evropských zdrojů.

Ve svém výzkumu se zpočátku věnoval aplikacím spline-funkcí a waveletů [37, 49], což vyústilo v zájem o metodu fiktivní oblasti [38, 29, 26, 27, 28, 43]. Postupně se začal zabývat kontaktními úlohami se třením, zejména pracoval na vývoji vhodných optimalizačních algoritmů. U úloh ve dvou prostorových dimenzích (2D) [10, 23, 12] se výpočty prováděly algoritmy pro oboustranná jednoduchá omezení (*box constraints*) ([7, 9]). Ve třech prostorových dimenzích (3D) se objevil jiný typ omezení – kvadratické nerovnosti, které v izotropním případě představují kruhy (*spherical constraints*) a v ortotropním případě elipsy (*elliptical constraints*). První metoda řešení těchto úloh používala polygonální aproximaci kruhů [30], což ale vedlo k nárůstu dimenze úlohy. Na základě geometrické analýzy podmínek optimality byl navržen algoritmus vhodný pro řešení rozsáhlých úloh [39] a pro upravenou variantu tohoto algoritmu byla dokázána R-lineární rychlost konvergence [40]. Algoritmus posloužil při řešení úloh s různými modely tření [50, 32, 31] a v úlohách tvarové optimalizace [3]. Byl také implementován do knihovny MatSol vyvíjené v rámci Národního superpočítačového centra IT4Innovations [51]. Dále byl upraven pro kuželová omezení (*conical constraints*) a testován v úlohách kontaktní dynamiky sypkých hmot ([33, 55]). Dalším směrem výzkumu byla analýza semi-hladké Newtonovy metody při řešení kontaktních úloh ([56]). V pracích [46, 47] bylo uká-

záno, že nepřesná implementace této metody pro Trescův model tření vede na výše zmíněné optimalizační algoritmy. Mechanická interpretace vnitřních lineárních soustav pro Coulombův model tření umožnila navrhnout kontinuální techniku, kterou lze vypočítat vícenásobná řešení [25]. V souvislosti s řešením kontaktních úloh se také podílel na vývoji paralelních algoritmů založených na metodě rozložení oblasti FETI ([19]). Je spoluautorem článku [11], v němž byla navržena T(otal)FETI metoda. Algoritmus, založený na rozšířeném Lagrangianu umožňující použít FETI metody při řešení 3D kontaktních úloh, byl analyzován v [13]. Jedním ze stěžejních kroků FETI metod je eliminace primárních neznámých pomocí singulárních matic. V pracích [44, 42] byla provedena analýza akcí zobecněných inverzí pro rozsáhlé singulární matice ([6, 15]), která ukázala, že každá zobecněná inverze použita v jisté variantě metody Schurova doplnku vede na akci More-Penrosovy pseudoinverze. V současné době se autor věnuje úlohám proudění se skluzovými podmínkami ([20]), které jsou v nejjednodušším případě analogické Trescovu modelu tření. Výpočty lze provádět optimalizačními algoritmy vyvinutými pro kontaktní úlohy, jejich efektivitu však výrazně ovlivňuje velký počet tlakových neznámých bez omezení [41]. Jako účinný se zde projevil algoritmus založený na metodě vnitřních bodů [45], který měl u kontaktních úloh pouze okrajový význam. Řešení fyzikálně realistických úloh zahrnuje modelování moderních hydrofobních materiálů [48].

Poznámka: Odkazy původních prací jsou citovány v hranatých závorkách, odkazy na práce jiných autorů jsou uvedeny navíc v závorkách kulatých.

1 Úvod

V tezích se budeme věnovat kontaktním úlohám lineární pružnosti se třením a úlohám proudění se skluzovými okrajovými podmínkami. Zaměříme se zejména na výpočetní efektivitu optimalizačních algoritmů pro řešení algebraických úloh, které dostaneme po aproximaci metodou konečných prvků. Duální formulace úloh vedou v nejjednodušších případech na minimalizaci ryze kvadratického funkcionálu s omezeními pro vektor duálních neznámých, kterými jsou Lagrangeovy multiplikátory. Ve složitějších případech, kdy například mez skluzu závisí na řešení nebo se vyskytuje další nelinearita, je potřeba minimalizační úlohu řešit opakovaně. Algebraické úlohy jsou přitom zpravidla velmi rozsáhlé, kdy počty neznámých i počty omezení mohou dosahovat řádově miliónů. Máme-li být schopni tyto úlohy řešit dostatečně efektivně, je potřeba při návrhu algoritmu zohlednit konkrétní tvar omezení. Ta se podstatně liší ve dvou (2D) a ve třech (3D) prostorových dimenzích, což vysvětlíme pro Trescův model tření.

Nechť Ω je omezená oblast v \mathbb{R}^d , $d = 2, 3$, představující těleso, jehož pružné chování popisuje Lameho systém parciálních diferenciálních rovnic, které uvedeme později. Nechť $\mathbf{u} : \Omega \rightarrow \mathbb{R}^d$ je pole posunutí a $\boldsymbol{\sigma} : \Omega \rightarrow \mathbb{R}^{d \times d}$ je pole symetrických tenzorů napětí. Nechť γ_c je neprázdná otevřená část hranice $\partial\Omega$, kde uvažujeme kontakt s dokonale tuhým překážkou a Trescův model tření s mezí skluzu $g : \gamma_c \rightarrow \mathbb{R}_+$. O hranici $\partial\Omega$ a o funkcích $\mathbf{u} = \mathbf{u}(\mathbf{x})$, $\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{x})$, $\mathbf{x} \in \bar{\Omega}$ a $g = g(\mathbf{x})$, $\mathbf{x} \in \bar{\gamma}_c$, předpokládáme, že jsou dostatečně hladké. Symbolem $\mathbf{n} = \mathbf{n}(\mathbf{x}) \in \mathbb{R}^d$ označujeme jednotkový vektor vnější normály k $\partial\Omega$ v bodě $\mathbf{x} \in \partial\Omega$. Ve 2D existuje v bodě $\mathbf{x} \in \partial\Omega$ jeden tečný směr k $\partial\Omega$. Jednotkový tečný vektor $\mathbf{t} = \mathbf{t}(\mathbf{x}) \in \mathbb{R}^2$ je proto určen jednoznačně až na orientaci, a dvojice $\{\mathbf{n}, \mathbf{t}\}$ tvoří lokální ortonormální bázi v \mathbb{R}^2 . Ve 3D existuje v bodě $\mathbf{x} \in \partial\Omega$ nekonečně mnoho tečných směrů k $\partial\Omega$. U izotropních úloh volíme libovolně dva jednotkové na sebe kolmé tečné vektory $\mathbf{t}_j = \mathbf{t}_j(\mathbf{x}) \in \mathbb{R}^3$, $j = 1, 2$. Trojice $\{\mathbf{n}, \mathbf{t}_1, \mathbf{t}_2\}$ tvoří lokální ortonormální bázi v \mathbb{R}^3 . U anizotropních úloh je více možností. V ortotropním případě jsou tečné vektory na sebe kolmé, jejich volba ale není libovolná, je určena směry s nejmenší a největší mezí skluzu [31]. V dalším textu se omezíme na izotropní případ. Budeme předpokládat, že existuje m kontaktních uzlů $\mathbf{x}_i \in \bar{\gamma}_c$ a $\mathcal{M} = \{1, \dots, m\}$ je množina jejich lokálních indexů.

Pro $\mathbf{x} \in \partial\Omega$ definujeme ve 2D normálové složky posunutí a napětí jako skalární veličiny $u_n = \mathbf{n}(\mathbf{x})^\top \mathbf{u}(\mathbf{x})$ a $\sigma_n = \mathbf{n}(\mathbf{x})^\top \boldsymbol{\sigma}(\mathbf{x}) \mathbf{n}(\mathbf{x})$. Podobně definujeme také tečné složky posunutí a napětí: $u_t = \mathbf{t}(\mathbf{x})^\top \mathbf{u}(\mathbf{x})$ a $\sigma_t = \mathbf{t}(\mathbf{x})^\top \boldsymbol{\sigma}(\mathbf{x}) \mathbf{n}(\mathbf{x})$. Trescův model tření má ve 2D tento tvar:

$$\left. \begin{aligned} |\sigma_t| &\leq g \\ |\sigma_t| < g &\Rightarrow u_t = 0 \\ |\sigma_t| = g &\Rightarrow \exists c_t \geq 0 : u_t = -c_t \sigma_t \end{aligned} \right\} \text{ na } \gamma_c, \quad (1)$$

kde $|\cdot|$ je absolutní hodnota. V duální algebraické úloze hledáme vektor Lagrangeových multiplikátorů $\boldsymbol{\lambda} \in \mathbb{R}^{2m}$. Pro jeho složky λ_{m+i} aproximující hodnoty $-\sigma_t(\mathbf{x}_i)$ dostaneme z (1) oboustranná jednoduchá omezení $|\lambda_{m+i}| \leq g_i$, kde g_i je aproximace $g(\mathbf{x}_i)$. Duální algebraickou úlohu lze zapsat následovně:

$$\begin{aligned} &\text{minimalizuj } \frac{1}{2} \boldsymbol{\lambda}^\top \mathbf{A} \boldsymbol{\lambda} - \boldsymbol{\lambda}^\top \mathbf{b}, \\ &\text{za podmíněk } 0 \leq \lambda_i, |\lambda_{m+i}| \leq g_i, i \in \mathcal{M}, \end{aligned} \quad (2)$$

kde $\mathbf{A} \in \mathbb{R}^{2m \times 2m}$ je symetrická pozitivně definitní matice, $\mathbf{b} \in \mathbb{R}^{2m}$ a jednostranná omezení reprezentují podmínku nepronikání.

Pro $\mathbf{x} \in \partial\Omega$ ve 3D jsou normálové složky posunutí a napětí formálně stejné jako ve 2D. Tečné složky posunutí a napětí jsou ale vektorové veličiny: $\mathbf{u}_t = (u_{t_1}(\mathbf{x}), u_{t_2}(\mathbf{x}))^\top \in \mathbb{R}^2$ a $\boldsymbol{\sigma}_t = (\sigma_{t_1}(\mathbf{x}), \sigma_{t_2}(\mathbf{x}))^\top \in \mathbb{R}^2$, kde $u_{t_j}(\mathbf{x}) = \mathbf{t}_j(\mathbf{x})^\top \mathbf{u}(\mathbf{x})$ a $\sigma_{t_j}(\mathbf{x}) = \mathbf{t}_j(\mathbf{x})^\top \boldsymbol{\sigma}(\mathbf{x}) \mathbf{n}(\mathbf{x})$, $j = 1, 2$. Trescův model tření má ve 3D tento tvar:

$$\left. \begin{aligned} \|\boldsymbol{\sigma}_t\| &\leq g \\ \|\boldsymbol{\sigma}_t\| < g &\Rightarrow \mathbf{u}_t = \mathbf{0} \\ \|\boldsymbol{\sigma}_t\| = g &\Rightarrow \exists c_t \geq 0 : \mathbf{u}_t = -c_t \boldsymbol{\sigma}_t \end{aligned} \right\} \text{ na } \gamma_c, \quad (3)$$

kde $\|\cdot\|$ označuje euklidovskou normu v \mathbb{R}^2 . Vektor Lagrangeových multiplikátorů $\boldsymbol{\lambda} \in \mathbb{R}^{3m}$ je sestaven tak, že $\boldsymbol{\lambda}_{t,i} = (\lambda_{m+2i-1}, \lambda_{m+2i})^\top \in \mathbb{R}^2$ aproximuje hodnoty $-\boldsymbol{\sigma}_t(\mathbf{x}_i)$. Z podmíněk (3) dostaneme separovatelná sférická omezení $\|\boldsymbol{\lambda}_{t,i}\| \leq g_i$. Duální algebraická úloha má ve 3D tuto podobu:

$$\begin{aligned} &\text{minimalizuj } \frac{1}{2} \boldsymbol{\lambda}^\top \mathbf{A} \boldsymbol{\lambda} - \boldsymbol{\lambda}^\top \mathbf{b}, \\ &\text{za podmíněk } 0 \leq \lambda_i, \lambda_{m+2i-1}^2 + \lambda_{m+2i}^2 \leq g_i^2, i \in \mathcal{M}, \end{aligned} \quad (4)$$

kde $\mathbf{A} \in \mathbb{R}^{3m \times 3m}$ je symetrická pozitivně definitní matice a $\mathbf{b} \in \mathbb{R}^{3m}$. Význam jednostranných omezení je stejný jako ve 2D případě.

Vlastnosti duálních úloh mají rozhodující vliv na průběh iteračního výpočet. U úlohy (2) existuje konečně mnoho mezních stavů, v nichž jsou některá omezení splněna jako rovnosti. Nalezneme-li mezní stav odpovídající řešení, redukuje se minimalizační úloha na soustavu lineárních rovnic pro zbývající proměnné. Některé iterační algoritmy jsou proto schopny řešení vypočítat v konečném počtu kroků. U úlohy (4) je situace složitější. Již pro jediné sférické omezení existuje mezních stavů nespočetně mnoho (vektor $\boldsymbol{\lambda}_{t,i}$ může ležet libovolně na hranici příslušné kružnice). Výpočet řešení v konečném počtu kroků proto obecně nelze očekávat, což má za důsledek vyšší výpočetní nároky. Problematická je také citlivost na počítačovou aritmetiku. Potřebujeme-li například ve výpočtu rozpoznat, zda $\boldsymbol{\lambda}_{t,i}$ leží na příslušné kružnici, pak nestačí testovat rovnost $\|\boldsymbol{\lambda}_{t,i}\| = g_i$, protože k jejímu splnění v důsledku zaokrouhlovacích chyb téměř nikdy nedojde. Abychom mohli o některých vektorech $\boldsymbol{\lambda}_{t,i}$ prohlásit, že na příslušné kružnici leží (to je potřeba například pro stanovení aktivní množiny), volíme v okolí kružnice „pás“, jehož šířka je určena parametrem odvozeným od hodnoty ukončovací přesnosti. Optimální šířku tohoto pásu je potřeba testovat experimentálně.

2 Formulace úloh

V tomto odstavci uvedeme formulace kontaktních úloh lineární pružnosti se třením a úloh proudění se skluzovou okrajovou podmínkou. Provedeme jejich aproximace metodou konečných prvků a odvodíme duální algebraické úlohy.

2.1 Kontaktní úlohy se třením

Kontaktní úlohy budeme formulovat ve 3D pro izotropní Trescův a Coulombův model tření.

2.1.1 Spojitá formulace

Uvažujme dvě lineárně pružná tělesa zadaná jako dvě omezené disjunktní oblasti $\Omega^k \subset \mathbb{R}^3$ s hranicemi $\partial\Omega^k$, $k = 1, 2$. Hranice rozdělíme na tři neprázdné disjunktní části γ_u^k , γ_p^k a γ_c^k tak, že $\partial\Omega^k = \bar{\gamma}_u^k \cup \bar{\gamma}_p^k \cup \bar{\gamma}_c^k$, $k = 1, 2$. Na γ_u^k uvažujeme nulové posunutí a na γ_p^k povrchové napětí $\mathbf{p}^k : \gamma_p^k \rightarrow \mathbb{R}^3$, $k = 1, 2$. Dvojice γ_c^1 a γ_c^2 představuje kontaktní rozhraní mezi tělesy, kde předepíšeme tři kontaktní podmínky: nepronikání těles, přenos kontaktních napětí a Trescovo nebo Coulombovo tření. Dále předpokládáme, že na každé těleso působí objemové síly $\mathbf{f}^k : \Omega^k \rightarrow \mathbb{R}^3$, $k = 1, 2$.

Pole posunutí $\mathbf{u}^k : \Omega^k \rightarrow \mathbb{R}^3$ splňuje Lameho systém parciálních diferenciálních rovnic a Dirichletovy a Neumannovy okrajové podmínky:

$$\left. \begin{aligned} \operatorname{div} \boldsymbol{\sigma}^k + \mathbf{f}^k &= \mathbf{0} & \text{v} & \Omega^k, \\ \mathbf{u}^k &= \mathbf{0} & \text{na} & \gamma_u^k, \\ \boldsymbol{\sigma}^k \mathbf{n}^k - \mathbf{p}^k &= \mathbf{0} & \text{na} & \gamma_p^k, \end{aligned} \right\} k = 1, 2, \quad (5)$$

kde \mathbf{n}^k je jednotkový vektor vnější normály k $\partial\Omega^k$. Tensor napětí je určen Hookovým zákonem $\boldsymbol{\sigma}^k = c_1^k \operatorname{tr}(\boldsymbol{\epsilon}^k) \mathbf{I} + 2c_2^k \boldsymbol{\epsilon}^k$, kde $\boldsymbol{\epsilon}^k = 1/2(\nabla \mathbf{u}^k + (\nabla \mathbf{u}^k)^\top)$ je linearizovaný tenzor deformace, „tr“ je stopa matice, $\mathbf{I} \in \mathbb{R}^{3 \times 3}$ je matice jednotková a $c_1^k, c_2^k > 0$ jsou Lameho konstanty, $k = 1, 2$.

Kontaktní podmínky zapíšeme pomocí vzájemně jednoznačného přechodového zobrazení $\boldsymbol{\chi} : \gamma_c^1 \rightarrow \gamma_c^2$, které určuje počáteční vzdálenost mezi tělesy a kritický směr. Počáteční vzdálenost $d : \gamma_c^1 \rightarrow \mathbb{R}_+$ je definována jako $d(\mathbf{x}) = \|\boldsymbol{\chi}(\mathbf{x}) - \mathbf{x}\|$, $\mathbf{x} \in \gamma_c^1$. Kritický směr $\boldsymbol{\nu} : \gamma_c^1 \rightarrow \mathbb{R}^3$ je definován následovně: je-li $d(\mathbf{x}) \neq 0$, pak $\boldsymbol{\nu}(\mathbf{x}) = (\boldsymbol{\chi}(\mathbf{x}) - \mathbf{x})/d(\mathbf{x})$, je-li $d(\mathbf{x}) = 0$, pak $\boldsymbol{\nu}(\mathbf{x}) = \mathbf{n}^1(\mathbf{x})$, $\mathbf{x} \in \gamma_c^1$. Dále uvažujeme na γ_c^1 relativní kontaktní posunutí $u_\nu(\mathbf{x}) = \boldsymbol{\nu}(\mathbf{x})^\top (\mathbf{u}^1(\mathbf{x}) - \mathbf{u}^2(\boldsymbol{\chi}(\mathbf{x})))$ a kontaktní napětí $\sigma_\nu(\mathbf{x}) = \boldsymbol{\nu}(\mathbf{x})^\top \boldsymbol{\sigma}^1(\mathbf{x}) \mathbf{n}^1(\mathbf{x})$ v kritickém směru $\boldsymbol{\nu}(\mathbf{x})$, $\mathbf{x} \in \gamma_c^1$. Pomocí těchto veličin zapíšeme linearizovanou podmínku nepronikání:

$$u_\nu - d \leq 0, \quad \sigma_\nu \leq 0, \quad \sigma_\nu(u_\nu - d) = 0 \quad \text{na} \quad \gamma_c^1. \quad (6)$$

V rovině kolmé na $\boldsymbol{\nu} = \boldsymbol{\nu}(\mathbf{x})$ zvolíme dva tečné vektory $\mathbf{t}_j = \mathbf{t}_j(\mathbf{x})$, $j = 1, 2$ takové, že trojice $\{\boldsymbol{\nu}, \mathbf{t}_1, \mathbf{t}_2\}$ tvoří lokální ortonormální bázi v \mathbb{R}^3 s počátkem v $\mathbf{x} \in \gamma_c^1$. Koefficient tření na γ_c^1 budeme označovat \mathcal{F} , kde $\mathcal{F} : \gamma_c^1 \rightarrow \mathbb{R}_+$. Podmínku pro Coulombovo tření lze zapsat takto:

$$\|\boldsymbol{\sigma}_t\| \leq -\mathcal{F}\sigma_\nu, \quad \boldsymbol{\sigma}_t^\top \mathbf{u}_t - \mathcal{F}\sigma_\nu \|\mathbf{u}_t\| = 0 \quad \text{na} \quad \gamma_c^1, \quad (7)$$

kde $\mathbf{u}_t = (u_{t_1}, u_{t_2})^\top$ je relativní tečné kontaktní posunutí a $\boldsymbol{\sigma}_t = (\sigma_{t_1}, \sigma_{t_2})^\top$ je tečné kontaktní napětí se složkami $u_{t_j}(\mathbf{x}) = \mathbf{t}_j(\mathbf{x})^\top (\mathbf{u}^1(\mathbf{x}) - \mathbf{u}^2(\boldsymbol{\chi}(\mathbf{x})))$ a $\sigma_{t_j}(\mathbf{x}) = \mathbf{t}_j(\mathbf{x})^\top \boldsymbol{\sigma}^1(\mathbf{x}) \mathbf{n}^1(\mathbf{x})$, $\mathbf{x} \in \gamma_c^1$, $j = 1, 2$. Poslední kontaktní podmínkou je přenos kontaktních napětí:

$$\boldsymbol{\sigma}^1 \boldsymbol{\nu} = (\boldsymbol{\sigma}^2 \circ \boldsymbol{\chi}) \boldsymbol{\nu} \quad \text{na} \quad \gamma_c^1, \quad (8)$$

kde $(\sigma^2 \circ \chi)(\mathbf{x}) = \sigma^2(\chi(\mathbf{x}))$ pro $\mathbf{x} \in \gamma_c^1$.

U úlohy s Trescovým třením nahradíme napětí $-\sigma_\nu$ v podmínce (7) zadanou mezí skluzu $g : \gamma_c^1 \rightarrow \mathbb{R}_+$. Podmínku Trescova tření lze zapsat následovně:

$$\|\boldsymbol{\sigma}_t\| \leq \mathcal{F}g, \quad \boldsymbol{\sigma}_t^\top \mathbf{u}_t + \mathcal{F}g\|\mathbf{u}_t\| = 0 \quad \text{na} \quad \gamma_c^1. \quad (9)$$

Vztahy (9) jsou ekvivalentní (3) pro $g := \mathcal{F}g$ a $\gamma_c := \gamma_c^1$.

Ve slabých formulacích použijeme tyto množiny funkcí:

$$\begin{aligned} \mathbb{V} &= \{\mathbf{v} = (\mathbf{v}^1, \mathbf{v}^2) \in (H^1(\Omega^1))^3 \times (H^1(\Omega^2))^3 \mid \mathbf{v}^k = \mathbf{0} \text{ na } \gamma_u^k, k = 1, 2\}, \\ \mathbb{K} &= \{\mathbf{v} \in \mathbb{V} \mid v_\nu - d \leq 0 \text{ na } \gamma_c^1\}, \\ X_\nu &= \{\varphi \in L^2(\gamma_c^1) \mid \exists \mathbf{v} \in \mathbb{V} : \varphi = v_\nu \text{ na } \gamma_c^1\}, \\ X_{\nu+} &= \{\varphi \in X_\nu \mid \varphi \geq 0 \text{ na } \gamma_c^1\}, \\ X_{t+} &= \{\varphi \in L^2(\Gamma_c^1) \mid \exists \mathbf{v} \in \mathbb{V} : \varphi = \|\mathbf{v}_t\| \text{ na } \gamma_c^1\}. \end{aligned}$$

Symbolem X'_ν označíme duální prostor k X_ν a $X'_{\nu+}$ bude kužel všech nezáporných prvků z X'_ν . Dualitu mezi X'_ν a X_ν zapisujeme symbolem $\langle \cdot, \cdot \rangle$. Dále budeme předpokládat, že $X_{t+} \subseteq X_\nu$, takže $\langle \mathcal{F}g, \|\mathbf{v}_t\| \rangle$ bude definováno pro každé $\mathcal{F}g \in X'_\nu$ a $\mathbf{v} \in \mathbb{V}$. Existence slabých řešení kontaktních úloh je zajištěna při dostatečné hladkosti zadaných funkcí: $\mathbf{f}^k \in (L^2(\Omega^k))^3$, $\mathbf{p}^k \in (L^2(\gamma_p^k))^3$, $k = 1, 2$ a $d \in X_{\nu+}$. O koeficientu tření $\mathcal{F} \in L^2(\gamma_c^1)$ navíc předpokládáme, že je omezený: $\mathcal{F}_{\min} \leq \mathcal{F} \leq \mathcal{F}_{\max}$ na γ_c^1 , kde $0 < \mathcal{F}_{\min} < \mathcal{F}_{\max}$ jsou dané konstanty. Nakonec definujeme funkcionály:

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &= \sum_{k=1}^2 \int_{\Omega^k} \boldsymbol{\sigma}(\mathbf{u}^k) : \boldsymbol{\epsilon}(\mathbf{v}^k) \, d\mathbf{x}, \quad \mathbf{u}, \mathbf{v} \in \mathbb{V}, \\ b(\mathbf{v}) &= \sum_{k=1}^2 \left(\int_{\Omega^k} (\mathbf{f}^k)^\top \mathbf{v}^k \, d\mathbf{x} + \int_{\Gamma_p^k} (\mathbf{p}^k)^\top \mathbf{v}^k \, ds \right), \quad \mathbf{v} \in \mathbb{V}, \\ j(\mathbf{v}) &= \langle \mathcal{F}g, \|\mathbf{v}_t\| \rangle, \quad \mathbf{v} \in \mathbb{V}. \end{aligned}$$

Je-li navíc $\mathcal{F}g \in L^2(\gamma_c^1)$, můžeme dualitu $\langle \cdot, \cdot \rangle$ zapsat jako skalární součin na $L^2(\gamma_c^1)$. Funkcionál $j(\mathbf{v})$ je pak reprezentován křivkovým integrálem přes γ_c^1 .

Klasickým řešením kontaktní úlohy s Trescovým třením je dvojice polí posunutí \mathbf{u}^k v Ω^k , $k = 1, 2$ vyhovujících vztahům (5), (6), (9) a (8). Slabé řešení této úlohy je určeno variační nerovnicí: pro dané $g \in X'_{\nu+}$ najdi $\mathbf{u} = \mathbf{u}(g) \in \mathbb{K}$ takové, že

$$a(\mathbf{u}, \mathbf{v} - \mathbf{u}) + j(\mathbf{v}) - j(\mathbf{u}) \geq b(\mathbf{v} - \mathbf{u}) \quad \forall \mathbf{v} \in \mathbb{K}. \quad (10)$$

Ekvivalentně lze (10) zapsat jako minimalizační úlohu: pro dané $g \in X'_{\nu+}$ najdi $\mathbf{u} = \mathbf{u}(g) \in \mathbb{K}$ takové, že

$$J(\mathbf{u}) \leq J(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbb{K}, \quad (11)$$

kde $J(\mathbf{v}) = \frac{1}{2}a(\mathbf{v}, \mathbf{v}) - b(\mathbf{v}) + j(\mathbf{v})$. Úlohy (10) a (11) mají jediné řešení, které je u obou úloh stejné ([24]).

Klasickým řešením kontaktní úlohy s Coulombovým třením je dvojice polí posunutí \mathbf{u}^k v Ω^k , $k = 1, 2$ vyhovujících vztahům (5), (6), (7) a (8). Nechť $\sigma_\nu = \sigma_\nu(g) \in X'_{\nu+}$ je kontaktní napětí na γ_c^1 odpovídající řešení $\mathbf{u}(g)$ kontaktní úlohy s Trescovým třením (11) pro $g \in X'_{\nu+}$. Pak lze definovat zobrazení $\Psi : X'_{\nu+} \rightarrow X'_{\nu+}$ předpisem:

$$\Psi : g \mapsto -\sigma_\nu(g), \quad g \in X'_{\nu+}. \quad (12)$$

Slabým řešením kontaktní úlohy s Coulombovým třením rozumíme řešení $\mathbf{u} = \mathbf{u}(g) \in \mathbb{K}$ úlohy (11), pro které platí

$$\Psi(-\sigma_\nu(g)) = -\sigma_\nu(g).$$

Jinými slovy $-\sigma_\nu(g)$ je pevný bod zobrazení Ψ na $X'_{\nu+}$. Pro dostatečně malý koeficient tření \mathcal{F} je dokázána existence alespoň jednoho slabého řešení ([17]).

2.1.2 Aproximace metodou konečných prvků

Na oblastech Ω^k vytvoříme dělení na čtyřstěny \mathcal{T}_h^k , která jsou regulární a kompatibilní s dělením hranic $\partial\Omega^k$ na γ_u^k , γ_p^k a γ_c^k , $k = 1, 2$. Pro jednoduchost budeme předpokládat, že uzly ležící na $\overline{\gamma_c^2}$ jsou obrazem uzlů z $\overline{\gamma_c^1}$ v zobrazení χ . Prostor \mathbb{V} aproximujeme pomocí lineárních konečných prvků:

$$\begin{aligned} \mathbb{V}_h &= \mathbb{V} \cap (\mathbb{V}_h^1 \times \mathbb{V}_h^2), \\ \mathbb{V}_h^k &= \{\mathbf{v}_h^k \in (C(\overline{\Omega}^k))^3 : \mathbf{v}_{h|T}^k \in (P_1(T))^3 \quad \forall T \in \mathcal{T}_h^k\}, \quad k = 1, 2, \end{aligned}$$

kde $P_1(T)$ je množina všech lineárních polynomů na T . Označme $n_u = \dim \mathbb{V}_h$. Množinu přípustných posunutí \mathbb{K} aproximujeme předpisem:

$$\mathbb{K}_h = \{\mathbf{v}_h \in \mathbb{V}_h : v_{h\nu}(\mathbf{x}_i^1) - d(\mathbf{x}_i^1) \leq 0 \quad \forall i \in \mathcal{M}\},$$

kde \mathbf{x}_i^1 , $i \in \mathcal{M}$ jsou uzly ležící na $\overline{\gamma_c^1} \setminus \overline{\gamma_u^1}$, m je jejich počet a $\mathcal{M} = \{1, \dots, m\}$. Aproximaci funkcionálu $j(\mathbf{v})$ provedeme numerickou integrací:

$$\int_{\gamma_c^1} \mathcal{F}g \|\mathbf{v}_{ht}\| ds \approx \sum_{i \in \mathcal{M}} \omega_i \mathcal{F}(\mathbf{x}_i^1) g(\mathbf{x}_i^1) \|\mathbf{v}_{ht}(\mathbf{x}_i^1)\| =: j_h(\mathbf{v}_h),$$

kde ω_i jsou integrační váhy. Ve výpočtech používáme složenou lichoběžníkovou formuli. Aproximace úlohy (11) metodou konečných prvků má tvar: najdi $\mathbf{u}_h = \mathbf{u}_h(g) \in \mathbb{K}_h$ takové, že

$$J_h(\mathbf{u}_h) \leq J_h(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in \mathbb{K}_h, \quad (13)$$

kde $J_h(\mathbf{v}_h) = \frac{1}{2}a(\mathbf{v}_h, \mathbf{v}_h) - b(\mathbf{v}_h) + j_h(\mathbf{v}_h)$. Úlohu (13) lze zapsat jako algebraickou minimalizační úlohu s nerovnostním omezením:

$$\begin{aligned} & \text{minimalizuj} \quad \frac{1}{2} \mathbf{u}^\top \mathbf{K} \mathbf{u} - \mathbf{u}^\top \mathbf{f} + \sum_{i \in \mathcal{M}} \mathcal{F}_i g_i \|\mathbf{u}_{t,i}\|, \\ & \text{za podmínky} \quad \mathbf{N} \mathbf{u} - \mathbf{d} \leq \mathbf{0}, \end{aligned} \quad (14)$$

kde $\mathbf{u} \in \mathbb{R}^{n_u}$ je vektor uzlových posunutí, $\mathbf{K} \in \mathbb{R}^{n_u \times n_u}$ je symetrická pozitivně definitní matice, $\mathbf{f} \in \mathbb{R}^{n_u}$, $\mathbf{d} \in \mathbb{R}_+^m$ má prvky $d_i = d(\mathbf{x}_i^1)$, $\mathcal{F}_i = \mathcal{F}(\mathbf{x}_i^1)$, $g_i = \omega_i g(\mathbf{x}_i^1)$, $\mathbf{u}_{t,i} = ((\mathbf{T} \mathbf{u})_{2i-1}, (\mathbf{T} \mathbf{u})_{2i})^\top \in \mathbb{R}^2$ pro $i \in \mathcal{M}$. Matice $\mathbf{N} \in \mathbb{R}^{m \times n_u}$ a $\mathbf{T} \in \mathbb{R}^{2m \times n_u}$ jsou definovány následovně: i -tý řádek \mathbf{N} obsahuje na vhodných pozicích nejvýše tři nenulové prvky vektoru $\boldsymbol{\nu}(\mathbf{x}_i^1)$, $i \in \mathcal{M}$; $(2i-1)$ -ní a $(2i)$ -tý řádek \mathbf{T} obsahují na vhodných pozicích nejvýše tři nenulové prvky vektorů $\mathbf{t}_1(\mathbf{x}_i^1)$ a $\mathbf{t}_2(\mathbf{x}_i^1)$, $i \in \mathcal{M}$. Nakonec označme $\mathbf{g} = (g_1, \dots, g_m)^\top \in \mathbb{R}_+^m$ a $\mathcal{F} = \text{diag}(\mathcal{F}_1, \dots, \mathcal{F}_m) \in \mathbb{R}^{m \times m}$.

2.1.3 Duální algebraické úlohy

Duální formulací úlohy (14) odstraníme nediferencovatelnost minimalizované funkce a zjednodušíme tvar nerovnostních omezení. Nejprve napíšeme sedlobodovou formulaci.

Nechť $L : \mathbb{R}^{n_u} \times \Lambda_\nu \times \Lambda_t(\mathcal{F}g) \rightarrow \mathbb{R}$ je Lagrangeova funkce úlohy (14):

$$L(\mathbf{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) = \frac{1}{2} \mathbf{u}^\top \mathbf{K} \mathbf{u} - \mathbf{u}^\top \mathbf{f} + \boldsymbol{\lambda}_\nu^\top (\mathbf{N} \mathbf{u} - \mathbf{d}) + \boldsymbol{\lambda}_t^\top \mathbf{T} \mathbf{u},$$

kde $\Lambda_\nu = \mathbb{R}_+^m$, $\Lambda_t(\mathcal{F}g) = \Lambda_{t,1}(\mathcal{F}_1 g_1) \times \cdots \times \Lambda_{t,m}(\mathcal{F}_m g_m)$ jsou množiny Lagrangeových multiplikátorů a dílčí množiny $\Lambda_{t,i}(\mathcal{F}_i g_i)$ odpovídající jednotlivým kontaktním uzlům jsou definovány následovně: $\Lambda_{t,i}(\mathcal{F}_i g_i) = \{\boldsymbol{\lambda}_{t,i} \in \mathbb{R}^2 : \|\boldsymbol{\lambda}_{t,i}\| \leq \mathcal{F}_i g_i\}$, $\boldsymbol{\lambda}_{t,i} = (\lambda_{t,2i-1}, \lambda_{t,2i})^\top$, $i \in \mathcal{M}$. Snadno lze ukázat, že řešení úlohy (14) je první složkou sedlového bodu $(\mathbf{u}^*, \boldsymbol{\lambda}_\nu^*, \boldsymbol{\lambda}_t^*) \in \mathbb{R}^{n_u} \times \Lambda_\nu \times \Lambda_t(\mathcal{F}g)$, který je řešením úlohy:

$$L(\mathbf{u}^*, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \leq L(\mathbf{u}^*, \boldsymbol{\lambda}_\nu^*, \boldsymbol{\lambda}_t^*) \leq L(\mathbf{u}, \boldsymbol{\lambda}_\nu^*, \boldsymbol{\lambda}_t^*) \quad \forall (\mathbf{u}, \boldsymbol{\lambda}_\nu, \boldsymbol{\lambda}_t) \in \mathbb{R}^{n_u} \times \Lambda_\nu \times \Lambda_t(\mathcal{F}g). \quad (15)$$

Složky Lagrangeových multiplikátorů $\boldsymbol{\lambda}_{\nu,i}^*$ a $\boldsymbol{\lambda}_{t,i}^*$ aproximují kontaktní napětí $-\sigma_\nu(\mathbf{x}_i^1)$ a $-\sigma_t(\mathbf{x}_i^1)$, $i \in \mathcal{M}$. Sedlový bod je jednoznačně určen KKT (Karush-Kuhn-Tuckerovými) podmínkami optimality:

$$\mathbf{K} \mathbf{u}^* + \mathbf{N}^\top \boldsymbol{\lambda}_\nu^* + \mathbf{T}^\top \boldsymbol{\lambda}_t^* - \mathbf{f} = \mathbf{0}, \quad (16)$$

$$\mathbf{N} \mathbf{u}^* - \mathbf{d} \leq \mathbf{0}, \quad \boldsymbol{\lambda}_\nu^* \geq \mathbf{0}, \quad \boldsymbol{\lambda}_\nu^{*\top} (\mathbf{N} \mathbf{u}^* - \mathbf{d}) = 0, \quad (17)$$

$$\left. \begin{array}{l} \|\boldsymbol{\lambda}_{t,i}^*\| \leq \mathcal{F}_i g_i \\ \|\boldsymbol{\lambda}_{t,i}^*\| < \mathcal{F}_i g_i \Rightarrow \mathbf{u}_{t,i}^* = \mathbf{0} \\ \|\boldsymbol{\lambda}_{t,i}^*\| = \mathcal{F}_i g_i \Rightarrow \exists c_i \geq 0 : \mathbf{u}_{t,i}^* = c_i \boldsymbol{\lambda}_{t,i}^* \end{array} \right\} \quad i \in \mathcal{M}. \quad (18)$$

Ze vztahu (16) můžeme vyjádřit první složku sedlového bodu:

$$\mathbf{u}^* = \mathbf{K}^{-1} (\mathbf{f} - \mathbf{N}^\top \boldsymbol{\lambda}_\nu^* - \mathbf{T}^\top \boldsymbol{\lambda}_t^*). \quad (19)$$

Dosazením do první nerovnosti v (15) dostaneme úlohu zapsanou výhradně pomocí Lagrangeových multiplikátorů. Nejdříve zavedeme vhodné značení.

Definujme kvadratickou funkci $q : \mathbb{R}^{3m} \rightarrow \mathbb{R}$ předpisem

$$q(\boldsymbol{\lambda}) = \frac{1}{2} \boldsymbol{\lambda}^\top \mathbf{A} \boldsymbol{\lambda} - \boldsymbol{\lambda}^\top \mathbf{b}, \quad (20)$$

kde $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_\nu^\top, \boldsymbol{\lambda}_t^\top)^\top$, $\mathbf{A} = \mathbf{B} \mathbf{K}^{-1} \mathbf{B}^\top$ je symetrická pozitivně definitní matice, $\mathbf{B} = (\mathbf{N}^\top, \mathbf{T}^\top)^\top$ je matice s plnou řádkovou hodnotou, $\mathbf{b} = \mathbf{B} \mathbf{K}^{-1} \mathbf{f} - \mathbf{c}$ a $\mathbf{c} = (\mathbf{d}^\top, \mathbf{0}^\top)^\top$. Gradient funkce q označený $\mathbf{r} : \mathbb{R}^{3m} \rightarrow \mathbb{R}^{3m}$ je určen výrazem

$$\mathbf{r}(\boldsymbol{\lambda}) = \mathbf{A} \boldsymbol{\lambda} - \mathbf{b}.$$

Dále označíme $\boldsymbol{\lambda}^* = (\boldsymbol{\lambda}_\nu^{*\top}, \boldsymbol{\lambda}_t^{*\top})^\top$ a $\Lambda(\mathcal{F}g) = \Lambda_\nu \times \Lambda_t(\mathcal{F}g)$. Dosazením \mathbf{u}^* z (19) můžeme první nerovnost v (15) upravit do tvaru:

$$(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*)^\top \mathbf{r}(\boldsymbol{\lambda}^*) \geq 0 \quad \forall \boldsymbol{\lambda} \in \Lambda(\mathcal{F}g).$$

Protože $\boldsymbol{\lambda}^*$, které vyhovuje této variační nerovnosti, minimalizuje q na $\Lambda(\mathcal{F}g)$ ([54]), dostáváme duální podobu algebraické kontaktní úlohy s Trescovým třením:

$$\boldsymbol{\lambda}^* = \arg \min_{\boldsymbol{\lambda} \in \Lambda(\mathcal{F}g)} q(\boldsymbol{\lambda}). \quad (21)$$

Řešení této úlohy lze vypočítat optimalizačními algoritmy ze Sekce 3. Po výpočtu $\boldsymbol{\lambda}^*$ určíme vektor uzlových posunutí \mathbf{u}^* podle vztahu (19). Poznamenejme ještě, že úlohu (21) lze zapsat ve tvaru (4) pro $g_i := \mathcal{F}_i g_i$, $i \in \mathcal{M}$.

Ve zbývající části odstavce ukážeme hlavní myšlenku řešení algebraické úlohy s Coulombovým třením. Definujme zobrazení $\Psi : \Lambda_\nu \rightarrow \Lambda_\nu$ jako diskrétní analogii zobrazení (12) předpisem:

$$\Psi : \mathbf{g} \mapsto \boldsymbol{\lambda}_\nu^*, \quad \mathbf{g} \in \Lambda_\nu,$$

kde $\boldsymbol{\lambda}_\nu^*$ je první složka řešení úlohy (21) pro dané \mathbf{g} . Algebraickým řešením duální úlohy s Coulombovým třením rozumíme jakékoliv řešení úlohy (21), pro které je odpovídající $\boldsymbol{\lambda}_\nu^*$ pevným bodem zobrazení Ψ , tj. platí: $\Psi(\boldsymbol{\lambda}_\nu^*) = \boldsymbol{\lambda}_\nu^*$. Existence alespoň jednoho pevného bodu pro libovolný koeficient tření je dokázána v ([24]). Výpočet pevného bodu lze provést metodou postupných aproximací:

$$\boldsymbol{\lambda}_\nu^{(0)} \in \Lambda_\nu, \quad \boldsymbol{\lambda}_\nu^{(k+1)} = \Psi(\boldsymbol{\lambda}_\nu^{(k)}), \quad k = 0, 1, 2, \dots \quad (22)$$

Posloupnost $\{\boldsymbol{\lambda}_\nu^{(k)}\}$ vytvářena tímto iteračním předpisem konverguje k pevnému bodu $\boldsymbol{\lambda}_\nu^*$, jestliže Ψ je kontraktivní zobrazení v Λ_ν . V takovém případě je pevný bod určen jednoznačně. Kontraktivita zobrazení Ψ byla dokázána v ([22]) pro dostatečně malý koeficient tření. Horní mez \mathcal{F}_{\max} garantující kontraktivitu Ψ ale závisí na normě sítě konečných prvků. V [32] byl tento výsledek zobecněn pro případ ortotropního tření. Při výpočtu iterací (22) vypočítáme $\Psi(\boldsymbol{\lambda}_\nu^{(k)})$ jako řešení úlohy (21) pro $\mathbf{g} = \boldsymbol{\lambda}_\nu^{(k)}$. Pro optimální nastavení ukončovacích kritérií je výpočet pevného bodu srovnatelně náročný s vyřešením jediné úlohy s Trescovým třením.

2.2 Úlohy proudění s podmínkou skluzu

V tomto odstavci ukážeme použití podmínek podobných Trescovu tření v úlohách proudění se skluzem kapaliny podél stěny. Pro jednoduchost se omezíme na statickou úlohu ve 2D zapsanou pomocí Stokesových rovnic. Skluzová podmínka bude kombinací klasické Navierovy podmínky [52] a podmínky Trescova typu.

2.2.1 Spojitá formulace

Nechť Ω je omezená oblast v \mathbb{R}^2 s dostatečně hladkou hranicí $\partial\Omega$, která je rozdělena na tři disjunktní části: $\partial\Omega = \bar{\gamma}_D \cup \bar{\gamma}_N \cup \bar{\gamma}_C$. V oblasti Ω uvažujeme proudění viskózní nestlačitelné newtonovské kapaliny modelované následujícím systémem parciálních diferenciálních rovnic a okrajových podmínek:

$$\left. \begin{aligned} -\nu\Delta\mathbf{u} + \nabla p &= \mathbf{f} && \text{v } \Omega, \\ \nabla \cdot \mathbf{u} &= 0 && \text{v } \Omega, \\ \mathbf{u} &= \mathbf{u}_D && \text{na } \gamma_D, \\ \boldsymbol{\sigma} &= \boldsymbol{\sigma}_N && \text{na } \gamma_N, \\ u_n &= 0 && \text{na } \gamma_C, \\ u_t = 0 &\Rightarrow |\sigma_t| \leq g && \text{na } \gamma_C \\ \sigma_t u_t + g|u_t| + \kappa u_t^2 &= 0 && \text{na } \gamma_C. \end{aligned} \right\} \quad (23)$$

Hledáme rychlostní pole $\mathbf{u} : \Omega \rightarrow \mathbb{R}^2$ a tlakové pole $p : \Omega \rightarrow \mathbb{R}$. Zadané veličiny jsou kinematičká viskozita $\nu > 0$, síla působící na kapalinu $\mathbf{f} : \Omega \rightarrow \mathbb{R}^2$, předepsaná vtoková nebo výtoková rychlost $\mathbf{u}_D : \gamma_D \rightarrow \mathbb{R}^2$ a předepsané napětí $\boldsymbol{\sigma}_N : \gamma_N \rightarrow \mathbb{R}^2$. Na $\partial\Omega$ uvažujeme jednotkový vektor vnější normály $\mathbf{n} = \mathbf{n}(\mathbf{x}) \in \mathbb{R}^2$ a jednotkový tečný vektor $\mathbf{t} = \mathbf{t}(\mathbf{x}) \in \mathbb{R}^2$ pro $\mathbf{x} \in \partial\Omega$. Dále necht' $u_n(\mathbf{x}) = \mathbf{n}(\mathbf{x})^\top \mathbf{u}(\mathbf{x})$ je normálová a $u_t(\mathbf{x}) = \mathbf{t}(\mathbf{x})^\top \mathbf{u}(\mathbf{x})$ je tečná složka rychlosti v $\mathbf{x} \in \partial\Omega$. Napětí je definováno předpisem $\boldsymbol{\sigma} = \nu \frac{d\mathbf{u}}{dn} - p\mathbf{n}$ a $\sigma_t(\mathbf{x}) = \mathbf{t}(\mathbf{x})^\top \boldsymbol{\sigma}(\mathbf{x})$ je jeho smyková složka, $\mathbf{x} \in \gamma_C$. Konečně $g : \gamma_C \rightarrow \mathbb{R}_+$ je mez skluzu a $\kappa : \gamma_C \rightarrow \mathbb{R}_+$ je koeficient přílnavosti.

Pro $\kappa = 0$ má okrajová podmínka předepsaná pro tečné komponenty na γ_C tvar Trescova modelu tření z úloh lineární pružnosti. Pro $\kappa \neq 0$ ji můžeme do tohoto tvaru převést ve slabé formulaci.

Zaveďme množinu funkcí pro rychlostní pole:

$$V_{\mathbf{u}_D}(\Omega) = \{\mathbf{v} \in (H^1(\Omega))^2 : \mathbf{v} = \mathbf{u}_D \text{ na } \gamma_D, v_n = 0 \text{ na } \gamma_C\}$$

a funkciónály

$$a_\kappa(\mathbf{v}, \mathbf{w}) = \nu \int_{\Omega} \nabla \mathbf{v} : \nabla \mathbf{w} \, d\mathbf{x} + \int_{\gamma_C} \kappa v_t w_t \, ds, \quad b(\mathbf{v}, q) = - \int_{\Omega} q(\nabla \cdot \mathbf{v}) \, d\mathbf{x},$$

$$l(\mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, d\mathbf{x} + \int_{\gamma_N} \boldsymbol{\sigma}_N \cdot \mathbf{v} \, ds, \quad j(\mathbf{v}) = \int_{\gamma_C} g|v_t| \, ds,$$

kde $\nabla \mathbf{v} : \nabla \mathbf{w} = \nabla v_1 \cdot \nabla w_1 + \nabla v_2 \cdot \nabla w_2$, $\mathbf{v} = (v_1, v_2)^\top$, $\mathbf{w} = (w_1, w_2)^\top$. Slabá formulace úlohy (23) má tento tvar: najdi $(\mathbf{u}, p) \in V_{\mathbf{u}_D}(\Omega) \times L^2(\Omega)$ takové, že

$$\left. \begin{aligned} a_\kappa(\mathbf{u}, \mathbf{v} - \mathbf{u}) + b(\mathbf{v} - \mathbf{u}, p) + j(\mathbf{v}) - j(\mathbf{u}) &\geq l(\mathbf{v} - \mathbf{u}) \quad \forall \mathbf{v} \in V_0(\Omega), \\ b(\mathbf{u}, q) &= 0 \quad \forall q \in L^2(\Omega). \end{aligned} \right\} \quad (24)$$

Existenci a jednoznačnost obou složek řešení lze dokázat za předpokladu, že všechny tři části hranice γ_D , γ_N a γ_C jsou neprázdné a zadané funkce jsou dostatečně hladké: $\mathbf{f} \in (L^2(\Omega))^2$, $\mathbf{u}_D \in (H^{1/2}(\gamma_D))^2$, $\boldsymbol{\sigma}_N \in (L^2(\gamma_N))^2$ a $g, \kappa \in L^2_+(\gamma_C)$ ([20]). Protože koeficient přilnavosti κ zde vystupuje v bilineární formě a_κ , je sublineární člen j stejný jako v případě Trescova tření. V konečném důsledku to pak znamená, že můžeme použít stejné algoritmy.

2.2.2 Aproximace metodou smíšených konečných prvků

Na oblasti Ω vytvoříme dělení na trojúhelníky \mathcal{T}_h , které je regulární a kompatibilní s dělením hranice $\partial\Omega$ na γ_D , γ_N a γ_C . Pomocí metody smíšených konečných prvků definujeme na \mathcal{T}_h aproximaci $V_{\mathbf{u}_D, h}$ množiny $V_{\mathbf{u}_D}(\Omega)$ pro rychlostní složku řešení a aproximaci W_h prostoru $L^2(\Omega)$ pro tlakovou složku řešení. Používáme dvojici konečných prvků označovanou jako *P1-bubble/P1* ([36]), která splňuje *inf-sup* podmínku stability (nazývanou také Ladyženskaja-Babuška-Brezziho podmínka) ([4, 18]).

Aproximace úlohy (24) má tento tvar: najdi $(\mathbf{u}_h, p_h) \in V_{\mathbf{u}_D, h} \times W_h$ takové, že

$$\left. \begin{aligned} a_\kappa(\mathbf{u}_h, \mathbf{v}_h - \mathbf{u}_h) + b(\mathbf{v}_h - \mathbf{u}_h, p_h) + j_h(\mathbf{v}_h) - j_h(\mathbf{u}_h) &\geq l(\mathbf{v}_h - \mathbf{u}_h) \quad \forall \mathbf{v}_h \in V_{0, h}, \\ b(\mathbf{u}_h, q_h) &= 0 \quad \forall q_h \in W_h, \end{aligned} \right\} \quad (25)$$

kde j_h je aproximace funkciónálu j . V našich výpočtech používáme aproximaci vytvořenou pomocí složené lichoběžníkové formule:

$$\int_{\gamma_C} g|v_{ht}| \, ds \approx \sum_{\mathbf{x}_i \in N_C} \omega_i g(\mathbf{x}_i) |v_{ht}(\mathbf{x}_i)| =: j_h(\mathbf{v}_h),$$

kde ω_i jsou integrační váhy a N_C je množina integračních uzlů $\mathbf{x}_i \in \bar{\gamma}_C \setminus \bar{\gamma}_D$, které jsou shodné s uzly triangulace. Uvedená aproximace zajišťuje stabilitu Lagrangeových multiplikátorů, které používáme v duální algebraické úloze ([1]).

Úlohu (25) lze ekvivalentně zapsat jako algebraickou variační nerovnici: najdi $(\mathbf{u}, \mathbf{p}, \boldsymbol{\lambda}_n) \in \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \times \mathbb{R}^{n_c}$ takové, že

$$\left. \begin{aligned} & \mathbf{u}^\top \mathbf{A}_\kappa (\mathbf{v} - \mathbf{u}) - (\mathbf{q} - \mathbf{p})^\top \mathbf{E} \mathbf{p} + (\mathbf{v} - \mathbf{u})^\top \mathbf{B}^\top \mathbf{p} + (\mathbf{v} - \mathbf{u})^\top \mathbf{N}^\top \boldsymbol{\lambda}_n + \\ & \quad + \mathbf{g}^\top (|\mathbf{T}\mathbf{v}| - |\mathbf{T}\mathbf{u}|) \geq \mathbf{l}^\top (\mathbf{v} - \mathbf{u}) + \mathbf{c}^\top (\mathbf{q} - \mathbf{p}) \quad \forall (\mathbf{v}, \mathbf{q}) \in \mathbb{R}^{n_u} \times \mathbb{R}^{n_p}, \\ & \mathbf{B}\mathbf{u} = \mathbf{0}, \quad \mathbf{N}\mathbf{u} = \mathbf{0}, \end{aligned} \right\} \quad (26)$$

kde $\mathbf{A}_\kappa \in \mathbb{R}^{n_u \times n_u}$ je symetrická pozitivně definitní matice tuhosti, $\mathbf{E} \in \mathbb{R}^{n_p \times n_p}$ je symetrická pozitivně semidefinitní stabilizační matice (pro dvojici prvků *P1-bubble/P1*), $\mathbf{B} \in \mathbb{R}^{n_p \times n_u}$ je matice reprezentující operátor divergence, $\mathbf{T}, \mathbf{N} \in \mathbb{R}^{m \times n_u}$ jsou matice sestavené z tečných a normálových vektorů na $\bar{\gamma}_C \setminus \bar{\gamma}_D$, $\mathbf{l} \in \mathbb{R}^{n_u}$, $\mathbf{c} \in \mathbb{R}^{n_p}$, $\mathbf{g} \in \mathbb{R}_+^m$ a $|\mathbf{x}| = (|x_1|, \dots, |x_m|)^\top$ pro $\mathbf{x} \in \mathbb{R}^m$; n_p označuje počet uzlů triangulace \mathcal{T}_h , m je počet uzlů \mathbf{x}_i ležících na $\bar{\gamma}_C \setminus \bar{\gamma}_D$ a n_u je celkový počet rychlostních složek. Splnění podmínky neprostupnosti hranice (23)₅ je zajištěno pomocí Lagrangeova multiplikátoru $\boldsymbol{\lambda}_n$.

2.2.3 Duální algebraická úloha

Nechť $L : \mathbb{R}^{n_u} \times \boldsymbol{\Lambda} \rightarrow \mathbb{R}$ je Lagrangeova funkce úlohy (26):

$$L(\mathbf{u}, \boldsymbol{\lambda}) = \frac{1}{2} \mathbf{u}^\top \mathbf{A}_\kappa \mathbf{u} - \frac{1}{2} \mathbf{p}^\top \mathbf{E} \mathbf{p} - \mathbf{u}^\top \mathbf{l} - \mathbf{p}^\top \mathbf{c} + \boldsymbol{\lambda}^\top \mathbf{C} \mathbf{u},$$

kde $\boldsymbol{\Lambda} = \{\boldsymbol{\lambda}_t \in \mathbb{R}^m : |\lambda_t| \leq \mathbf{g}\} \times \mathbb{R}^{m+n_p}$ označuje množinu Lagrangeových multiplikátorů, $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_t^\top, \boldsymbol{\lambda}_n^\top, \mathbf{p}^\top)^\top \in \boldsymbol{\Lambda}$ a $\mathbf{C} = (\mathbf{T}^\top, \mathbf{N}^\top, \mathbf{B}^\top)^\top$. Snadno lze ukázat, že řešení úlohy (26) je ekvivalentní určení sedlového bodu $(\mathbf{u}^*, \boldsymbol{\lambda}^*) \in \mathbb{R}^{n_u} \times \boldsymbol{\Lambda}$, který splňuje:

$$L(\mathbf{u}^*, \boldsymbol{\lambda}) \leq L(\mathbf{u}^*, \boldsymbol{\lambda}^*) \leq L(\mathbf{u}, \boldsymbol{\lambda}^*) \quad \forall (\mathbf{u}, \boldsymbol{\lambda}) \in \mathbb{R}^{n_u} \times \boldsymbol{\Lambda}. \quad (27)$$

Složky $\lambda_{t,i}^*$ Lagrangeova multiplikátoru $\boldsymbol{\lambda}_t^*$ aproximují smyková napětí $-\sigma_t(\mathbf{x}_i)$ při $\kappa = 0$. Odvození duální algebraické úlohy je podobné jako v Sekci 2.1.3. Ze druhé nerovnosti v (27) dostaneme vztah $\mathbf{u}^* = \mathbf{A}_\kappa^{-1}(\mathbf{l} - \mathbf{C}^\top \boldsymbol{\lambda}^*)$, který použijeme pro eliminaci \mathbf{u}^* v první nerovnosti. Po úpravách dospějeme k minimalizační úloze pro duální veličiny:

$$\boldsymbol{\lambda}^* = \arg \min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} q(\boldsymbol{\lambda}), \quad (28)$$

kde $q(\boldsymbol{\lambda}) = \frac{1}{2} \boldsymbol{\lambda}^\top \mathbf{A} \boldsymbol{\lambda} - \boldsymbol{\lambda}^\top \mathbf{b}$, $\mathbf{A} = \mathbf{C} \mathbf{A}_\kappa^{-1} \mathbf{C}^\top + \text{diag}(\mathbf{E}, \mathbf{0}, \mathbf{0})$ je symetrická pozitivně definitní matice a $\mathbf{b} = \mathbf{C} \mathbf{A}_\kappa^{-1} \mathbf{l} - (\mathbf{c}^\top, \mathbf{0}^\top, \mathbf{0}^\top)^\top$. Úloha (28) obsahuje velký počet neznámých bez omezení (odpovídajících tlakové složce řešení). Jak uvidíme později, tento charakter úlohy významně ovlivňuje výběr vhodného optimalizačního algoritmu.

3 Optimalizační algoritmy

Budeme uvažovat takovou formulaci minimalizační úlohy, aby zahrnovala úlohy (21) a (28):

$$\boldsymbol{\lambda}^* = \arg \min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} q(\boldsymbol{\lambda}), \quad (29)$$

kde $q(\boldsymbol{\lambda}) = \frac{1}{2} \boldsymbol{\lambda}^\top \mathbf{A} \boldsymbol{\lambda} - \boldsymbol{\lambda}^\top \mathbf{b}$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ je symetrické pozitivně definitní matice a $\boldsymbol{\lambda}, \mathbf{b} \in \mathbb{R}^n$. Přípustná množina $\boldsymbol{\Lambda}$ bude konvexní a separovatelná:

$$\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_1 \times \dots \times \boldsymbol{\Lambda}_m,$$

kde jednotlivé množiny $\Lambda_i = \{\boldsymbol{\lambda}_i \in \mathbb{R}^{n_i} : f_i(\boldsymbol{\lambda}_i) \leq 0\}$, $i \in \mathcal{M}$ jsou definovány pomocí spojitě diferencovatelných konvexních funkcí $f_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$, $n_i \geq 1$ a $\sum_{i \in \mathcal{M}} n_i = n$. Symbolem $\mathcal{M} = \{1, 2, \dots, m\}$ značíme opět indexovou množinu. Za předpokladu, že $\Lambda \neq \emptyset$, existuje jediné řešení $\boldsymbol{\lambda}^* \in \Lambda$ ([54]). Abychom se vyhnuli zbytečným komplikacím, budeme předpokládat, že množiny Λ_i mají neprázdný vnitřek, na němž nabývají funkce f_i záporných hodnot.

Vhodnou volbou funkcí f_i lze vyjádřit všechna omezení z předchozích odstavců.

(i) Pro $n_i = 1$, $f_i(\lambda_i) = l_i - \lambda_i$ a $l_i \in \mathbb{R}$, dostáváme jednostranné jednoduché omezení:

$$l_i \leq \lambda_i.$$

V úloze (21) tato omezení zajišťují nepronikání těles.

(ii) Pro $n_i = 2$, $f_i(\boldsymbol{\lambda}_i) = \lambda_{i,1}^2 + \lambda_{i,2}^2 - g_i^2$ a $g_i \in \mathbb{R}_+$, dostáváme sférické omezení:

$$\lambda_{i,1}^2 + \lambda_{i,2}^2 \leq g_i^2.$$

Tato omezení zajišťují v úloze (21) splnění podmínky Trescova tření.

(iii) Pro $n_i = 1$, $f_i(\lambda_i) = (\lambda_i + g_i)(\lambda_i - g_i)$ a $g_i \in \mathbb{R}_+$, dostáváme oboustranné jednoduché omezení:

$$-g_i \leq \lambda_i \leq g_i.$$

V úloze (28) se pomocí těchto omezení modeluje přilnutí nebo skluz tekutiny podél stěny.

(iv) Je-li $n_i = 1$ a $f_i(\lambda_i) = -1$, pak λ_i není podrobno žádnému omezení. V úloze (28) tato situace nastává pro tlakové proměnné a pro proměnné zajišťující neproniknutelnost stěny.

Formulace úlohy (29) zahrnuje řadu dalších typů omezení. Například (ii) lze pozměnit pro případ ortotropního tření, které vede na nerovnosti představující elipsy. Snadno lze také provést „posuny“ omezení, které jsou důsledkem homogenizace u FETI metod rozložení oblasti.

Budeme používat následující značení. Symbolem $\|\cdot\|_{\mathbf{A}}$ budeme značit energetickou normu $\|\boldsymbol{\lambda}\|_{\mathbf{A}} = (\boldsymbol{\lambda}^\top \mathbf{A} \boldsymbol{\lambda})^{1/2}$, $\boldsymbol{\lambda} \in \mathbb{R}^n$. Euklidovskou normu a indukovanou maticovou normu budeme značit $\|\cdot\|$. Spektrální číslo podmíněnosti bude definováno výrazem $\kappa(\mathbf{A}) = \sigma_{\max}/\sigma_{\min}$, kde σ_{\min} a σ_{\max} označuje nejmenší a největší vlastní číslo matice \mathbf{A} .

3.1 Metoda aktivních množin

Metody z tohoto odstavce jsou založeny na hledání tzv. „aktivních“ nerovnostních omezení, která jsou pro řešení $\boldsymbol{\lambda}^*$ splněna jako rovnosti. Tyto rovnosti určují příslušné složky vektoru $\boldsymbol{\lambda}^*$. Zbývající složky vyhovují soustavě lineárních rovnic, kterou lze řešit metodou sdružených gradientů (CG). Hledání aktivních omezení a iterace metody CG se vzájemně kombinují. Odstavec navazuje na výsledky prací [39, 40].

3.1.1 Podmínky optimality

Gradient $\mathbf{r} = \mathbf{r}(\boldsymbol{\lambda})$ funkce q v bodě $\boldsymbol{\lambda}$ má tvar

$$\mathbf{r}(\boldsymbol{\lambda}) = \nabla q(\boldsymbol{\lambda}) = \mathbf{A} \boldsymbol{\lambda} - \mathbf{b}.$$

Řešení $\boldsymbol{\lambda}^*$ úlohy (29) je určeno KKT podmínkami optimality ([54]), které lze zapsat ve tvaru:

$$\mathbf{r}_i^* + \mu_i^* \nabla f_i(\boldsymbol{\lambda}_i^*) = \mathbf{0}, \quad f_i(\boldsymbol{\lambda}_i^*) \leq 0, \quad \mu_i^* \geq 0, \quad \mu_i^* f_i(\boldsymbol{\lambda}_i^*) = 0, \quad i \in \mathcal{M},$$

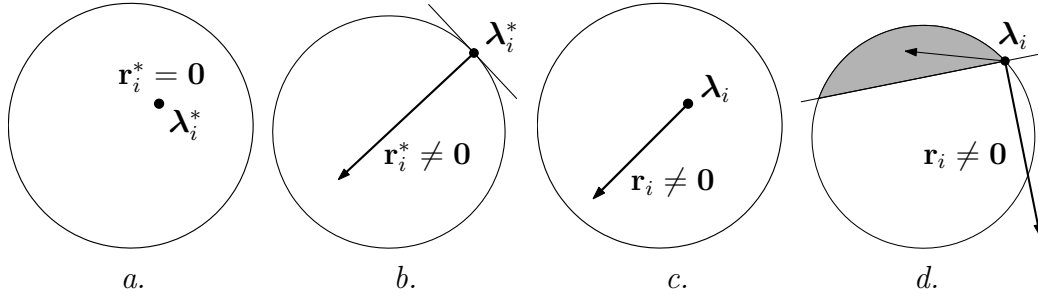
kde μ_i^* , $i \in \mathcal{M}$ jsou Lagrangeovy multiplikátory odpovídající řešení $\boldsymbol{\lambda}^*$ a \mathbf{r}_i^* označuje i -tý subvektor vektoru $\mathbf{r}^* = \mathbf{r}(\boldsymbol{\lambda}^*)$. Eliminací Lagrangeových multiplikátorů dostaneme tento výsledek.

Lemma 1 [39] Vektor $\lambda^* \in \Lambda$ je řešením úlohy (29) právě, když pro každé $i \in \mathcal{M}$ platí:

$$f_i(\lambda_i^*) < 0 \implies \mathbf{r}_i^* = \mathbf{0},$$

$$f_i(\lambda_i^*) = 0 \implies \mathbf{r}_i^* + \frac{\|\mathbf{r}_i^*\|}{\|\nabla f_i(\lambda_i^*)\|} \nabla f_i(\lambda_i^*) = \mathbf{0}.$$

Lemma popisuje rovnovážný stav dosažený v řešení λ^* , který znázorňuje Obrázek 1.a-b. Pro $\lambda_i^* \in \text{Int } \Lambda_i$ odpovídá rovnovážnému stavu nulový vektor \mathbf{r}_i^* . Pro $\lambda_i^* \in \partial \Lambda_i$ je rovnováha dosažena, je-li \mathbf{r}_i^* vektorem vnitřní normály k $\partial \Lambda_i$ v bodě λ_i^* (protože $\nabla f_i(\lambda_i^*)$ je vektor vnější normály), nebo když \mathbf{r}_i^* je vektor nulový.

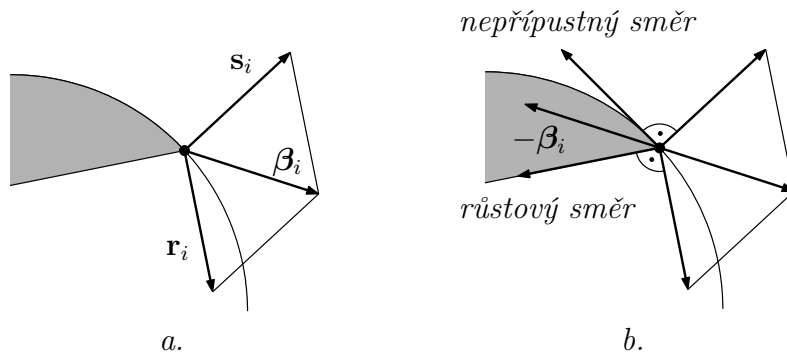


Obrázek 1: Rovnováha a nerovnováha pro kruh Λ_i .

Nechť $\lambda \in \Lambda$ není řešením úlohy (29). Pak pro aspoň jedno $i \in \mathcal{M}$ nastane jeden ze dvou případů: buď je $\lambda_i \in \text{Int } \Lambda_i$ a $\mathbf{r}_i \neq \mathbf{0}$, viz Obrázek 1.c, nebo je $\lambda_i \in \partial \Lambda_i$ a $\mathbf{r}_i \neq \mathbf{0}$ není vektorem vnitřní normály, viz Obrázek 1.d. Subvektory λ_i odpovídající prvnímu případu lze použít pro start metody CG. Ve druhém případě lze vykonat přípustný minimalizační krok směřující do množiny vyznačené na obrázcích šedě. Lze přitom použít vektory

$$\beta_i = \mathbf{r}_i + \frac{\|\mathbf{r}_i\|}{\|\nabla f_i(\lambda_i)\|} \nabla f_i(\lambda_i),$$

kteřé jsou znázorněny na Obrázku 2. V práci [39] jsou tato geometrická pozorování použita pro sestavení algoritmu, který je zobecněním algoritmu pro jednoduchá oboustranná omezení ([7]). Jistou nevýhodou je nespojitost funkce, kterou představují KKT podmínky z Lemmatu 1. Pro návrh algoritmu proto použijeme podmínku optimality založenou na projekci.



Obrázek 2: a.) Definice vektoru β_i , kde $\mathbf{s}_i = (\|\mathbf{r}_i\|/\|\nabla f_i(\lambda_i)\|)\nabla f_i(\lambda_i)$; b.) protože $\|\mathbf{s}_i\| = \|\mathbf{r}_i\|$, vektor $-\beta_i$ pólí úhel mezi nepřípustnými směry a směry vedoucími k růstu hodnoty q .

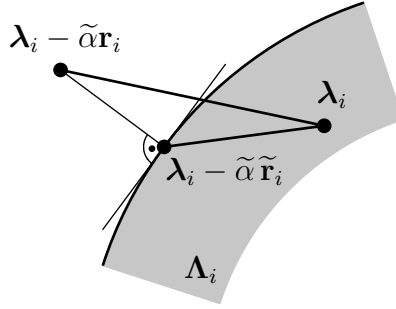
Nechť $\mathbf{P}_\Lambda : \mathbb{R}^n \rightarrow \Lambda$ je ortogonální projekce na Λ . Protože množina Λ je separovatelná, můžeme \mathbf{P}_Λ zapsat pomocí ortogonálních projekcí $\mathbf{P}_{\Lambda_i} : \mathbb{R}^{n_i} \rightarrow \Lambda_i$, $i \in \mathcal{M}$:

$$\begin{aligned}\mathbf{P}_\Lambda(\boldsymbol{\lambda}) &= (\mathbf{P}_{\Lambda_1}(\boldsymbol{\lambda}_1)^\top, \mathbf{P}_{\Lambda_2}(\boldsymbol{\lambda}_2)^\top, \dots, \mathbf{P}_{\Lambda_m}(\boldsymbol{\lambda}_m)^\top)^\top, \\ \mathbf{P}_{\Lambda_i}(\boldsymbol{\lambda}_i) &= \arg \min_{\boldsymbol{\mu} \in \Lambda_i} \|\boldsymbol{\mu} - \boldsymbol{\lambda}_i\|, \quad \boldsymbol{\lambda}_i \in \mathbb{R}^{n_i}.\end{aligned}$$

Redukovaný gradient $\tilde{\mathbf{r}} = \tilde{\mathbf{r}}(\boldsymbol{\lambda})$ v bodě $\boldsymbol{\lambda} \in \Lambda$ je pro pevný krok $\tilde{\alpha} > 0$ definován předpisem

$$\tilde{\mathbf{r}}(\boldsymbol{\lambda}) = \frac{1}{\tilde{\alpha}} (\boldsymbol{\lambda} - \mathbf{P}_\Lambda(\boldsymbol{\lambda} - \tilde{\alpha} \tilde{\mathbf{r}})), \quad (30)$$

viz Obrázek 3. Nulový redukovaný gradient představuje podmínku optimality úlohy (29).



Obrázek 3: Geometrické znázornění redukovaného gradientu.

Lemma 2 [40] Vektor $\boldsymbol{\lambda}^* \in \Lambda$ je řešením úlohy (29) právě, když $\tilde{\mathbf{r}}(\boldsymbol{\lambda}^*) = \mathbf{0}$.

Množinu \mathcal{M} rozložíme v bodě $\boldsymbol{\lambda} \in \Lambda$ na volnou množinu $\mathcal{F}(\boldsymbol{\lambda})$ a na aktivní množinu $\mathcal{A}(\boldsymbol{\lambda})$:

$$\mathcal{F}(\boldsymbol{\lambda}) = \{i \in \mathcal{M} : f_i(\boldsymbol{\lambda}_i) < 0\}, \quad \mathcal{A}(\boldsymbol{\lambda}) = \{i \in \mathcal{M} : f_i(\boldsymbol{\lambda}_i) = 0\}.$$

Redukovaný gradient pak rozložíme na redukovaný volný gradient $\tilde{\boldsymbol{\varphi}} = \tilde{\boldsymbol{\varphi}}(\boldsymbol{\lambda})$ a redukovaný hraniční gradient $\tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}(\boldsymbol{\lambda})$ tak, že

$$\begin{aligned}\tilde{\boldsymbol{\varphi}}_i &= \tilde{\mathbf{r}}_i \quad \text{pro } i \in \mathcal{F}(\boldsymbol{\lambda}), \quad \tilde{\boldsymbol{\varphi}}_i = \mathbf{0} \quad \text{pro } i \in \mathcal{A}(\boldsymbol{\lambda}), \\ \tilde{\boldsymbol{\beta}}_i &= \mathbf{0} \quad \text{pro } i \in \mathcal{F}(\boldsymbol{\lambda}), \quad \tilde{\boldsymbol{\beta}}_i = \tilde{\mathbf{r}}_i \quad \text{pro } i \in \mathcal{A}(\boldsymbol{\lambda}).\end{aligned}$$

Dále definujeme volný gradient $\boldsymbol{\varphi} = \boldsymbol{\varphi}(\boldsymbol{\lambda})$ předpisem

$$\boldsymbol{\varphi}_i = \mathbf{r}_i \quad \text{pro } i \in \mathcal{F}(\boldsymbol{\lambda}), \quad \boldsymbol{\varphi}_i = \mathbf{0} \quad \text{pro } i \in \mathcal{A}(\boldsymbol{\lambda}).$$

3.1.2 Algoritmus KPRGP

Algoritmus používá pevnou délku kroku $\tilde{\alpha} \in (0, 2\sigma_{\max}^{-1})$ a kombinuje tři kroky, které vytvářejí posloupnost $\{\boldsymbol{\lambda}^{(k)}\}$ aproximující řešení $\boldsymbol{\lambda}^*$:

– *expanzivní krok* může přidávat indexy do aktivní množiny:

$$\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} - \tilde{\alpha} \tilde{\boldsymbol{\varphi}}(\boldsymbol{\lambda}^{(k)});$$

– *proporcionalizační krok* může uvolňovat indexy z aktivní množiny:

$$\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} - \tilde{\alpha} \tilde{\boldsymbol{\beta}}(\boldsymbol{\lambda}^{(k)});$$

– *krok sdružených gradientů* má za cíl dosáhnout co největšího poklesu funkční hodnoty funkce q na množině

$$W_{\mathcal{A}(\boldsymbol{\lambda}^{(s)})} = \{\boldsymbol{\lambda} \in \boldsymbol{\Lambda} : \lambda_i = \lambda_i^{(s)}, i \in \mathcal{A}(\boldsymbol{\lambda}^{(s)})\},$$

kde $\boldsymbol{\lambda}^{(s)}$ je výsledkem posledního expanzivního nebo proporcionalizačního kroku, nebo je to počáteční aproximace $\boldsymbol{\lambda}^{(0)}$. Provádí se pomocí jednoho kroku z obvyklé rekurence metody (CG) ([21]):

$$\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} - \alpha_{cg}^{(k)} \mathbf{p}^{(k)}, \quad \alpha_{cg}^{(k)} = \frac{\mathbf{r}(\boldsymbol{\lambda}^{(k)})^\top \mathbf{p}^{(k)}}{(\mathbf{p}^{(k)})^\top \mathbf{A} \mathbf{p}^{(k)}}, \quad (31)$$

kde $\mathbf{p}^{(k)}$ jsou sdružené směry, u nichž jsou nulové složky odpovídající indexům aktivní množiny $\mathcal{A}(\boldsymbol{\lambda}^{(s)})$, tj. $\mathbf{p}_i^{(k)} = \mathbf{0}$ pro $i \in \mathcal{A}(\boldsymbol{\lambda}^{(s)})$. Výpočet sdružených směrů provádíme tak, že začneme s $\mathbf{p}^{(s)} = \boldsymbol{\varphi}(\boldsymbol{\lambda}^{(s)})$ a pak počítáme podle vztahů:

$$\mathbf{p}^{(k)} = \boldsymbol{\varphi}(\boldsymbol{\lambda}^{(k)}) - \gamma^k \mathbf{p}^{(k-1)}, \quad \gamma^{(k)} = \frac{\boldsymbol{\varphi}(\boldsymbol{\lambda}^{(k)})^\top \mathbf{A} \mathbf{p}^{(k-1)}}{(\mathbf{p}^{(k-1)})^\top \mathbf{A} \mathbf{p}^{(k-1)}}, \quad k > s. \quad (32)$$

Rekurence (31), (32) probíhá souvisle dokud nedojde k jejímu přerušení, buď při pokusu vygenerovat nepřipustný krok (ležící vně $\boldsymbol{\Lambda}$), nebo na základě vyhodnocení uvolňovacího kritéria.

Uvolňovacím kritériem nazýváme nerovnost

$$\tilde{\boldsymbol{\beta}}(\boldsymbol{\lambda}^{(k)})^\top \mathbf{r}(\boldsymbol{\lambda}^{(k)}) < \Gamma^2 \tilde{\boldsymbol{\varphi}}(\boldsymbol{\lambda}^{(k)})^\top \mathbf{r}(\boldsymbol{\lambda}^{(k)}), \quad (33)$$

kteřá porovnává aktuální chybu odpovídající aktivní množině (vlevo) s chybou odpovídající volné množině (vpravo) pomocí daného parametru $\Gamma > 0$. Jestliže nerovnost (33) platí, nazýváme aproximaci $\boldsymbol{\lambda}^{(k)}$ *ostře proporcionalní*. Pro ostře proporcionalní $\boldsymbol{\lambda}^{(k)}$ se snažíme pokračovat v iteracích metody CG anebo rozšířit aktivní množinu. V opačném případě se snažíme aktivní množinu zúžit.

Algoritmus uvedeme v jeho zjednodušené podobě, detailní implementace je popsána v [40].

ALGORITMUS KPRGP (*KKT – Proportioning with Reduced Gradient Projections*) Nechť je dáno $\boldsymbol{\lambda}^{(0)} \in \boldsymbol{\Lambda}$, $\Gamma > 0$ a $\tilde{\alpha} \in (0, 2\sigma_{\max}^{-1})$, kde σ_{\max} je největší vlastní číslo matice \mathbf{A} . Pro $k \geq 0$ a známé $\boldsymbol{\lambda}^{(k)}$ urči $\boldsymbol{\lambda}^{(k+1)}$ podle jednoho z následujících pravidel:

- (1°) Je-li $\tilde{\mathbf{r}}(\boldsymbol{\lambda}^{(k)}) = \mathbf{0}$, polož $\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)}$.
- (2°) Je-li $\boldsymbol{\lambda}^{(k)}$ ostře proporcionalní, navrhní $\boldsymbol{\lambda}^{(k+1)}$ krokem sdružených gradientů. Jestliže toto $\boldsymbol{\lambda}^{(k+1)}$ leží v množině $\boldsymbol{\Lambda}$, pak ho přijmi, jinak vypočti $\boldsymbol{\lambda}^{(k+1)}$ expanzivním krokem.
- (3°) Není-li $\boldsymbol{\lambda}^{(k)}$ ostře proporcionalní, vypočti $\boldsymbol{\lambda}^{(k+1)}$ proporcionalizačním krokem.

Uvedené algoritmické schema bylo navrženo v ([16]) pro úlohy s jednostrannými jednoduchými omezeními. Jeho použití pro úlohy s obecnými konvexními separovatelnými omezeními v [40] si vyžádalo úpravu původního uvolňovacího kritéria do podoby (33). Základem analýzy konvergence je následující lemma.

Lemma 3 ([5]) *Nechť σ_{\max} je největší a σ_{\min} je nejmenší vlastní číslo matice \mathbf{A} . Nechť množiny $\boldsymbol{\Lambda}_i$, $i \in \mathcal{M}$ jsou subsymetrické a nechť $\boldsymbol{\lambda}^*$ je řešením úlohy (29). Pak platí:*

$$q(\mathbf{P}_{\boldsymbol{\Lambda}}(\boldsymbol{\lambda} - \tilde{\alpha} \mathbf{r}(\boldsymbol{\lambda}))) - q(\boldsymbol{\lambda}^*) \leq \eta_1(\tilde{\alpha}) (q(\boldsymbol{\lambda}) - q(\boldsymbol{\lambda}^*)) \quad \forall \boldsymbol{\lambda} \in \boldsymbol{\Lambda},$$

kde

$$\eta_1(\tilde{\alpha}) = \begin{cases} 1 - \tilde{\alpha} \sigma_{\min} & \text{pro } \tilde{\alpha} \in [0, \sigma_{\max}^{-1}], \\ 1 - (2\sigma_{\max}^{-1} - \tilde{\alpha}) \sigma_{\min} & \text{pro } \tilde{\alpha} \in [\sigma_{\max}^{-1}, 2\sigma_{\max}^{-1}]. \end{cases}$$

Pro první polovinu intervalu byl důkaz proveden v [40] bez předpokladu o subsymetrii množin Λ_i a podobně byla dokázána i věta o konvergenci. Pomocí Lemmatu 3 lze konvergenční tvrzení formulovat pro $\tilde{\alpha}$ z intervalu dvojnásobné délky.

Věta 1 *Nechť je dáno $\lambda^{(0)} \in \Lambda$, $\Gamma > 0$, $\tilde{\alpha} \in (0, 2\sigma_{\max}^{-1})$, σ_{\max} je největší a σ_{\min} je nejmenší vlastní číslo matice \mathbf{A} . Nechť množiny Λ_i , $i \in \mathcal{M}$, jsou subsymetrické, $\lambda^* \in \Lambda$ je řešením úlohy (29) a $\hat{\Gamma} = \max\{\Gamma, \Gamma^{-1}\}$. Nechť $\{\lambda^k\}$ je posloupnost počítaná algoritmem KPRGP. Pak platí*

$$q(\lambda^{(k+1)}) - q(\lambda^*) \leq \eta_2(\tilde{\alpha}) (q(\lambda^{(k)}) - q(\lambda^*)),$$

kde

$$\eta_2(\tilde{\alpha}) = 1 - \frac{\tilde{\alpha}\sigma_{\min}}{2 + 2\hat{\Gamma}^2} < 1.$$

Odhad chyby v energetické normě má tento tvar:

$$\|\lambda^{(k)} - \lambda^*\|_{\mathbf{A}}^2 \leq 2\eta_2(\tilde{\alpha})^k (q(\lambda^{(0)}) - q(\lambda^*)).$$

Snadno lze také ukázat, že nejmenší hodnota konvergenčního faktoru je

$$\eta_2(\tilde{\alpha}) = 1 - \frac{1}{4}\kappa(\mathbf{A})^{-1}$$

a je jí dosaženo pro $\Gamma = \hat{\Gamma} = 1$ a $\tilde{\alpha} = \sigma_{\max}^{-1}$.

Všechny množiny Λ_i v úlohách (21) a (28) jsou subsymetrické. Příkladem množiny, která subsymetrická není, je například kužel. Konvergence algoritmu KPRGP je u nesubsymetrických množin zaručena pro $\tilde{\alpha} \in (0, \sigma_{\max}^{-1}]$. Analýza konvergence však ošetřuje i nejnepříznivější situaci, při níž musí nastat pokles funkční hodnoty funkce q ve směru vlastního vektoru odpovídajícího největšímu vlastnímu číslu σ_{\max} . Výpočet kroku v tomto směru, obzvláště u rozsáhlých úloh, je velmi málo pravděpodobný, takže algoritmus zpravidla konverguje i při mnohem větších hodnotách $\tilde{\alpha}$ než dovoluje teorie.

V práci [13] se algoritmus KPRGP používá pro řešení vnitřních úloh v metodě nazvané SMALSE-M (*semimonotonic augmented Lagrangians for separable and equality constraints*). Algoritmus SMALSE-M minimalizuje funkci q na přípustné množinou Λ obohacené o rovnostní vazbu a slouží k paralelnímu řešení 3D kontaktních úloh se třením pomocí FETI metod rozložení oblasti [11].

3.1.3 Projekce

Výpočet ortogonálních projekcí \mathbf{P}_{Λ_i} lze ve většině praktický případů provést pomocí explicitních vzorců. Ukážeme tři příklady, které jsou obsaženy v úlohách (21) a (28):

(i) je-li $\Lambda_i = \{\lambda_i \in \mathbb{R} : l_i \leq \lambda_i\}$, $l_i \in \mathbb{R}$, pak

$$\mathbf{P}_{\Lambda_i}(\lambda_i) = \begin{cases} l_i & \text{pro } \lambda_i < l_i, \\ \lambda_i & \text{pro } l_i \leq \lambda_i; \end{cases}$$

(ii) je-li $\Lambda_i = \{\lambda_i \in \mathbb{R}^2 : \|\lambda_i\| \leq g_i\}$, $g_i \in \mathbb{R}_+$, pak

$$\mathbf{P}_{\Lambda_i}(\lambda_i) = \begin{cases} \frac{g_i}{\|\lambda_i\|} \lambda_i & \text{pro } \|\lambda_i\| > g_i, \\ \lambda_i & \text{pro } \|\lambda_i\| \leq g_i; \end{cases}$$

(iii) je-li $\Lambda_i = \{\lambda_i \in \mathbb{R} : |\lambda_i| \leq g_i\}$, $g_i \in \mathbb{R}_+$, pak

$$\mathbf{P}_{\Lambda_i}(\lambda_i) = \begin{cases} -g_i & \text{pro } \lambda_i < -g_i, \\ \lambda_i & \text{pro } |\lambda_i| \leq g_i, \\ g_i & \text{pro } \lambda_i > g_i. \end{cases}$$

U úloh s ortotropním třením je $\Lambda_i = \{\lambda_i \in \mathbb{R}^2 : (\lambda_{i,1}/a_{i,1})^2 + (\lambda_{i,2}/a_{i,2})^2 \leq g_i^2\}$, kde $a_{i,1}, a_{i,2}, g_i \in \mathbb{R}_+$. Ortogonální projekce \mathbf{P}_{Λ_i} je pak určena kořenem jistého polynomu čtvrtého stupně, který dostaneme při zápisu elipsy v polárních souřadnicích. Pro výpočet této projekce se používá vhodná numerická metoda [31].

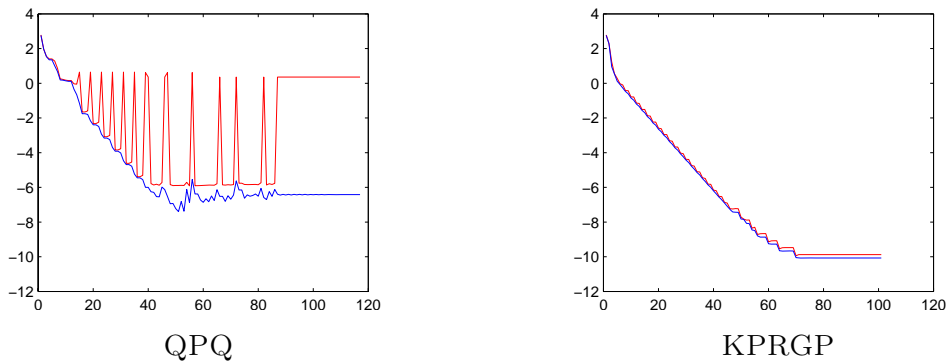
3.1.4 Numerické experimenty

Vlastnosti algoritmu KPRGP představíme na dvou akademických příkladech. Řešení kontaktní úlohy bude ukázáno v dalších odstavcích.

Příklad 1 Uvažujme úlohu (29), kde $\mathbf{A} = \text{tridiag}(-1, 4, -1) \in \mathbb{R}^{12 \times 12}$, $\mathbf{b} = \mathbf{A}\mathbf{y}$, $\mathbf{y} = (2, 1, 0.5, 0, 0, 11, 10^{-5}, -1, \sqrt{2}, -0.1, 4.1 \cdot 10^{-4}, 143)^\top$, $\mathbf{g} = (2, 1, 0.5, 2, 10^{-3}, 154)^\top$ a $\Lambda = \{\lambda \in \mathbb{R}^{12} : \lambda_{2i-1}^2 + \lambda_{2i}^2 \leq g_i^2, i = 1, \dots, 6\}$. V řešení $\lambda^* \in \mathbb{R}^{12}$ jsou aktivní tři omezení: $\mathcal{A}(\lambda^*) = \{2, 3, 5\}$. Nejprve porovnáme výkonnost algoritmu KPRGP s algoritmem QPQ [39], který používá nespojité KKT podmínky optimality z Lemmatu 1. V Tabulce 1 sledujeme počet násobení matice-vektor potřebných pro dosažení srovnatelně přesných výsledků. Algoritmus KPRGP je efektivnější ve většině případů, zejména pro $\Gamma = 1$. Obrázek 4 ukazuje normu redukovaného gradientu (modře) a normu KKT podmínek optimality z Lemmatu 1 (červeně). Je zřejmé, že algoritmus QPQ začíná stagnovat dříve než algoritmus KPRGP a že KKT podmínky optimality nejsou schopny rozpoznat řešení v důsledku nepřesností způsobených počítačovou aritmetikou (první graf, červená čára).

Tabulka 1: Srovnání algoritmů QPQ a KPRGP.

| Γ | 100 | 10 | 5 | 1 | 0.5 | 0.4 | 0.2 | 0.1 | 0.05 | 0.01 | 0.001 |
|----------|-----|----|----|----|-----|-----|-----|-----|------|------|-------|
| QPQ | 43 | 41 | 39 | 32 | 37 | 40 | 48 | 41 | 38 | 36 | 36 |
| KPRGP | 30 | 24 | 21 | 19 | 20 | 20 | 22 | 22 | 26 | 37 | 50 |

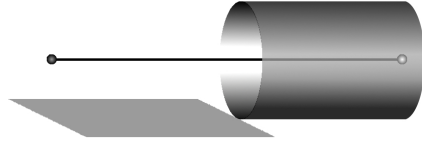


Obrázek 4: Iterační historie pro $\lambda^{(0)} = \mathbf{0}$, $\Gamma = 1$ a $\tilde{\alpha} = \sigma_{\max}^{-1}$.

Příklad 2 Ve druhém příkladu budeme řešit úlohu:

$$\text{minimalizuj } \frac{1}{2} \int_0^1 \|\mathbf{y}'(t)\|^2 dt - \int_0^1 \mathbf{y}(t)^\top \mathbf{f}(t) dt$$

pro $\mathbf{y} = (y_1, y_2)^\top \in K$, kde $K = \{\mathbf{y} \in (H_0^1(0, 1))^2 : y_2(t) \geq l \text{ na } (0, 0.5), \|\mathbf{y}(t)\| \leq g \text{ na } (0.5, 1)\}$ a $\mathbf{f}(t) = (36\pi^2 \sin 6\pi t, -4\pi^2 \sin 2\pi t)^\top$. Úloha popisuje zjednodušený model elastické struny, která je umístěna ve výšce l nad rovinnou překážkou a pak je vsunuta do trubice o poloměru g , viz Obrázek 5. Aproximace metodou konečných prvků pro ekvidistantní síť a $4m$ stupňů volnosti vede na úlohu (29) s přípustnou množinou $\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^{4m} : \lambda_i \geq l, \lambda_{m+i}^2 + \lambda_{2m+i}^2 \leq g^2, i \in \mathcal{M}\}$. V Tabulce 2 uvádíme počet násobení matice-vektor a informaci o aktivních omezeních: $n_{b,A} : n_{b,F} / n_{c,A} : n_{c,F}$, kde $n_{b,A}$, $n_{b,F}$, $n_{c,A}$ a $n_{c,F}$ je počet po řadě aktivních a volných jednoduchých omezení a počet aktivních a volných sférických omezení. Výsledky ukazují, že výkonnost algoritmu je vyšší pro úzkou trubici, kdy je velký počet aktivních omezení. U široké trubice je počet aktivních omezení malý, jsou však téměř ve stavu „slabého kontaktu“, kdy algoritmus před nalezením řešení vykonává dlouhou posloupnost oscilujících iterací.



Obrázek 5: Geometrie elastické struny před deformací.

Tabulka 2: Výpočet pro $l = 0$; jednoduchá i sférická omezení jsou aktivní.

| $4m$ | $g = 1.4$ | $g = 1$ | $g = 0.5$ | $g = 0.3$ | $g = 0.01$ | $g = 0.001$ |
|------|---------------|---------------|---------------|---------------|--------------|-------------|
| | 5:3/2:6 | 6:2/4:4 | 7:1/5:3 | 7:1/5:3 | 8:0/8:0 | 8:0/8:0 |
| 32 | 89 | 80 | 54 | 42 | 17 | 19 |
| | 10:6/2:14 | 11:5/5:11 | 13:3/6:10 | 14:2/9:7 | 16:0/16:0 | 16:0/16:0 |
| 64 | 205 | 240 | 132 | 90 | 23 | 27 |
| | 20:12/4:28 | 22:10/5:27 | 26:6/10:22 | 29:3/16:16 | 32:0//31:1 | 32:0/32:0 |
| 128 | 608 | 620 | 324 | 239 | 96 | 41 |
| | 39:25/4:60 | 45:19/8:56 | 52:12/18:46 | 57:7/26:38 | 64:0/58:6 | 64:0/64:0 |
| 256 | 1677 | 1764 | 951 | 695 | 215 | 121 |
| | 77:51/4:124 | 89:39/12:116 | 104:24/33:95 | 114:14/49:79 | 127:1/111:17 | 128:0/126:2 |
| 512 | 4117 | 5248 | 3755 | 1960 | 646 | 309 |
| | 155:101/6:250 | 177:79/22:234 | 208:48/60:196 | 228:28/95:161 | 253:3/219:37 | 256:0/249:7 |
| 1024 | 12084 | 16851 | 10516 | 7065 | 2051 | 793 |

3.2 Metoda vnitřních bodů

Algoritmy založené na metodě aktivních množin nemění při souvislém běhu iterací metody CG proměnné, které odpovídají aktivní množině, tj. leží na hranicích $\partial\Lambda_i$. Tyto proměnné stagnují nejméně do příštího restartu metody CG. Metoda vnitřních bodů udržuje aproximace řešení uvnitř přípustné množiny, takže se neustále obměňují všechny proměnné. Odstavec shrnuje výsledky práce [45].

3.2.1 Podmínky optimality

Budeme uvažovat úlohu (29) s přípustnou množinou

$$\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^{3m} : \lambda_i \geq l_i, \lambda_{i+m}^2 + \lambda_{i+2m}^2 \leq g_i^2, i \in \mathcal{M}\},$$

kteřá odpovídá 3D kontaktní úloze s Trescovým třením (21) po permutaci neznámých. Důvodem speciální volby množiny Λ je použití Jakobiho matice, jejíž struktura je ovlivněna tvarem omezení.

Nechť $L : \mathbb{R}^n \times \mathbb{R}^{2m} \rightarrow \mathbb{R}$ je Lagrangeova funkce naší úlohy:

$$L(\boldsymbol{\lambda}, \boldsymbol{\nu}) = q(\boldsymbol{\lambda}) + \sum_{i \in \mathcal{M}} \nu_i (l_i - x_i) + \sum_{i \in \mathcal{M}} \nu_{i+m} (\lambda_{i+m}^2 + \lambda_{i+2m}^2 - g_i^2),$$

kde $\boldsymbol{\nu} = (\nu_1, \dots, \nu_{2m})^\top \in \mathbb{R}^{2m}$ je vektor Lagrangeových multiplikátorů. Řešení $\boldsymbol{\lambda}^*$ je první složkou sedlového bodu $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \in \mathbb{R}^n \times \mathbb{R}^{2m}$ funkce L . Sedlový bod je jednoznačně určen následujícím systémem KKT podmínek optimality ([54]):

$$\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{0}, \quad \nabla_{\boldsymbol{\nu}} L(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq \mathbf{0}, \quad \boldsymbol{\nu} \geq \mathbf{0}, \quad \boldsymbol{\nu}^\top \nabla_{\boldsymbol{\nu}} L(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{0}. \quad (34)$$

Zavedením pomocné proměnné $\mathbf{z} = -\nabla_{\boldsymbol{\nu}} L(\boldsymbol{\lambda}, \boldsymbol{\nu})$, $\mathbf{z} \in \mathbb{R}^{2m}$, můžeme (34) zapsat jako

$$\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \mathbf{0}, \quad \nabla_{\boldsymbol{\nu}} L(\boldsymbol{\lambda}, \boldsymbol{\nu}) + \mathbf{z} = \mathbf{0}, \quad \boldsymbol{\nu}^\top \mathbf{z} = \mathbf{0}, \quad \boldsymbol{\nu} \geq \mathbf{0}, \quad \mathbf{z} \geq \mathbf{0}. \quad (35)$$

Rovnosti ze vztahů (35) zapíšeme pomocí funkce $\mathbf{F} : \mathbb{R}^{n+4m} \rightarrow \mathbb{R}^{n+4m}$ definované předpisem

$$\mathbf{F}(\mathbf{v}) = (\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{\lambda}, \boldsymbol{\nu})^\top, (\nabla_{\boldsymbol{\nu}} L(\boldsymbol{\lambda}, \boldsymbol{\nu}) + \mathbf{z})^\top, \mathbf{e}^\top \mathbf{N} \mathbf{Z})^\top,$$

kde $\mathbf{v} = (\boldsymbol{\lambda}^\top, \boldsymbol{\nu}^\top, \mathbf{z}^\top)^\top \in \mathbb{R}^{n+4m}$, $\mathbf{N} = \text{diag}(\boldsymbol{\nu})$, $\mathbf{Z} = \text{diag}(\mathbf{z})$ a $\mathbf{e} = (1, \dots, 1)^\top \in \mathbb{R}^{2m}$. Následující lemma uvádí podmínky optimality ve tvaru, který je základem pro sestavení algoritmu.

Lemma 4 [45] *Řešení $\boldsymbol{\lambda}^*$ úlohy (29) je první složkou vektoru $\mathbf{v}^* = (\boldsymbol{\lambda}^{*\top}, \boldsymbol{\nu}^{*\top}, \mathbf{z}^{*\top})^\top$, který je jednoznačně určen následujícím systémem vztahů*

$$\mathbf{F}(\mathbf{v}) = \mathbf{0}, \quad \boldsymbol{\nu} \geq \mathbf{0}, \quad \mathbf{z} \geq \mathbf{0}. \quad (36)$$

Jakobiho matice $\mathbf{J} = \mathbf{J}(\mathbf{v})$ funkce \mathbf{F} v bodě \mathbf{v} má tento tvar:

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} & \mathbf{0} \\ \mathbf{J}_{21} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{Z} & \mathbf{N} \end{pmatrix},$$

kde

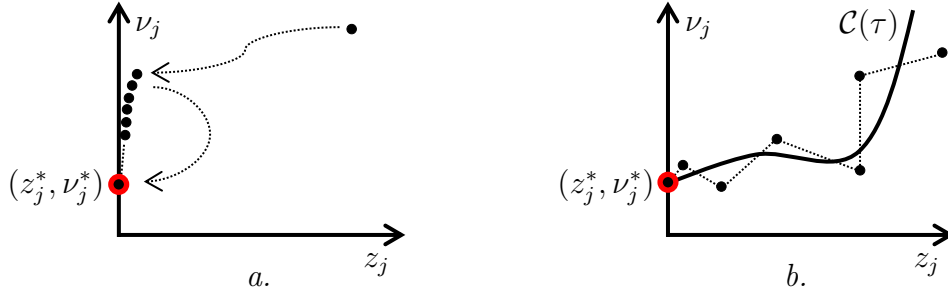
$$\mathbf{J}_{11} = \mathbf{A} + \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{N}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 2\mathbf{N}_2 \end{pmatrix}, \quad \mathbf{J}_{12} = \mathbf{J}_{21}^\top = \begin{pmatrix} -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{X}_2 \\ \mathbf{0} & 2\mathbf{X}_3 \end{pmatrix},$$

$\mathbf{N}_2 = \text{diag}(\nu_{m+1}, \dots, \nu_{2m})$ a $\mathbf{X}_k = \text{diag}(\lambda_{(k-1)m+1}, \dots, \lambda_{km})$, $k = 2, 3$. Snadno lze ukázat, že pro každé $\boldsymbol{\nu} > \mathbf{0}$, $\mathbf{z} > \mathbf{0}$ je matice \mathbf{J} regulární.

Nechť $\mathbf{v}^{(k)} = (\boldsymbol{\lambda}^{(k)\top}, \boldsymbol{\nu}^{(k)\top}, \mathbf{z}^{(k)\top})^\top$, $\boldsymbol{\nu}^{(k)} > \mathbf{0}$, $\mathbf{z}^{(k)} > \mathbf{0}$ je známá aproximace řešení \mathbf{v}^* . Newtonova metoda s tlumením spočívá ve výpočtu newtonovského směru $\Delta \mathbf{v}^{(k+1)}$ ze soustavy lineárních rovnic s Jakobiho maticí a následně v určení nové iterace $\mathbf{v}^{(k+1)}$ pomocí délky kroku $\alpha_k \in (0, 1]$:

$$\mathbf{J}(\mathbf{v}^{(k)}) \Delta \mathbf{v}^{(k+1)} = -\mathbf{F}(\mathbf{v}^{(k)}), \quad \mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \alpha_k \Delta \mathbf{v}^{(k+1)}. \quad (37)$$

Vhodně zvolenou hodnotou α_k můžeme zaručit, aby složky $\boldsymbol{\nu}^{(k+1)}$, $\mathbf{z}^{(k+1)}$ vektoru $\mathbf{v}^{(k+1)}$ byly kladné. Iterační výpočet založený na rovnostech (37) však zpravidla po několika prvních iteracích přiblíží vektory $\boldsymbol{\nu}^{(k)}$, $\mathbf{z}^{(k)}$ k hranici prvního hyperkvadrantu, což způsobí, že v dalších iteracích už lze vykonávat jen krátké kroky α_k . Konvergence je v takovém případě velmi pomalá; viz schematické znázornění na Obrázku 6.a. Algoritmus sestavený v tomto odstavci bude udržovat jednotlivé aproximace řešení hlouběji uvnitř přípustné množiny, takže kroky α_k budou relativně dlouhé po celou dobu výpočtu, jak znázorňuje Obrázek 6.b.



Obrázek 6: Řešení je znázorněno červeně, jednotlivé iterace černě; a.) Newtonova metoda s tlumením; b.) metoda sledování centrální cesty.

3.2.2 Algoritmus PF

Nejprve pozměníme úlohu (36) takto:

$$\mathbf{F}(\mathbf{v}) = (\mathbf{0}^\top, \mathbf{0}^\top, \tau \mathbf{e}^\top)^\top, \quad \boldsymbol{\nu} > \mathbf{0}, \quad \mathbf{z} > \mathbf{0}, \quad (38)$$

kde $\tau > 0$. Řešení \mathbf{v}^τ pomocných úloh (38) definují v \mathbb{R}^{n+4m} křivku $\mathcal{C}(\tau)$, kterou budeme nazývat *centrální cestou*; viz Obrázek 6.b. Tato křivka směřuje do bodu \mathbf{v}^* , jestliže parametr τ konverguje k nule. V našem algoritmu budeme kombinovat Newtonovu metodu aplikovanou na rovnici (38) s poklesem hodnot τ tak, aby iterace sledovaly $\mathcal{C}(\tau)$ a konvergovaly k \mathbf{v}^* . Jednotlivé iterace budou ležet v *okolí* centrální cesty definovaném předpisem

$$\begin{aligned} \mathcal{N}(\gamma, \beta) &= \{ \mathbf{v} = (\boldsymbol{\lambda}^\top, \boldsymbol{\nu}^\top, \mathbf{z}^\top)^\top \in \mathbb{R}^{n+4m} : \|\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{\lambda}, \boldsymbol{\nu})\| \leq \beta \vartheta, \\ &\quad \|\nabla_{\boldsymbol{\nu}} L(\boldsymbol{\lambda}, \boldsymbol{\nu}) + \mathbf{z}\| \leq \beta \vartheta, \quad \boldsymbol{\nu} \geq \mathbf{0}, \quad \mathbf{z} \geq \mathbf{0}, \quad \nu_i z_i \geq \gamma \vartheta, \quad i \in \mathcal{M}_2 \}, \end{aligned}$$

kde $\mathcal{M}_2 = \{1, 2, \dots, 2m\}$, $\beta \geq 0$, $\gamma \in (0, 1]$ a $\vartheta = \vartheta(\mathbf{v}) = (\boldsymbol{\nu}^\top \mathbf{z})/2m$ je *míra duality*. Snadno lze ukázat, že platí $\mathcal{N}(\gamma, \beta) \supseteq \mathcal{C}(\tau)$ a $\mathcal{N}(1, 0) = \mathcal{C}(\tau)$. Dále lze ukázat, že $\mathbf{v} \in \mathcal{N}(\gamma, \beta)$ je řešením (36), pokud je $\nu_i = 0$ nebo $z_i = 0$ pro aspoň jedno $i \in \mathcal{M}_2$.

V k -té iteraci algoritmu sledování centrální cesty pozměníme hodnotu τ pomocí součinu míry duality $\vartheta_k = \vartheta(\mathbf{v}^{(k)})$ a *centrujícího parametru* $c_k \in [c_{\min}, c_{\max}]$, $0 \leq c_{\min} \leq c_{\max} \leq 1$. Při volbě $c_k = 0$ dochází k výpočtu standardního newtonovského směru, zatímco $c_k = 1$ určuje plně centrováný směr. Algoritmus dále používá podmínku *Armijova typu* (40) ([54]) zajišťující pokles hodnot posloupnosti $\{\vartheta_k\}$, což umožňuje dokázat konvergenci.

ALGORITMUS PF (*Path-following*) Nechť $\gamma \in (0, 1]$, $\beta \geq 1$, $0 < c_{\min} \leq c_{\max} \leq 1/2$, $\omega \in (0, 1)$ a $\epsilon \geq 0$. Dále nechť $\mathbf{v}^{(0)} \in \mathcal{N}(\gamma, \beta)$ a $k := 0$.

(1°) Vyber $c_k \in [c_{\min}, c_{\max}]$.

(2°) Je-li $\boldsymbol{\nu}^{(k)} > \mathbf{0}$ a $\mathbf{z}^{(k)} > \mathbf{0}$, vyřeš soustavu lineárních rovnic

$$\mathbf{J}(\mathbf{v}^{(k)}) \Delta \mathbf{v}^{(k+1)} = -\mathbf{F}(\mathbf{v}^{(k)}) + (\mathbf{0}^\top, \mathbf{0}^\top, c_k \vartheta_k \mathbf{e}^\top)^\top, \quad (39)$$

jinak polož $\Delta \mathbf{v}^{(k+1)} = \mathbf{0}$.

(3°) Vypočti

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \alpha_k \Delta \mathbf{v}^{(k+1)}$$

pomocí největšího $\alpha_k \in (0, 1]$, které splňuje $\mathbf{v}^{(k+1)} \in \mathcal{N}(\gamma, \beta)$ a

$$\vartheta_{k+1} \leq (1 - \alpha_k \omega (1 - c_k)) \vartheta_k. \quad (40)$$

(4°) Za výsledek výpočtu vezmi $\bar{\mathbf{v}} = \mathbf{v}^{(k+1)}$, jestliže

$$err^{(k)} := \|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\| / \|\mathbf{v}^{(k+1)}\| \leq \epsilon,$$

jinak polož $k := k + 1$ a pokračuj krokem (1°).

Iteraci algoritmu PF lze vždy vykonat, protože pro $\nu^{(k)} > 0$ a $\mathbf{z}^{(k)} > 0$ je Jakobiho matice regulární a existenci α_k vyhovujícího všem požadavkům z kroku (3°) dokazuje Lemma 5. Algoritmus navazuje na postupy řešení úloh lineárního programování ([57]). Hlavní změna spočívá ve struktuře soustav lineárních rovnic (39), jejichž vlastnosti závisí na matici \mathbf{A} , která se u úloh lineárního programování nevyskytuje. Důležitou roli zde hraje předpomínění, protože Jakobiho matice jsou v okolí řešení \mathbf{v}^* velmi špatně podmíněné a $\mathbf{J}(\mathbf{v}^*)$ může být singulární.

Lemma 5 [45] *Nechť $\gamma \in (0, 1]$, $\beta \geq 1$, $0 < c_{\min} \leq c_{\max} \leq 1/2$, $\omega \in (0, 1)$ a $\hat{\mathbf{v}} \in \mathcal{N}(\gamma, \beta)$, $\hat{\mathbf{v}} \neq \mathbf{v}^*$. Pak existuje $\hat{\delta} > 0$ a $\hat{\alpha} \in (0, 1]$ takové, že délka kroku α_k v kroku (3°) algoritmu PF splňuje $\hat{\alpha} \leq \alpha_k \leq 1$ pro každé $\mathbf{v}^{(k)} \in \mathcal{N}(\gamma, \beta) \cap \mathcal{B}(\hat{\mathbf{v}}, \hat{\delta})$.*

Lemma je základem pro důkaz následující konvergenční věty.

Věta 2 [45] *Nechť $\gamma \in (0, 1]$, $\beta \geq 1$, $0 < c_{\min} \leq c_{\max} \leq 1/2$, $\omega \in (0, 1)$ a $\epsilon = 0$. Nechť posloupnost $\{\mathbf{v}^{(k)}\}$ počítaná algoritmem PF je omezená. Potom je posloupnost $\{\mathbf{v}^{(k)}\}$ buďto konečná a její poslední člen $\bar{\mathbf{v}}$ je řešením úlohy (36), nebo je nekonečná a řešením úlohy (36) je její limitní bod.*

Toto konvergenční tvrzení je slabší než tvrzení analogické Věty 1 pro algoritmus KPRGP. Pro posouzení správnosti vypočítaného řešení je však dostačující.

3.2.3 Předpokládání

Výpočetní efektivita algoritmu závisí ve značné míře na způsobu řešení soustav lineárních rovnic (39). Pro větší přehlednost zde budeme vynechávat iterační index k . Soustavu (39) pak můžeme zapsat jako

$$\begin{pmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} & \mathbf{0} \\ \mathbf{J}_{21} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{Z} & \mathbf{N} \end{pmatrix} \begin{pmatrix} \Delta\lambda \\ \Delta\nu \\ \Delta\mathbf{z} \end{pmatrix} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix}, \quad (41)$$

kde $\mathbf{r}_1 \in \mathbb{R}^n$, $\mathbf{r}_2, \mathbf{r}_3 \in \mathbb{R}^{2m}$ jsou složky vektoru pravé strany. Ukážeme dvě metody řešení této soustavy založené na použití Schurova doplňku ([2]).

První metoda pracuje se symetrickou pozitivně definitní maticí $\mathbf{J}_{SC} = \mathbf{J}_{SC}(\mathbf{v})$ definovanou předpisem

$$\mathbf{J}_{SC} = \mathbf{J}_{11} + \mathbf{J}_{12}\mathbf{D}_{22}^{-1}\mathbf{J}_{21},$$

kde $\mathbf{D}_{22} = \mathbf{N}^{-1}\mathbf{Z}$. Jedná se o matici redukované soustavy, která vznikne z (41) eliminací neznámých $\Delta\nu$ a $\Delta\mathbf{z}$ podle vztahu

$$\begin{pmatrix} \Delta\nu \\ \Delta\mathbf{z} \end{pmatrix} = \begin{pmatrix} -\mathbf{D}_{22}^{-1} & \mathbf{Z}^{-1} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \left(\begin{pmatrix} \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix} - \begin{pmatrix} \mathbf{J}_{21} \\ \mathbf{0} \end{pmatrix} \Delta\lambda \right). \quad (42)$$

Redukovaná soustava pro neznámou $\Delta\lambda$ má tvar

$$\mathbf{J}_{SC}\Delta\lambda = \mathbf{r}_1 - \mathbf{J}_{12}(\mathbf{Z}^{-1}\mathbf{r}_3 - \mathbf{D}_{22}^{-1}\mathbf{r}_2). \quad (43)$$

Druhá metoda používá symetrickou indefinitní matici $\mathbf{J}_{AM} = \mathbf{J}_{AM}(\mathbf{v})$ definovanou předpisem

$$\mathbf{J}_{AM} = \begin{pmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & -\mathbf{D}_{22} \end{pmatrix},$$

což je matice redukované soustavy, kterou z (41) dostaneme eliminací neznámé $\Delta\mathbf{z}$:

$$\Delta\mathbf{z} = \mathbf{N}^{-1}\mathbf{r}_3 - \mathbf{D}_{22}\Delta\boldsymbol{\nu}. \quad (44)$$

Redukovaná soustava pro neznámé $\Delta\boldsymbol{\lambda}$ a $\Delta\boldsymbol{\nu}$ má tvar

$$\mathbf{J}_{AM} \begin{pmatrix} \Delta\boldsymbol{\lambda} \\ \Delta\boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 - \mathbf{N}^{-1}\mathbf{r}_3 \end{pmatrix}. \quad (45)$$

Řešení soustavy (41) provedeme ve dvou krocích: nejdříve vhodnou iterační metodou vyřešíme (43) respektive (45) a pak vypočítáme zbývající neznámé pomocí (42) respektive (44). Výpočet řešení soustav (43) respektive (45) chceme provést metodou CG, která vyžaduje dobře podmíněnou symetrickou pozitivně definitní matici ([21]). Protože špatná podmíněnost Jakobihovy matice se přenáší na matice \mathbf{J}_{SC} a \mathbf{J}_{AM} , je nutné soustavy rovnic předpokládat. Pro \mathbf{J}_{AM} používáme indefinitní předpodmiňovač

$$\mathbf{P}_{AM} = \begin{pmatrix} \mathbf{D}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & -\mathbf{D}_{22} \end{pmatrix},$$

kde $\mathbf{D}_{11} = \text{diag}(\mathbf{J}_{11})$ je matice, která nutně obsahuje $\mathbf{D} = \text{diag}(\mathbf{A})$. Pro \mathbf{J}_{SC} používáme jako předpodmiňovač matici

$$\mathbf{P}_{SC} = \mathbf{D}_{11} + \mathbf{J}_{12}\mathbf{D}_{22}^{-1}\mathbf{J}_{21}.$$

Akce inverzí \mathbf{P}_{AM}^{-1} a \mathbf{P}_{SC}^{-1} , které vyžaduje předpokládaná metoda CG, jsou výpočetně náročné, protože \mathbf{P}_{AM} sestává z diagonálních bloků a \mathbf{P}_{SC} je dokonce diagonální. Analýzou zobecněných úloh na vlastní čísla lze dokázat, že spektra předpokládaných matic $\mathbf{P}_{AM}^{-1}\mathbf{J}_{AM}$ a $\mathbf{P}_{SC}^{-1}\mathbf{J}_{SC}$ jsou omezena nezávisle na iteraci $\mathbf{v} = \mathbf{v}^{(k)}$. To je klíčový výsledek pro praktickou využitelnost algoritmu PF.

Věta 3 [45] *Všetchna vlastní čísla matice $\mathbf{P}_{AM}^{-1}\mathbf{J}_{AM}$ jsou kladná. Vlastní číslo 1 má násobnost $2m$ a zbývajících n vlastních čísel jsou vlatními čísly matice $\mathbf{P}_{SC}^{-1}\mathbf{J}_{SC}$.*

Věta 4 [45] *Nechť $\sigma_{\min}(\mathbf{A})$ a $\sigma_{\max}(\mathbf{A})$ jsou nejmenší a největší vlastní číslo matice \mathbf{A} . Nechť $\sigma_{\min}(\mathbf{D}) = \min_i\{a_{ii}\}$ a $\sigma_{\max}(\mathbf{D}) = \max_i\{a_{ii}\}$ jsou nejmenší a největší vlastní číslo matice \mathbf{D} , kde a_{ii} jsou diagonální prvky matice \mathbf{A} . Všetchna vlastní čísla matice $\mathbf{P}_{SC}^{-1}\mathbf{J}_{SC}$ leží v intervalu*

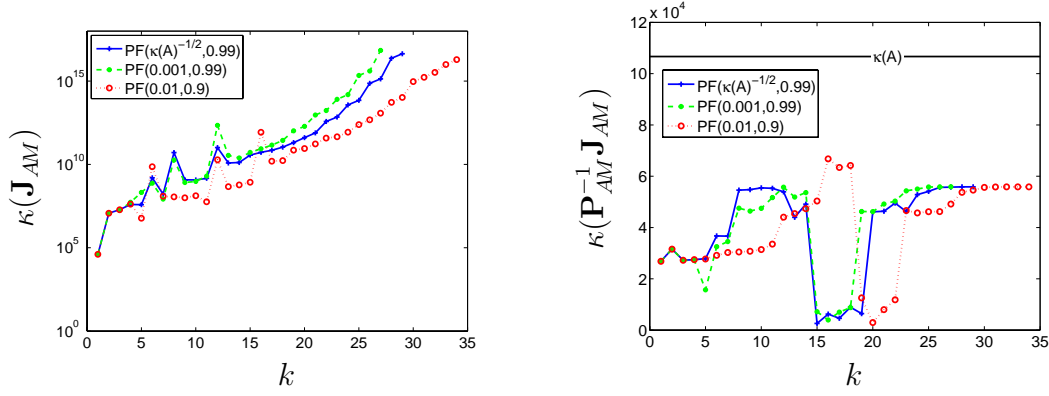
$$[\sigma_{\min}(\mathbf{A})/\sigma_{\max}(\mathbf{D}), \sigma_{\max}(\mathbf{A})/\sigma_{\min}(\mathbf{D})].$$

Z Věty 4 dostaneme horní odhad pro číslo podmíněnosti předpokládané matice \mathbf{J}_{SC} :

$$\kappa(\mathbf{P}_{SC}^{-1}\mathbf{J}_{SC}) \leq \frac{\sigma_{\max}(\mathbf{A})\sigma_{\max}(\mathbf{D})}{\sigma_{\min}(\mathbf{A})\sigma_{\min}(\mathbf{D})} = \kappa(\mathbf{A})\kappa(\mathbf{D}).$$

Z Věty 3 a nerovností $\sigma_{\min}(\mathbf{A})/\sigma_{\max}(\mathbf{D}) \leq 1 \leq \sigma_{\max}(\mathbf{A})/\sigma_{\min}(\mathbf{D})$ vyplývá stejný odhad i pro číslo podmíněnosti $\kappa(\mathbf{P}_{AM}^{-1}\mathbf{J}_{AM})$.

Ověření tohoto výsledku provedeme při řešení úlohy z Příkladu 2 pro $l = 0$, $g = 1.4$ a $n = 4m = 1024$. Obrázek 7 ukazuje, že čísla podmíněnosti nepředpokládaných matic v průběhu iteračního výpočtu rostou do extrémně vysokých hodnot, zatímco u předpokládaných matic zůstávají pod teoretickým odhadem $\kappa(\mathbf{A}) \doteq 1.07 \times 10^5$ (protože $\kappa(\mathbf{D}) = 1$).



Obrázek 7: Předpodmínění matice \mathbf{J}_{AM} při různých nastaveních vnitřního adaptivního ukončovacího kritéria; zápis $\text{PF}(r_{tol}, c_{fact})$ obsahuje hodnoty parametrů ze Sekce 3.2.4 bod (ii).

3.2.4 Implementace

Optimální implementace algoritmu PF závisí na řadě parametrů, jejichž volbu je potřeba posoudit podle výsledků numerických experimentů [45].

(i) Počáteční vektor pro metodu CG v k -té iteraci algoritmu PF se volí jako výsledný vektor z iterace předchozí. Vnitřní iterace metody CG se ukončují adaptivní ukončovací přesností $\epsilon_{cgm} = \epsilon_{cgm}^{(k)} \times \|\mathbf{r}^{(k)}\|$, kde $\mathbf{r}^{(k)}$ je vektor pravé strany v (43) nebo (45). Hodnota $\epsilon_{cgm}^{(k)}$ se nastavuje proporcionalně vzhledem k aktuální přesnosti pro vnější iterace $err^{(k-1)}$ a pokud to nevede k dostatečnému zpřesnění, zvýší se přesnost $\epsilon_{cgm}^{(k-1)}$ z předchozí iterace:

$$\epsilon_{cgm}^{(k)} = \min\{r_{tol} \times err^{(k-1)}, c_{fact} \times \epsilon_{cgm}^{(k-1)}\},$$

kde $0 < r_{tol} < 1$, $0 < c_{fact} < 1$, $err^{(-1)} = 1$ a $\epsilon_{cgm}^{(-1)} = r_{tol}/c_{fact}$ (např. $r_{tol} = 0.1$ a $c_{fact} = 0.9$).

(ii) Hodnota $c_k \in [c_{\min}, c_{\max}]$ se volí adaptivně pomocí následujícího předpisu ([53]):

$$c_k = \min\{c_{\max}, \max\{c_{\min}, c_s \times ((1 - \xi_k)/\xi_k)^3\}\},$$

kde $\xi_k = \min_{i \in \mathcal{M}_2} \{\nu_i^{(k)} z_i^{(k)}\} / \vartheta_k$ a $c_s > 0$. Parametr ξ_k leží v intervalu $[\gamma, 1]$, protože $\nu_i^{(k)} z_i^{(k)} \geq \gamma \vartheta_k$. Hodnotu $\xi_k = 1$ dostaneme, je-li $\nu_i^{(k)} z_i^{(k)} = \vartheta_k$ pro všechna $i \in \mathcal{M}_2$, což nastane pro $\mathbf{v}^{(k)}$ ležící blízko centrální cesty. V takovém případě $c_k = c_{\min}$ a $\Delta \mathbf{v}^{(k+1)}$ bude blízko standardnímu newtonovskému směru. Naopak při $\xi_k = \gamma$ je vektor $\mathbf{v}^{(k)}$ na hranici množiny $\mathcal{N}(\gamma, \beta)$ a hodnota $c_k > c_{\min}$ způsobí centrování. Kritická hodnota, kdy $c_k = c_{\max}$, je $\xi_k^{crit} = 1/(1 + \sqrt[3]{c_{\max}/c_s})$. Ve výpočtech používáme $c_{\min} = 10^{-30}$, $c_{\max} = 0.5$ a $c_s = 1.25 \times 10^{-5}$ ($\xi_k^{crit} \doteq 0.0284$).

(iii) Volbou délky kroku α_k v kroku (3°) algoritmu PF musíme zaručit splnění všech nerovností definujících množinu $\mathcal{N}(\gamma, \beta)$ a (40). Hodnota α_k se vypočítá zkracováním kroku:

- Zvol $\rho \in (0, 1)$, $\alpha_k^0 \in (0, 1]$ a $q := 0$.
- Pokud $ineq(\alpha_k^q)$ neplatí, polož $\alpha_k^{q+1} = \rho \alpha_k^q$, $q := q + 1$ a opakuj.
- Jako výsledek vezmi $\alpha_k = \alpha_k^q$.

Počáteční hodnota α_k^0 se určí podle předpisu

$$\alpha_k^0 = \min_{\Delta \nu_i^{(k+1)}, \Delta z_i^{(k+1)} < 0} \{1, -\delta \nu_i^{(k)} / \Delta \nu_i^{(k+1)}, -\delta z_i^{(k)} / \Delta z_i^{(k+1)}\},$$

kde $\delta \in (0, 1)$, což zaručí $\nu^{k+1} > 0$, $\mathbf{z}^{k+1} > 0$. Splnění zbývajících nerovností dosáhneme tak, že v cyklu (b) nahradíme $ineq(\alpha_k^q)$ příslušnou nerovností. Testování provádíme od výpočetně nejméně náročné nerovnosti postupně ke složitějším. Používáme přitom různé hodnoty ρ .

Nejdříve zaručíme splnění $\nu_i^{(k+1)} z_i^{(k+1)} \geq \gamma \vartheta_{k+1}$ pro všechna $i \in \mathcal{M}_2$ s $\rho = \rho_1$, pak (40) s $\rho = \rho_2$ a nakonec $\|\nabla_{\lambda} L(\boldsymbol{\lambda}^{(k+1)}, \boldsymbol{\nu}^{(k+1)})\| \leq \beta \vartheta_{k+1}$ a $\|\nabla_{\nu} L(\boldsymbol{\lambda}^{(k+1)}, \boldsymbol{\nu}^{(k+1)}) + \mathbf{z}^{(k+1)}\| \leq \beta \vartheta_{k+1}$ s $\rho = \rho_3$. Vlastnost ukončení jednotlivých cyklů plyne z Lemmatu 5. Ve výpočtech používáme tyto hodnoty parametrů: $\delta = 0.999$, $\rho_1 = \rho_2 = 0.9$ a $\rho_3 = 0.5$.

(iv) Volbou $\omega \in (0, 1)$ v (40) dosáhneme jistého kompromisu mezi délkou kroku α_k a poklesem hodnot posloupnosti $\{\vartheta_k\}$. Pro ω blízké jedné vznikají krátké kroky α_k , zatímco pro ω blízké nule nastane malý pokles mezi ϑ_k a ϑ_{k+1} . Volíme $\omega = 0.1$.

(v) Nakonec ukážeme jak zajistit, aby $\mathbf{v}^{(0)} = (\boldsymbol{\lambda}^{(0)\top}, \boldsymbol{\nu}^{(0)\top}, \mathbf{z}^{(0)\top})^\top$ bylo prvkem množiny $\mathcal{N}(\gamma, \beta)$. Libovolně zvolíme $\boldsymbol{\lambda}^{(0)} \in \mathbb{R}^n$, např. $\boldsymbol{\lambda}^{(0)} = \mathbf{0}$. Vektory $\boldsymbol{\nu}^{(0)} \in \mathbb{R}^{2m}$ a $\mathbf{z}^{(0)} \in \mathbb{R}^{2m}$ zvolíme pomocí zadaných kladných konstant c_ν a c_z , tj. $\boldsymbol{\nu}^{(0)} = c_\nu \mathbf{e}$ a $\mathbf{z}^{(0)} = c_z \mathbf{e}$. Platí $\nu_i^{(0)} z_i^{(0)} = c_\nu c_z \geq \gamma c_\nu c_z = \gamma \vartheta_0$ pro všechna $i \in \mathcal{M}_2$ a pro každé $\gamma \in (0, 1]$, takže volba γ nepodléhá žádnému omezení. Volbu parametru β je potřeba přizpůsobit počáteční aproximaci. Nejdříve vypočítáme

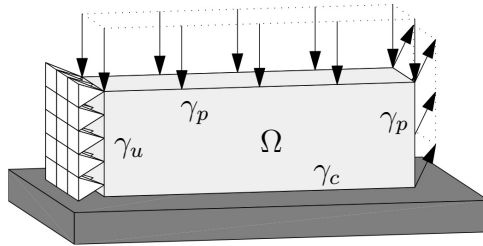
$$\bar{\beta} = \beta_0 \times \max\{\|\nabla_{\lambda} L(\boldsymbol{\lambda}^{(0)}, \boldsymbol{\nu}^{(0)})\|/\vartheta_0, \|\nabla_{\nu} L(\boldsymbol{\lambda}^{(0)}, \boldsymbol{\nu}^{(0)}) + \mathbf{z}^{(0)}\|/\vartheta_0\}$$

pro dané $\beta_0 \geq 1$. Protože vyžadujeme $\beta \geq 1$, zvolíme $\beta = \max\{1, \bar{\beta}\}$. Volbou vysoké hodnoty β snížíme výpočetní nároky, protože nejnáročnější testy pro derivace Lagrangeovy funkce budou většinou splněny a priori. Používáme: $c_\nu = c_z = 1$, $\gamma = 0.001$ a $\beta_0 = 10^9$.

3.2.5 Numerické experimenty

V prvním příkladě budeme řešit 3D kontaktní úlohu s Trescovým třením a posoudíme výkonnost algoritmu PF(r_{tol}, c_{fact}) s různým nastavením r_{tol}, c_{fact} a algoritmu KPRGP [45]. Druhý příklad bude věnován 2D úloze proudění se skluzovou okrajovou podmínkou (23) [41].

Příklad 3 Uvažujme $\Omega = (0, 3) \times (0, 1) \times (0, 1)$ jako ocelovou cihlu ležící na podložce. Dekompozice hranice $\partial\Omega$ je následující: $\gamma_u = \{0\} \times (0, 1) \times (0, 1)$, $\gamma_c = (0, 3) \times (0, 1) \times \{0\}$ a $\gamma_p = \partial\Omega \setminus (\bar{\gamma}_u \cup \bar{\gamma}_c)$, viz Obrázek 8, z něhož je také patrné zadání povrchového napětí. Objemové síly jsou nulové. Mez skluzu má hodnotu $g = 0.3$. Algebraická struktura úlohy je stejná jako v případě formulace pro dvě tělesa (21). Tabulka 3 uvádí počty násobení matice-vektor n_A , vnější (newtonovské) iterace $iter$, kterými dosáhneme relativní přesnosti výsledků na stejné úrovni 10^{-6} , a výpočetní čas CPU v sekundách. Výsledky ukazují na vyšší efektivitu algoritmu PF ve všech případech.



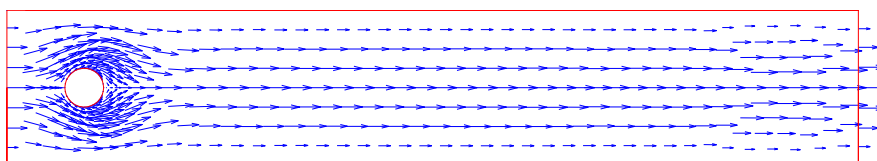
Obrázek 8: Ocelová cihla na podložce.

Příklad 4 Úlohou (23) budeme modelovat proudění v kanále s kruhovou překážkou, na níž je předepsána skluzová podmínka. Nechť $\Omega = (-0.2, 2) \times (-0.2, 0.2) \setminus \mathcal{C}$, kde \mathcal{C} je kruh se středem v počátku o poloměru 0.05. Hranice $\partial\Omega$ je rozdělena na části s různými okrajovými podmínkami takto: $\gamma_D = \gamma_{D,1} \cup \gamma_{D,2} \cup \gamma_{D,3}$, $\gamma_{D,1} = (-0.2, 2) \times \{-0.2\}$, $\gamma_{D,2} = (-0.2, 2) \times \{0.2\}$, $\gamma_{D,3} = \{-0.2\} \times (-0.2, 0.2)$, $\gamma_N = \{2\} \times (-0.2, 0.2)$ a $\gamma_C = \partial\mathcal{C}$. Úlohu řešíme pro data $\mathbf{f} = \mathbf{0}$, $\nu = 1$, $\mathbf{u}_D|_{\gamma_{D,1} \cup \gamma_{D,2}} = \mathbf{0}$, $\mathbf{u}_D|_{\gamma_{D,3}} = 7.5(0.04 - y^2, 0)$, $y \in (-0.2, 0.2)$ a $\boldsymbol{\sigma}_N = \mathbf{0}$. Rychlostní pole

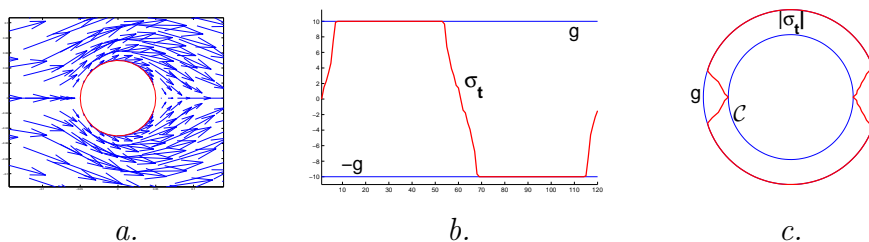
Tabulka 3: Trescovo tření, vnitřní řešič používá matici \mathbf{J}_{SC} .

| n_u/m | PF(0.3, 0.99) | | PF(0.1, 0.9) | | PF(0.01, 0.99) | | KPRGP | |
|------------|---------------|--------|--------------|--------|----------------|--------|-------|--------|
| | $iter/n_A$ | CPU | $iter/n_A$ | CPU | $iter/n_A$ | CPU | n_A | CPU |
| 900/180 | 15/87 | 0.28 | 14/92 | 0.28 | 14/119 | 0.36 | 187 | 0.48 |
| 2646/378 | 13/85 | 1.61 | 13/96 | 1.78 | 12/105 | 1.89 | 181 | 2.70 |
| 5832/648 | 15/91 | 6.43 | 14/104 | 7.11 | 15/130 | 8.47 | 187 | 10.14 |
| 10890/990 | 17/120 | 23.20 | 15/116 | 22.17 | 17/166 | 29.72 | 203 | 30.64 |
| 18252/1404 | 16/106 | 48.22 | 15/116 | 51.51 | 15/147 | 62.04 | 230 | 78.13 |
| 28350/1890 | 17/121 | 111.18 | 16/125 | 113.40 | 16/151 | 131.90 | 254 | 177.31 |
| 41616/2448 | 18/130 | 226.19 | 17/130 | 224.11 | 17/175 | 285.51 | 259 | 344.97 |

pro mez skluzu $g = 10$ je znázorněno na Obrázku 9. Na kruhové překážce dochází k přilnutí i k prokluzu, viz Obrázek 10. V Tabulce 4 zaznamenáváme počet násobení matice-vektor potřebných pro výpočet srovnatelně přesných výsledků. Poznamenejme, že pro $g = 30$ nastává pouze přilnutí. Zatímco výpočetní nároky algoritmu založeného na metodě aktivních množin AS výrazně rostou pro jemnější sítě, u algoritmu PF jsou stále na přijatelné úrovni. Rozdílné chování porovnávaných algoritmů lze vysvětlit velkým počtem duálních neznámých, na něž není kladeno žádné omezení. V našem případě je $n = m = 2n_c + n_c^2$, avšak počet omezení je pouze n_c . Algoritmus AS je varianta algoritmu KPRGP pro oboustranná jednoduchá omezení.



Obrázek 9: Obtékání kruhové překážky.



Obrázek 10: a.) Rychlostní pole v okolí kruhové překážky. b.- c.) rozložení hodnot σ_t podél skluzové hranice.

3.3 Semi-hladká Newtonova metoda

Použití semi-hladké Newtonovy metody ukážeme pro 3D kontaktní úlohu s Trescovým třením. Na rozdíl od předchozích optimalizačních algoritmů se tato metoda aplikuje na primárně-duální formulaci kontaktní úlohy (16)-(18). Pro duální implementaci metody však ihned ukážeme souvislost s metodou aktivních množin ze Sekce 3.1, což také umožňuje dokázat podobný konvergenční výsledek. Tento odstavec se opírá o práce [46, 47].

Tabulka 4: Výkonnost algoritmů AS a PF (v závorkách je výpočetní čas v sekundách).

| $n_u/n_p/n_c$ | $g = 10$ | | $g = 30$ | |
|-----------------|----------------|------------|-------------|------------|
| | AS | PF | AS | PF |
| 862/526/30 | 6867(7.2) | 208(0.6) | 682(1.2) | 181(0.2) |
| 3566/1972/60 | 22579(118.6) | 283(12.4) | 1060(8.4) | 268(1.2) |
| 14494/7624/120 | 130615(4110.4) | 374(12.4) | 1816(59.2) | 389(12.7) |
| 58430/29968/240 | >200000 | 459(135.4) | 3083(849.1) | 454(130.7) |

3.3.1 Podmínky optimality

Budeme navazovat na Sekci 2.1.3 s následující výhradou: namísto součinu $\mathcal{F}_i g_i$ budeme psát pouze g_i , $i \in \mathcal{M}$. Budeme také používat množiny Lagrangeových multiplikátorů

$$\Lambda_\nu = \mathbb{R}_+^m, \quad \Lambda_t = \Lambda_{t,1}(g_1) \times \cdots \times \Lambda_{t,m}(g_m),$$

kde $\Lambda_{t,i}(g_i) = \{\lambda_{t,i} \in \mathbb{R}^2 : \|\lambda_{t,i}\| \leq g_i\}$, $\lambda_{t,i} = (\lambda_{t,2i-1}, \lambda_{t,2i})^\top$, $g_i \geq 0$, $i \in \mathcal{M}$ a ortogonální projekce na tyto množiny $\mathbf{P}_{\Lambda_\nu} : \mathbb{R}^m \rightarrow \Lambda_\nu$ a $\mathbf{P}_{\Lambda_t} : \mathbb{R}^{2m} \rightarrow \Lambda_t$. Protože obě množiny Λ_ν i Λ_t jsou separovatelné, můžeme projekce zapsat po složkách. Pro projekci \mathbf{P}_{Λ_ν} jsou její složky ortogonálními projekcemi na \mathbb{R}_+ :

$$\mathbf{P}_{\Lambda_\nu, i}(\lambda_\nu) = \max\{0, \lambda_{\nu, i}\}, \quad i \in \mathcal{M}.$$

Projekci \mathbf{P}_{Λ_t} lze zapsat pomocí ortogonálních projekcí ψ_{g_i} na kruhy $\Lambda_{t,i}(g_i)$, $i \in \mathcal{M}$:

$$\psi_{g_i}(\lambda_{t,i}) = \begin{cases} \lambda_{t,i} & \text{pro } \|\lambda_{t,i}\| \leq g_i, \\ \frac{g_i}{\|\lambda_{t,i}\|} \lambda_{t,i} & \text{pro } \|\lambda_{t,i}\| > g_i, \end{cases}$$

takže $\mathbf{P}_{\Lambda_t}(\lambda_t) = (\psi_{g_1}(\lambda_{t,1})^\top, \dots, \psi_{g_m}(\lambda_{t,m})^\top)^\top$.

Definujme funkci $\mathbf{G} : \mathbb{R}^{n_u+3m} \rightarrow \mathbb{R}^{n_u+3m}$ předpisem

$$\mathbf{G}(\mathbf{y}) = \begin{pmatrix} \mathbf{K}\mathbf{u} + \mathbf{N}^\top \lambda_\nu + \mathbf{T}^\top \lambda_t - \mathbf{f} \\ \lambda_\nu - \mathbf{P}_{\Lambda_\nu}(\lambda_\nu + \rho(\mathbf{N}\mathbf{u} - \mathbf{d})) \\ \lambda_t - \mathbf{P}_{\Lambda_t}(\lambda_t + \rho\mathbf{T}\mathbf{u}) \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} \mathbf{u} \\ \lambda_\nu \\ \lambda_t \end{pmatrix} \in \mathbb{R}^{n_u+3m}, \quad (46)$$

kde $\rho > 0$ je libovolný pevný parametr. Snadno lze ukázat, že KKT podmínky optimality (16)-(18) jsou splněny pro $\mathbf{y}^* = (\mathbf{u}^{*\top}, \lambda_\nu^{*\top}, \lambda_t^{*\top})^\top$, jestliže platí

$$\mathbf{G}(\mathbf{y}^*) = \mathbf{0}. \quad (47)$$

Funkce \mathbf{G} je nediferencovatelná. Je však „semi-hladká“ ve smyslu následujícího odstavce.

3.3.2 Diferencovatelnost „slantly“

Protože český překlad „nakloněná diferencovatelnost“ se nepoužívá, budeme mluvit o funkcích diferencovatelných „slantly“. Jedná se o typ diferencovatelnosti, kdy funkce představující derivaci musí nabývat „rozumných“ hodnot na okolí bodu, ale její hodnota v bodě samotném je nepodstatná. Připomíná to situaci u Lebesgueova integrálu, kdy nezáleží na množině míry nula. V konečném důsledku to pak znamená, že pomocí newtonovských iterací lze řešit rovnici (47), kde funkce \mathbf{G} je na množině míry nula nediferencovatelná. Hlavní pojmy zavedeme abstraktně ([35]).

Nechť Y, Z jsou Banachovy prostory s normami $\|\cdot\|_Y, \|\cdot\|_Z$ a $\mathcal{L}(Y, Z)$ označuje prostor všech ohraničených lineárních zobrazení z Y do Z s normou $\|\cdot\|_{\mathcal{L}(Y, Z)}$. Nechť $U \subseteq Y$ je otevřená množina a $G : U \rightarrow Z$ je daná funkce.

Definice 1 (i) Funkce G se nazývá slantly diferencovatelná v bodě $y \in U$, existuje-li zobrazení $G^\circ : U \rightarrow \mathcal{L}(Y, Z)$ takové, že $\{G^\circ(y+h)\}$ jsou stejnoměrně omezené pro dostatečně malá $h \in Y$ a platí

$$\lim_{h \rightarrow 0} \frac{\|G(y+h) - G(y) - G^\circ(y+h)h\|_Z}{\|h\|_Y} = 0.$$

Funkce G° se nazývá slanting funkce pro G v bodě y .

(ii) Funkce G se nazývá slantly diferencovatelná na U , existuje-li zobrazení $G^\circ : U \rightarrow \mathcal{L}(Y, Z)$ takové, že G° je slanting funkcí pro G v každém bodě $y \in U$. Funkce G° se nazývá slanting funkce pro G na U .

Věta 5 ([35, 34]) Nechť G je slantly diferencovatelná na U se slanting funkcí G° . Nechť $y^* \in U$ je řešení rovnice $G(y) = 0$. Jestliže $G^\circ(y)$ je regulární pro všechna $y \in U$ a $\{\|G^\circ(y)^{-1}\|_{\mathcal{L}(Y, Z)} : y \in U\}$ je omezená množina, pak newtonovské iterace

$$y^{(k+1)} = y^{(k)} - G^\circ(y^{(k)})^{-1}G(y^{(k)}) \quad (48)$$

konvergují superlineárně k y^* , pokud $y^{(0)}$ je dostatečně blízko y^* .

Následující dva příklady použijeme při řešení rovnice (47). Ortogonální projekce $\phi : \mathbb{R} \rightarrow \mathbb{R}_+$ je slantly diferencovatelná na \mathbb{R} a za její slanting funkci vezmeme:

$$\phi^\circ(y) = \begin{cases} 1 & \text{pro } y \geq 0, \\ 0 & \text{pro } y < 0. \end{cases} \quad (49)$$

Ortogonální projekce $\psi_g : \mathbb{R}^2 \rightarrow \mathcal{C}(g)$, kde $\mathcal{C}(g) = \{\mathbf{y} \in \mathbb{R}^2 : \|\mathbf{y}\| \leq g\}$ je kruh o poloměru $g \geq 0$, je slantly diferencovatelná a za její slanting funkci vezmeme:

$$\psi^\circ(\mathbf{y}) = \begin{cases} \mathbf{I}_{2 \times 2} & \text{pro } \|\mathbf{y}\| \leq g, \\ \frac{g}{\|\mathbf{y}\|} \left(\mathbf{I}_{2 \times 2} - \frac{1}{\|\mathbf{y}\|^2} \mathbf{y}\mathbf{y}^\top \right) & \text{pro } \|\mathbf{y}\| > g. \end{cases} \quad (50)$$

Lze ukázat, že matice $\mathbf{I}_{2 \times 2} - 1/(\|\mathbf{y}\|^2) \mathbf{y}\mathbf{y}^\top$ je ortogonální projekce na $(\text{Im } \mathbf{y})^\perp$, takže platí

$$\psi^\circ(\mathbf{y})\hat{\mathbf{y}} = \mathbf{0} \quad (51)$$

pro každé $\hat{\mathbf{y}} \in \text{Im } \mathbf{y}$ a $\|\mathbf{y}\| > g$.

Nyní vytvoříme slanting funkci pro \mathbf{G} z rovnice (47). Funkci \mathbf{G} rozložíme podle jejího blokového členění na tři části:

$$\mathbf{G}(\mathbf{y}) = (\mathbf{G}_1(\mathbf{y})^\top, \mathbf{G}_2(\mathbf{y})^\top, \mathbf{G}_3(\mathbf{y})^\top)^\top.$$

Komponenta \mathbf{G}_1 je diferencovatelná a její klasická derivace je slanting funkcí:

$$\mathbf{G}_1^\circ(\mathbf{y}) = (\mathbf{K}, \mathbf{N}^\top, \mathbf{T}^\top).$$

Vytvoření slanting funkce pro \mathbf{G}_2 se opírá o (49). Použijeme přitom terminologii aktivních množin. Aktivní a neaktivní množinu $\mathcal{A}_\nu, \mathcal{I}_\nu \subseteq \mathcal{M}$ definujeme takto:

$$\mathcal{A}_\nu = \{i \in \mathcal{M} : 0 \leq \lambda_{\nu, i} + \rho(\mathbf{N}\mathbf{u} - \mathbf{d})_i\}, \quad \mathcal{I}_\nu = \mathcal{M} \setminus \mathcal{A}_\nu. \quad (52)$$

Budeme také používat indikátorovou matici, která je pro $\mathcal{S} \subseteq \mathcal{M}$ definována následovně: $\mathbf{D}(\mathcal{S}) = \text{diag}(s_1, s_2, \dots, s_m)$, kde $s_i = 1$ pro $i \in \mathcal{S}$ a $s_i = 0$ pro $i \notin \mathcal{S}$. Potom můžeme psát $\mathbf{G}_2(\mathbf{y}) = \boldsymbol{\lambda}_\nu - \mathbf{D}(\mathcal{A}_\nu)(\boldsymbol{\lambda}_\nu + \rho(\mathbf{N}\mathbf{u} - \mathbf{d}))$. Slanting funkci získáme obvyklým derivováním:

$$\bar{\mathbf{G}}_2^o(\mathbf{y}) = (-\rho\mathbf{D}(\mathcal{A}_\nu)\mathbf{N}, \mathbf{D}(\mathcal{I}_\nu), \mathbf{0}).$$

Podobně vytvoříme na základě (50) slanting funkci pro \mathbf{G}_3 . Aktivní a neaktivní množinu $\mathcal{A}_t, \mathcal{I}_t \subseteq \mathcal{M}$ v bodě $\mathbf{y} \in \mathbb{R}^{n_u+3m}$ definujeme takto:

$$\mathcal{A}_t = \{i \in \mathcal{M} : \|\boldsymbol{\lambda}_{t,i} + \rho\mathbf{u}_{t,i}\| \leq g_i\}, \quad \mathcal{I}_t = \mathcal{M} \setminus \mathcal{A}_t. \quad (53)$$

Potom $\mathbf{G}_3(\mathbf{y}) = -\rho\mathbf{D}(\bar{\mathcal{A}}_t)\mathbf{T}\mathbf{u} + \mathbf{D}(\bar{\mathcal{I}}_t)(\boldsymbol{\lambda}_t - \mathbf{P}_{\Lambda_t}(\boldsymbol{\lambda}_t + \rho\mathbf{T}\mathbf{u}))$, kde $\bar{\mathcal{A}}_t = \{2i-1, 2i : i \in \mathcal{A}_t\}$ a $\bar{\mathcal{I}}_t = \{2i-1, 2i : i \in \mathcal{I}_t\}$, $\bar{\mathcal{A}}_t, \bar{\mathcal{I}}_t \subseteq \mathcal{M}_2$ pro $\mathcal{M}_2 = \{1, 2, \dots, 2m\}$. Slanting funkci pro \mathbf{G}_3 zapíšeme po složkách:

$$\mathbf{G}_3^o(\mathbf{y}) = (\mathbf{G}_{31}^o(\mathbf{y}), \mathbf{G}_{32}^o(\mathbf{y}), \mathbf{G}_{33}^o(\mathbf{y})),$$

kde jednotlivé složky odvodíme obvyklým derivováním podle \mathbf{u} , $\boldsymbol{\lambda}_\nu$ a $\boldsymbol{\lambda}_t$. Dostaneme

$$\begin{aligned} \mathbf{G}_{31}^o(\mathbf{y}) &= -\rho\mathbf{D}(\bar{\mathcal{A}}_t)\mathbf{T} - \rho\mathbf{D}(\bar{\mathcal{I}}_t)\mathbf{P}_{\Lambda_t}^o(\boldsymbol{\lambda}_t + \rho\mathbf{T}\mathbf{u})\mathbf{T}, \\ \mathbf{G}_{32}^o(\mathbf{y}) &= \mathbf{0}, \\ \mathbf{G}_{33}^o(\mathbf{y}) &= \mathbf{D}(\bar{\mathcal{I}}_t)(\mathbf{I} - \mathbf{P}_{\Lambda_t}^o(\boldsymbol{\lambda}_t + \rho\mathbf{T}\mathbf{u})), \end{aligned}$$

kde $\mathbf{P}_{\Lambda_t}^o(\boldsymbol{\lambda}_t + \rho\mathbf{T}\mathbf{u}) = \text{diag}(\psi_{g_1}^o(\mathbf{y}_1), \dots, \psi_{g_m}^o(\mathbf{y}_m)) \in \mathbb{R}^{2m \times 2m}$ a $\mathbf{y}_i = \boldsymbol{\lambda}_{t,i} + \rho\mathbf{u}_{t,i}$, $i \in \mathcal{M}$. Shrnutím všech výsledků dostaneme slanting funkci pro \mathbf{G} na \mathbb{R}^{n_u+3m} :

$$\mathbf{G}^o(\mathbf{y}) = \begin{pmatrix} \mathbf{K} & \mathbf{N}^\top & \mathbf{T}^\top \\ -\rho\mathbf{D}(\mathcal{A}_\nu)\mathbf{N} & \mathbf{D}(\mathcal{I}_\nu) & \mathbf{0} \\ \mathbf{G}_{31}^o(\mathbf{y}) & \mathbf{0} & \mathbf{G}_{33}^o(\mathbf{y}) \end{pmatrix}.$$

V každé iteraci (48) musíme řešit soustavu lineárních rovnic s maticí $\mathbf{G}^o(\mathbf{y}^{(k)})$ a pravou stranou $\mathbf{G}^o(\mathbf{y}^{(k)})\mathbf{y}^{(k)} - \mathbf{G}(\mathbf{y}^{(k)})$ pro neznámou $\mathbf{y}^{(k+1)}$. Vektor pravé strany lze s pomocí (51) zjednodušit na tvar, který používáme v (54).

ALGORITMUS SSN (*Semi-smooth Newton method*) Nechť $\mathbf{y}^{(0)} = ((\mathbf{u}^{(0)})^\top, (\boldsymbol{\lambda}_\nu^{(0)})^\top, (\boldsymbol{\lambda}_t^{(0)})^\top)^\top \in \mathbb{R}^{n_u+3m}$, $\rho > 0$, $\varepsilon \geq 0$, $err^{(0)} = 1$ a $k := 0$.

(1°) Je-li $err^{(k)} \leq \varepsilon$, vezmi za výsledek výpočtu $\bar{\mathbf{y}} = \mathbf{y}^{(k)}$, jinak pokračuj krokem (2°).

(2°) Sestav aktivní a neaktivní množiny podle (52), (53) v $\mathbf{y} = ((\mathbf{u}^{(k)})^\top, (\boldsymbol{\lambda}_\nu^{(k)})^\top, (\boldsymbol{\lambda}_t^{(k)})^\top)^\top$.

(3°) Vyřeš soustavu lineárních rovnic:

$$\mathbf{G}^o(\mathbf{y}^{(k)})\mathbf{y}^{(k+1)} = \begin{pmatrix} \mathbf{f} \\ -\rho\mathbf{D}(\mathcal{A}_\nu)\mathbf{d} \\ \mathbf{D}(\bar{\mathcal{I}}_t)\mathbf{P}_{\Lambda_t}(\boldsymbol{\lambda}_t^{(k)} + \rho\mathbf{T}\mathbf{u}^{(k)}) \end{pmatrix}. \quad (54)$$

(4°) Polož $err^{(k+1)} = \|\mathbf{y}^{(k+1)} - \mathbf{y}^{(k)}\| / \|\mathbf{y}^{(k+1)}\|$, $k := k + 1$ a pokračuj krokem (1°).

Výsledek Věty 5 o superlineární rychlosti konvergence platí za předpokladu, že vnitřní soustavy lineárních rovnic (54) jsou řešeny přesně. U rozsáhlých úloh je však nerealistické přesné řešení očekávat. Další potíž spočívá v nesymetrii matic $\mathbf{G}^o(\mathbf{y}^{(k)})$.

3.3.3 Duální implementace

Třetí blokovou rovnici v (54) lze rozštěpit vzhledem k aktivní a neaktivní množině:

$$-\rho \mathbf{T}_{\bar{\mathcal{A}}_t} \mathbf{u}^{(k+1)} = \mathbf{0}, \quad \boldsymbol{\lambda}_{t,i}^{(k+1)} - \boldsymbol{\psi}_{g_i}^o(\mathbf{y}_i^{(k)}) \mathbf{y}_i^{(k+1)} = \boldsymbol{\psi}_{g_i}(\mathbf{y}_i^{(k)}), \quad i \in \mathcal{I}_t.$$

Jestliže výpočet konverguje, pak členy $\boldsymbol{\psi}_{g_i}^o(\mathbf{y}_i^{(k)}) \mathbf{y}_i^{(k+1)}$ konvergují podle (51) k nule. Tyto členy proto vypustíme, což je ekvivalentní nahrazení $\mathbf{G}^o(\mathbf{y})$ maticí:

$$\tilde{\mathbf{G}}^o(\mathbf{y}) = \begin{pmatrix} \mathbf{K} & \mathbf{N}^\top & \mathbf{T}^\top \\ -\rho \mathbf{D}(\mathcal{A}_\nu) \mathbf{N} & \mathbf{D}(\mathcal{I}_\nu) & \mathbf{0} \\ -\rho \mathbf{D}(\bar{\mathcal{A}}_t) \mathbf{T} & \mathbf{0} & \mathbf{D}(\bar{\mathcal{I}}_t) \end{pmatrix}. \quad (55)$$

Následující lemma ukazuje, že získáme stejné řešení.

Lemma 6 [47] *Nechť \mathbf{G} respektive $\tilde{\mathbf{G}}^o$ je definováno rovností (46) respektive (55) a necht' \mathbf{y}^* je řešením $\mathbf{G}(\mathbf{y}) = \mathbf{0}$. Jestliže modifikované newtonovské iterace*

$$\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} - \tilde{\mathbf{G}}^o(\mathbf{y}^{(k)})^{-1} \mathbf{G}(\mathbf{y}^{(k)}) \quad (56)$$

konvergují, pak jako limitní bod dostaneme \mathbf{y}^ .*

Pro nepřesnou implementaci (56) lze dokázat globální R-lineární rychlost konvergence. Nejdříve ale odvodíme duální implementaci (56) vyloučením posloupnosti $\{\mathbf{u}^k\}$.

Každou iteraci (56) lze rozložit na dvě části vztahující se k neaktivním množinám:

$$\boldsymbol{\lambda}_{\nu, \mathcal{I}_\nu}^{(k+1)} = \mathbf{0}, \quad \boldsymbol{\lambda}_{t,i}^{(k+1)} = \boldsymbol{\psi}_{g_i}(\mathbf{y}_i^{(k)}), \quad i \in \mathcal{I}_t, \quad (57)$$

a k aktivním množinám:

$$\begin{pmatrix} \mathbf{K} & \mathbf{N}_{\mathcal{A}_\nu}^\top & \mathbf{T}_{\bar{\mathcal{A}}_t}^\top \\ \mathbf{N}_{\mathcal{A}_\nu} & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_{\bar{\mathcal{A}}_t} & \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u}^{(k+1)} \\ \boldsymbol{\lambda}_{\nu, \mathcal{A}_\nu}^{(k+1)} \\ \boldsymbol{\lambda}_{t, \bar{\mathcal{A}}_t}^{(k+1)} \end{pmatrix} = \begin{pmatrix} \mathbf{f} - \mathbf{T}_{\bar{\mathcal{I}}_t}^\top \boldsymbol{\lambda}_{t, \bar{\mathcal{I}}_t}^{(k+1)} \\ \mathbf{d}_{\mathcal{A}_\nu} \\ \mathbf{0} \end{pmatrix}. \quad (58)$$

Pro zjednodušení zápisu označíme:

$$\mathcal{A} = \mathcal{A}_\nu \cup \{i + m : i \in \bar{\mathcal{A}}_t\} \quad \mathcal{I} = \{1, 2, \dots, 3m\} \setminus \mathcal{A}. \quad (59)$$

Dále označíme $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_\nu^\top, \boldsymbol{\lambda}_t^\top)^\top$, $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_\nu \times \boldsymbol{\Lambda}_t$ a použijeme matice \mathbf{A} , \mathbf{B} a vektory \mathbf{b} , \mathbf{c} zavedené v (20). Je-li $\mathcal{A} = \emptyset$, potom je vektor $\boldsymbol{\lambda}^{(k+1)}$ plně určen vztahy (57). Necht' \mathcal{A} je neprázdná množina. Druhá a třetí bloková rovnice v (58) dává $\mathbf{B}_{\mathcal{A}} \mathbf{u}^{(k+1)} = \mathbf{c}_{\mathcal{A}}$ a z první rovnice dostaneme $\mathbf{u}^{(k+1)} = \mathbf{K}^{-1}(\mathbf{f} - \mathbf{B}_{\mathcal{A}}^\top \boldsymbol{\lambda}_{\mathcal{A}}^{(k+1)} - \mathbf{B}_{\mathcal{I}}^\top \boldsymbol{\lambda}_{\mathcal{I}}^{(k+1)})$, kde $\boldsymbol{\lambda}_{\mathcal{I}}^{(k+1)}$ je určeno vztahy (57). Dosazením snadno zjistíme, že $\boldsymbol{\lambda}_{\mathcal{A}}^{(k+1)}$ řeší následující lineární soustavu

$$\mathbf{A}_{\mathcal{A}\mathcal{A}} \boldsymbol{\lambda}_{\mathcal{A}}^{(k+1)} = \hat{\mathbf{b}}_{\mathcal{A}}$$

pro $\hat{\mathbf{b}}_{\mathcal{A}} = \mathbf{b}_{\mathcal{A}} - \mathbf{A}_{\mathcal{A}\mathcal{I}} \boldsymbol{\lambda}_{\mathcal{I}}^{(k+1)}$. Matice $\mathbf{A}_{\mathcal{A}\mathcal{A}}$ je symetrická pozitivně definitní principální podmatice matice \mathbf{A} , takže $\boldsymbol{\lambda}^{(k+1)}$ minimalizuje kvadratickou funkci q z (20) vzhledem k omezujícím podmínkám (57). Nakonec si všimněme, že gradient \mathbf{r} funkce q splňuje

$$\mathbf{r}^{(k)} = \mathbf{r}(\boldsymbol{\lambda}^{(k)}) = \mathbf{B} \mathbf{K}^{-1} \mathbf{B}^\top \boldsymbol{\lambda}^{(k)} - \mathbf{B} \mathbf{K}^{-1} \mathbf{f} + \mathbf{c} = -\mathbf{B} \mathbf{u}^{(k)} + \mathbf{c} = \begin{pmatrix} -\mathbf{N} \mathbf{u}^{(k)} + \mathbf{d} \\ -\mathbf{T} \mathbf{u}^{(k)} \end{pmatrix}.$$

V definici aktivních a neaktivních množin lze proto nahradit $\mathbf{u} = \mathbf{u}^k$ gradientem:

$$(\mathbf{N}\mathbf{u}^{(k)} - \mathbf{d})_i = -r_i^{(k)}, \quad \mathbf{u}_{t,i}^{(k)} = -\mathbf{r}_i^{(k)} = -(r_{2i-1}^{(k)}, r_{2i}^{(k)})^\top, \quad i \in \mathcal{M}.$$

Algoritmus (56) můžeme zapsat výhradně pomocí duálních veličin. Jako ukončovací kritérium použijeme redukovaný gradient (30) pro $\tilde{\alpha} = \rho$.

ALGORITMUS MSSN (*Modified SSN method*) Nechť $\boldsymbol{\lambda}^{(0)} \in \mathbb{R}^{3m}$, $\rho > 0$, $\varepsilon \geq 0$, $err^{(0)} = \|\tilde{\mathbf{r}}_\rho(\mathbf{P}_\Lambda(\boldsymbol{\lambda}^{(0)}))\|$ a $k := 0$.

(1°) Je-li $err^{(k)} \leq \varepsilon$, vezmi za výsledek výpočtu $\bar{\boldsymbol{\lambda}} = \mathbf{P}_\Lambda(\boldsymbol{\lambda}^{(k)})$, jinak pokračuj krokem (2°).

(2°) Sestav aktivní a neaktivní množiny v bodě $\boldsymbol{\lambda}^{(k)}$:

$$\mathcal{A}_\nu = \{i \in \mathcal{M} : \lambda_i^{(k)} - \rho r_i^{(k)} \geq 0\}, \quad \mathcal{I}_\nu = \mathcal{M} \setminus \mathcal{A}_\nu, \quad (60)$$

$$\mathcal{A}_t = \{i \in \mathcal{M} : \|\boldsymbol{\lambda}_{t,i}^{(k)} - \rho \mathbf{r}_i^{(k)}\| \leq g_i\}, \quad \mathcal{I}_t = \mathcal{M} \setminus \mathcal{A}_t. \quad (61)$$

(3°) Najdi $\boldsymbol{\lambda}^{(k+1)}$ takové, že

$$\boldsymbol{\lambda}^{(k+1)} = \arg \min q(\boldsymbol{\lambda}) \quad \text{za podmínky (57).}$$

(4°) Polož $err^{(k+1)} = \|\tilde{\mathbf{r}}_\rho(\mathbf{P}_\Lambda(\boldsymbol{\lambda}^{(k+1)}))\|$, $k := k + 1$ a pokračuj krokem (1°).

Následující konvergenční výsledek plyne z Lemmatu 6.

Věta 6 [47] *Nechť $\boldsymbol{\lambda}^{(0)} \in \mathbb{R}^{3m}$, $\rho > 0$, $\varepsilon = 0$ a $\boldsymbol{\lambda}^* = (\boldsymbol{\lambda}_\nu^{*\top}, \boldsymbol{\lambda}_t^{*\top})^\top$ označuje druhou a třetí složku řešení rovnice (47). Jestliže posloupnost $\{\boldsymbol{\lambda}^{(k)}\}$ počítána algoritmem MSSN konverguje, pak jejím limitním bodem je $\boldsymbol{\lambda}^*$.*

Ve zbytku odstavce navrhne výpočetně efektivní implementaci algoritmu MSSN, která je založena na provedení několik iterací metody CG pro přibližný výpočet minima v kroku (3°). Přibližné minimum označíme opět $\boldsymbol{\lambda}^{(k+1)}$. Na metodu CG se budeme odkazovat zápisem

$$\boldsymbol{\lambda}^{(k+1)} = \text{CG}(\mathbf{A}, \mathbf{b}, \mathcal{A}, \boldsymbol{\lambda}^{(k+1,0)}, tol^{(k+1)}), \quad (62)$$

kde $\boldsymbol{\lambda}^{(k+1,0)}$ je počáteční iterace a $tol^{(k+1)}$ je ukončovací přesnost. Volba $tol^{(k+1)}$ je založena na technice popsané v Sekci 3.2.4 bod (ii).

ALGORITMUS IMSSN (*Inexact MSSN method*) Nechť $\boldsymbol{\lambda}^{(0)} \in \mathbb{R}^{3m}$, $\rho > 0$, $\varepsilon \geq 0$, $r_{tol}, c_{fact} \in (0, 1)$, $err^{(0)} = \|\tilde{\mathbf{r}}_\rho(\mathbf{P}_\Lambda(\boldsymbol{\lambda}^{(0)}))\|$, $tol^{(0)} = r_{tol}/c_{fact}$ a $k := 0$.

(1°) Je-li $err^{(k)} \leq \varepsilon$, vezmi za výsledek výpočtu $\bar{\boldsymbol{\lambda}} = \mathbf{P}_\Lambda(\boldsymbol{\lambda}^{(k)})$, jinak pokračuj krokem (2°).

(2°) Sestav aktivní a neaktivní množiny v bodě $\boldsymbol{\lambda}^{(k)}$ podle (60)-(61) a (59).

(3.1°) $tol^{(k+1)} = \min\{r_{tol} \times err^{(k)}/err^{(0)}, c_{fact} \times tol^{(k)}\}$.

(3.2°) $\boldsymbol{\lambda}_{\mathcal{A}}^{(k+1,0)} = \boldsymbol{\lambda}_{\mathcal{A}}^{(k)}$, $\boldsymbol{\lambda}_{\nu, \mathcal{I}_\nu}^{(k+1,0)} = \mathbf{0}$, $\boldsymbol{\lambda}_{t,i}^{(k+1,0)} = \boldsymbol{\psi}_{g_i}(\boldsymbol{\lambda}_{t,i}^{(k)} - \rho \mathbf{r}_i^{(k)})$, $i \in \mathcal{I}_t$.

(3.3°) $\boldsymbol{\lambda}^{(k+1)} = \text{CG}(\mathbf{A}, \mathbf{b}, \mathcal{A}, \boldsymbol{\lambda}^{(k+1,0)}, tol^{(k+1)})$.

(4°) Polož $err^{(k+1)} = \|\tilde{\mathbf{r}}_\rho(\mathbf{P}_\Lambda(\boldsymbol{\lambda}^{(k+1)}))\|$, $k := k + 1$ a pokračuj krokem (1°).

Algoritmus lze interpretovat jako restartovanou metodu CG pro výpočet minima úlohy (29) při speciální volbě přípustné množiny Λ .

3.3.4 Globalizace

Modifikací algoritmu IMSSN vytvoříme globálně konvergující algoritmus tak, že zajistíme, aby posloupnost $\{q(\boldsymbol{\lambda}^{(k)})\}$ byla monotónně klesající. Nejdříve uvedeme nutnou a postačující podmínku, která zaručí, že restart metody CG nezvýší hodnotu funkce q .

Lemma 7 [46, 47] *Nechť $\boldsymbol{\lambda}^{(k)} \in \boldsymbol{\Lambda}$ a $\rho \in (0, 2\sigma_{\max}^{-1})$, kde σ_{\max} je největší vlastní číslo matice \mathbf{A} . Pro vektor $\boldsymbol{\lambda}^{(k+1,0)}$ z kroku (3.2°) algoritmu IMSSN platí: $q(\boldsymbol{\lambda}^{(k)}) \geq q(\boldsymbol{\lambda}^{(k+1,0)})$.*

Vzhledem k vlastnostem metody CG z Lemmatu 7 vyplývá, že posloupnost $\{q(\boldsymbol{\lambda}^{(k)})\}$ bude nerostoucí. Další lemma interpretuje krok (3.2°) jako část projekce na množinu $\boldsymbol{\Lambda}$.

Lemma 8 [46, 47] *Nechť vektor $\boldsymbol{\lambda}^{(k+1,0)}$ je určen krokem (3.2°) algoritmu IMSSN. Pak platí:*

$$\boldsymbol{\lambda}_{\mathcal{I}}^{(k+1,0)} = \mathbf{P}_{\boldsymbol{\Lambda}, \mathcal{I}}(\boldsymbol{\lambda}^{(k)} - \rho \mathbf{r}(\boldsymbol{\lambda}^{(k)})). \quad (63)$$

Algoritmus IMSSN pozměníme podle tří bodů. (i) Pro parametr ρ použijeme horní odhad z Lemmatu 7. (ii) Zajistíme, aby všechny iterace ležely v množině $\boldsymbol{\Lambda}$: zvolíme $\boldsymbol{\lambda}^{(0)} \in \boldsymbol{\Lambda}$, metodu CG ukončíme při pokusu vygenerovat iteraci vně $\boldsymbol{\Lambda}$ a poslední iteraci doplníme „půlkrokem“, který skončí na hranici některé separabilní části $\boldsymbol{\Lambda}$. Místo CG pak budeme psát CG_{feas} . (iii) Projekci (63) v kroku (3.2°) rozšíříme i na komponenty aktivní množiny.

ALGORITMUS GIMSSN (*Globally convergent IMSSN method*) Nechť $\boldsymbol{\lambda}^{(0)} \in \boldsymbol{\Lambda}$, $\rho \in (0, 2\sigma_{\max}^{-1})$, $\varepsilon \geq 0$, $r_{tol}, c_{fact} \in (0, 1)$, $err^{(0)} = \|\tilde{\mathbf{r}}_{\rho}(\boldsymbol{\lambda}^{(0)})\|$, $tol^{(0)} = r_{tol}/c_{fact}$ a $k := 0$, kde σ_{\max} je největší vlastní číslo matice \mathbf{A} .

(1°) Je-li $err^{(k)} \leq \varepsilon$, vezmi za výsledek výpočtu $\bar{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^{(k)}$, jinak pokračuj krokem (2°).

(2°) Sestav aktivní a neaktivní množiny v bodě $\boldsymbol{\lambda}^{(k)}$ podle (60)-(61) a (59).

(3.1°) $tol^{(k+1)} = \min\{r_{tol} \times err^{(k)}/err^{(0)}, c_{fact} \times tol^{(k)}\}$.

(3.2°) $\boldsymbol{\lambda}^{(k+1,0)} = \mathbf{P}_{\boldsymbol{\Lambda}}(\boldsymbol{\lambda}^{(k)} - \rho \mathbf{r}(\boldsymbol{\lambda}^{(k)}))$.

(3.3°) $\boldsymbol{\lambda}^{(k+1)} = \text{CG}_{feas}(\mathbf{A}, \mathbf{b}, \mathcal{A}, \boldsymbol{\lambda}^{(k+1,0)}, tol^{(k+1)})$.

(4°) Polož $err^{(k+1)} = \|\tilde{\mathbf{r}}_{\rho}(\boldsymbol{\lambda}^{(k+1)})\|$, $k := k + 1$ a pokračuj krokem (1°).

Formálně je tento algoritmus shodný s algoritmem GISSNM z [46] pro kontaktní úlohy ve 2D, rozdíl je ve způsobu výpočtu projekce $\mathbf{P}_{\boldsymbol{\Lambda}}$. Konvergenční výsledek, který je obsahem následující věty, je stejný pro úlohy ve 2D i ve 3D.

Věta 7 [46, 47] *Nechť $\boldsymbol{\lambda}^{(0)} \in \boldsymbol{\Lambda}$, $\varepsilon = 0$, $\rho \in (0, 2\sigma_{\max}^{-1})$, $r_{tol}, c_{fact} \in (0, 1)$, σ_{\min} je nejmenší a σ_{\max} je největší vlastní číslo matice \mathbf{A} a $\boldsymbol{\lambda}^* = (\boldsymbol{\lambda}_v^{*\top}, \boldsymbol{\lambda}_t^{*\top})^{\top}$ označuje druhou a třetí složku řešení rovnice (47). Nechť $\{\boldsymbol{\lambda}^{(k)}\}$ je posloupnost počítaná algoritmem GIMSSN. Pak platí:*

(i) *posloupnost $\{q(\boldsymbol{\lambda}^{(k)})\}$ je klesající tak, že*

$$q(\boldsymbol{\lambda}^{(k+1)}) - q(\boldsymbol{\lambda}^*) \leq \eta(\rho) (q(\boldsymbol{\lambda}^{(k)}) - q(\boldsymbol{\lambda}^*)),$$

kde

$$\eta(\rho) = \begin{cases} 1 - \rho\sigma_{\min} & \text{pro } \rho \in (0, \sigma_{\max}^{-1}], \\ 1 - (2\sigma_{\max}^{-1} - \rho)\sigma_{\min} & \text{pro } \rho \in [\sigma_{\max}^{-1}, 2\sigma_{\max}^{-1}); \end{cases}$$

(ii) *je-li $\{\boldsymbol{\lambda}^{(k)}\}$ konečná posloupnost, pak jejím posledním prvkem je $\boldsymbol{\lambda}^*$;*

(iii) *je-li $\{\boldsymbol{\lambda}^{(k)}\}$ nekonečná posloupnost, pak konverguje R-lineárně k $\boldsymbol{\lambda}^*$ tak, že*

$$\|\boldsymbol{\lambda}^{(k)} - \boldsymbol{\lambda}^*\| \leq C \eta(\rho)^{k/2},$$

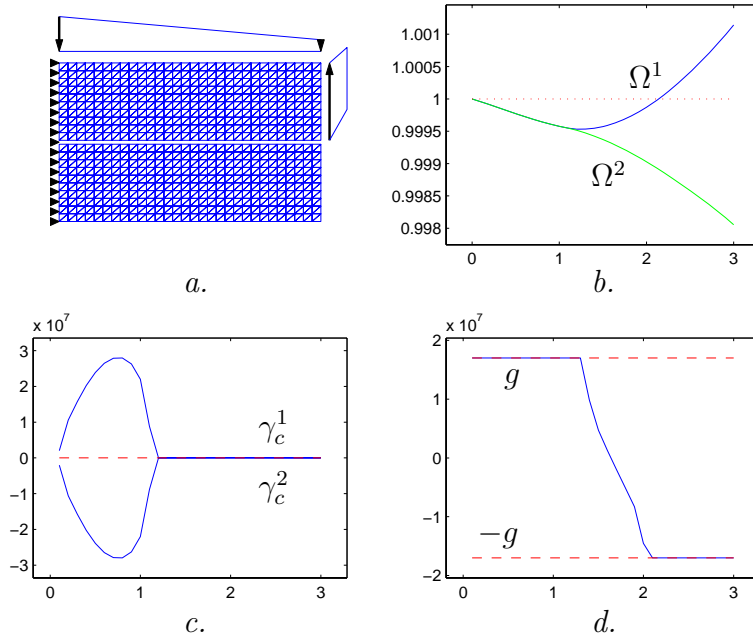
kde $C = \sqrt{2(q(\boldsymbol{\lambda}^0) - q(\boldsymbol{\lambda}^*))}/\sigma_{\min}$.

Porovnání s Větou 1 ukazuje, že algoritmy GIMSSM a KPRGP jsou si velmi podobné. Algoritmus GIMSSM dovoluje mírné pronikání hranice přípustné množiny Λ .

3.3.5 Numerické experimenty

Použití semi-hladké Newtonovy metody budeme testovat pro kontaktní úlohu ve 2D a ve 3D.

Příklad 5 Uvažujme 2D kontaktní úlohu s Trescovým třením pro dvě pružná tělesa $\Omega^1 = (0, 3) \times (1, 2)$, $\Omega^2 = (0, 3) \times (0, 1)$ s Youngovým modulem $E^k = 21.19 \cdot 10^{10}$ a Poissonovou konstantou $\nu^k = 0.277$ (ocel), kde $c_1^k = E^k \nu^k / (1 - (\nu^k)^2)$, $c_2^k = E^k / (2(1 + \nu^k))$, $k = 1, 2$. Dekompozice hranic $\partial\Omega^1$ a $\partial\Omega^2$ je následující: $\gamma_u^1 = \{0\} \times (1, 2)$, $\gamma_c^1 = (0, 3) \times \{1\}$, $\gamma_p^1 = \partial\Omega^1 \setminus \overline{\gamma_u^1 \cup \gamma_c^1}$ a $\gamma_u^2 = \{0\} \times (0, 1)$, $\gamma_c^2 = (0, 3) \times \{1\}$, $\gamma_p^2 = \partial\Omega^2 \setminus \overline{\gamma_u^2 \cup \gamma_c^2}$. Objemové síly jsou nulové pro obě tělesa. Nenulová povrchová napětí $\mathbf{p}^1 = (p_x^1, p_y^1)$ na γ_p^1 jsou zadána takto: $p_y^1(s, 2) = p_{y,L}^1 + p_{y,R}^1 s$, $s \in (0, 3)$, $p_x^1(3, s) = p_{x,B}^1(2 - s) + p_{x,U}^1(s - 1)$, $s \in (1, 2)$, $p_y^1(3, s) = p_{y,B}^1(2 - s) + p_{y,U}^1(s - 1)$, $s \in (1, 2)$, kde $p_{y,L}^1 = -6 \cdot 10^7$, $p_{y,R}^1 = -1 \cdot 10^7$, $p_{x,B}^1 = p_{x,U}^1 = p_{y,U}^1 = 2 \cdot 10^7$ a $p_{y,B}^1 = 4 \cdot 10^7$. Mez skluzu má hodnotu $g = 1.7 \cdot 10^7$ na γ_c^1 .

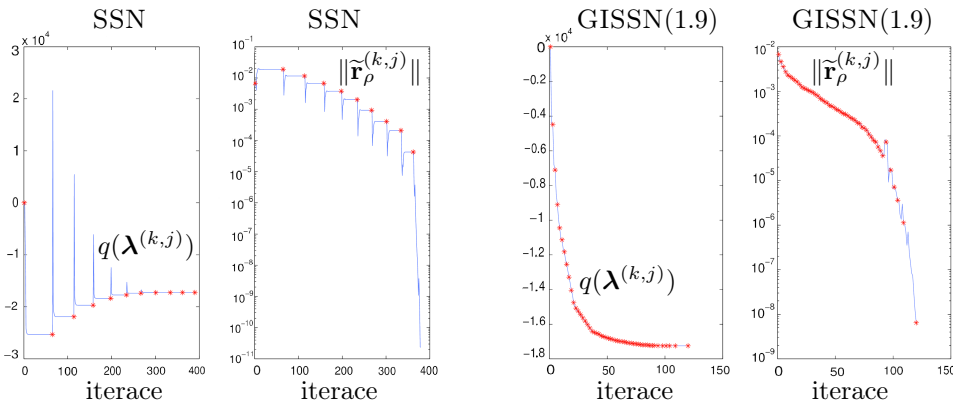


Obrázek 11: a.) Tělesa Ω^1 a Ω^2 , povrchové napětí a triangulace. b.) Kontaktní rozhraní po deformaci. c.) Normálové kontaktní napětí. d.) Tečné kontaktních napětí.

Protože řešíme úlohu ve 2D, testujeme variantu algoritmu KPRGP pro oboustranná jednoduchá omezení označenou AS, která je blízká algoritmu MPRGP z ([16]). V algoritmu označeném AS-S použijeme definici aktivních množin ze semi-hladké Newtonovy metody. Konvergenční výsledek Věty 1 platí pro AS i pro AS-S. Tabulka 5 uvádí počty násobení matice-vektor n_A , kterými bylo všemi algoritmy dosaženo výsledků v relativní přesnosti na úrovni 10^{-6} . Délku kroku v algoritmech AS a AS-S volíme $\tilde{\alpha} = \rho = \beta \cdot \sigma_{\max}^{-1}$, kde ρ je parametr z algoritmu GISSN. Věty 1 a 7 zaručují konvergenci pro $0 < \beta < 2$. Výkonnost algoritmů GISSN a AS je podobná. U algoritmu AS-S došlo ke stabilizaci, která se projevila nejlepší výkonností pro $\beta = 20$, kdy algoritmus AS osciloval. Obrázek 12 ukazuje průběh iterační historie. Hodnoty funkce q pro všechny iterace $\bigcup_k \{\lambda^{(k,j)}\}$ metody CG jsou znázorněny modře, zatímco hodnoty v newtonovských iteracích $\{\lambda^{(k)}\}$ jsou označeny červenou hvězdičkou. První graf SSN ilustruje restartování minimalizačního procesu. První graf GISSN zobrazuje monotónní pokles funkčních hodnot, což je plně v souladu s teorií.

Tabulka 5: Srovnání různých algoritmů ve 2D: $\rho = \beta \cdot \sigma_{\max}^{-1}$

| | GIMSSN | AS-S | AS | GIMSSN | AS-S | AS |
|------------|--------|-------|-------|--------|-------|-------|
| β | 1.9 | 1.9 | 1.9 | 15 | 20 | 15 |
| n_u/m | n_A | n_A | n_A | n_A | n_A | n_A |
| 43920/360 | 203 | 262 | 198 | 86 | 85 | 89 |
| 77760/480 | 240 | 323 | 227 | 101 | 95 | 89 |
| 121200/600 | 270 | 339 | 253 | 96 | 95 | 105 |
| 174240/720 | 279 | 361 | 265 | 113 | 106 | 106 |
| 204360/780 | 290 | 381 | 286 | 108 | 109 | 110 |



Obrázek 12: Iterační historie ve 2D.

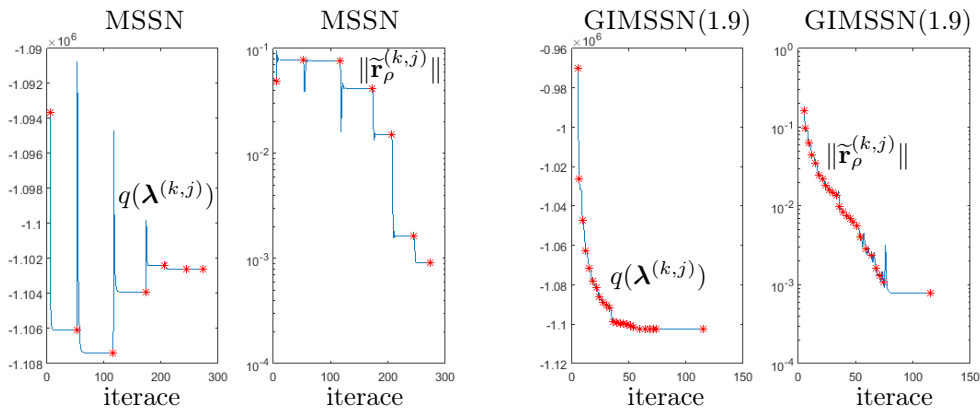
Příklad 6 Uvažujme dvě prostorová pružná tělesa (cihly z různých druhů oceli s rozměry v [mm]) $\Omega^1 = (0, 2000) \times (0, 1000) \times (1000, 2000)$ a $\Omega^2 = (0, 2000) \times (0, 1000) \times (0, 1000)$ s Youngovým modulem 50000 a Poissonovou konstantou 0.277 pro Ω^1 a 0.35 pro Ω^2 . Mez skluzu má hodnotu $g = 0.3$. Objemové síly jsou nulové. Nenulová povrchová napětí $\mathbf{p} = (p_1, p_2, p_3)$ působí na $L_8 = (0, 2000) \times (0, 1000) \times \{2000\}$ a $L_6 = \{2000\} \times (0, 1000) \times (0, 1000)$, kde je $p_1 = 0$, $p_2 = 0$, $p_3 = -1$. Tabulka 6 uvádí počty násobení matice-vektor n_A potřebné pro relativní přesnost výsledků na úrovni 10^{-6} a počty newtonovských iterací *iter*. Algoritmus SSN řeší vnitřní nesymetrické lineární soustavy (54) nepřesně metodou BICGSTAB se stejným ukončováním vnitřních iterací jako algoritmus IMSSN. Algoritmus IMSSN je desetkrát efektivnější. Opět volíme $\rho = \beta \cdot \sigma_{\max}^{-1}$. Obrázek 13 zobrazuje iterační historii stejným způsobem jako Obrázek 12 a obdobné jsou i závěry.

4 Závěr

Kontaktní úlohy se třením jsou intenzivně studovány v posledních několika desetiletích. Postupně se začala věnovat pozornost vývoji výkonných algoritmů pro numerický výpočet řešení. Ačkoliv různých algoritmů je známo velké množství, ne všechny lze s úspěchem vždy použít. V těchto tezích jsou představeny tři optimalizační algoritmy vyvinuté speciálně pro řešení kontaktních úloh. Algoritmy byly nebo jsou implementovány do několika počítačových knihoven v rámci Národního superpočítačového centra IT4I na VŠB – TUO. V tezích jsou algoritmy formulovány v podobě, která umožňuje dokázat jejich teoretické vlastnosti a vytvořit představu

Tabulka 6: Srovnání různých algoritmů ve 3D: $\rho = 1.9 \cdot \sigma_{\max}^{-1}$

| | SSN | IMSSN | GIMSSN |
|-------------|------------|------------|------------|
| β | 1.9 | 1.9 | 1.9 |
| n_u/m | $iter/n_A$ | $iter/n_A$ | $iter/n_A$ |
| 24576/256 | 9/1180 | 10/161 | 41/322 |
| 55566/441 | 9/1428 | 9/143 | 47/525 |
| 105456/676 | 11/2506 | 10/185 | 52/643 |
| 178746/961 | 10/2162 | 10/190 | 57/716 |
| 279936/1296 | 10/2298 | 9/167 | 65/835 |
| 413536/1681 | 10/2586 | 10/216 | 76/783 |



Obrázek 13: Iterační historie ve 3D.

o jejich vzájemných souvislostech. Pro řešení velmi rozsáhlých úloh z inženýrské praxe je potřeba tyto základní algoritmy kombinovat s dalšími technikami, jako jsou například některé varianty metod rozložení oblasti umožňující paralelní implementaci. Pomocí standardního komerčního softwaru lze kontaktní úlohy řešit jen v omezené míře, protože množství různých variant modelů tření, přilnavosti či skluzu neumožňuje postihnout vše jednotným obecným přístupem. Důvodem pro studium specializovaných algoritmů je také rozvoj výpočetních techniky, která sice nabízí stále sofistikovanější prostředky, ale jejich efektivní využití vyžaduje navrhovat nové výpočetní metody. Konečně hlavní motivací je uspokojování požadavků praktických inženýrských aplikací, které jako součást vývoje moderních technologií vedou k vytváření nových matematických modelů. Typickým příkladem je modelování hydrofobních povrchů, u nichž se teprve nedávno zjistilo, že jejich popis lze provést pomocí podobných postupů, jaké byly vyvíjeny pro klasické kontaktní úlohy.

Použitá literatura

- [1] AYADI M., GDOURA M., SASSI T.: *Mixed formulation for Stokes problem with Tresca friction*, C. R. Acad. Sci. Paris, Ser. I, 348(2010), 1069-1072.
- [2] BENZI M., GOLUB G., LIESSEN J.: *Numerical solution of saddle point systems*, Acta Numerica, **14** (2005), 1-137.

- [3] BEREMLIJSKI P., HASLINGER J., KOČVARA M., **KUČERA R.**, OUTRATA J.: *Shape optimization in three-dimensional contact problems with Coulomb friction*, SIAM J. Optim., 20(2009), 1, 416-444.
- [4] BREZZI F., FORTIN M.: *Mixed and hybrid finite element methods*, Series in Computational Mathematics 15, Springer-Verlag, New York, 1991.
- [5] BOUCHALA J., DOSTÁL Z., VODSTRČIL P.: *Separable spherical constraints and the decrease of a quadratic function in the gradient projection step*, J. Optim. Theory Appl., 157(2013), 1, 132-140.
- [6] BRZOBOHATÝ T., DOSTÁL Z., KOZUBEK T., KOVÁŘ P., MARKOPOULOS A.: *Cholesky decomposition with fixing nodes to stable computation of a generalized inverse of the stiffness matrix of a floating structure*, Internat. J. Numeric. Meth. Engrg., 88(2011), 5, 493-509.
- [7] DOSTÁL Z.: *Box constrained quadratic programming with proportioning and projections*, SIAM J. Optim., 7(1997), 3, 871-887.
- [8] DOSTÁL Z.: *Inexact semimonotonic augmented Lagrangians with optimal feasibility convergence for convex bound equality constrained quadratic programming*, SIAM J. Numer. Anal., 43(2005), 6, 96-115.
- [9] DOSTÁL Z.: *Optimal quadratic programming algorithms: with applications to variational inequalities*, SOIA 23, Springer US, New York, 2009.
- [10] DOSTÁL Z., HASLINGER J., **KUČERA R.**: *Implementation of fixed point method for duality based solution of contact problems with friction*, J. Comput. Appl. Math., 140(2002), 1-2, 245-256.
- [11] DOSTÁL Z., HORÁK D., **KUČERA R.**: *Total FETI - an easier implementable variant of the FETI method for numerical solution of elliptic PDE*, Comm. Numer. Meth. Engrg., 22(2006), 12, 1155-1162.
- [12] DOSTÁL Z., HORÁK D., **KUČERA R.**, VONDRÁK V., HASLINGER J., DOBIÁŠ J., PTÁK S.: *FETI based algorithms for contact problems: scalability, large displacements and 3D Coulomb friction*, Comput. Meth. Appl. Mech. Engrg., 194(2005), 2-5, 395-409.
- [13] DOSTÁL Z., **KUČERA R.**: *An optimal algorithm for minimization of quadratic functions with bounded spectrum subject to separable convex inequality and linear equality constraints*, SIAM J. Optim., 20(2010), 6, 2913-2938.
- [14] DOSTÁL Z., KOZUBEK T.: *An optimal algorithm and superrelaxation for minimization of a quadratic function subject to separable convex constraints with applications*, Math. Prog., 135(2011), 1-2, 195-220.
- [15] DOSTÁL Z., KOZUBEK T., MARKOPOULOS A., MENŠÍK M.: *Cholesky decomposition of a positive semidefinite matrix with known kernel*, Appl. Math. Comput., 217(2011), 13, 6067-6077.
- [16] DOSTÁL Z., SCHÖBERL J.: *Minimizing quadratic functions over non-negative cone with the rate of convergence and finite termination*, Comput. Optim. Appl., 30(2005), 1, 23-44.

- [17] ECK C., JARUŠEK J., KRBEČ M.: *Unilateral contact problems: Variational methods and existence theorems*, Taylor and Francis, 2005.
- [18] ELMAN H., SILVESTER D., WATHEN A.: *Finite elements and fast iterative solvers with applications in incompressible fluid dynamics*, Oxford University Press, Oxford, 2005.
- [19] FARHAT C., MANDEL J., ROUX F.: *Optimal convergence properties of the FETI domain decomposition method*, Comput. Meth. Appl. Mech. Engrg., 115(1994), 3-4, 367-388.
- [20] FUJITA H.: *A mathematical analysis of motions of viscous incompressible fluid under leak or slip boundary conditions*, Mathematical Fluid Mechanics and Modeling, RIMS Kokyuroko, 888(1994), Kyoto University, Kyoto, 199-216.
- [21] GOLUB G., VAN LOAN C.: *Matrix computations*. MD: The Johns Hopkins University Press, Baltimore 1996.
- [22] HASLINGER J.: *Approximation of the Signorini problem with friction, obeying Coulomb law*, Math. Meth. Appl., 5(1983), 422-437.
- [23] HASLINGER J., DOSTÁL Z., KUČERA R.: *On a splitting type algorithm for the numerical realization of contact problems with Coulomb friction*, Comput. Meth. Appl. Mech. Engrg., 191(2002), 2261-2281.
- [24] HASLINGER J., HLAVÁČEK I., NEČAS J.: *Numerical methods for unilateral problems in solid mechanics*, in Handbook of Numerical Analysis, vol. IV, Ciarlet, P.G. and Lions, J.L. eds., North-Holland, Amsterdam (1996), 313-485.
- [25] HASLINGER J., JANOVSKÝ V., KUČERA R., MOTYČKOVÁ K.: *Nonsmooth continuation of parameter dependent static contact problems with Coulomb friction*, přijato v Math. Comput. Simul. (2016).
- [26] HASLINGER J., KOZUBEK T., KUČERA R.: *On a fictitious domain method for unilateral problems*, In: Numerical Mathematics and Advanced Applications (EUMATH 2007), eds.: K. Kunisch, G. Of, O. Steinbach, Springer 2008, 803-810.
- [27] HASLINGER J., KOZUBEK T., KUČERA R.: *Fictitious domain formulation of unilateral problems: analysis and algorithms*, Computing, 84(2009), 1, 69-96.
- [28] HASLINGER J., KOZUBEK T., KUČERA R.: *Fictitious domain method for linear elasticity*, In: Proceedings SNA'09, Ostrava, February 2009, 39-42.
- [29] HASLINGER J., KOZUBEK T., KUČERA R., PEICHL G.: *Projected Schur complement method for solving non-symmetric saddle-point systems arising from fictitious domain approach*, Num. Lin. Algebra Appl., 14(2007), 9, 713-739.
- [30] HASLINGER J., KUČERA R., DOSTÁL Z.: *An algorithm for the numerical realization of 3D contact problems with Coulomb friction*, J. Comput. Appl. Math., 164-165(2004), 387-408.
- [31] HASLINGER J., KUČERA R., KOZUBEK T.: *Convex programming with separable ellipsoidal constraints: application in contact problems with orthotropic friction*, Optimization with PDE Constraints, ESF Networking Program 'OPTPDE', Springer Series: LNCSE, Vol. 101, Hoppe, R.(Ed.), 2014, XII, 221-242.

- [32] HASLINGER J., KUČERA R., LIGURSKÝ T.: *Qualitative analysis of 3D elastostatic contact problems with orthotropic Coulomb friction and solution-dependent coefficients of friction*, J. Comput. Appl. Math., 235(2011), 12, 3464-3480.
- [33] HEYN T., ANITESCU M., TASORA A., NEGRUT D.: *Using Krylov subspace and spectral methods for solving complementarity problems in many-body contact dynamics simulation*, Internat. J. Numeric. Meth. Engrg., 95(2013), 7, 541-561.
- [34] HINTERMÜLLER M., ITO K., KUNISCH K.: *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim., 13(2003), 3, 865-888.
- [35] CHEN X., NASHED Z., QI L.: *Smoothing methods and semismooth methods for non-differentiable operator equations*, SIAM J. Numer. Anal., 38(2000), 4, 1200-1216.
- [36] KOKO J.: *Vectorized Matlab codes for the Stokes problem with P1-bubble/P1 finite element*, at: <http://www.isima.fr/~jkoko/Codes.html>, 2012.
- [37] KUČERA R.: *Interpolating and smoothing biquadratic spline*, Appl. Math., 40(1995), 5, 339-356.
- [38] KUČERA R.: *Complexity of an algorithm for solving saddle-point systems with singular blocks arising in wavelet-Galerkin discretizations*, Appl. Math., 50(2005), 3, 291-308.
- [39] KUČERA R.: *Minimizing quadratic functions with separable quadratic constraints*, Optim. Meth. Soft., 22(2007), 3, 453-467.
- [40] KUČERA R.: *Convergence rate of an optimization algorithm for minimizing quadratic functions with separable convex constraints*, SIAM J. Optim., 19(2008), 2, 846-862.
- [41] KUČERA R., HASLINGER J., ŠÁTEK V., JAROŠOVÁ M.: *Efficient methods for solving the Stokes problem with slip boundary conditions*, přijato v Mathematics and Computers in Simulation (2016).
- [42] KUČERA R., KOZUBEK T., MARKOPOULOS A.: *On large-scale generalized inverses in solving two-by-two block linear systems*, Lin. Algebra Appl., 438(2013), 7, 3011-3029.
- [43] KUČERA R., KOZUBEK T., MARKOPOULOS A., HASLINGER J., MOCEK L.: *Projected Krylov methods for solving non-symmetric two-by-two block linear systems arising from fictitious domain formulations*, AEEE, 12(2014), 2, 131-143.
- [44] KUČERA R., KOZUBEK T., MARKOPOULOS A., MACHALOVÁ J.: *On the Moore-Penrose inverse in solving saddle-point systems with singular diagonal blocks*, Num. Lin. Algebra Appl., 19(2012), 4, 677-699.
- [45] KUČERA R., MACHALOVÁ J., NETUKA H., ŽENČÁK P.: *An interior point algorithm for the minimization arising from 3D contact problems with friction*, Optim. Meth. Soft., 28(2013), 6, 1195-1217.
- [46] KUČERA R., MOTYČKOVÁ K., MARKOPOULOS A.: *The R-linear convergence rate of an algorithm arising from the semi-smooth Newton method applied to 2D contact problems with friction*, Comput. Optim. Appl., 2(2015), 61, 437-461.
- [47] KUČERA R., MOTYČKOVÁ K., MARKOPOULOS A.: *An analysis of the semi-smooth Newton method for solving 3D contact problems with the Tresca friction*, v přípravě (2016).

- [48] **KUČERA R.**, ŠÁTEK V., HASLINGER J., FIALOVÁ S., POCHYLÝ F.: *Modelling of hydrophobic surfaces by the Stokes problem with the stick-slip boundary conditions*, přijato v ASME. J. Fluids Eng. (2016), doi:10.1115/1.4034199.
- [49] **KUČERA R.**, VLČEK J., VLČEK K.: *DSP implementation of image compression by multiresolutional analysis*, Radioengineering, 7(1998), 1, 7-9.
- [50] LIGURSKÝ T., HASLINGER J., **KUČERA R.**: *Approximation and numerical realization of 3D contact problems with Coulomb friction and a solution-dependent coefficient of friction*, Internat. J. Numer. Meth. Engrg., 82(2010), 9, 1180-1206.
- [51] MatSol na <http://industry.it4i.cz/produkty/matsol/>
- [52] NAVIER C. L. M. H.: *Mémoire sur les lois du mouvement des fluides*, Mém. de lâ Acad. R. Sci. Paris, 6(1823), 389-416.
- [53] NOCEDAL J., WÄCHTER A., WALTZ R. A.: *Adaptive barrier strategies for nonlinear interior methods*, TR RC 23563, IBM T.J. Watson Research Center, 2005.
- [54] NOCEDAL J., WRIGHT S. J.: *Numerical optimization*, Springer-Verlag, New York, 1999.
- [55] POSPÍŠIL L.: *Development of algorithms for solving minimizing problems with convex quadratic function on special convex sets and applications*, Disertace, VŠB–TUO, 2015.
- [56] WOHLMUTH B. I.: *Variationally consistent discretization schemes and numerical algorithms for contact problems*, Acta Numerica, 20(2011), 569-734.
- [57] WRIGHT S. J.: *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia 1997.

Abstract

In last decades, the number of numerical methods for the solution of contact problems with different friction laws have been developed. It includes, among others, active set strategies, Newton-type methods, or interior point methods. Algorithms based on the mentioned ideas are discussed in the theses, especially for contact problems in 3D. In the case of the Tresca friction law, the dual algebraic problem arising from the finite element approximation leads to the minimization of a strictly quadratic cost function subject to separable convex constraints, spherical for isotropic surfaces or elliptical for orthotropic ones. The active set algorithm is based on an geometrical interpretation of optimality conditions showing how to combine conjugate gradient method with reduced gradient steps. The algorithm enjoys the R-linear convergence rate as for box constraints. Its typical feature consists in the fact that the unknowns corresponding to the active set are "dead" during one continuous conjugate gradient loop. The interior point algorithm is principally different. It keeps all iterations inside the feasible region so that no constraint is active during iterative computations. In other word, all unknowns are "always living" that may positively influence computations although theoretical results are not very strong. The last algorithm discussed in the thesis uses the semi-smooth Newton method that enables us to treat non-differentiability of contact problems by the slant differentiability concept. The inexact implementation based again on an active set strategy is closely related to the previous active set algorithm including the R-linear convergence rate.